

Parameter estimation of a predator-prey model using a genetic algorithm

Jackelyne Gómez Restrepo
Ingeniería Matemática
Universidad EAFIT
Medellín, Colombia
jgomezr7@eafit.edu.co

Carlos Mario Vélez Sánchez
Ingeniería Matemática
Universidad EAFIT
Medellín, Colombia
cmvelez@eafit.edu.co

Abstract— This paper presents the parameter estimation of a biological system with real data extracted from literature, and different model structures: Lotka-Volterra model (basic predator-prey model) and saturation predator-prey model. A third model is proposed and tested in simulation due to lack of appropriated real data; it includes inputs that excite the system and makes the estimation process more manageable. The model parameters were estimated using a genetic algorithm, which gives a combination of parameters used in simulation to compare outputs with real data and decide, using a cost function, which parameters are better. Comparing the models, the Lotka-Volterra model provides better adjustment but with unrealistic assumptions, while saturation model represents a system with real assumptions but the fit is not very high.

Keywords- predator-prey model; parameter estimation; genetic algorithms.

I. INTRODUCTION

Parameter estimation is a useful method to obtain the parameters of a mathematical model through knowledge of data on a given system [1]. It focuses on adjusting the model parameters to get a final model which delivers outputs similar to the observed data; however, in many cases the process can gain complexity in the presence of disturbances or badly taken data. The application to a nonlinear system, as will be done in this paper, is a bit more complicated than doing so with linear systems, hence heuristic methods are useful in these situations [2].

Mathematical modeling tools can be applied to problems of any topic; for example, sciences like biology are full with hundreds of interesting problems to be modeled [3], like the predator-prey model, which has been a topic of much research because of its characteristics and flexibility [4].

In this paper parameter estimation for three proposed predator-prey models is performed with genetic algorithms for obtaining outputs similar to real data. One of the models includes inputs which excite the system; this variation in the Lotka-Volterra model assumes two inputs which adds and subtracts prey or predators in some intervals of time.

The paper is organized as follows. Section two explains the Lotka-Volterra and saturation models from a mathematical point of view; section three describes the genetic algorithm and its implementation; section four shows the parameter estimation process and results; and section five presents a summary with some conclusions and future work.

II. PREDATOR-PREY MODELS

A. The Lotka-Volterra Model

The predator-prey model is a representation of the interaction between two species of animals that live in the same environment, and the quantity of each group of animals depends on two things: the birth or death rate and the successful meetings [5]. This model is based on important assumptions which restrict actual conditions in the system [6]:

- Prey has innumerable quantity of food
- Predators eat prey
- Both predators and prey have meetings that are proportional to the product between both populations

The quantity of species in each group is represented with time-dependent variables: $P(t)$ for prey and $D(t)$ for predators. Equations (1) and (2) represent this system and Fig. 1 implements them in simulation including inputs (see section IV-C).

$$\frac{dP}{dt} = \alpha P - \beta PD \quad (1)$$

$$\frac{dD}{dt} = -\gamma D + \delta PD \quad (2)$$

Parameters α , β , γ and δ are positive constants. Parameters α and β indicate the birth rate of prey and death rate of predators respectively, and γ and δ represent the growth and decrease of each population. This model has a major impact in several areas, but it is imprecise because of its simplicity. Because of that, more realistic models have been proposed to limit the

number of species, saturate the populations or include the hunting times.

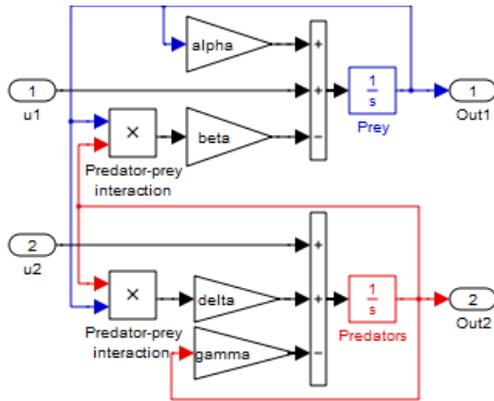


Figure 1. Predator-Prey Model

B. Saturation Model

One of the changes of Lotka-Volterra model is the addition of mathematical expressions that saturate the number of prey [7], as shown in (3) and (4).

$$\frac{dP}{dt} = rP \left(1 - \frac{P}{K} \right) - aPD \quad (3)$$

$$\frac{dD}{dt} = -cD + dPD \quad (4)$$

Parameter r is the growth rate of prey, c is the mortality rate of predators, d describes the efficiency of the predator hunting, while a measures the inefficiency of the prey population and K is the limit for prey. This model is more realistic than the original model because prey have not an exponential growth (which is an assumption of the Lotka-Volterra model), but it is not totally adjusted to the real world, mainly because its parameters are constants all the time and it does not add hunting times as other models.

III. GENETIC ALGORITHMS

The parameter estimation in this paper is performed with a genetic algorithm (GA) which searches system parameters in a way that the fit of the outputs obtained with the model were near to 100%. Genetic algorithms are adaptive heuristics based on learning; they work on a population of individuals which represents the search space [8]. In each cycle (equivalent to one generation) a new set of solutions is created through the crossing of certain parents in the current population; in many cases the children acquire new features because of their parents or their probability of mutation, which allows variability and thus the algorithm does not remain in local optimums. These algorithms attempt to replicate natural behaviors where stronger individuals, with greater ease of adaptation, survive, so that the final set of solutions contains the best solutions. However, GA cannot find the optimal

solution but it can find feasible solutions in short times. The implemented algorithm is:

```

Read Number_of_generations (iterations)
Generate solution space
While n<= Number_of_generatios
    Evaluate each possible solution in the solution space
    Choose the k best solutions and put them as parents
    For i=1:k/2
        Cross parents (Two children are obtained in each crossing)
    End_For
    For j=1:k
        Mutate children with a probability
    End_For
    Evaluate each mutated children
    Remove the k worst solutions from the solution space and add the
    k mutated children
    n=n+1
End_While
Choose the best solution
    
```

The implemented GA chooses the parents in a deterministic way, which means that the chosen parents are the best solutions in the current space; later, those parents are crossed by one point, one child is going to have a part from his father and the other one from his mother, and the other child is going to have the remaining parts. For getting variability, a probability of mutation is assigned to each child: it means that any child can change some of its characteristics. Then, a fixed number of those children will be introduced into the search space. This algorithm stops until the maximum number of iterations is reached. Finally, the best solution is extracted from the last search space.

For choosing the best solution with the least possible error, the algorithm calculates a cost function which determines the magnitude of the prediction errors or, which is the same, the difference between the outputs from the estimated model and the real data. The cost function for estimation in the present work is:

$$V = \frac{1}{2} \sum_{i=1}^n \left[(y_{1i} - \hat{y}_{1i})^2 + (y_{2i} - \hat{y}_{2i})^2 \right] \quad (5)$$

Variables y_{1i} and y_{2i} represent the real data (prey and predators, respectively), and \hat{y}_{1i} and \hat{y}_{2i} represent the output from the identified system (prey and predators, respectively).

IV. PARAMETER ESTIMATION

In this chapter parameters of three models are estimated using the same data: rabbits (prey) and lynxes (predators). The information was extracted from a research made by Hudson Bay Company [9].

A. Parameter estimation with a Lotka-Volterra Model

The first proposed model is the Lotka-Volterra model which has an output fit of 76.99% for the prey and 75.15% for predators. This result is considered a good solution because the fit is better than the results with analytical methods: solution of differential equations and approximations using growth and death assumptions like exponential behaviors.

These outputs were the product of 150 generations (iterations) with 80 possible solutions in the search space and 70 parents; the children had a probability of mutation of 0.6.

Fig. 2 shows the behavior of each group of species: the continue graphs are the estimated populations and the points are the real data; the estimated outputs describe a behavior similar to the real outputs, which means that the model can represent the biological system even when the Lotka-Volterra model works with unreal assumptions.

Table I compares the parameter values obtained in the estimation and the values obtained with approximations using differential equations (analytical methods).

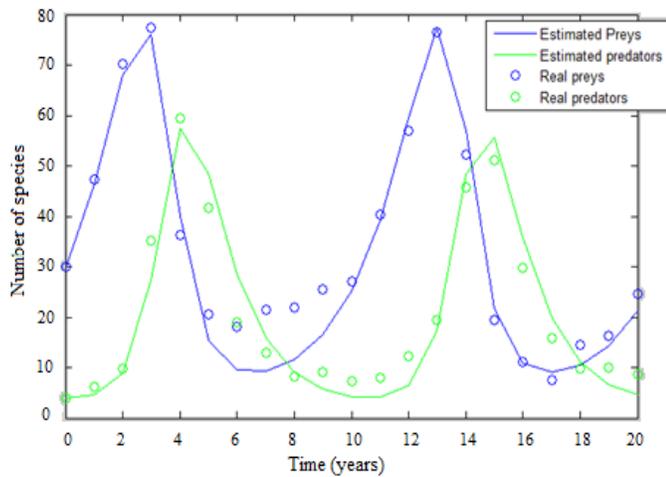


Figure 2. Real and estimated outputs

TABLE I. ESTIMATED PARAMETERS WITH A LOTKA-VOLTERRA MODEL

Parameters	Results	
	Results with GA	Results with analytical methods
α	0.55	0.4
β	0.027	0.018
γ	0.83	0.8
δ	0.026	0.023
Fit of prey	76.99%	-32.51%
Fit of predators	75.15%	-59.71%

B. Parameter estimation with a saturation model

In the parameter estimation of the saturation predator-prey model, the output fit for prey is 71.14% and 64.34% for predators. For this model the fit is lower than previous because

of the complexity of the model when the number of prey is assumed as finite; when the number of parameters increases, the estimation process becomes more complex. Both models can describe the same behavior of the real data; however, each model has different assumptions and number of parameters, so that it is possible to find different and consistent models which are able to estimate the parameters of same system.

The parameter estimation with a GA was made using 60 generations (iterations), 80 possible solutions in the search space, 70 parents, and a probability of mutation of 0.6 for their children. Fig. 3 shows the behavior of each group of species. Continue line is the estimated model and the points are the real data.

According to the results in Table I and Table II, the Lotka-Volterra model has a better adjustment than complex models like the saturation model, which describes more realistic systems with a good adjustment.

C. Including inputs into the model

The parameter estimation with systems that include input data should be easier and better as the inputs would excite the system persistently [10] and it could be manipulated through them. However, such data were not found in the literature and test with inputs was not possible in this work.

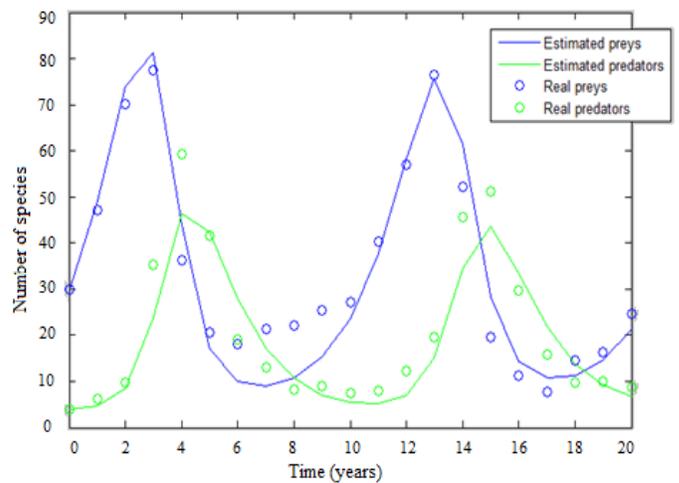


Figure 3. Real and Estimated Model

TABLE II. ESTIMATED PARAMETERS WITH A SATURATION MODEL

Parameters	Results with GA
α	0.65
β	0.034
γ	0.68
δ	0.021
K	71247
Fit of prey	71.14%
Fit of predators	64.34%

An input (u_1 or u_2) in the predator-prey model is represented by a pulse signal which indicates the rate of addition or subtraction of species in some time interval: if the input is a positive number, the system is acquiring prey or predators; if the input is a negative number, the system is losing members. For example, a pulse input of magnitude N in one time unit adds N species in the population in that time; if different pulse signals are used as inputs, the number of species in the populations would change constantly (see Fig. 4). On the other hand, systems with inputs can be controlled [11]; implementing a control for a predator-prey system would keep stable each group of species and check extreme changes in the population of the ecosystem which may lead to overpopulation or under population.

Equations (5) and (6) show the model with inputs:

$$\frac{dP}{dt} = \alpha P - \beta PD + u_1 \quad (5)$$

$$\frac{dD}{dt} = -cD + dPD + u_2 \quad (6)$$

Parameter estimation with inputs was made in simulation using the algorithm implemented over last models, and the fit was as good as in the other cases (Table III). The favorable results from the simulation set up a new and interesting perspective in biological field.

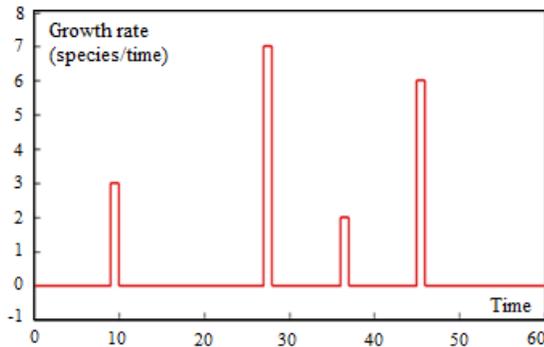


Figure 4. Input shape in predator-prey model

TABLE III. ESTIMATED PARAMETERS IN SIMULATION TESTS

Parameters	Results with GA
α	0.61
β	0.030
γ	0.79
δ	0.025
Fit of prey	71.06%
Fit of predators	78.88%

V. CONCLUSIONS

The parameter estimation using genetic algorithms with a Lotka-Volterra model has a favorable fit but it requires a large number of iterations (generations) to find the best solution. The saturation model shows a lower fit but in a smaller number of iterations. Comparing both models, the Lotka-Volterra model provides better fit but with unrealistic assumptions, while saturation model represents a system with real assumptions but its fit is not very high; the estimation can be improved with more time consuming.

In review, both models reflect the same real behavior between two species, but they deliver different parameters values because they are built with different assumptions. However, those results can be improved if the GA acquires complexity in some of its processes or even if other predator-prey models are used in the estimation.

Although the parameter estimation with a model that includes inputs was made in a simulation context, it is a interesting possibility for future researches, where the biological system can be manipulated with persistently excited inputs or controlled to keep a stable system. An interesting application of estimation with a model with inputs is the estimation of pulse parameters (amplitude and width), which can help to identify the time and magnitude of a change in population due to an external event; that is, identify input shape in Fig. 4 given a set of output data.

REFERENCES

- [1] C. Jost, and R. Arditi, "From pattern to process: Identifying predator-prey models from time-series data", *Population Ecology*, Vol. 43, No. 43, pp 229-243, 2001.
- [2] M. Mitchell, "Genetic Algorithms: An Overview", *Complexity*, Vol.1, No. 1, pp 31-39, 1995
- [3] C. Jeffries, *Mathematical modeling in Ecology*. Boston: Birkhäuser, 1989.
- [4] C. Katz, "A Nonequilibrium Marine Predator-Prey Interaction", *Ecology*, Vol. 66, No. 5, pp 1426-1438, 1985.
- [5] D. Griffiths, "Prey availability and the food of predators", *Ecology*, Vol. 56, No. 5, pp 1209-1214, 1975.
- [6] J. Stewart, *Calculus - Concepts and Contexts*. Fourth edition. International Thomson Editors, 1998.
- [7] B. McArdle, and J. Lutton, "Effects of Prey-size and Predator-instar on the Predation of *Daphnia* by *Notonecta*", *Ecological Entomology*, No. 4, pp 267-275, 1979.
- [8] S. Sait, and H. Youssef, *Iterative Computer Algorithms with Applications in Engineering*. Los Alamitos: IEEE Computer Society Press, pp 53-59, 1999
- [9] E. R. Leigh. "The ecological role of Volterra's equations, in *Some Mathematical Problems in Biology - a modern discussion using Hudson's Bay Company data on lynx and hares in Canada from 1847 to 1903*". 1968
- [10] L. Ljung, *System Identification Theory for the User*. Second edition. New Jersey: Prentice Hall, pp 412-414, 1999.
- [11] B. Kuo, *Automatic Control Systems*. Fourth edition. New Jersey: Prentice Hall, 1996.