Article

Controlling ecological and biological networks via evolutionary modelling

Alessandro Ferrarini

Department of Evolutionary and Functional Biology, University of Parma, Via G. Saragat 4, I-43100 Parma, Italy
E-mail: sgtpm@libero.it, alessandro.ferrarini@unipr.it

Abstract

The controllability of network-like systems is a topical key-issue in many disciplines, including ecology and biology. It relies on the ability to guide a system's behaviour towards the desired state through the appropriate handling of a few input variables. To date, controllability of networks is based on the identification of the set of driver nodes that can guide the system's dynamics. I introduce here a new framework for the controllability of both network nodes and edges based on the use of evolutionary modelling, and provide an exemplification of its application.

Keywords edges control; genetic algorithms; network control; nodes control.

1 Introduction

Network controllability is the ability to guide a system's behaviour towards the desired state through the appropriate handling of a few input variables (Luenberger, 1979; Slotine and Li, 1991). The controllability of network-like systems is becoming a trendy key-issue in many disciplines, including ecology (Caldarelli, 2007) and biology (Dorogovtsev and Mendes, 2003; Kim and Motter, 2009; Marucci et al., 2009).

In Nature, Liu et al. (2011) proposed analytical tools for the controllability of complex networks, identifying the set of driver nodes that can guide the system's dynamics. Their way to the controllability of networks is based on the identification of the set of driver nodes that can guide the system's dynamics, in other words on the choice of a subset of nodes that are selected to be permanently controlled. This assumption seems motivated by real-world networks observation, where a decentralized control action is often applied only to part of the nodes.

In 2011, I remarked the need for five key-improvements to this usual approach in networks controllability (Ferrarini, 2011). In particular, I noticed that several tools, like genetic optimization, could reasonably lower the seemingly intractable problem of edges control.

Hence, I introduce here a new framework for network controllability based on the use of evolutionary

modelling, and provide an exemplification of its application.

2 Mathematical Formulation

An ecological (or biological) dynamical system of n interacting species is as follows:

$$\frac{d\vec{S}}{dt} = \varphi(\vec{S}, t)$$

(1)

where S_i is the number of individuals (or the total biomass) of the generic i-th species. If we also consider inputs (e.g. species reintroductions) and outputs (e.g. hunting or road mortality) from outer universe, we must write:

$$\frac{d\vec{S}}{dt} = \varphi(\vec{S}, t) + \vec{I}(t) + \vec{O}(t)$$

(2)

As noted by numerous authors (e.g., Liu et al., 2011; Slotine and Li, 1991) most real systems' dynamics can be modelled and simulated using a system of canonical, linear equations, as follows:

$$\begin{cases} \frac{dS_1}{dt} = a_{11}S_1 + ... + a_{1n}S_n + I_1 + O_1 \\ ... \\ \frac{dS_n}{dt} = a_{n1}S_1 + ... + a_{nn}S_n + I_n + O_n \end{cases}$$

(3)

that can be written in the compact form:

$$\frac{d\vec{S}}{dt} = A*\vec{S} + \vec{I} + \vec{O}$$

(4)

with initial values

$$\vec{S}_0 = <S_1(0), S_2(0)...S_n(0)>$$

(5)

and co-domain limits

$$\begin{cases} S_{1min} \leq S_1(t) \leq S_{1max} \\ ... \\ S_{nmin} \leq S_n(t) \leq S_{nmax} \end{cases}$$

(6)

and where

$$A = \begin{pmatrix} a_{11} & ... & a_{1n} \\ ... & ... & ... \\ a_{n1} & ... & a_{nn} \end{pmatrix}$$

(7)

is the matrix of the per unit-of-time effect on S_i due to unitary S_j.

3 Network Control via Evolutionary Modelling

Genetic algorithms (GAs) consist of optimization procedures based on principles inspired by natural selection. GAs involves 'chromosomal' representations of proposed problem solutions which undergo genetic operations

such as selection, crossover and mutation (Holland, 1975; Goldberg, 1989, Parolo et al., 2009; Ferrarini, 2012).
Under genetic optimization, equation (3) becomes:

$$\begin{cases} \left(\dfrac{dS_1}{dt}\right)_{OPT} = a_{11*}S_1^{\,*} + ... + a_{1n*}S_n^{\,*} + I_{1*} + O_{1*} \\ ... \\ \left(\dfrac{dS_n}{dt}\right)_{OPT} = a_{n1*}S_1^{\,*} + ... + a_{nn*}S_n^{\,*} + I_{n*} + O_{n*} \end{cases} \qquad (8)$$

where asterisks stand for the optimization of edges (i.e., coefficients of interaction among variables) or nodes (i.e., initial stocks), that is the modification of their values at the beginning of the network dynamics in order to get a certain goal (e.g., maximization of the final value of a certain variable).

4 An Applicative Example

Fig. 1 depicts an ecological network borrowed with modification from Ferrarini (2012b). Greenish nodes represent positive actors or events for the goal of the network control, i.e. the increase of individuals of the target species (centre of the network). Reddish nodes represent ecological actors or events with negative impact on the target species. Blueish nodes represent resources needed by the target species. The goal is to preserve target species' occurrence in the study area. Stocks stand for the actual amounts of individuals or biomass. Updates stand for yearly internal dynamics (i.e., intraspecific gains due to births and/or immigration rates minus losses due to deaths and/or emigration rates). Minimum and maximum values stand for lowest and highest values of stock values. For the sake of simplicity, the maximum possible value for each actor (in italic hereafter) has been set to 100. The percent value associated to links represent the percentage of the receiver that is yearly consumed by the transmitter at the beginning of the network simulation. Road mortality and re-introductions account for 18 and 10 individuals per year respectively.
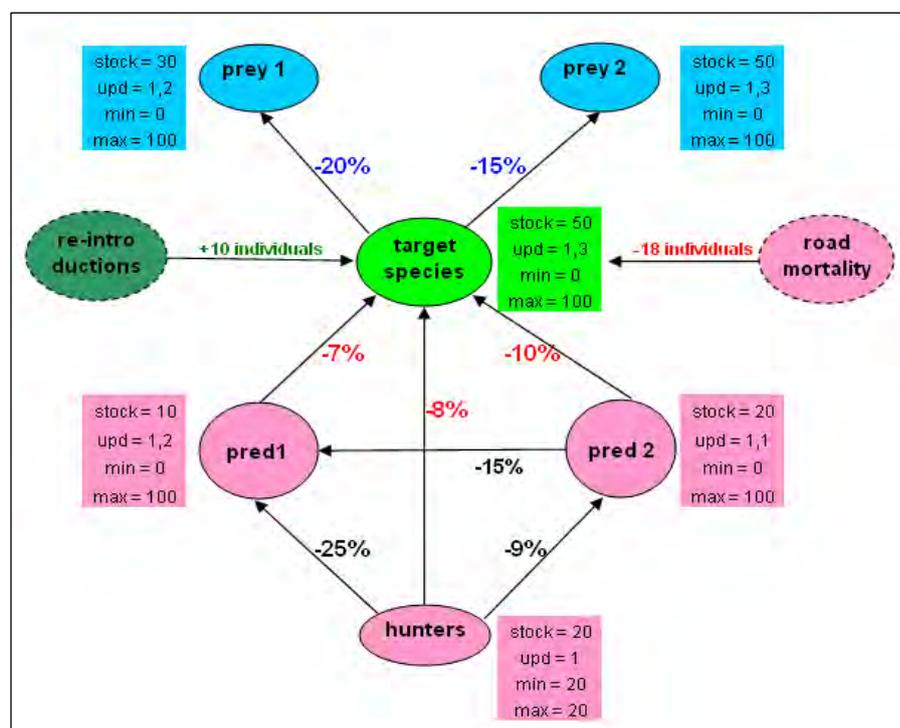
Fig. 1 The ecological network on which evolutionary control has been applied.

Since data are yearly-based, I express equation (8) using a system of difference recurrent equations, instead of differential ones:

$$
\begin{cases}
\left(S_1\right)_{t+1\,OPT} = a_{11*}\left(S_1\right)_t + ... + a_{1n*}\left(S_n\right)_t + \left(I_1\right)_{t*} + \left(O_1\right)_{t*} \\
... \\
\left(S_n\right)_{t+1\,OPT} = a_{n1*}\left(S_1\right)_t + ... + a_{nn*}\left(S_n\right)_t + \left(I_n\right)_{t*} + \left(O_n\right)_{t*}
\end{cases}
\tag{9}
$$

For instance, *prey1*'s dynamic (Fig. 2) is given by:

$$
\frac{d(prey1)}{dt} = +0.2 * prey1 + a * target\_species
\tag{10}
$$

where $a = \dfrac{-20}{100} * \dfrac{30}{50} = -0.12$ (11)

or, expressed using a difference equation:

$$
\begin{cases}
prey1_{t+1} = +0.2*prey1_t - 0.12*target\_species_t \\
prey1(0) = 30
\end{cases}
\tag{12}
$$

INTERACTIONS MATRIX (per-cycle changes to receivers per-unit of transmitters)										
		Receivers								
		target	prey1	prey2	pred1	pred2	hunters	var G	var H	var I	var J
Transmitters	target	1.3	-0.12	-0.15	0	0	0	0	0	0	0
	prey1	0	1.2	0	0	0	0	0	0	0	0
	prey2	0	0	1.3	0	0	0	0	0	0	0
	pred1	-0.35	0	0	1.2	0	0	0	0	0	0
	pred2	-0.2	0	0	-0.075	1.1	0	0	0	0	0
	hunters	-0.2	0	0	-0.125	-0.09	1	0	0	0	0
	var G	0	0	0	0	0	0	0	0	0	0
	var H	0	0	0	0	0	0	0	0	0	0
	var I	0	0	0	0	0	0	0	0	0	0
	var J	0	0	0	0	0	0	0	0	0	0

Fig. 2 The interaction matrix relative to the ecological network of Fig. 1. The matrix has been calculated using Quant-Lab (Ferrarini, in preparation).

The previous ecological network has the following inertial dynamics (Fig. 3), with the *target species* (green line) going extinct after 7 years.

5 Network Control Using GAs

Since GAs are based on random searching for solutions, I performed iterative processes for determining appropriate parameters. Previous research revealed that the optimal solution may be to search at a high rate of crossover, a low rate of mutation and proper population size (Kuo et al., 2000). In this study, crossover was set at a probability of 60% while mutations occur with a probability of 5%. This low setting helps to avoid getting trapped local optima during the search (D'heygere et al., 2006). The initial population consisted of 500 chromosomes that were evolved over minimal 10,000 generations. These parameters were set after preliminary experiments. I applied a steady-state genetic algorithm with a one-point crossover operator (Wall, 1996) to accomplish crossover. In this case the parent genome strings are cut at some random position to produce two "head" and two "tail" segments. The "tail" segments are swapped to produce two new genomes. For parent selection the roulette wheel selection method was used (Goldberg, 1989), where the likelihood of selection is

proportionate to the fitness score given by the performance criterion. After crossover and mutation, the individuals with the lowest fitness scores were removed. Hence, I was able to find a set of feasible solutions to the goal of preserving the *target species*.
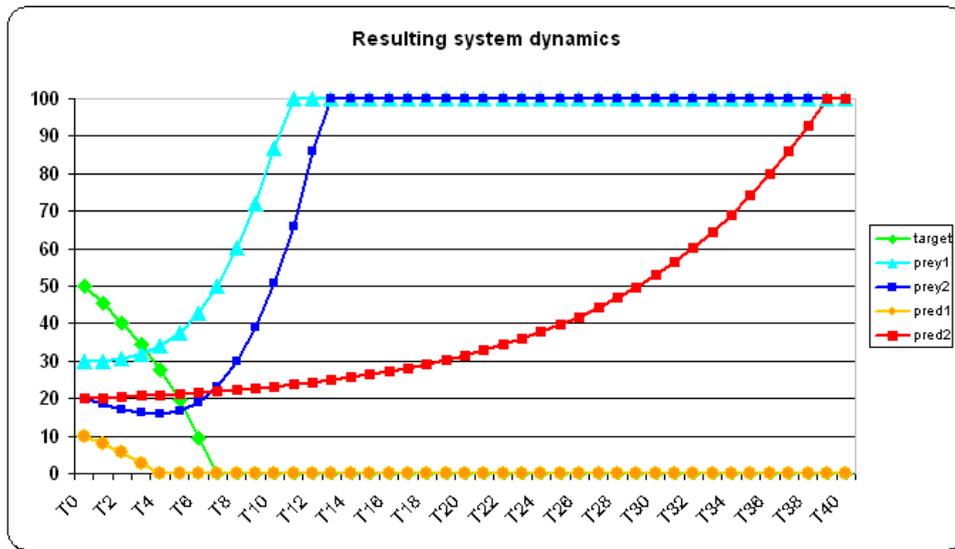
Fig. 3 Resulting dynamics for the network of Fig. 1. X-axis measures time in years. Dynamics have been calculated using Quant-Lab (Ferrarini, in preparation).

Being all other simulation parameters equal, by imposing *pred2*'s initial stock to be <8 (instead of 20), I achieved the following dynamics (Fig. 4). Both *target species* and *prey1* are stably safe.
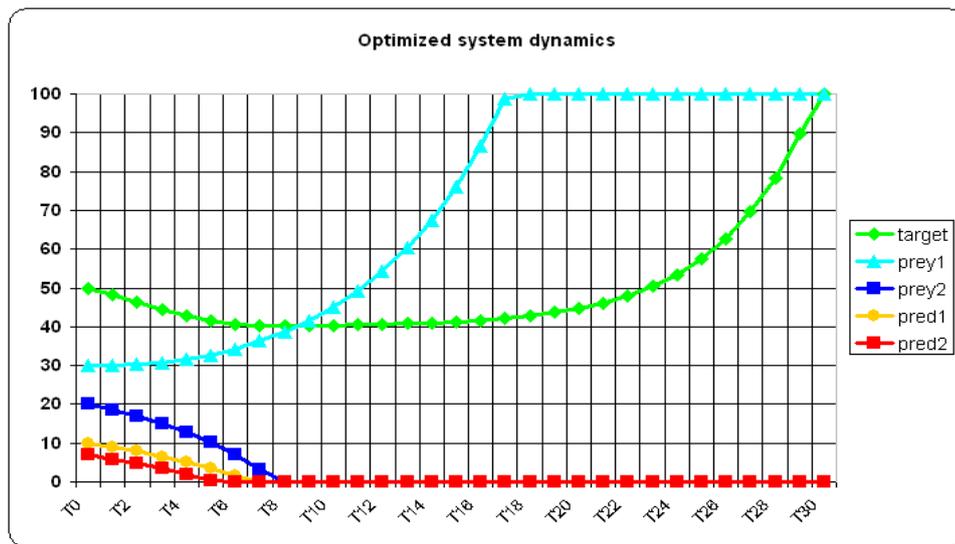
Fig. 4 Solution 1: acting on *pred2*'s initial stock. It is set here to 7 instead of 20. Dynamics have been calculated using Quant-Lab (Ferrarini, in preparation).

Interestingly, with *pred2*'s initial stock ≤5 the stocks of both preys (*prey1* and *prey2*) would go to 0 (Fig. 5), thus only solutions with 6 ≤ *prey2*'s stock <8 are feasible for long-term conservation of the *target species*. Otherwise, a successive prey's re-stocking would be needed.
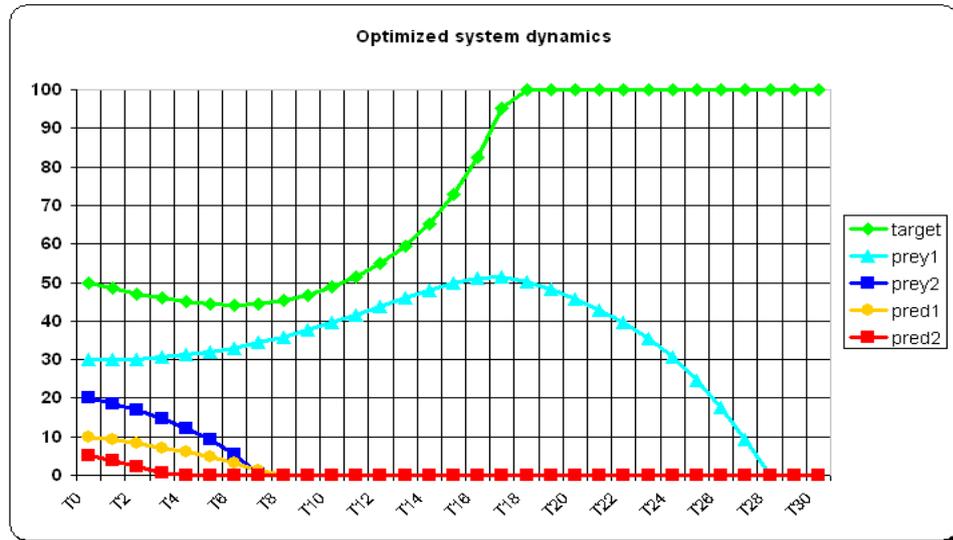
Fig. 5 Resulting dynamics with *prey2*'s initial stock set to 5.

Another solution is possible by increasing *target species*' intraspecific coefficient above the value 1.36 (instead of the actual value 1.3; Fig. 6). This would also require *prey1*'s (or *prey2*) re-stocking after 28 years in order to preserve some prey for the *target species*.
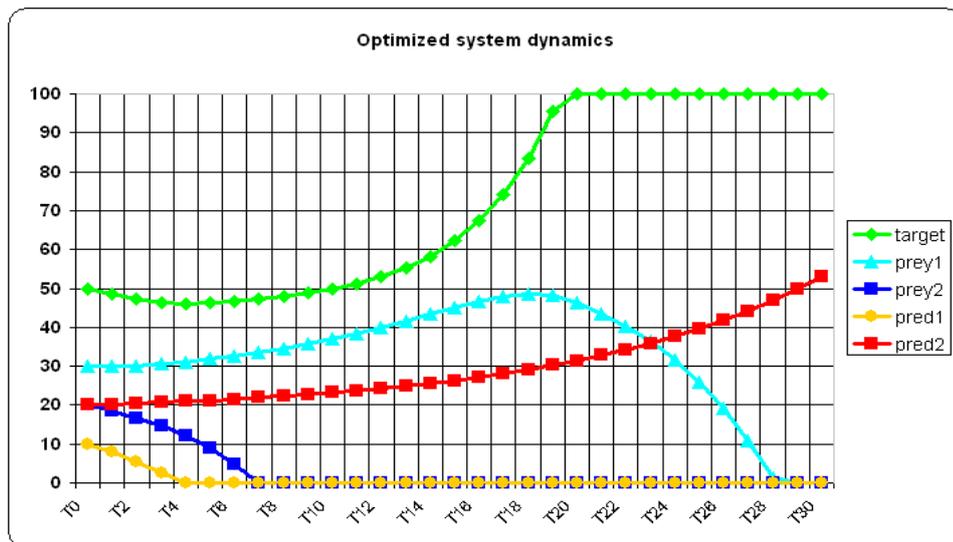
Fig. 6 Solution 2. Resulting dynamics with *target species*' intraspecific coefficient set to 1.361.

A further solution is possible by increasing input-output balance (*reintroductions* and *road mortality*). Actually it's equal to -8 in the study area (10 reintroductions and 18 road casualties per year), but if it were >-6 it would determine the following dynamics (Fig. 7). This could be achieved by increasing reintroduction activities, or by limiting road casualties with proper road signals. *Prey1*'s restocking would be also required after 21 years.
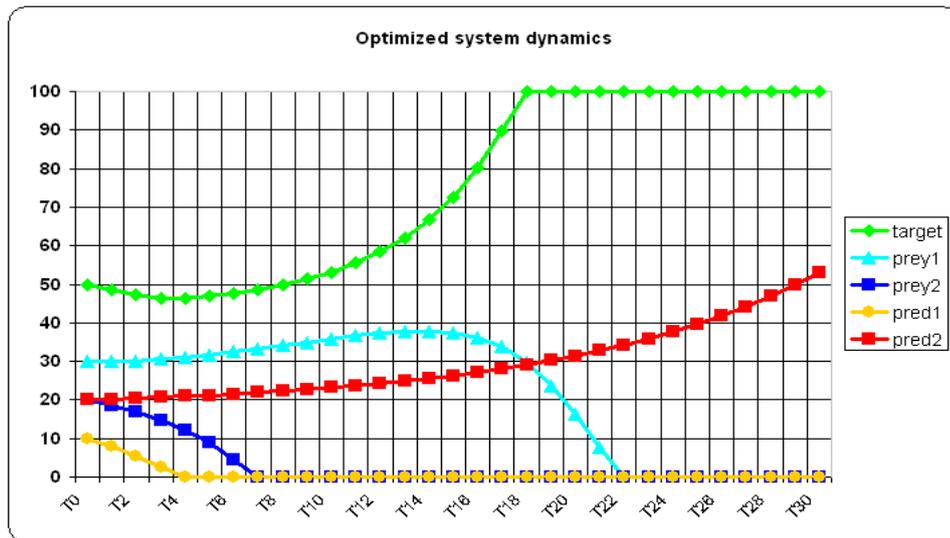
Fig. 7 Solution 3. Resulting dynamics with input-output balance set to -5.

Another solution is given by yearly subtraction of *pred1* and *pred2's* individuals. At least 5 individuals of both species must be yearly removed in order to preserve the *target species* (Fig. 8).
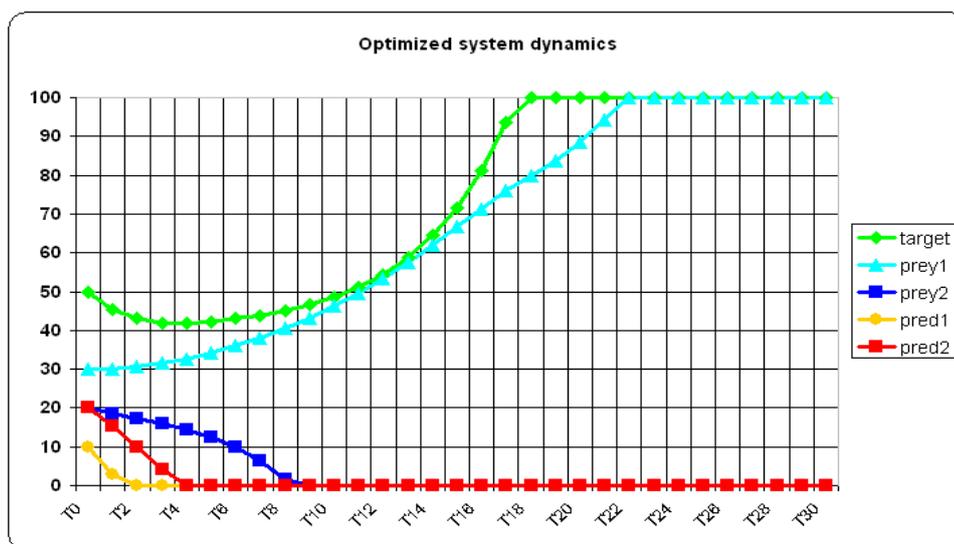
Fig. 8 Solution 4. Resulting dynamics with yearly removal of *pred1* and *pred2's* individuals (5 individuals of both species).

Interestingly, the removal of *hunters* (Fig. 9) at the beginning of the simulation (i.e., *hunters*(0)=0) doesn't allow to achieve the proposed goal. In fact, although in the study area hunters kill *target species*' individuals, they also subtract individuals to the two species of predators.
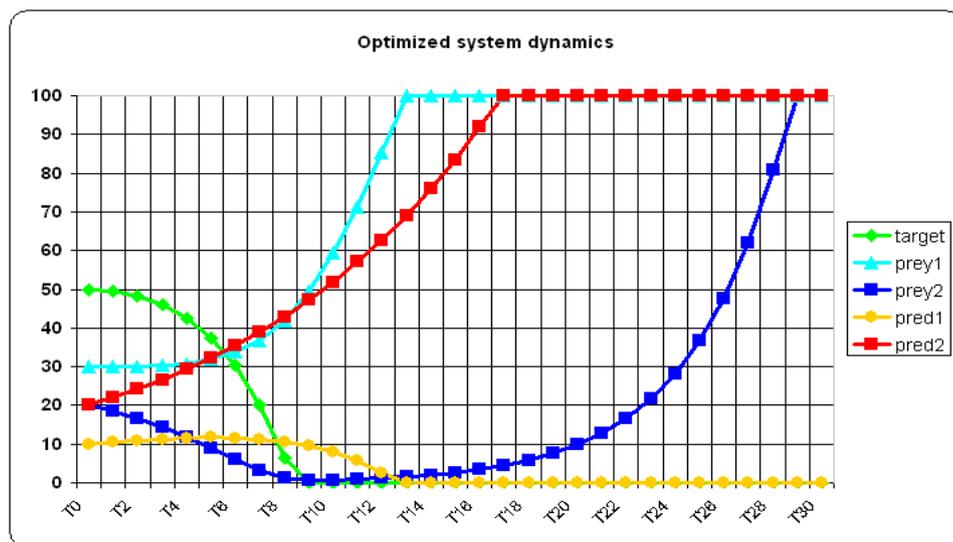
Fig. 9 Resulting dynamics achieved by removing hunters at T=0 (i.e., year 2012).

Since many alternative solutions can be found by applying evolutionary modelling to the control of ecological and biological networks, I also propose the following decision framework:

- goal setting (e.g. *target species* maximization);
- network setup;
- application of GAs to network dynamics;
- multiple solutions detection via evolutionary modelling;
- setup of cost-benefit ratio (CBR) for each detected solution;
- choice of the action (or actions) that minimize CBR.

A step forward could be given by *a priori* incorporating CBRs into goal setting, i.e. the goal of the network control would be to minimize CBR while maximizing *target species*' stock.

6 Conclusions

To date, the controllability of ecological and biological networks is based on the identification of the set of driver nodes that can guide the system's dynamics, in other words on the choice of a subset of nodes that are selected to be permanently controlled. In addition, it is commonly supposed that, even if all network weights are known, the control based on a brute-force algorithm would be computationally prohibitive for large networks.

Although a *n*-node network could bear up to *n*(n-1)* links (or *n*n* if we also consider self-links) among nodes, this could be conceived as a better chance of network control, not just like a computational difficulty. The higher the number of switches on which one can act, the higher the chance to commute the actual network into the desired one. The application of genetic algorithms to network dynamics allows to act on the highest number of switches, while maintaining the computational effort to tractable levels. Genetic algorithms can reasonably lower the seemingly intractable problem of edge control. In addition I showed here that they allowed to control multiple nodes and edges at the same time.

The framework proposed here might also be applied to semi-quantitative and qualitative networks (Ferrarini 2011b; Ferrarini 2011c). In addition, I suggest that it could be applied to find out how to counteract the negative impacts of climate change on biodiversity. A full understanding of the effects of climate-induced

scenarios on biotic communities necessarily necessitates the consideration of the whole interaction network among species (Ferrarini, 2012c). This requires a $n \times n$ "climate-induced" interaction matrix whose off-diagonal elements w_{ij} take into account the inter-specific effects of species i upon species j as primed by a climate change scenario, while elements w_{jj} take into account intra-specific effects (on growth, phenology and reproduction) on single species. Outcomes of such network of "climate-induced" interactions could predict undesirable effects on biodiversity that can be counterbalanced by management actions that can be properly decided by applying evolutionary modelling to such dynamics.

References

Caldarelli G. 2007. Scale-Free Networks: Complex Webs in Nature and Technology. Oxford University Press, UK

D'heygere T, Goethals PLM, Niels De Pauw N. 2006. Genetic algorithms for optimisation of predictive ecosystems models based on decision trees and neural networks. Ecological Modelling, 195: 20–29

Dorogovtsev SN, Mendes JFF. 2003. Evolution of Networks: From Biological Nets to the Internet and WWW. Oxford University Press, UK

Ferrarini A. 2011. Some thoughts on the controllability of network systems. Network Biology, 1: 186-188

Ferrarini A. 2011b. Fuzzy cognitive maps outmatch loop analysis in dynamic modeling of ecological systems. Computational Ecology and Software, 1(1): 55-59

Ferrarini A. 2011c. Some steps forward in semi-quantitative network modelling. Network Biology, 1(1): 72-78

Ferrarini A. 2012. Betterments to biodiversity optimal sampling. Proceedings of the International Academy of Ecology and Environmental Sciences, 2(4): 246-250

Ferrarini A. 2012b. Network di biodiversità in alcune aree della regione Emilia-Romagna. Technical relation (in Italian)

Ferrarini A. 2012c. I think different: models of climate change impact on plant species are unrealistic. Environmental Skeptics and Critics, 1(2): 30-33

Ferrarini A. 201X. Quant-Lab: a software for solving Quantitative Ecological Networks. In preparation

Goldberg D.E. 1989. Genetic Algorithms in Search Optimization and Machine Learning. Addison-Wesley, Reading, USA

Holland J.H. 1975. Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence. University of Michigan Press, Ann Arbor, USA

Kim DH, Motter AE. 2009. Slave nodes and the controllability of metabolic networks. New Journal of Physics, 11: 113047

Kuo SF, Merkley GP, Liu, C.W. 2000. Decision support for irrigation project planning using a genetic algorithm. Agriculture and Water Management, 45: 243–266

Liu YY, Slotine JJ, Barabasi AL. 2011. Controllability of complex networks. Nature, 473:167-173

Luenberger DG. 1979. Introduction to Dynamic Systems: Theory, Models, & Applications. Wiley, USA

Marucci L, Santini S, di Bernardo M. 2009. How to turn a genetic circuit into a synthetic tunable oscillator, or a bistable switch. PLoS ONE, 4: e8083

Parolo G, Ferrarini A, Rossi G. 2009. Optimization of tourism impacts within protected areas by means of genetic algorithms. Ecological Modelling, 220: 1138-1147

Slotine JJ, Li W. 1991. Applied Nonlinear Control. Prentice-Hall, USA

Wall M. 1996. GAlib: A C++ Library of Genetic Algorithm Components, version 2.4. Massachusetts Institute of Technology, USA