Jul 1st, 12:00 AM

Modelling the spread of an invasive crayfish population with fuzzy cellular automata

Stefano Marsili-Libelli

E. Giusti

F. Gherardi

Modelling the spread of an invasive crayfish population with fuzzy cellular automata

S. Marsili-Libelli*§, **E. Giusti*, F. Gherardi ****

*Department of Systems and Computers, University of Florence, Italy
** Department of Biology, University of Florence, Italy
§ Corresponding author, email: marsili@dsi.unifi.it

Abstract: This paper analyzes the spread of the invasive red swamp crayfish (*Procambarus clarkii*) with a Cellular Automata (CA) model in a mixed terrestrial/aquatic environment. The spatial domain is subdivided in a number of cells, each behaving as a cellular automaton with local rules governing inner growth and movements across the cells. The novelty of the CA approach consists of incorporating a fuzzy inference system in the local rules determining the habitat suitability of neighbouring cells, thus determining the mobility of crayfish. Though the study is in an early stage of development, the preliminary results presented here look encouraging enough to pursue further development.

Keywords: Cellular automata; Fuzzy inference systems; Population dynamics; Invasive population; Knowledge representation.

1. INTRODUCTION

The red swamp crayfish (*Procambarus clarkii, Girard, 1858*) is a widespread pest that represents a major threat to the biodiversity of the European freshwater ecosystems [Aquiloni et al., 2005]. It outcompetes the indigenous crayfish *Austropotamobius pallipes* due to its tolerance to habitat alterations and higher growth rate [Paglianti and Gherardi, 2004]. In addition to supplanting the indigenous crayfish population, the red swamp crayfish threatens the environment by its destructive feeding habits and intense burrowing [Barbaresi and Gherardi, 2006], often leading to the instability and decreased flood resistance of river banks.

The red swamp crayfish (henceforth referred to as crayfish for short) typically inhabits marshy, mixed aquatic-terrestrial environments. Individuals are highly mobile [Gherardi et al., 2002]: they alternate short bouts of intense locomotion to long periods of relatively short movements. Variability in their ranging behaviour has been studied and simulated with mono-directional random walk [Barbaresi et al., 2004].

This paper aims at extending the modelling of crayfish mobility in a two-dimensional heterogeneous domain using the cellular automata (CA) approach. Further, the transition rules governing the CA evolution are formalized in a series of fuzzy inference rules based on the existing literature and expert judgment.

The motivation behind the development of such a model is to help actions for the containment of this crayfish, e.g. by simulating the release of sterile males and/or reducing the appropriate sites for reproductive burrows (i.e. burrows where ovigerous females shelter until juvenile weaning).

2. CELLULAR AUTOMATA

Cellular Automata (CA) [Shiff, 2008] can be viewed as a collection of neighbouring and interacting spatial domains (cells), each ruled by local evolutionary dynamics. The three fundamental features defining a CA are:

- *Uniformity:* all the cells obey the same set of rules;

- *Synchronism:* all the cells are updated at the same time;

- *Locality:* the rules applied to one cell are valid for that cell only.

What makes CA appealing as a modelling tool is that complex behaviours can be obtained by coupling a collection of cells, each with a simple behaviour.
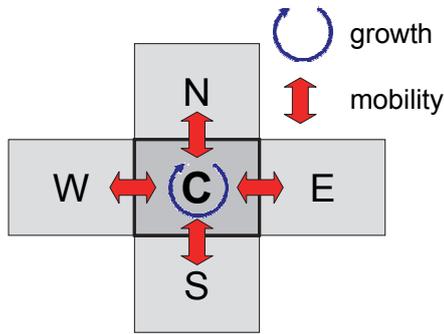

Figure 1. The von Neumann neighbourhood surrounding a central cell.

Since the strength of CA is the breakdown of a complex system into a collection of simple coordinated systems, the key point is the description of the interactions among cells. The Von Neumann neighbourhood (Figure 1) shows the generic cell (C) in a two-dimensional domain surrounded by the four interacting cells denoted as north (N), east (E), south (S), and west (W). The individuals living in a cell are considered a metapopulation [Hanski, 1998] and are described by a single dynamic variable. This approach is a convenient way to represent the spatial heterogeneity of a typically inhabited Mediterranean ecosystem [Touloumis and Stamou, 2009]. Inside each cell the population growth is assumed to follow the Ricker discrete-time logistic equation [Renshaw, 1991]

$$x_{t+1} = x_t \cdot R^{\left(1 - \frac{x_t}{K} \right)},$$
(1)

where R is the maximum growth rate and K is the carrying capacity of the metapopulation in each cell. The interactions with the adjoining cells are represented by mobility, assumed as the average voluntary and independent movement of the crayfish in C, x_t^C, excluding passive displacements due to water movements. The full cell dynamics is obtained by complementing (1) with mobility terms, i.e.

$$x_{t+1}^C = \underbrace{x_t^C R^{\left(1 - \frac{x_T^C}{K} \right)}}_{growth} + \underbrace{\sum_{j=N,E,S,W} m_t^j}_{mobility}.$$
(2)

Figure 1 shows the spatially indexed cell metapopulation x_t^C, with C representing the generic cell and N, E, S, W the neighbouring cells along the coordinate axes. The modelling exercise is now focussed on specifying the mobility term m_t based on the information obtained from the literature and from expert judgement and on

its formalization into a Fuzzy Inference System (FIS), as described in the next section.

3. CA TRANSITION RULES FOR THE CRAYFISH

The transition rules governing crayfish movements across neighbouring cells are organized as a Fuzzy Inference System (FIS) on the basis of the observed crayfish behaviour. Translating in mathematical terms all the aspects of the crayfish living habits is a formidable task and for the time being it suffices to outline the method of knowledge representation in an open-ended fuzzy formalization, which can be later augmented as more information is obtained from direct observations.

3.1. Habitat suitability

Crayfish move in both the aquatic and terrestrial environments in search of food and appropriate substrates to burrow, particularly to hatch or to avoid dehydration. It is supposed here that movements are motivated by the "attractiveness" of neighbouring cells, because they either have a lower population density or offer a more suitable habitat. Hence, the next step is to devise a mathematical characterization of the cell suitability, assuming that this is described (for the time being) by vegetation, water depth and population density. More habitat features can be added in the future without changing the general framework of the algorithm.

The suitability is computed through a set of Sugeno inference rules [Takagi and Sugeno, 1985]

$$R_i : if\ \left(veg\ is\ V\right)\ and\ \left(depth\ is\ D\right)\ and\ \left(Pop\ is\ P\right)\ then\ \ S_i = W\ \ \ i = 1,...,N$$

$$S = \frac{\sum_{i=1}^{N} \mu_i S_i}{\sum_{i=1}^{N} \mu_i} \tag{3}$$

Suitable membership functions, shown in Figure 2, have been defined for the antecedents (*veg, depth, Pop*) on a normalized support to describe the three possible states (Low, Medium, High), whereas the consequent singletons W={0, 0.5, 1} represent the suitability grades {*Unsuitable, Transition, Suitable*}. Fourteen inference rules were defined, based on the pertinent literature and personal communications with the aim of reproducing the observed behaviours.
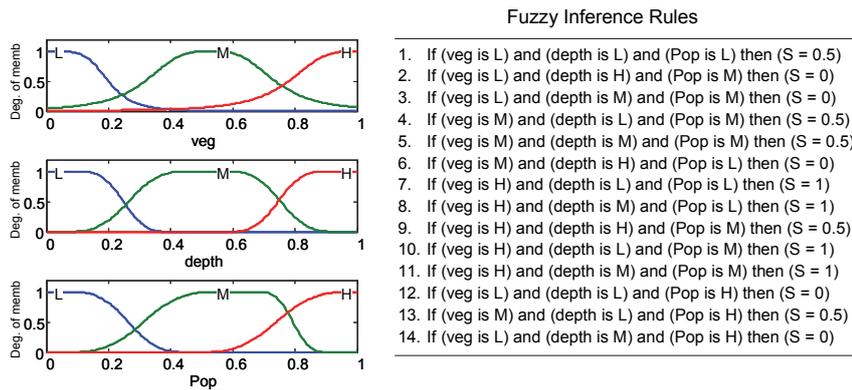

Fuzzy Inference Rules

1. If (veg is L) and (depth is L) and (Pop is L) then (S = 0.5)
2. If (veg is L) and (depth is H) and (Pop is M) then (S = 0)
3. If (veg is L) and (depth is M) and (Pop is M) then (S = 0)
4. If (veg is M) and (depth is L) and (Pop is M) then (S = 0.5)
5. If (veg is M) and (depth is M) and (Pop is M) then (S = 0.5)
6. If (veg is M) and (depth is H) and (Pop is L) then (S = 0)
7. If (veg is H) and (depth is L) and (Pop is L) then (S = 1)
8. If (veg is H) and (depth is M) and (Pop is L) then (S = 1)
9. If (veg is H) and (depth is H) and (Pop is M) then (S = 0.5)
10. If (veg is H) and (depth is L) and (Pop is M) then (S = 1)
11. If (veg is H) and (depth is M) and (Pop is M) then (S = 1)
12. If (veg is L) and (depth is L) and (Pop is H) then (S = 0)
13. If (veg is M) and (depth is L) and (Pop is H) then (S = 0.5)
14. If (veg is L) and (depth is M) and (Pop is H) then (S = 0)

Figure 2. Definition of the membership functions and inference rules for the suitability computation.

3.2. Mobility rules

The mobility term m_t depending on the cells suitability are obtained by the FIS of Figure 2. Whenever the population in a generic cell C reaches a population threshold x_{th} and the suitability of any of the adjacent cells is above a suitability

threshold S_{th}, then mobility occurs in proportion to the suitability gradient and causes a share α_t of the population in C $\left(x_t^C\right)$ to spread to the most suitable cell in the neighbouring set. This can be formalized with the following criterion

$$m_t^j \atop \{j=N,E,S,W\} = \begin{cases} = \alpha_t^j \cdot x_t^C \cdot \left(S_t^j - S_t^C\right) & if \left(x_t^C > x_{th}\right) \wedge \left(S_t^j > S_t^C\right) \\ = 0 & otherwise \end{cases}. \tag{4}$$

3.3. Cell dimension and update period

Normally the CA theory assumes that cells are abstract concepts, hosting rules and transitions. In this case, however, their purpose is also to represent physical boundaries containing metapopulations. Hence, the concepts of motion and cell update are related to the cell physical dimensions. Assuming that the individuals forming a metapopulation moves about with average speed v_m, a hypothetical crayfish starting from the cell centre would reach its boundary after a time

$$v_m = \frac{\delta}{2t_m} \quad \Rightarrow \quad t_m = \frac{\delta}{2v_m}, \tag{5}$$

where δ is the cell dimension. Now, if the update mechanism is to record possible motions outside the cell, the update period should be selected at least equal to t_m so that only movements above the average can be considered as meaningful motions as they cross the cell boundaries..

The problem is now to obtain a statistics for the crayfish movements. Barbaresi et al. [2004] recorded the movements of several individuals from a population and their observations are now used to fit a lognormal distribution from which stochastic realizations can be generated. A comparison between the recorded speed data and a realization from the fitted distribution, with mean $\mu = 3.5319 \pm 0.2014 \, (cm/h)$ and standard deviation $\sigma = 1.2736 \pm 0.1595 \, (cm/h)$, is shown in Figure 3.
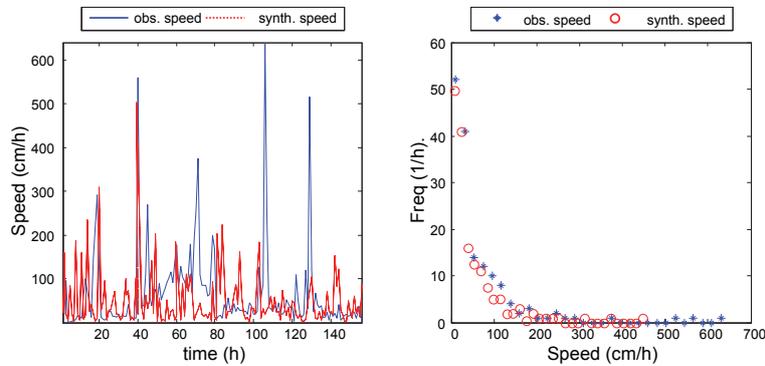

Figure 3. Fitting the speed data from [Barbaresi et al., 2004] to a lognormal distribution. The left plot shows the original data and a sample realization from the distribution, whereas the right plot compares their histograms.

Now the problem is how to generate as many movements as there are individuals in a cell. To do this, at time *t* a set of movements is generated from the fitted lognormal statistics with the dimension of the current population, i.e.

$$v_t = lognrdn\left(\mu, \sigma, x_t^C, 1\right), \tag{6}$$

where x_t^C is the population in the cell C at time *t* and 1 means that only the ensemble at time *t* is generated, so that v_t is a memoryless random vector of the

same dimension as x_t^C, whose upper limit is therefore given by the carrying capacity K. Taking the ensemble average speed $\bar{v}_t = <v_t>$ a relationship between the cell update period t_m and the cell size δ can be established from (5). Further, the average speed $\bar{v}_t$ is incorporated into the coefficient $\alpha_t = \beta \cdot \bar{v}_t$ to reflect the mobility fluctuations in (4).

3.4. The complete CA for crayfish spread

The overall algorithm for the computation of the mobility term m_t is shown in Figure 4. The state of the three environmental variables (population at time *t*, vegetation, and water depth) is used as antecedents in the FIS computing the cell suitability, which is then used in (4) to compute the outbound mobility.
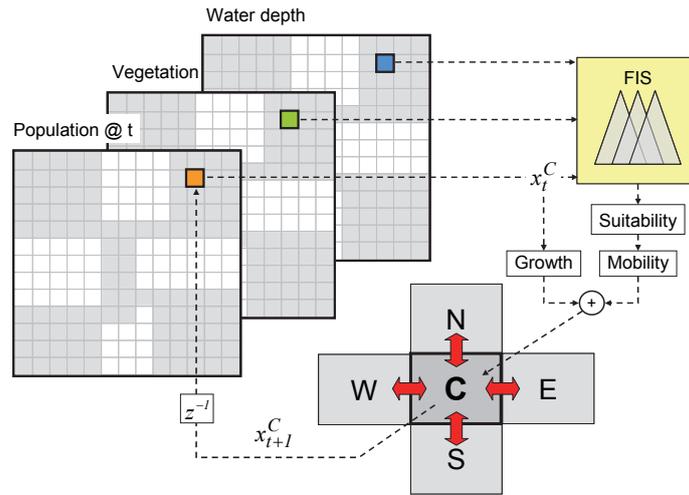

Figure 4. Fuzzy inference of the cell suitability coupled with the growth dynamics (2). The z^{-1} block represents the unit delay of the simulation time-step.
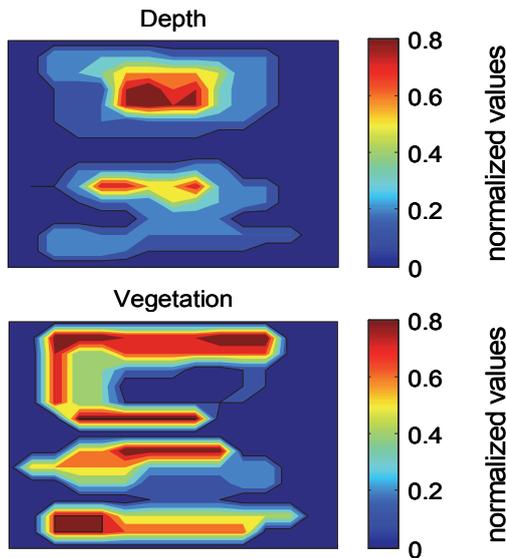
4. SIMULATION EXPERIMENTS

The hypothetical simulation domain of Figure 5 was considered for a preliminary test. It is composed of two ponds separated by an impassable barrier. The upper pond is deeper in the centre, whereas the vegetation is evenly distributed over the entire accessible area, except for a small area at the right end of the lower pond.

The simulation was initialized with two "seed" metapopulations, placed in similar position, given the isolation between the two sub-domains. The central part of the upper pond is rather deep and therefore its suitability is very low, though the vegetation is abundant. A fairly low growth rate was selected (R = 1.2) causing the population to evolve regularly to the carrying capacity, as shown in Figure 6. The limited habitat suitability of the bottom part of the lower pond and of the top left

Figure 5. Characteristics of the simulation domain in terms of water depth and vegetation.

corner of the upper pond discourages colonization even when the overall carrying capacity has been attained elsewhere.
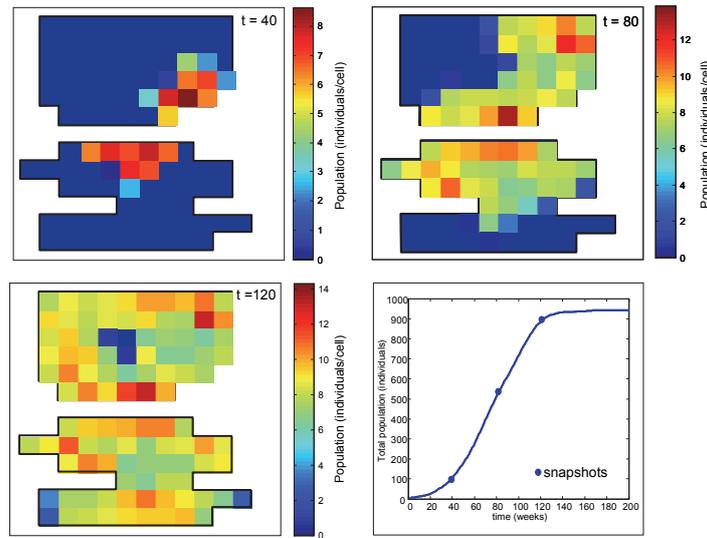

Figure 6. Spread of the crayfish population over the simulation domain.
The snapshots corresponding to the three maps are shown in the total population graph (bottom right).

5. APPLICATION TO AN INVADED AREA

A preliminary attempt to test the model in a real habitat is now considered. The northern boundary of the Massaciuccoli Lake, in northwest Tuscany (central Italy), was selected. This shallow lake and its surroundings have been invaded for several years by crayfish. The area considered for the simulation is 1000 x 1000 metres and it was modelled with a lattice of 200 x 200 cells. The area is a mix of clear water ponds, deep channels, and thick reed beds [Giusti et al., 2011a]. The water depth varies from few centimetres to one meter. The crayfish have colonized the marshy areas, subject to periodical floods as a consequence of the wide variation in lake water level [Giusti et al., 2011b]. Figure 7 indicates the relative water depth and abundance of vegetation.
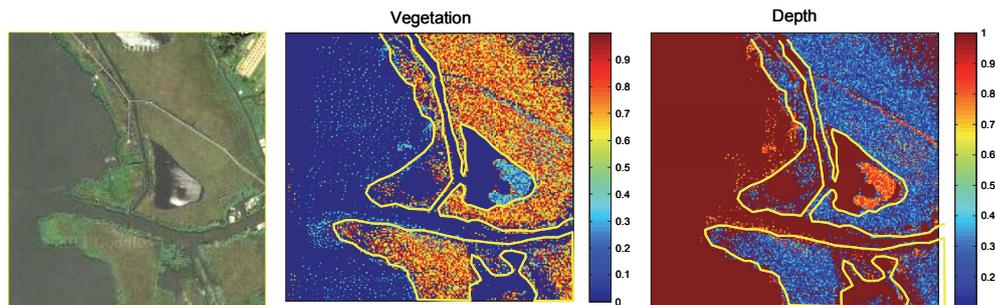

Figure 7. Satellite view of the Massaciuccoli area (left). Vegetation and depth of the area are shown in the middle and right graphs.

The simulation was seeded with two small colonies in the upper right corner of the area. Figure 8 shows the expansion of the population in the area, following the preferred habitat that is characterized by abundance of vegetation (bulrush and reed in this case) and shallow water. It can be seen that almost no colonization occurs in the clear water areas, where the water depth exceeds one metre. The simulations also show that expansion occurs preferably after the carrying capacity in the colonized cells has been attained and that the marshy areas are preferred. In fact, only for *t* = 300 a fraction of the population has expanded in the clear water area in the centre-right part.
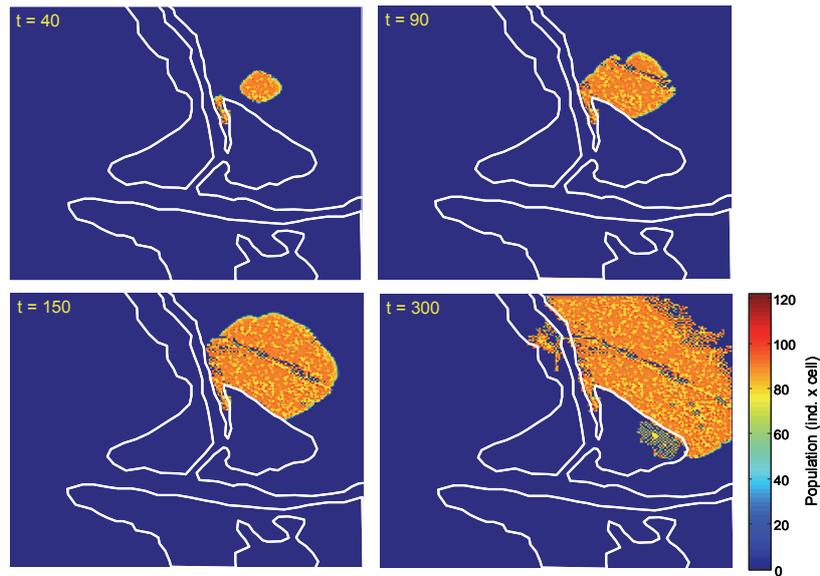
Figure 8. Simulated spread of the crayfish population in the northern boundary of the Massaciuccoli Lake. The colorbar indicates the cell population density whereas the time units are weeks.

6. CONCLUSION

This paper has presented some preliminary result of a novel attempt at describing the spread of an invasive crayfish population living in marshy environments. The innovative aspects of this model are:

a) *The cellular automata approach* used to describe the spatial-temporal dynamics of the crayfish population;

b) *The fuzzy formalization of the growth* in each cell, described by a simple Ricker logistic equation, and complemented with suitability-driven logic governing the movements across adjoining cells;

c) *The formalization of the cell suitability*, using a fuzzy inference system, provisionally taking into account relative crowding, food availability, and water depth.

Based on field observations, some behavioural rules have been formalized to assess the suitability of the cells and hence the attractiveness of the new habitat to the individuals in the neighbouring cells.

Previously, only deterministic models were available [Anastacio et al., 1999] and the population growth was not related to habitat conditions. To reproduce the variability of the crayfish motion patterns, alternating long quiescent periods to short bouts of long-range explorations, a lognormal statistics was fitted to the observed ranging behaviour and incorporated into the model. The attempt presented here is admittedly in its infancy and much work lies ahead. The aspects that will require further investigation are:

a) Improvement of the growth dynamics by considering the reproductive season and the average number of offspring, i.e. introducing an age-structured population model differentiating between young-of-the-year and adults [Renshaw, 1991];

b) Extension of the habitat suitability logic to other biological parameters, such as reproductive behaviour;

c) Model refinement by collecting more observation for a better tuning of the motion statistics, growth dynamics, and fuzzy rules.

Though the test domains presented here are fairly hypothetical or lack definitive data validation, nevertheless they reflect the typical habitats in which crayfish thrive and were used mainly to illustrate the methodology. The next effort will obviously

be to adapt the model to real habitats for which enough data can be obtained for reliable calibration and testing

7. ACKNOWLEDGEMENT

REFERENCES

Anastacio, P.M. and J.C. Marques, Population biology and production of the red swamp crayfish Procambarus clarkii (Girard) in the lower Mondego river valley, Portugal, *J. Crustacean Biology* 15(1), 156–168, 1995.

Anastacio, P.M., S.N. Nielsen and J.C. Marques, CRISP (crayfish and rice integrated system of production): 2. Modelling crayfish (Procambarus clarkii) population dynamics, *Ecological Modelling* 123 5–16, 1999.

Aquiloni L., M. Ilhéu and F. Gherardi, Habitat use and dispersal of the invasive crayfish Procambarus clarkii in ephemeral water bodies of Portugal, *Marine and Freshwater Behaviour and Physiology* 38(4), 225-236, 2005.

Barbaresi, S., G. Santini, E. Tricarico and F. Gherardi, Ranging behaviour of the invasive crayfish, *Procambarus clarkii* (Girard), *J. of Natural History* 38, 2821 - 2832, 2004.

Barbaresi, S. and F. Gherardi, experimental evidence for homing in the red swamp crayfish, procambarus clarkii, *Bull. Fr. Pêche Piscic.* 380-381, 1145-1154, 2006.

Craig, P.D., Imposed and inherent scales in cellular automata models of habitat, *Ecological Modelling* 221, 2425–2434, 2010.

Giusti, E., S. Marsili-Libelli and S. Mattioli, A fuzzy quality index for the environmental assessment of a restored wetland, *Water Sci. Tech.* **63** (9), 2061 – 2070, 2011a.

Giusti, E., S. Marsili-Libelli and A. Gualchieri, Modelling a coastal lake for flood and quality management. paper presented at *Watermatex 2011, 8th IWA Symposium on Systems Analysis and Integrated Assessment,* San Sebastian (SP), 2011b.

Gherardi, F., E. Tricarico and E. Ilhéu, Movement patterns of an invasive crayfish, *Procambarus clarkii*, in a temporary stream of southern Portugal, *Ethology Ecology & Evolution* 14, 183-197, 2002.

Guinot, V., Cellular automata for ecological modelling: deriving transition rules from physical parameters, paper presented at *Hydroinformatics 2002, Fifth International Conference on Hydroinformatics*, Cardiff, UK, 2002.

Hanski, I., Metapopulation Dynamics, *Nature* 396, 41-49, 1998.

Paglianti, A. and F. Gherardi, Combined effects of temperature and diet on growth and survival of young-of-year crayfish: a comparison between indigenous and invasive species, *J. Crustacean Biology*, 24(1), 140–148, 2004.

Renshaw, E., *Modelling Biological Populations in Space and Time,* Cambridge University Press, 403 pp. Cambridge, 1991.

Schiff, J.L., *Cellular Automata: A Discrete View of the World*, Wiley-Interscience, 280 pp., New York, 2008.

Takagi, T. and Sugeno, M. Fuzzy identification of systems and its applications to modeling and control. *IEEE Trans. on Systems, Man, and Cybernetics 15*, 116 – 132, 1985.

Touloumis, K. and G.P. Stamou, A metapopulation approach of the dynamics of arthropods from Mediterranean-type ecosystems, *Ecological Modelling* 220, 1105–1112, 2009.