Including Expert Knowledge in Finite Element Models by Means of Fuzzy Based Parameter Estimation

Oliver Krol*,[1], Nicolas Weiss[1], Thomas Bernard[1] and Felix Sawo[1]
[1]Fraunhofer Institute for Information and Data Processing IITB
*Corresponding author: Fraunhoferstr. 1, 76131 Karlsruhe, Germany, Oliver.Krol@iitb.fraunhofer.de

Abstract: In this paper we present a novel approach for modelling spatial distributed biochemical and environmental processes like the growth of plants and the related biochemical reactions. One of the main challenges for modelling of spatial distributed phenomena is the estimation of the model parameters. The physical phenomena like flow and mass transport can be described by PDEs of fluid dynamics, but for effects like growth rates often no analytic models are available. However, in many cases experts have knowledge about the system behaviour that can be formulated by a set of *if-then*-rules. As this kind of knowledge can easily be handled by so-called Fuzzy models we propose the coupling of FEM models with such Fuzzy models. By this means one or more parameters of the classical PDEs are estimated by Fuzzy Models. Besides the natural inclusion of expert knowledge a second benefit of this approach consists in the fact that Fuzzy Models can describe even very nonlinear phenomena. The proposed approach is applied for modelling the growth of algae of Orbetello lake in Italy.

Keywords: Finite Element Method, Fuzzy Model, Expert Knowledge, Eutrophication

1 Introduction

One of the main challenges for modelling of spatial distributed phenomena is the estimation of the model parameters. E.g. in a lot of biochemical or biophysical applications the physical phenomena like flow and mass transport can be described by PDEs of fluid dynamics, but for parameters like growth rates often no analytic models are available or the existing models require a large number of parameters which are hard to determine. However, in many cases experts have knowledge about the system behaviour that can be formulated by a set of *if-then*-rules. As this kind of knowledge can easily be handled by so-called Fuzzy models [11] we propose the coupling of finite element models with such Fuzzy models. By this means one or more parameters of the classical PDEs are estimated by Fuzzy Models. Besides the natural inclusion of expert knowledge a second benefit of this approach consists in the fact that Fuzzy Models can describe even very nonlinear phenomena.

2 Concept of Fuzzy based Parameter Estimation

The present concept assumes that spatially distributed models represented by PDEs (e.g. Navier-Stokes equations) can be numerically solved by finite element method. Parameters that can not be calculated by analytic models (like static or ordinary differential equations (ODEs)) can often be approximated by means of expert knowledge that is incorporated by *if-then*-rules. With an inference scheme of Fuzzy Logic [11] the expert knowledge can be transformed into physical crisp values. This approach allows the maximal use of expert knowledge and reduces limitations due to missing information and parameter values. Figure 1 shows the structure of a coupled Fuzzy and FEM model where the output variables of the FEM model is returned to the Fuzzy model as input variables. This model structure in general yields a nonlinear FEM model.

The calculation of the Fuzzy model consists in three steps [11]: 1. Fuzzyfication of input parameters, 2. inference of *if-then*-rules and 3. defuzzyfication. Within step one the considered input/output variables are transformed to linguistic variables, e.g. *small, medium, large* and maximum ranges of the considered variables are defined. In the inference unit (step two) the actual processing of the linguistic variables is performed whereby *if-then*-rules have to be formulated by experts that describe the interdependence of the input parameters to the output parameter(s). The defuzzyfication (step three) generates crisp physical output values of the Fuzzy model.

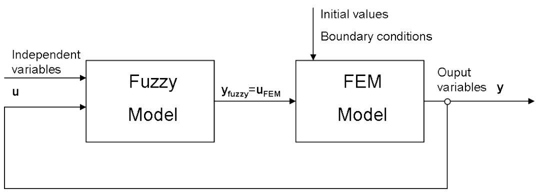

Figure 1: Structure of coupling spatially distributed FEM-models and Fuzzy models.

The concept of coupled Fuzzy and FEM models is demonstrated by the modelling and simulation of algae growth in flat water bodies and related eutrophication effects. The proposed approach is very suited for this problem as the hydrodynamics is well understood whereas the biochemical processes are still a topic of research and the developed models are very complex and require a big number of parameters [2]. Since in general the required parameters (e.g. growth rates) are not measurable for the entire model domain the idea raised to estimate them by taking the main variables (e.g. temperature, fluid velocity, nutrients concentration) into account and estimate the algae growth rate by Fuzzy model. The next section summarises the main results of our studies.

3 Case Study: Algae Growth

3.1 Eutrophication in flat water bodies

Effluents and alluvial deposits from agrarian areas transport great quantities of chemical nutrients that can easily reach nearby water bodies causing an increase in the concentration of nitrogen and phosphor compounds. A direct consequence of high concentrations of these elements can be observed in an excessive growth of algae. Its decomposition is an intensive oxygen consuming biochemical reaction that decreases the dissolved oxygen level concentration. The main factors driving algae growth are nutrient concentration, water temperature and light incidence [3], [9]. Additionally hydrodynamics play an important role since the mass distribution in flat water bodies is mainly determined by the fluid velocity [6]. Depending on the level of eutrophication the degradation of water quality can have harmful consequences for the ecosystem. If the dissolved oxygen concentration falls under a certain level, the decomposition of organic material via aerobe bacteria ceases, ammonia, methane and hydrogen sulphide concentrations produced by anaerobe bacteria rises, having deadly consequences for fauna and flora [8].

High levels of eutrophication can be avoided by artificially removing algae from the water, however, this measure raises important costs for the responsible agencies [7]. For a sophisticated water course management detailed information about where algae growth has to be expected is of great value. Therefore a detailed model that covers the underlying hydrodynamics as well as the biochemical process is required to achieve realistic predictions of such areas

In the following sections we will present a hydrodynamic model which was modular implemented in COMSOL using the Chemical Engineering Module (incompressible Navier-Stokes application mode with transport phenomena) and a Fuzzy model which contains the expert knowledge for the description of algae growth in MATLAB. Additionally we will describe the coupling of both models, achieving a novel approach for modelling spatial distributed and environmental processes. We will conclude with the results of an application example on the Orbetello Lagoon in Italy.

3.2 Hydrodynamic Model

The behaviour of a Newtonian fluid (velocity and pressure) is described by the Navier-Stokes equations (1) - (2). These build a system of coupled nonlinear partial differential equations that can be solved approximately by numerical methods like FEM.

$$\frac{\partial \rho}{\partial t} + \mathrm{div}(\rho \vec{u}) = 0 \quad (1)$$

$$\frac{\partial \vec{u}}{\partial t} + \vec{u} \cdot \nabla \vec{u} = \frac{1}{\rho} \nabla p - \vec{g} + \eta \nabla \mathrm{div} \vec{u} \quad (2)$$

$$\frac{\partial c_{\mathrm{B}}}{\partial t} + \vec{u} \cdot \nabla c_{\mathrm{B}} = \nabla(D_{\mathrm{diff}} \nabla c_{\mathrm{B}}) + \nabla(D_{\mathrm{disp}} \nabla c_{\mathrm{B}}) + R_{\mathrm{B}} \quad (3)$$

$$\frac{\partial c_{\mathrm{N}}}{\partial t} + \vec{u} \cdot \nabla c_{\mathrm{N}} = \nabla(D_{\mathrm{diff}} \nabla c_{\mathrm{N}}) + \nabla(D_{\mathrm{disp}} \nabla c_{\mathrm{N}}) + R_{\mathrm{N}} \quad (4)$$

In addition to the Navier-Stokes equations the transport of mass by a fluid can be described with the transport equations (3 - 4), as we can observe this is coupled via the fluid velocity with the Navier-Stokes Equations. The transport equations are presented here in a general form accounting for advection, diffusion and dispersion as transport mechanisms. However since the dominant transport mechanism in flat water bodies is advection [6], the diffusion and dispersion coefficients have been neglected during our studies. The subindices B and N stand for biomass (algae) and nutrients, these will be distributed depending on the growth and depletion rates R_B, R_N, and the fluid velocity $\vec{u}$ over the simulation model.

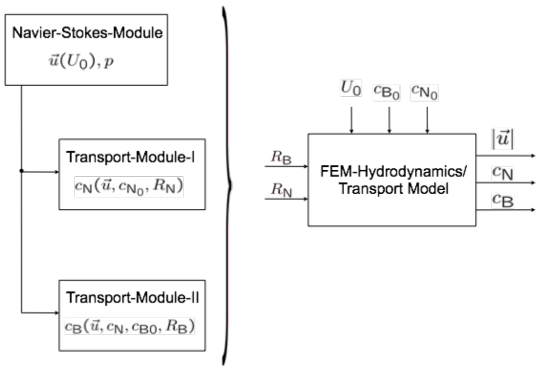

Figure 2: Interdependencies of the different simulation modules in COMSOL. $\vec{u}$ represents the fluid velocity, p the pressure, c_N the nutrients concentration, c_B the biomass concentration, R_N a nutrient depletion rate, R_B a biomass growth rate, and U_0, c_{N_0} and c_{B_0} the initial conditions for fluid velocity, nutrient and biomass concentration.

Equations (1)-(4) were modular implemented using the COMSOL Navier-Stokes and Transport Modules [1]. Figure 2 describes the variables that are simulated in the different modules and their interdependencies. The biomass growth rate R_B in equation 3 will be responsible for the coupling of the FEM hydrodynamic model and the expert system simulating algae growth. Additionally there will be a natural interaction between nutrients and biomass that will be represented by coupling both transport equations. The coupling of both models will be described in detail in section 3.4.

Since the aim of our studies was the exploration of a Fuzzy-FEM approach to improve spatial distributed modelling and not the exact simulation of flow-biological interactions, several model simplifications were established. First, nutrients and biomass were condensed to two magnitudes c_N and c_B that represent the concentrations of nutrients and algae in the water. Second, we confine our simulations to algae species that grow at the surface of water bodies (e.g. Nodularia Spumigena [8]) and to flat water bodies (i.e. circulation processes can be neglected [6]), simplifying a 3-dimensional problem to a 2-dimensional one. Last, we restrict our simulation to incompressible fluids.

3.3 Fuzzy Model

In the previous section we presented a hydrodynamics model that simulates the behaviour of flow and transport phenomena. The aim of this work is to introduce expert knowledge in the previous model. In order to achieve this we will present in this section a model based on expert knowledge and fuzzy logic to simulate algae growth, that was implemented using the MATLAB Fuzzy Logic Toolbox.

The first and perhaps most important step in the development of an expert system is knowledge acquisition [11]. In biology as well as in other scientific disciplines this is not an easy task to accomplish, since normally there is a great number of parameters that have to be taken into consideration and expert knowledge is not easily available. For our work we adapted a version of the HABES expert system that was developed as part of the EU Fifth Framework Programme. HABES is a Harmful Algal Blooms Expert System which aims to improve the understanding on the interaction between relevant factors for algae growth, particularly the interaction between physical and biological factors [10].

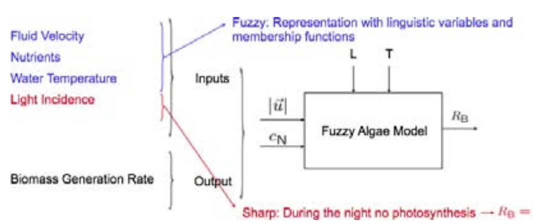

Figure 3: Fuzzy model for the simulation of algae growth. Inputs comply with the influence parameters defined in section 3.1. Output is a biomass growth rate that will later be coupled with equation (3).

Input	Linguistic Variable	Range
L	Light	[no Light , Light]
T [$^\circ C$]	Temperature	[5 , 30]
$\lvert\vec{u}\rvert$ [$\frac{m}{s}$]	Fluid Velocity	[0 , 2]
c_{N} [$\frac{g}{m^3}$]	Nutrients	[0 , 1]
Output	Linguistic Variable	Range
R_{B} [$\frac{g}{s \cdot m^3}$]	Biomass Growth Rate	[0 , 0.1]

Table 1: Fuzzy Variables.

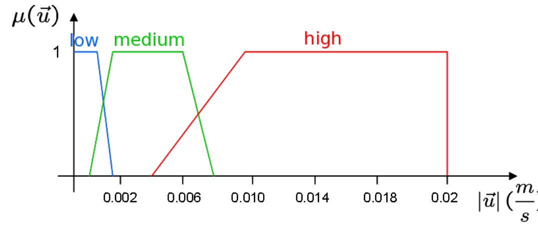

Figure 4: Membership function for the fluid velocity $|\vec{u}|$.

If	*temperature* = *high*,
and	*fluid velocity* = *low*,
and	*nutrients* = *medium*,
then	*growth rate* = *medium*.

Table 2: Example of an *if-then*-rule.

The model inputs are according to the influence parameters previously defined in section 3.1. These are light incidence, water temperature, fluid velocity and nutrients concentration. The output magnitude is a biomass growth rate R_{B} that will later serve as an input of the hydrodynamic model, more specifically of the Transport Equation. A block diagram of the fuzzy algae model is shown in Figure (3) and a summary of inputs, outputs, associated linguistic variables and ranges is listed in Table (1). Every model input except for light influence will be described with a certain membership value of the following three linguistic values: *low*, *medium* and *high*. For simplification reasons light influence will be represented as a binary magnitude (i.e. light or no light). As long as the simulation time is less than a predefined day length time the biomass growth rate R_{B} will be determined via fuzzy inference of the remaining inputs, in the other case it will be set equally to zero. This is based upon the fact that during the night hours no photosynthesis is possible [3].

Two elements are still missing to complete the knowledge base. First, the membership functions that map the sharp value of a given input into a membership value of *low*, *medium* and *high*, an example for the fluid velocity $|\vec{u}|$ is shown in Figure (4). Finally, a collection of 27 *if-then*-rules (out of the combination of three linguistic values) that describes the relationships between the different inputs and output fulfil the knowledge base, an example is shown in Table (2). During simulations the calculation of the growth rate R_{B} is calculated by interpolating non linear characteristic maps that result from the fuzzy modelling. In Figure (5) an example for a characteristic map of R_{B} against water temperature T and nutrient concentration c_{N} is shown.

At this point it is important to clarify where each of the inputs of the expert system is generated. The fluid velocity $|\vec{u}|$ and the nutrient concentration c_{N} are calculated in every simulation step by the hydrodynamic model described in section 3.2. The remaining inputs (light incidence L and water temperature T) have been implemented as look-up tables. This means that for any simulation time (e.g. a summer or winter day) the system can determine an approximate value of the light incidence and water temperature. The first variable is based on water temperature measurements of the Orbetello Lagoon [4] and the values for light incidence were determined via a simple trigonometric model [5].

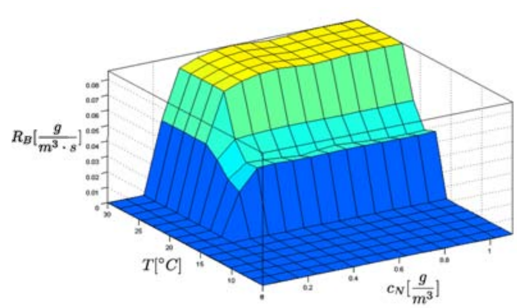

Figure 5: Non linear characteristic map for the biomass growth rate R_{B} against water temperature T and nutrient concentration c_{N}.

In the past two sections we presented a hydrodynamic model implemented in COMSOL for the simulation of a flat water body based on the Navier-Stokes-Equations and an expert system implemented in MATLAB for the simulation of algae growth. It is the main goal of the following section to discuss in detail the combination and simulation of both models.

3.4 Model Coupling

The first step in coupling both models is the aggregation of the hydrodynamic model in Figure (2) and the fuzzy algae model in Figure (3) to a complete model. It is known from the previous sections that the fuzzy algae model requires the fluid velocity $|\vec{u}|$ and the nutrient concentration c_{N} for the calculation of the growth rate R_{B}. On the other hand the hydrodynamic model demands the biomass growth rate R_{B} and the nutrient depletion rate R_{N} in Equations (6) and (7) (note that these are the simplified Transport Equations with only advection as a transport mechanism) for the determination of the current values of the nutrient concentration c_{N} and biomass concentration c_{B}.

$$\frac{\partial c_{\mathrm{N}}}{\partial t} = -\vec{u} \cdot \nabla c_{\mathrm{N}} + R_{\mathrm{N}}(t) \qquad (5)$$

$$\frac{\partial c_{\mathrm{B}}}{\partial t} = -\vec{u} \cdot \nabla c_{\mathrm{B}} + R_{\mathrm{B}}(t) \qquad (6)$$

$$R_{\mathrm{N}}(t) = -k \cdot R_{\mathrm{B}}(t) \qquad (7)$$

Additionally nutrients and biomass are also naturally coupled, i.e. in areas where biomass is produced (i.e. c_{B} increases) nutrients are depleted (i.e. c_{N} decreases). For the sake of simplification and since the main objective of our work lies not in the exact modelling of biological processes, we expressed this last interaction via a linear relationship in Equation (7). The complete system diagram is shown in Figure (6).

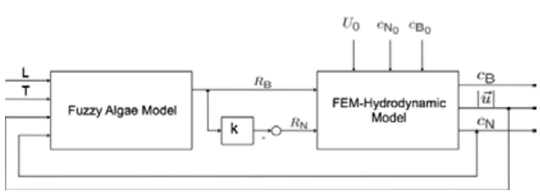

Figure 6: Complete system diagram.

On the software level coupling is based on the fact that COMSOL allows almost any function to be define and called during the simulation cycle [1]. For this purpose we implemented a MATLAB-function that receives as inputs the fluid velocity $\vec{u}$, the nutrient concentration c_{N} and simulation time t generated in COMSOL and has implemented the look-up tables for the remaining inputs light incidence L and water temperature T. It returns via execution of the MATLAB *evalfis*-function the value of the growth rate R_{B} based on the current inputs and the knowledge base defined in section 3.3.

3.5 Results

The Orbetello Lagoon exhibits a strong eutrophic state, for this reason and the fact that it has been the subject of several studies dealing with this problem [4], [7], we selected it as a demonstration example of the developed model. The lagoon is located in the west coast of Italy, it has a surface of 27 km^2 and is composed of two basins with inlets in the north-west and south-west corners. The outlet is located in the south-east corner. Both basins are connected via a small passage under a bridge that joins Orbetello with Monte Argentario. Its adjacent area is characterised by an intensive agricultural activity that discharges waters rich in nutrients into the lagoon, this having as a direct consequence an excessive growth of algae in the water surface [7]. In Figure (7) a chart of the lagoon is shown.

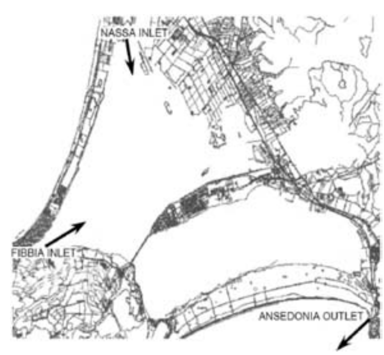

Figure 7: Chart of the Orbetello Lagoon from [7].

The simulation cycle is in accordance to the one described in section 3.4. The parameters for water temperature T and light incidence L have been fixed to values encouraging algae growth (i.e. warm water and long day length). First the fluid velocity $|\vec{u}|$ was simulated, the steady state profile is shown in Figure (8). It equates the structure of the lagoon with two inlets and one outlet as described in Figure (7) and one can clearly differentiate between areas with high and slow flow velocities. In areas with slow flow velocity algae growth will be expected. At this point it has to be noted that the parameterisation of the model does not reflect the exact nature of the lagoon since a scale factor has been introduced in order to avoid numerical instabilities (e.g. convergence) due to bad condition of the matrixes. This is induced by the fact that flow velocity in the Orbetello Lagoon is in the order of a few $\frac{cm}{s}$ and has often been neglected [7].

Secondly the nutrient concentration c_{N} with $R_{\mathrm{N}} = 0$ (i.e. the natural interaction between nutrients and biomass is neglected) was simulated. The steady state nutrient profile is shown in Figure (9) and it equates the previously simulated flow profile. In contrast to fluid velocity, algae growth will be expected in areas with high nutrient concentration.

At this point biomass concentration c_{B} and nutrient concentration c_{N}, related by $R_{\mathrm{N}} = -k \cdot R_{\mathrm{B}}$, were simulated and the results are shown in Figures (10) and (11) accordingly. A high algae concentration is observed in areas where the fluid velocity $|\vec{u}|$ is relative slow and the nutrient concentration c_{N} is relative high. This shows that expert knowledge has been successfully transferred to the FEM simulation of the hydrodynamic model. We have to notice here that the minimum of the mass concentrations is negative that of course contradicts the obvious fact that an element mass cannot be less than zero. However, reasons for this are given by inaccuracy of the FEM simulation and relies on the discontinuous nature of the boundary conditions for the nutrients and biomass concentration.

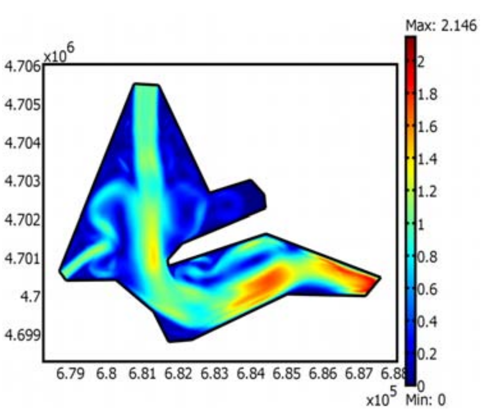

Figure 8: Fluid velocity $|\vec{u}|$ $[\frac{m}{s}]$.

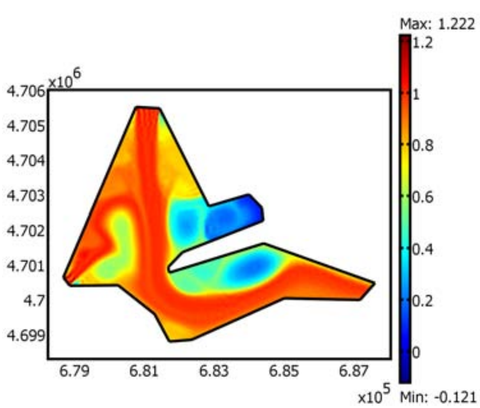

Figure 9: Nutrient concentration c_{N} $[\frac{g}{m^3}]$, $R_{\mathrm{N}} = 0$.

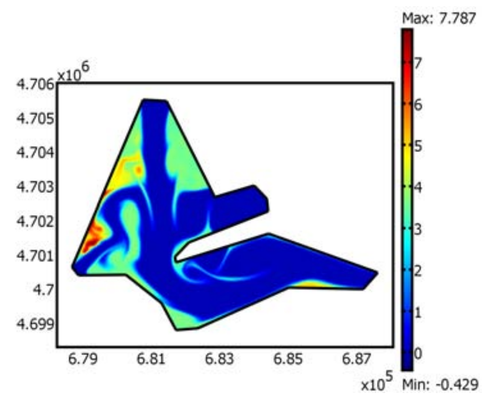

Figure 10: Biomass concentration c_{B} $[\frac{g}{m^3}]$.

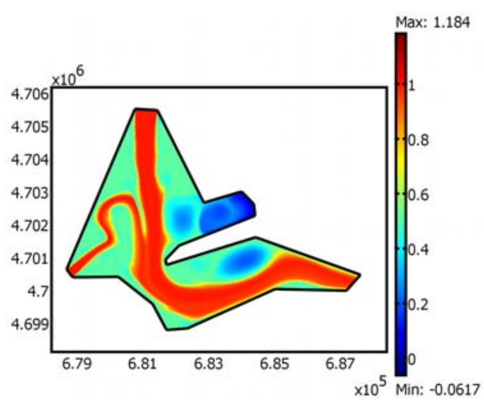

Figure 11: Nutrient concentration c_{N} $[\frac{g}{m^3}]$, $R_{\mathrm{N}} = -k \cdot R_{\mathrm{B}}$.

4 Conclusions

In this paper we presented a generic approach for modelling spatial distributed biological and environmental processes by introducing expert knowledge into finite element simulations. A hydrodynamic model was developed in COMSOL and an expert system based on fuzzy logic and a knowledge base was implemented in MATLAB. With the help of a coupling function both models were combined, showing promising results in the simulation of algae blooms in flat water bodies.

The developed model has a generic nature in the sense that it has not been tuned for a specific specie of algae. Furthermore, the developed approach can be applied to each spatially distributed problem where parameters can only be described by expert knowledge. Hence further research will focus on the application of the method to further processes.

References

[1] *Comsol multiphysics user's guide*, 2007.

[2] G. Coffaro and A. Sfriso, *Simulation model of ulva rigida growth in shallow water of the lagoon of venice*, Ecological Modelling **102** (1997), 55–66.

[3] R. H. Foy, C. E. Gibson, and R. V. Smith, *The influence of daylength , light intensity and temperature on the growth rates of planktonic blue-green algae*, Br. phycol. J. **11** (1976), 151–163.

[4] E. Giusti and S. Marsili-Libelli, *Modelling the interactions between nutrients and the submersed vegetation in the orbetello lagoon*, Ecological Modelling **184** (2005), 141–161.

[5] M. Ligr, *Calculation of the photoperiod length*, CABIOS **11** (1995), no. 2, 133–139.

[6] A. Malchereck, *Hydromechanik der oberflächengewässer*, Tech. report, Bundesanstalt fuer Wasserbau, Aussenstelle Kueste, 1995.

[7] S. Marsili-Libelli, *Fuzzy prediction of the algal blooms in the orbetello lagoon*, Environmental Modelling & Software **19** (2004), 799–808.

[8] S. Nehring, *Giftalgen - die cyanophycee nodularia spumigena*, Mikrokosmos **80** (1991), 321–327.

[9] C. Sorokin and R. W. Krauss, *The effects of light intensity on the growth rates of green algae*, Plant Physiology **33** (1958), 109–113.

[10] Harmful Algal Blooms Expert System, *http://www.wldelft.nl/proj/pdf/3uk00204 .scherm.pdf*, Tech. report, Delft Hydraulics, 2006.

[11] T. Takagi and M. Sugeno, *Fuzzy identification of systems and its application to modeling and control*, IEEE Transactions on Systems, Man and Cybernetics **15** (1985).