19th Mediterranean Conference on Control and Automation
Aquis Corfu Holiday Palace, Corfu, Greece
June 20-23, 2011

Agent-based modeling of an activated sludge process in a batch reactor

M. Pereda, J.M. Zamarreño

Abstract— **The aim of this work is to study the feasibility of using agent-based modeling to study the activated sludge process. A model in NetLogo has been proposed, and experiments have been developed comparing the model behavior with a classical modeling approximation for this process, the Monod model.**

I. Introduction

Water is one of the most fundamental natural resources, and together with air, land and energy are the four basic resources that support the development. This implies the need to develop processes for water purification, such as those that occur in wastewater treatment plants (WWTP), whose objectives are: to minimize water pollution, protect the environment, maintain quality of life of individuals and save energy.

This work is concern with the application of Agent-based modeling to study the activated sludge process, a wastewater treatment method in which the carbonaceous organic matter of wastewater provides an energy source for the production of new cells for a mixed population of microorganisms in an aquatic aerobic environment. The microbes convert carbon into cell tissue and oxidized end products that include carbon dioxide and water.

The chosen process (WWTP) is a suitable setting to develop agent-based models due to its features. This choice is motivated by the progressive awareness of the need for sustainable growth, ensuring environmental conservation. The difficulty of operation and control of the plant lies in the complexity and variability of the biological processes involved, among other factors, leading to an inefficient operation of real plants with the inevitable environmental consequences and high economic penalties. The improvements we could provide in the field of modeling would allow improvements in the control, supervision and operation in general.

Agents and multi-agent systems (MAS) have become one of the most active and lively research areas in computer science worldwide, from pure technical ones to even social ones. The systems are modeled by a series of entities with capacity of decision and action, autonomous software entities that exist within a software environment, which are called agents. Agent-based modeling (ABM) treats each individual component of a system as a single entity obeying its own pre-defined rules and reacting to its environment and neighboring agents accordingly. The models simulate the simultaneous operation and interactions between multiple agents in an attempt to recreate and predict the behavior of a complex phenomenon. It is said that the process to represent (macroscopic level) emerge from the microscopic level, i.e., the interaction of relatively simple behaviors leads to complex behavior from a global perspective. Phenomena such as flocks of birds, school of fish, and the complex biological systems of cells are good examples of how systems with simple goals can show complex emergent behaviors as a result of the communication with neighboring agents.

The applications of ABM began to develop around 1990, mainly due to the increase of computational capabilities of computers and the emergence of software that could run such models (Swarm [1], NetLogo [2], Repast [3], MASON [4], to name the most popular).

ABM has recently been applied to many different research areas, including economy [5], finance [6], natural resource management and ecology [7], [8], political science [9], biology [10], medicine [11], optimization of logistics and supply chain, analysis of traffic jams, etc.

In the field of process modeling, there are few references, but in recent years some researchers have published several works on this topic, such as modeling of chemical reactions [12], modeling of microbes [13] or applications in diagnosis of chemical processes [14].

This paper is organized as follows. In section two a brief introduction to the activated sludge process is given. Then a short introduction to agent-based modeling follows. In this section it is also described the model presented in this work. In section 4 the simulation results are shown. Finally, section 5 discusses points for further research and conclusions.

II. Description of the Activated Sludge Process

Conventionally, the WWTP processes are grouped into: water line: pretreatment, primary treatment, secondary and tertiary; sludge line: thickening, digestion, cooling, drying and disposal; and gas line for methane production.

The secondary treatment is responsible for removing biodegradable organic matter present in wastewater that has not been removed in the primary treatment. It is carried out by the growth of microorganisms capable of assimilating organic matter which is converted into other products and new microorganisms which can be easily removed from

Manuscript received January 27, 2011. This work was supported in part by the University of Valladolid under the Grant “Ayuda para la Formación de Personal Investigador”.

water by decantation. An activated sludge can be defined as a microbial mass produced when the wastewater is aerated continuously. That mass is made up of microorganisms that are able to break down and metabolize the main contaminants in the wastewater. The activated sludge process is a type of biological process that takes place in the secondary treatment of sewage at a wastewater treatment plant. The activated sludge process consists of: an aeration tank (bioreactor) where the activated sludge and food (contamination or substrate water content) is mixed and aerated for a certain period of time, a system of separation of activated sludge and treated effluent (secondary clarifier or clarifier), a recirculation of activated sludge to the aeration tank, and a system of treatment and disposal of the sludge produced (purge). The aeration tank contains a "sludge" which is what could be best described as a "mixed microbial culture", containing mostly bacteria, as well as protozoa, fungi, algae, etc. This sludge is constantly mixed and aerated either by compressed air bubblers located along the bottom, or by mechanical aerators on the surface. The wastewater to be treated enters the tank and mixes with the culture, which uses the organic compounds for growth-- producing more microorganisms-- and for respiration, which results mostly in the formation of carbon dioxide and water. The process can also be set up to provide biological removal of nitrogen and phosphorus. Refer to Fig. 1 for a simplified process flow sheet.

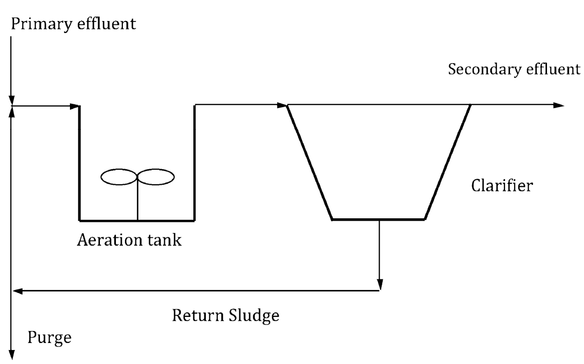

Fig. 1. Activated sludge process components.

After enough aeration time to reach the required level of treatment, the sludge is carried by the flow into the settling tank, or clarifier. The sludge collected at the bottom of the clarifier is then recycled to the aeration tank to consume more organic material. The term "activated" sludge is used, because by the time the sludge is returned to the aeration tank, the microorganisms have been in an environment depleted of "food" for some time, and are in a "hungry", or activated condition, eager to get busy biodegrading some more wastes. Since the amount of microorganisms, or biomass, increases as a result of this process, some must be removed on a regular basis for further treatment and disposal, adding to the solids produced in primary treatment.

The type of activated sludge system that has just been described is a continuous flow process. In this work we are modeling a simplified version of the process in a batch reactor, i.e., an isolated reactor in which an initial charge is introduced and provided with adequate ventilation conditions for the optimal development of the reactions of substrate-biomass interaction, in order to better focus on the interaction substrate-biomass.

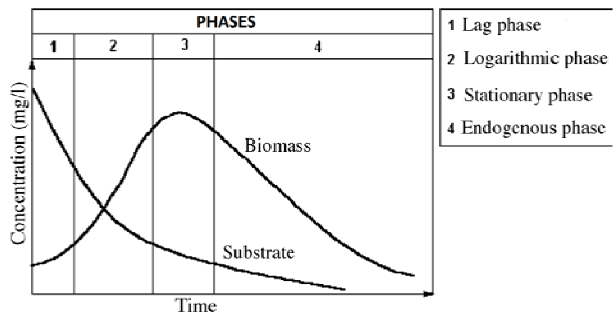

Fig. 2. Temporal evolution of the concentrations of substrate and biomass in an activated sludge batch reactor.

Under these conditions, the temporal evolution of the concentrations of substrate and biomass takes place as shown in Fig. 2. The growth curve is obtained by a count of the number of living cells (biomass) over time.

The biomass curve consists of several phases which are explained below. The microorganisms must first accommodate to their environment and available food. This accommodation period is called lag phase, and varies in size depending on the history of seeded microorganisms. If microorganisms are adapted to the environment, the lag phase will be very brief. Once growth has begun, it will continue rapidly. When maximum growth is occurring, the behavior is logarithmic, this is why this phase is called the logarithmic phase. Maximum growth cannot continue indefinitely. The food available may be ended up, environmental conditions may change (e.g., overpopulation, accumulation of waste products, etc.), and a population of predators can be developed. Cells that are unable to obtain food from outside sources initiate endogenous catabolism, i.e., they catabolize the protoplasm stored to maintain their energy. Other cells die or break releasing their protoplasm, which is added to the food available. This stage is represented in Fig. 2 with the name of stationary phase, and represents the time along which the production of new cell material is roughly compensated by death and endogenous respiration. Whereas in the stationary phase there is still some reproduction, the endogenous respiration and death dominate the fourth phase, called endogenous phase. In this last phase, the biomass decreases slowly, asymptotically approaching the horizontal axis.

There are several classic proposals for modeling this process, the most popular being the Monod model [15].

The *Monod model* is based on a mass balance. Therefore, there are terms associated with the contribution of biomass and substrate to the environment, and terms related to the removal of substrate and biomass from the environment. S represents the amount of soluble substrate (in milligrams per liter) and X represents the amount of biomass (in milligrams per liter), the rate of substrate consumption is dS/dt, and the rate of biomass growth is dX/dt. Equations of the model are as follows:

$$\frac{dX}{dt} = \mu \cdot X - K_d \cdot X \tag{1}$$

$$\frac{dS}{dt} = -\frac{\mu}{Y} \cdot X \tag{2}$$

$$\mu = \mu_{\max} \cdot \frac{S}{K_s + S} \tag{3}$$

where μ = specific growth rate coefficient; μ_{max} = maximum growth rate coefficient; K_s = Monod coefficient, also called the half-saturation coefficient because it corresponds to the concentration at which μ is half of its maximum; Y is the yield of the reaction; and K_d a parameter related to the death of biomass.

The biomass contribution term explains the process of 'birth' of microorganisms. The biomass loss term explains the process of 'death' of microorganisms. This model does not consider any term to describe a substrate contribution to the environment.

III. AGENT-BASED MODELING

Although the term agent is widely used, there is no single universally accepted definition. In the literature, different meanings of the term can be found, some of which are shown in [16] and [17]. Wooldridge and Jennings [18] define an intelligent agent as a hardware or (more often) a software entity situated in an environment that is capable of displaying flexible autonomous behavior, in order to achieve their own goals. As with the definition of agent, there is no universally accepted definition for the term multi-agent system. It could be said that a multi-agent system is the one formed by a group of agents.

One of the most important concepts associated with agent-based modeling is the concept of emergence. Emergent phenomena are macroscopic patterns arising from decentralized interactions of simple individual components [19]. The idea behind this definition is the possibility offered by agent-based systems to study the macro-properties that emerge in the system from the micro-definition of the behavior of the agents that compose it, i.e., to study micro-macro relations.

At present there are many software options for implementing models based on agents. These range from simple spreadsheets, to general purpose languages or libraries and more specific applications. In the literature, several comparative summaries of different tools of agent-based modeling can be found. In addition to [20], it is worth noting the Wikipedia entry under the name of Comparison of agent-based modeling software [21], and Getchell work [22]. For the first version of the model proposed in this paper, NetLogo was chosen as software platform, due to the easiness to implement agent-based models on it, the large number of examples available, the quality of the manual that is distributed with the tool, and, mainly because the modeling power of this tool is more than enough for the model to be implemented (in its first version).

The simulations presented in this work represent a genuine approach where evolution over time of the model is the result of the individual rules set to characterize the position in space, motion, uptake, metabolism, reproduction and death of biomass agents. The model describes all the actions of the agents and their interactions with their surroundings, and dynamics emerge from this.

The system to be modeled has been described in Section 2. It is a batch reactor for wastewater treatment.

The activated sludge process involves different species of microorganisms forming a food chain. In addition, many complex processes occur at the cellular level, such as substrate uptake, metabolism, growth, cell division, etc, that for each kind of organism is characterized by different parameters. This implies high complexity of modeling from the microscopic point of view. Therefore, it is necessary to reach a compromise decision between the descriptive ability of the model (detail) and simplicity, while taking into account the objective and scope of the investigation.

The following considerations are assumed: modeling a batch reactor, i.e., once the process starts there is no input nor output of biomass and substrate from the reactor, a single population of microorganisms is modeled, and it is assumed the existence of an aeration and agitation system that provides the oxygen necessary for aerobic bacterial action, which prevents the sedimentation of sludge in the reactor and allows the homogenization of activated sludge.

The agents of the model can be divided into two distinct groups: biomass agents and substrate agents. In both cases these are reactive agents. The position of each agent in the world is referenced to a Cartesian system of two axes, since it is a two-dimensional model.

Biomass agents have three properties: size (*size_b*): volume occupied by the agent in the environment, initialized to *initial_b_size*; age, i.e., number of periods the agent has lived, initially takes a value as a normal distribution N(*initial_age*, *sd_age*); and internal energy, property that indicates the number of iterations that the agent can exist without eating and not dying; initially setting a value based on a normal distribution N(*max_energy*, *sd_energy*), which is increased each time the agent eats by *i_energy_by_eating* parameter up to a maximum value of *max_energy*, and decremented at each iteration by *d_energy_by_starving*. Substrate agents have only one property, their size (*size_s*), i.e., volume occupied by the agent in the environment, initialized to *initial_s_size*.

Besides, the model is characterized by the following constants: *rep_size*, size from which the biomass is able to reproduce; *unit_growth*, amount by which the size of a biomass agent is increased when eating; *b_density*, biomass density; *s_density*, substrate density; *size_percentage*, percentage that applied to the size of a biomass agent indicates the amount that will decrease the size of a substrate agent eaten by that biomass agent; *sd_death_lmax*, *sd_death_gmax*, *max_age*, *mean_death_by_energy*, that will be explained later; and *max_energy*, *sd_energy, i_energy_by_eating, d_energy_by_starving,* all explained before.

Moreover, the model shows some degree of randomness due to different statistical arguments discussed below. Initially, agents are placed randomly on the environment as a uniform distribution. Biomass agents are created with initial values of energy as a normal distribution N(*max_energy*, *sd_energy*) , and initial age values as a normal distribution N(*initial_age*, *sd_age*). When biomass agents metabolize substrate, they preferably choose to eat substrate that are beneath them, and then in the nearby positions randomly selecting the substrate neighbor agent to eat. Biomass agents reproduce (to be split in half creating 2 individuals of the same characteristics as the parent, size half of the father) according to a normal distribution N(*rep_size*, 1). Biomass agents die by two causes: endogenous respiration and age. On the one hand, death by reaching the maximum age, which is an N(*max_age*, *sd_death_lmax*) if the age of biomass agent is less than the maximum age of death, and N(*max_age*, *sd_death_gmax*) if the age of agent biomass is greater than the maximum age of death, so an asymmetric Gaussian is set, with less variance for biomass agents older than the maximum age, because when they exceed this maximum age it is more likely to die over a narrow period. On the other hand, the other cause of death is the total depletion of the internal energy of biomass after the endogenous catabolism process, in which agents are unable to obtain food and they catabolize the protoplasm stored to maintain their energy. This death rate is modeled as an exponential distribution which mean is *mean_death_by_energy*.

Initialization of the model is as follows: assign values to the constants of the model, create the agents and place them in the world, and assign initial values of age and energy to biomass agents.

At each step the following actions are performed:

1) To shake. At each step of the program, agents in the world are arranged in a random position, thus modeling a system that operates in continuous stirred-tank reactor (CSTR) conditions, obtaining homogeneous characteristics in terms of substrate and biomass concentrations.
2) To eat. If a biomass agent has a substrate agent below, it will eat a portion of substrate agent equal to its size multiplied by *size_percentage*. If not, it will look at neighboring positions (north, south, east, west, northeast, northwest, southeast, southwest) and randomly pick one and eat from it. The biomass agents that have eaten, increase its size by the quantity it has eaten, and its internal energy by one. The substrate agent eaten decreases its size by a percentage of the eating biomass agent (*size_percentage* parameter is used here).
3) To reproduce. Biomass agents are divided in half creating 2 individuals of the same characteristics as the parent, and half the size of the parent, according to parameter *rep_size*.
4) To die. It is checked whether biomass agents will die from one of the two cases discussed above. If they meet the requirements for dying, they die, becoming their mass part of the food present in the reactor.
5) At each step of the program, the internal energy of each biomass agent is decremented by *d_energy_by_starving* and age increases by one.

This five actions are repeated iteratively until a number of predefined steps is reached.

IV. SIMULATION RESULTS

The simulations presented below show that an agent-based model is able to represent the dynamics of certain type of processes with results as good as traditional modeling paradigms, such as, in this case, the Monod model. The aim of this work is, in terms of its microscopic constituents, to understand the complex behavior of the macroscopic system.

For the Monod model parameterization, we use typical values from [23], as Table I shows. The parameter values for the agent-based model were chosen using the trial and error technique. They were chosen so the model response behaves in a similar way as Monod model response. The responses of both models are shown in Fig. 3. Moreover, Fig. 3 reproduces the different phases of the activated sludge process that were explained before, except for the lag phase, because the microorganisms in this model are initially accommodated to the environment.

TABLE I
MONOD MODEL PARAMETERS

Parameter	Value	Parameter	Value
S0 (mg/l)	50	*μmax (day-1)*	1.04
X0 (mg/l)	10	*Ks (mg/l)*	100
Y	0.55	*Kd (day-1)*	0.055

Fig. 4 and Fig. 5 show the evolution over time of biomass reproduction process and biomass deaths, respectively. The reproduction process takes place from the twenty-three simulation step to the forty-seven, as Fig. 4 shows. In Fig. 5 the two causes of death can be observed.

TABLE II
PROPOSED MODEL PARAMETERS

Parameter	Value	Parameter	Value
X0	1090	*initial_s_size*	10
S0	545	*unit_growth*	0.25
initial_b_size	1	*b_density*	1
initial_age	1	*s_density*	1
sd_age	1	*size_percentage*	0.6
max_energy	10	*sd_death_lmax*	20
sd_energy	2	*sd_death_gmax*	10
i_energy_by_eating	1.7	*max_age*	50
d_energy_by_starving	0.7	*mean_death_by_energy*	2

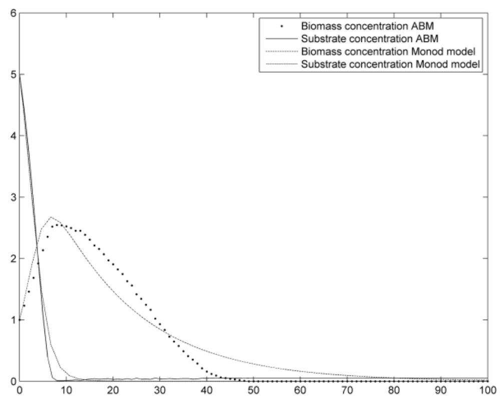

Fig. 3. Temporal evolution of the concentrations of substrate and biomass.

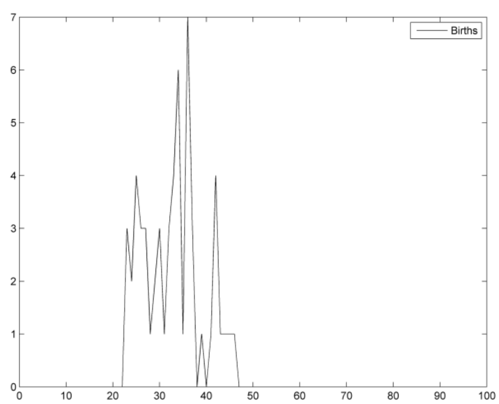

Fig. 4. Temporal evolution of biomass reproduction process.

To test the model, a set of experiments were done, comparing the experimental results with the expected behavior and respect to the initial case (Fig. 3).

For example, if the *size_percentage* (percentage that applied to the size of a biomass agent indicates the amount that will decrease the size of a substrate agent to be eaten by that biomass agent) is decreased by 50%, the evolution of substrate and biomass concentrations is as shown in Fig. 6. The biomass agents eat a smaller portion of substrate each time they eat, thus the food present in the environment (substrate) last longer than in the initial case. For this reason, the biomass concentration reaches a higher maximum value than in the first case, being in the environment more agents (if agents reach the reproduction size, they can reproduce, so the more food, the more biomass agents). When the substrate ends, stationary phase starts, followed by the endogenous phase. The evolution of these two phases is slower than in the initial case because there are more biomass agents.

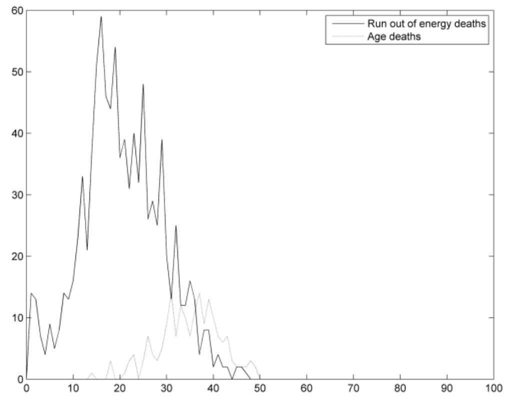

Fig. 5. Temporal evolution of biomass deaths.

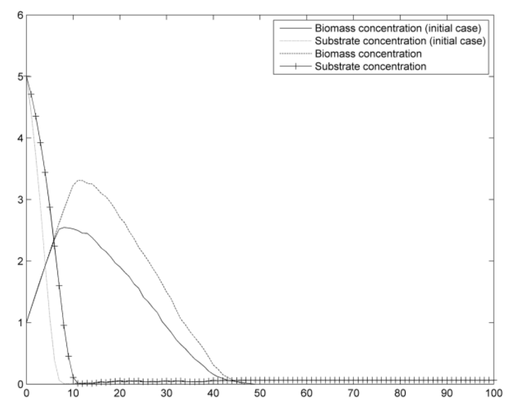

Fig. 6. Temporal evolution of the concentrations of substrate and biomass.

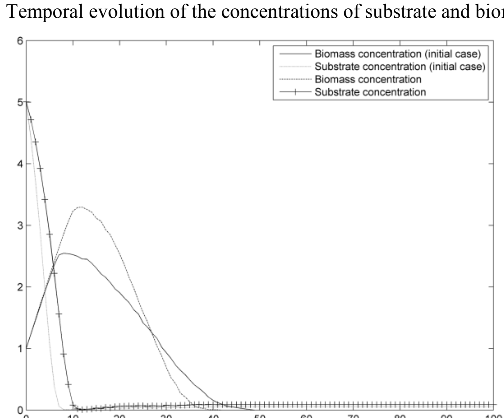

Fig. 7. Temporal evolution of the concentrations of substrate and biomass.

If a 20% of decreasing is performed in *max_age*, so biomass agents may die younger than in the initial case, the results are shown in Fig. 7 and Fig. 8.

As it can be observed in Fig. 8, biomass agents start dying earlier than in the initial case, as expected. This causes the food (substrate) to last longer and the logarithmic phase ends later. Moreover, the substrate concentration reaches zero (all substrate agents were eaten) earlier than in first case, as Fig. 7 shows.

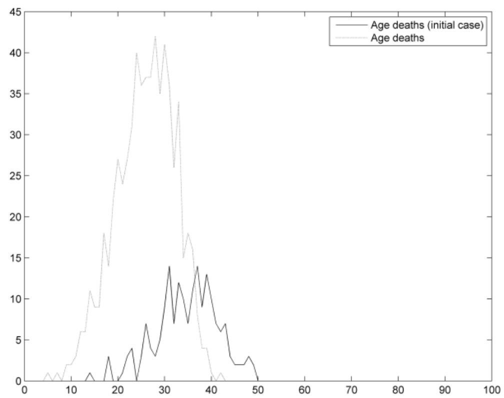

Fig. 8. Temporal evolution of biomass deaths.

V. Conclusion

As this paper has shown, Agent-Based Modeling allows the researcher to develop a bottom-up approximation to the study of complex problems such as the phenomena occurring in a wastewater treatment plant, and in particular, secondary treatment by activated sludge. An agent-based model for the activated sludge process in a batch reactor has been proposed, which allows to analyzing and a better understanding of the phenomena that occur in the system while varying some of its characteristics.

As a continuation of this project, the following lines arise: to model the activated sludge process in a continuous reactor and to incorporate real data to adjust model parameters, in order to be able to rehearse the various developments and techniques in a real WWTP; to incorporate into the model other factors that influence the activated sludge process, such as temperature, pH, etc ...; to build parallel implementations of the model to reduce the computational load; and to integrate the agent-based models with process control.

Acknowledgment

The authors gratefully acknowledge the support of the CICYT through the project "Operación óptima de estaciones depuradoras de aguas residuales (EDARs)", DPI2009-14410-C02-02.

References

[1] N. Minar, R. Burkhart, C. Langton, and M. Askenazi. (2011, January 24). The Swarm simulation system: a toolkit for building multi-agent simulations. [Online]. Available: http://ideas.repec.org/p/wop/safiwp/96-06-042.html

[2] U. Wilensky. (1999). NetLogo. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL. Available: http://ccl.northwestern.edu/netlogo/

[3] Collier, N.T., and M.J. North, "Repast SC++: A Platform for Large-scale Agent-based Modeling," in W. Dubitzky, K. Kurowski, and B. Schott, eds., Large-Scale Computing Techniques for Complex System Simulations, Wiley (In Press 2011).

[4] S. Luke, G. C. Balan, L. Panait, C. Cioffi-Revilla, and S. Paus, "MASON: A Java Multi-Agent Simulation Library," in Proceedings of the Agent 2003 Conference.

[5] L. Tesfatsion, "Agent-based computational economics: growing economies from the bottom up," Artificial life 8(1), pp. 55-82, 2002.

[6] B. Lebaron, "Agent Based Computational Finance: Suggested Readings and Early Research," Journal of Economic Dynamics, Control 24(5-7), pp. 679-702, 2000.

[7] F. Bousquet, C. Le Page, "Multi-agent simulations and ecosystem management: A review," Ecological Modelling 176(3-4), pp. 313-332, 2004.

[8] A. López-Paredes, C. Hernández Iglesias, *Agent Based Modelling in Natural Resource Management*. España: Insisoc.

[9] R. Axelrod, *The complexity of cooperation: Agent based models of competition and collaboration*. Princeton NJ: Princeton University Press, 1997.

[10] R. Paton, R. Gregory, C. Vlachos, J. Saunders, H. Wu , "Evolvable social agents for bacterial systems modeling," *IEEE Transactions on Nanobioscience*, vol 3, pp. 208-216, 2004.

[11] Y. Mansury, T.S. Deisboeck, *Simulating the time series of a selected gene expression profile in an agent-based tumor model*. Physica D: Nonlinear Phenomena 196, pp. 193-204, 2004.

[12] M. Pogson, R. Smallwood, E. Qwarnstrom, M. Holcombe, "Formal agent-based modelling of intracellular chemical interactions," *Biosystems*, vol. 85, pp. 37-45, 2006.

[13] S. Katare, V. Venkatasubramanian, "An agent-based learning framework for modeling microbial growth," *Engineering Applications of Artificial Intelligence,* vol. 14, pp. 715–726, 2001.

[14] S.Y. Eo, T.S. Chang, D. Shin, E.S. Yoon, "Cooperative problem solving in diagnostic agents for chemical processes," *Computers and Chemical Engineering*, vol. 24, pp. 729-734, 2000.

[15] J. Monod, "The growth of bacterial cultures," *Ann. Rev. Microbiol*., vol. 3, pp. 371-394, 1949.

[16] S. Franklin, A. Graesser, "Is it an agent or just a program? A Taxonomy for Autonomous Agents," in J.P Müller, M.J. Wooldridge, N.R. Jennings, "Intelligents Agents III", in *Proceedings of the Third International Workshop on Agent Theories, Architectutes, and Languages*, Berlin, Germany: Springer-Verlag, 1997.

[17] A. López-Paredes, *Ingeniería de Sistemas Sociales. Diseño, Modelado y Simulación de Sociedades Artificiales de Agentes*. Valladolid, España: Servicio de Publicaciones de la Universidad de Valladolid, 2004.

[18] M. Wooldridge and N. R. Jennings, "Intelligent agents: Theory and practice," *The Knowledge Engineering Review*, vol. 10,pp. 10115–152, 1995.

[19] J.H. Holland, *Emergence. From chaos to order*. Reading, MA: Addison- Wesley, 1998.

[20] J. M. Galán, "Evaluación integradora de políticas de agua: modelado y simulación con sociedades artificiales de agentes," Departamento de Ingeniería Civil, Universidad de Burgos, 2007.

[21] Wikipedia, (2011, January 19). Comparison of agent-based modeling software. Available: http://en.wikipedia.org/wiki/Comparison_of_agent-based_modeling_software

[22] A. Getchell (2008). Agent-based Modeling. Department of Physics, University of California, Davis. Available: http://www2.econ.iastate.edu/tesfatsi/AgentBasedModeling.AdamGetchell.phy250.Report.pdf

[23] D.L. Russell , "Practical Wastewater Treatment". John Wiley and Sons: Hoboken, New Jersey, 2006.