International Congress on Environmental Modelling and Software

1st International Congress on Environmental Modelling and Software - Lugano, Switzerland - June 2002

Jul 1st, 12:00 AM

Fuzzy Prediction of the Algal Blooms in the Orbetello Lagoon

Stefano Marsili-Libelli

G. Pacini

C. Barresi

Fuzzy Prediction of the Algal Blooms in the Orbetello Lagoon

S. Marsili-Libelli, G. Pacini, C. Barresi

Department of Systems and Computers, University of Florence
Via S. Marta, 3 – 50139 Firenze, ITALY
Email: marsili@ingfi1.ing.unifi.it

Abstract: The Orbetello lagoon is a shallow brackish waterbody subject to intense and diverse eutrophication (phytoplancton, macroalgae and macrophytes). Periodically a large amount of algae must be artificially removed, their collection and disposal representing a considerable management cost. This paper describes the design of a bloom predictor based on the daily fluctuations of simple water quality parameters such ad dissolved oxygen, oxidation-reduction potential, pH and temperature. The task of the fuzzy predictor is to recognise the possibility that a bloom of the macroalgae population is about to occur based on the changing daily pattern of these variables. The fuzzy predictor is based on a number of fuzzy rules derived from experimental observations and expert knowledge. A whole year of hourly data was analysed and used to form the initial knowledge-base. The tests show that the inferential engine has good predictive capabilities, which could be improved when more data becoming available.

Keywords: *Eutrophication, Fuzzy systems; Pattern Recognition; Decision Support Systems.*

1 INTRODUCTION

Shallow water bodies constitute environmentally sensitive areas and their conservation is becoming more and more important under the ecological and socio-economic point of view. On the other hand, their management is becoming increasingly complex and needs decision support tools in guiding the management of the resources. This paper describes a fuzzy predictor to forecast algal blooms in coastal lagoon and decide the harvesting strategy accordingly.

1.1 - The Orbetello Lagoon

The Orbetello lagoon, is situated along Italy west coast, has an extension of about 27 km^2 and is composed of two coastal ponds with two water inlets at each end of the western pond, and one outlet from the eastern pond, at the bottom-right end. The two ponds communicate though a narrow passage under a bridge connecting Orbetello with Mount Argentario. Natural flow, induced by tide and wind, is negligible and insufficient to provide the necessary turnover. As a result, in the last few years lack of water movement has resulted in a deterioration of water quality leading to anoxic conditions and algal blooms. Two water-quality monitoring stations have been deployed in the middle of each pond and transmit hourly data to the management office in Grosseto, 30 km north of the lagoon. They are indicated by the dots in Figure. 1.

Fig. 1 - Map of the Orbetello lagoon: the dots show the location of the two water quality monitoring stations.

1.2 The algae bloom management problem

Given the negligible water movement and the massive nutrient input, the eutrophication events in the lagoon are basically induced by benthic organic matter degradation occurring in the sewage-enriched sediments (Cartei *et al.,*[1998]). In analysing these bloom, the following facts should be taken into account:

- the nutrient release mechanism from the sediment, controlled by the chemical equilibrium at the water - sediment interface.
- the light competition between algae and macrophytes

Macroalgae growth, if unchecked, may produce a considerable amount of biomass whose decay would constitute a major environmental problem and eventually lead to the silting of the lagoon. For this reason the algae collection should start relatively early, before the disposal of their bio-mass represents a major waste problem. Moreover, harvesting should start when macrophytes are not yet fully developed to avoid cropping their stalks. In fact seagrass are beneficial for two reasons:

- their shallow roots transport oxygen in the upper layer of the sediment helping to maintain the aerobic condition necessary for phosphorus immobilisation;
- the shading provided by their floating leaves limits algae growth.

It should also be noticed that, given the shallow depth of the lagoon, the harvesting boats induce a considerable upwelling of the sediment (Lenzi and Mattei [1998]), where phosphorus is available, given its anoxic condition . When this material is dispersed in the aerobic water column, phosphorus release occurs before it is oxidised, adding to the eutrophic conditions. Again, this fact underlines the importance of optimising the harvesting effort to produce a timely, minimum impact cropping policy. This motivates the development of a predictive tool to support the harvesting decisions.

1.3 The basic interactions in the lagoon ecosystem

The prevailing macroalgae species in the lagoon are *Chaetomorpha linum, Cladophora vagabunda, Ulva rigida, Gracilaria verrucosa.* They are mostly floating and compete for nutrient and light with rooted macrophytes such as *Ruppia* (seagrass). Their growth dynamics are quite different, with macrophytes following a well-defined degree-days cumulative scheme, and macroalgae depending on nutrient availability for their blooms. As a consequence, the macrophytes development is quite predictable, whereas the algae, which are mostly phosphorus-limited, may bloom almost throughout the year, as their growth depends on the chemical equilibrium of the water/sediment interface.

The basic interactions between algae and macrophytes and between the biota and the sediment are shown in Figure. 2. The key factor driving macroalgae development is nutrient availability. In aerobic conditions inorganic phosphorus in the sediment is largely iron- or calcium-bound (Gomez *et al.* [1999]) and is therefore not available for algal growth. However oxygen depletion may occur as a consequence of diminished photosynthetic activity and increased sediment oxygen demand.

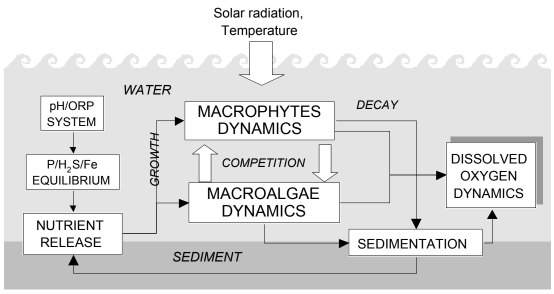

Fig. 2 - Basic interactions between the sediment chemical system and the biotic components (algae and macrophytes).

This situation may result in anaerobic conditions at the sediment interface, with phosphorus release (Cioffi and Gallerano [2000]). This condition is likely to trigger an algae bloom (Gomez *et al.* [1999]) if the amount of phosphorus bound in the sediment is large.

2 A FUZZY INFERENTIAL ENGINE FOR BLOOM RECOGNITION

An inference engine is now designed to detect water conditions favourable to macroalgal bloom. This is particularly difficult since no clear oxic/anoxic pattern can be discerned at any one time during the year, with the bottom remaining always in oxic conditions, though the sediment may occasionally become anaerobic, providing the conditions for P-release.

The idea at the basis of this inferential engine is to observe the daily pattern of easily measurable physico-chemical parameters such as dissolved oxygen (DO), oxidation - reduction potential (ORP), pH and temperature.

Hourly data from October 2000 to October 2001 were analysed. *The basic assumption is that some daily patterns of the chemical parameters are more likely than others to trigger the build-up of algal blooms.* Therefore daily and seasonal variations are considered the key to understanding the combinations of physico-chemical conditions responsible for the blooms.

As an example, it was noticed that before a spring bloom, the DO level is usually rather high and follows the typical daily pattern with an early morning sag and an afternoon peak. On the contrary, during the winter months hardly any daily cycle is apparent. In late summer, the lowest DO concentrations are found, indicating a high oxygen demand from the sediment. Figure. 3 and Figure. 4 compare the daily patterns of DO in spring and in winter.

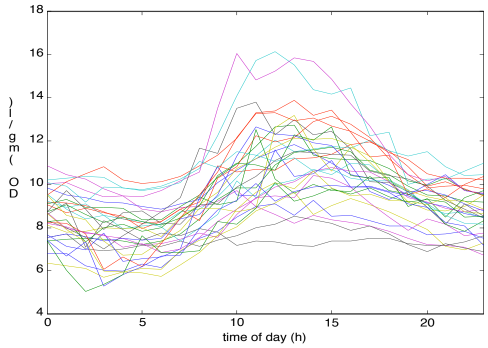

Fig. 3 - Diurnal DO pattern during April 2001.

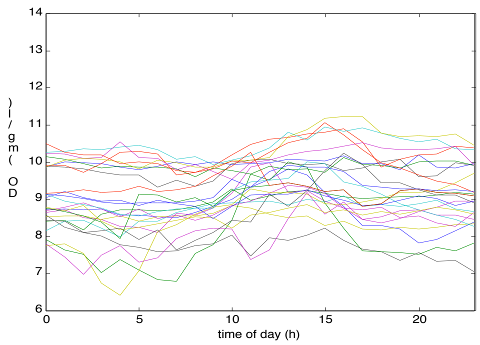

Fig. 4 - Diurnal DO pattern during January 2001.

Figure 4 differs from Figure 3 in the lack of DO photosynthetic increase during the light hours. The corresponding ORP patterns are shown in Figures 5 and 6.

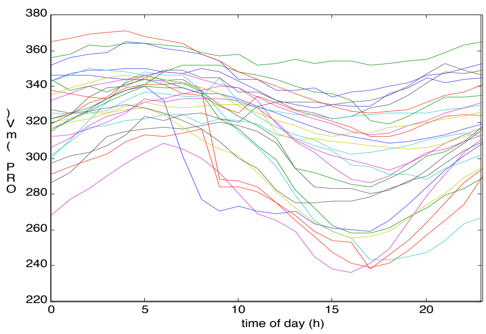

Fig. 5 - Diurnal ORP pattern during April 2001.

From these patterns a representative daily mask was computed for each significant period of the year.

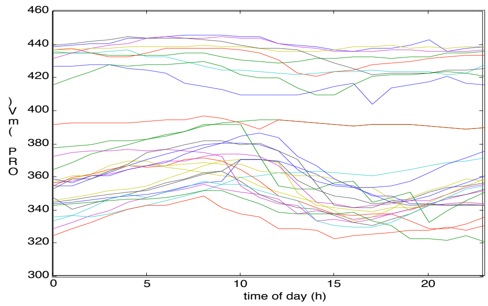

Fig. 6 - Diurnal ORP pattern during January 2001.

It was obtained by fitting a gaussian distribution to the hourly histogram of each variable, as shown in Figure 7 for DO. The same was done for ORP, pH and temperature.

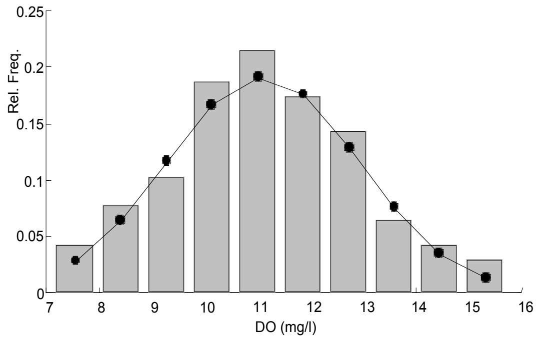

Fig. 7 - An example of fitting a gaussian distribution to the DO histogram (15^{th} hour for each day in April 2001).

After processing the whole day-length (0 - 23 hours) the average profile is determined, together with the upper and lower tolerance limits obtained by cumulating the statistics of Figure 7. In practice the upper and lower 5^{th} percentiles were considered. The result is shown in Figure 8 where the original daily patterns of Figure 3 are superimposed with the average and upper and lower 5^{th} percentiles.

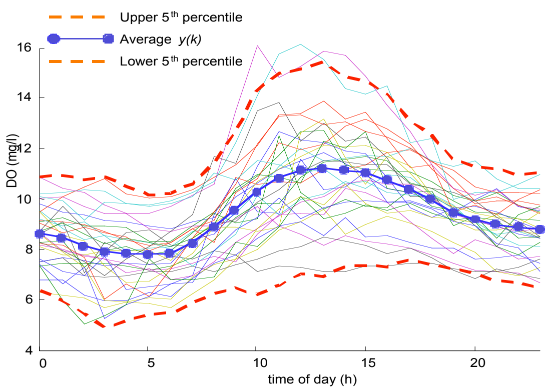

Fig. 8 - The daily patterns of Figure 4 with the superimposed statistical characterization.

The daily average of Figure. 8 was then approximated with a fuzzy curve. This approach to fuzzy modelling was introduced by Lin and Cunningham [1995] and Lin *et al.* [1996] as a way to approximate a set of data with a minimum number of fuzzy rules. In this context the fuzzy curve was adopted to smooth the pattern and to keep into account the hourly correlation during the day. The fuzzy radial basis function (FRBF) on which the fuzzy curve is constructed is a set of gaussian membership function $h(k) \in (0,1)$

$$h_k(x) = exp\left(-\frac{(x - t(k))^2}{2b^2}\right) \tag{1}$$

where x is the time of the day, $\{t(k) = 0, 1, \ldots, 23\}$ represents the k^{th} hour (with membership = 1) and b is a parameter controlling the "spread" of the membership function. Normally this parameter is arbitrarily chosen between 10% and 20% of the variable range (Lin and Cunningham [1995]). However, in this case, a more rational choice is introduced. The partial correlation function of the residuals between the smoothed curve and the actual data was used as a guideline to select b as in a simple AR(1) model. In practice, if ρ is the first sample (lag = 1 hour) of the autocorrelation function, from eq. (1) the relation holds

$$\rho = exp\left(-\frac{1}{2b^2}\right) \tag{2}$$

Solving for b yields

$$b = 1 / \sqrt{2\, ln\left(\frac{1}{\rho}\right)} \tag{3}$$

With the water quality data of the previous figures, it was found that $\rho = 1$, which yields $b = 0.85$. The daily FRBF of eq. (1) with this value of b is shown in Figure. 9.

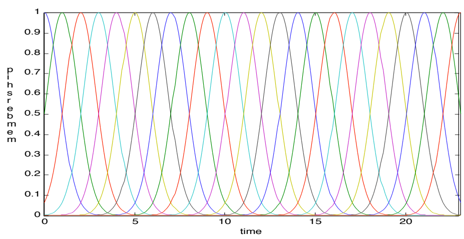

Fig. 9 - Fuzzy radial basis function for the daily fuzzy curve.

Once the FRBF (1) is established, the fuzzy curve is constructed as a Sugeno model (Takagi and Sugeno [1985], Lin and Cunningham [1995]) as the weighted sum of the FRBF over the 24 hours. Combining the time-approximating fuzzy curve (4) with the statistical characterisation of Figure. 8,

$$c(k) = \frac{\sum_{k=0}^{23} t(k) \cdot h_k(k)}{\sum_{k=0}^{23} h(k)} \tag{4}$$

yields the three dimensional fuzzy membership function of Figure. 10, where the gaussian hourly distributions are rescaled to serve as fuzzy membership functions

$$\mu_{DO}(k) = exp\left(-\frac{(DO(t(k)) - c(k))^2}{2 \cdot \sigma^2_{DO}(k)}\right) \tag{5}$$

In this way the fuzzy membership function of the DO daily pattern for April 2001 is obtained: for each hour of the day, the most expected DO value is represented by the fuzzy curve on top of the figure $c(k)$ and any other values has a decreasing degree of truth represented by the transversal gaussian curves controlled by the variance $\sigma^2_{DO}(k)$.

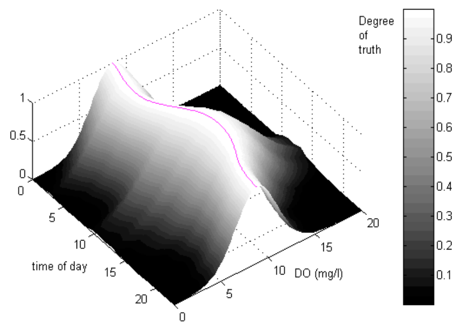

Fig. 10 - Three-dimensional membership function (value vs. daytime) for the daily DO pattern in April 2001.

In this manner a family of membership functions is derived to describe the most relevant behaviours of any particular day. The hourly degrees of truth are averaged over the day to yield the similarity of a generic daily pattern to the predefined behaviours. The idea of applying fuzzy pattern recognition techniques to environmental problems is not new (Marsili-Libelli and Muller, [1996]), but the novelty of this approach is in considering the features of the daily pattern and producing three-dimensional fuzzy memberships. Figure. 11 show the classification mechanism using the membership functions of Figure 10 for a generic daily fuzzy curve. The dots represent the degree of truth for each hour of the day. The example refers to the classification of a winter day DO against the DO spring mask. Obviously the degree of truth is rather low. The average degree of truth over the whole day is then considered for the inferential engine.

2.1 Fuzzy inferential engine

The procedure leading to Figure. 11 (fuzzy curve and pattern classification) represents a recognition tool for the daily pattern. This classifier constitutes the antecedent of a fuzzy inferential engine to yield the desired bloom forecast. To set the whole inference, however, an indication of the algal blooms over the same period (Oct. 2000 - Oct. 2001) is required.

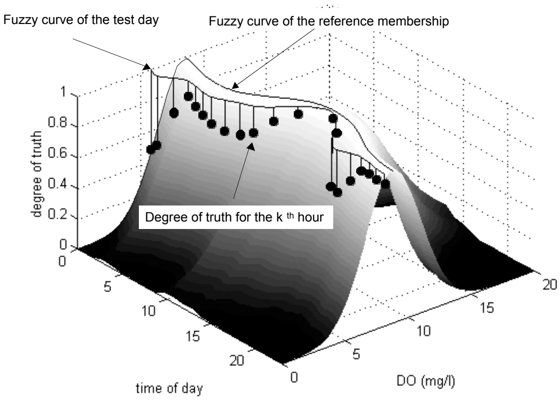

Fig. 11 - Classification of a test day against a three-dimensionalfuzzy mask. The dots represent the degrees of truth.

Only qualitative bloom data are so far available, but they suffice to set-up a fuzzy knowledge-base. From the algae collection records, the bloom behaviour of Table 1 has been reconstructed. The third column shows the partition of these bloom conditions into three classes, labelled *Low*, *Medium* and *High,* to be used in the fuzzy engine. The last entry (Oct. 2001) was left as a validation month, and therefore it was not included in the knowledge base.

Table 1
Algal bloom experimental classification assumed as the knowledge base.

Month	Bloom	Class
Oct. 2000	0.17	Low
Nov. 2000	0.26	Low
Dec. 2000	0.40	Medium
Jan. 2001	0.50	Medium
Febr. 2001	0.55	High
March 2001	0.67	High
April 2001	1.00	High
May 2001	0.80	High
July 2001	0.17	Low
Aug. 2001	0.07	Low
Sept. 2001	0.16	Low
Oct. 2001	0.56	Validation

The daily patterns of a particular class were grouped and processed together to form the three knowledge base prototypes for each parameter. This information was represented by the hourly mean $m_i(k)$ and variance $\sigma_i^2(k)$ of the fuzzy gaussian membership functions,

$$\mu_i(k) = exp\left(-\frac{\left(f(k)-m_i(k)\right)^2}{2\cdot\sigma_i^2(k)}\right) \qquad i='high','medium','low' \tag{6}$$

By inspection of the yearly bloom behaviour and the information extracted from the data, the set of inferential rules in Table 2 were established. The function *f(k)* appearing in eq. (6) is the fuzzy curve of the generic day being classified with respect to the *'high'*, *'medium'*, and *'low'* prototypes. The rationale behind the rules is quite simple: if all parameters are classified as *low*, then the bloom likelihood is also *low*, and the same applies for the *medium*, and *high* conditions.

Table 2
Rules for the fuzzy bloom predictor

Rule #1	Rule #2	Rule #3
IF DO is DO _low	IF DO is DO _medium	IF DO is DO _high
AND ORP is ORP _low	AND ORP is ORP _medium	AND ORP is ORP _high
AND pH is pH _low	AND pH is pH _medium	AND pH is pH _high
AND T is T _low	AND T is T _medium	AND T is T _high
THEN Bloom = B low	THEN Bloom = B medium	THEN Bloom = B high

The hourly degrees of truth provided by eq. (6) are averaged over the day length before entering the rules of Table 2

$$\bar{\mu}_i^j = \frac{1}{24}\sum_{k=0}^{23}\mu_i^j(k) \qquad \begin{cases} j = DO, ORP, pH, Temp \\ i='low','medium','high' \end{cases} \tag{7}$$

which then become

$$\begin{aligned} R_1 &: \bar{\mu}_{low}^{DO} \wedge \bar{\mu}_{low}^{ORP} \wedge \bar{\mu}_{low}^{pH} \wedge \bar{\mu}_{low}^{Temp} \rightarrow B_{low} \\ R_2 &: \bar{\mu}_{medium}^{DO} \wedge \bar{\mu}_{medium}^{ORP} \wedge \bar{\mu}_{medium}^{pH} \wedge \bar{\mu}_{medium}^{Temp} \rightarrow B_{medium} \\ R_3 &: \bar{\mu}_{high}^{DO} \wedge \bar{\mu}_{high}^{ORP} \wedge \bar{\mu}_{high}^{pH} \wedge \bar{\mu}_{high}^{Temp} \rightarrow B_{high} \end{aligned} \tag{8}$$

The *AND* symbol $\wedge$ may be implemented with any triangular norm and $\rightarrow$ denotes the fuzzy implication *THEN* operator, which should also be a T-norm operator. In this study the *product* operator was found to be the most reliable T-norm for both. The resulting bloom likelihood for each month is obtained by Sugeno defuzzification, i.e. averaging the bloom singletons B_i with their respective degrees of truth.

$$Bloom = \frac{\sum_i \mu_i \cdot B_i}{\sum_i \mu_i} \qquad i='low','medium','high' \tag{9}$$

The three bloom consequents appearing in eqs. (8) and (9) $\boldsymbol{B} = \left(Bloom_{low}, Bloom_{medium}, Bloom_{high}\right)$ are in principle unknown and must be estimated in order to adapt the algorithm to the knowledge-base. This was done by minimising the performance index (10) over the eleven month test period

$$B^* = \arg\min_{B} \; E(B) = \sum_{m=1}^{11} \left(\frac{Bloom(m) - Bloom_{exp}(m)}{Bloom_{exp}(m)} \right)^2 \quad (10)$$

The minimisation of (10) was obtained through a modified version of the Simplex search method (Marsili-Libelli [1992]) and the result is shown in Table 3 and Figure. 12.

Table 3
Estimated values of Bloom consequents

$Bloom_{low}$	$Bloom_{medium}$	$Bloom_{high}$
-0.213	0.587	2.134

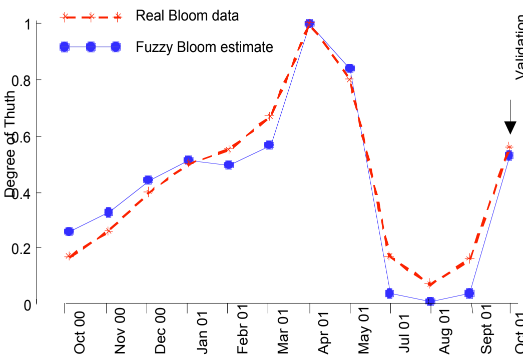

Fig. 12 - Comparison between observed and predicted blooms. The uncalibrated Oct. 2001 observation was used for validation.

It can be concluded that the proposed fuzzy inferential engine can indeed recognise the bloom conditions with acceptable accuracy, at least given the limited knowledge available. Its consistency is confirmed by the good performance with the validation month (Oct. 2001) where a secondary bloom occurred. The authors are well aware that few validation data have been used, but the measurement campaign has just started and only these data are presently available.

3 CONCLUSION

The Orbetello lagoon is subject to frequent, seemingly unpredictable blooms of macroalgae which represent a major environmental problem. Its management relies on early harvesting of the biomass and a forecasting tool is required to assist in this. This paper has presented the design of a fuzzy bloom estimator based on simple water quality parameters. The novelty of the algorithm is in the combination of daily patterns and the fuzzy significance of its statistical properties. The result is a set of three-dimensional fuzzy member-ship functions which are used in a Sugeno inferential engine to estimate the bloom likelihood given the current physico-chemical premises. The engine has been calibrated with merely qualitative data obtained from the lagoon administration and will be refined when more bloom data become available. But still from this preliminary version, its estimation capabilities appear adequate to forecast a bloom and support a pre-emptive cropping effort.

4 Acknowledgements

This research is supported by the Committee for the Management of the Environmental Emergency in the Orbetello Lagoon, under the supervision of the Tuscany Regional Government. The authors wish to thank Dr. Mauro Lenzi, biologist with the Laboratory of Lagoon Ecology and Aquaculture (LEALab) Orbetello, and Prof. Bendoricchio, University of Padova, for their support.

5 References

Cartei, P., Innamorati, M., Melillo, C., Trophic homeostasis sea modulated and lagoon autochtonous hypertrophy, *Biologia Marina Mediterranea*, **5**(1), 41 - 46, 1998.

Cioffi, F. and F. Gallerano, Response of Lake Piediluco to the change of hydrodynamic conditions and nutrient load reductions, *Ecological Modelling* **135**, 199 - 229, 2000.

Gomez, E., Durillon, C., Rofes, G. and B. Picot, Phosphate adsorption and release from sediments of brackish lagoons: pH, O_2 and loading influence, *Water Research,* **33**(10), 2437 - 2447, 1999.

Lenzi, M., and Mattei, M., Risultati di quattro anni di gestione dell'ecosistema lagunare di Orbetello attraverso la raccolta delle biomasse algali, *Biologi Italiani*, **2**, 7 - 12, 1998.

Lin, Y., Cunningham III, G.A., A new approach to fuzzy-neural system modeling, *IEEE Trans. Fuzzy Systems,* **3**(2), 190 - 198, 1995.

Lin, Y., Cunningham III, G.A., Coggeshall, S.V., Input variable identification - Fuzzy curves and fuzzy surfaces, *Fuzzy Sets and Systems,* **82**, 65 - 71, 1996.

Marsili-Libelli, S., Parameter estimation of ecological systems, *Ecological Modelling,* **62**, 233 - 258, 1992.

Marsili-Libelli S. and Müller A., Adaptive fuzzy pattern recognition in the anaerobic digestion process. *Pattern Recognition Letters*, **17**, 651 - 659, 1996.

Takagi, T., Sugeno, M., Fuzzy identification of systems and its applications to modeling and control, *IEEE Trans. Systems, Man, and Cybernetics,* **15**(1), 116 - 132, 1985.