A Qualitative Modeling Approach to Algal Bloom Prediction

Ulrich Heller
Dept. of Computer Science
Technical University of Munich
Orleansstr. 34, 81667 München, Germany
email: heller@informatik.tu-muenchen.de

Peter Struss
Dept. of Computer Science
Technical University of Munich
Orleansstr. 34, 81667 München, Germany
email: struss@informatik.tu-muenchen.de

François Guerrin
Unité de Biométrie et d'Intelligence Artificielle
INRA Toulouse
Castanet-Tolosan, France
email: guerrin@toulouse.inra.fr

Waldir Roque
Instituto de Matemática
Federal University of Rio Grande do Sul
Porto Alegre, Brazil
email: roque@mat.ufrgs.br

Abstract

Some problems typical for modeling ecological systems were encountered in the initial phase of a project that is concerned with the Rio Guaíba in Southern Brazil. This paper presents first results in modeling the phenomenon of algal bloom in a qualitative process-oriented modeling language. We addressed two problems. First, the spatial distribution of parameters and processes has to be taken into account which leads us to locate processes in or between compartments, the elements of a topological partitioning of the area. Second, the various processes involved act with speeds of different orders of magnitude (e.g. chemical reactions vs. changes in fish population) which requires techniques of time-scale abstraction.

1 Introduction

Knowledge-based systems supporting analysis and decision making in the environmental domain require a representation of our knowledge about the involved processes. Because of the very nature of these processes, our knowledge about them, and the information available, this is a great challenge for qualitative modeling.

In a collaboration between researchers from Brazil, Germany and France, development of model-based systems has been started that are targeted to support decision making concerning environmental problems of the Rio Guaíba in Southern Brazil. In this project, we started modeling the phenomenon of algal bloom. This paper presents some preliminary results of this work, a process-oriented description of some of the essential mechanisms contributing to algal bloom.

In particular, two problems have to be addressed that are typical for modeling ecological systems. First, the spatial distribution of parameters and processes has to be taken into account which leads us to locate processes in or between compartments, the elements of a topological partitioning of the area. Second, the various processes involved act with speeds of different orders of magnitude (e.g. chemical reactions vs. changes in fish population), which requires techniques of time-scale abstraction.

In the following sections, we give a brief description of the Rio Guaíba and the environmental issues raised, with a focus on the algal bloom phenomenon, which will be further analyzed in section 3. Section 4 presents our approach to modeling the interactions involved in this phenomenon in a process-oriented language, QPC (4.1), the handling of compartments and their interaction (4.2), and the application of time-scale abstraction in the composition of a scenario model (4.3). Finally, we discuss some open issues and tasks for future work.

2 The Rio Guaíba and the Algal Bloom Phenomenon

The Rio Guaíba passes the city of Porto Alegre, the capital of the southernmost state of Brazil. Calling it a river is just a convention, since its bays broaden up to 40 Km, while it stretches in its entire length only some 100 Km, thus partially behaving like a lake. Nevertheless, the flow in the center, called the navigation channel, reaches a speed of more than 1000 m^3/s, which is caused by the Rio Jacui, the main affluent of the Rio Guaíba. Further complicated by the wind conditions and the connection to a large lagoon (Lagoa dos Patos), that is coupled with the ocean, the hydrodynamics within the water body is rather complex and some phenomena like the temporary inversion of the flow, the so-called "reflux", are still not satisfactorily examined.

There are multiple sources of pollution, mostly city sewage of more than 1.2 million inhabitants and some industrial waste water of chemical plants and factories nearby. This adds to the organic pollution of the four affluent rivers, which drain industrial and agricultural regions. Because of the threat for the drinking water supply for the city and the increasing and severe dangers for the health of people swimming in bays near the city, the municipal department of water and sewage (DMAE) has been monitoring various parameters of the ecosystem in a number of locations over a period of more than 15 years.

However, the administration is mostly concerned with the suitability of the river for drinking water treatment, and therefore concentrates on a subset of chemical constituents and the important number of fecal coliform bacteria stemming from household sewage, which establishes the main danger of diseases by contact with or drinking of the water.

Only recently, the general impact on the ecosystem has been studied, and, alarmed by the repeated occurence of local algal blooms, the city department started to probe and monitor the different species of algae in the phytoplankton and also in the sediment layers.

An algal bloom is a significant and steep increase in total algal biomass, while at the same time one species of algae becomes significantly predominant. The number of organisms per liter, which is about 1000 under oligo- and mesotrophic conditions (i.e. very low or average level of nutrients, respectively) can increase up to 1 million organisms/liter during the bloom. Typically, the single predominant species covers more than 95 percent of the biomass.

In the Rio Guaíba, such phenomena have been observed in various locations with slow flow speeds, low water levels and the hydrodynamic conditions for the accumulation of nutrients. The occurences typically lasted between 4 and 10 days and were limited to a spatial extent of about one square kilometer. All observed incidents happened in the summer, when high temperatures coincide with low water levels. Typically blue-green algae like *Anacystis* or diatoms like *Cyclotella* or *Melosira* have been found to dominate the blooms.

One of the locations of particular interest, where algal blooms have occurred frequently, is the bay of Praia do Lami, situated some 30 Km south of the city center of Porto Alegre. The area is still sparsely populated, but according to plans of the municipal administration, infra-structure including water supply and a sewage collecting pipe system will be built to enable the expansion of the residential areas down the river.

Up to now the bay is one of the few shore regions, where swimming and other aquatic recreational activities are still possible. The area is therefore attractive to well-off inhabitants as well as illegal settlers. This imposes a serious problem, for the slums grow incontrollably and with few possibilities of restricting the discharge of waste into the river. In addition, there is a large number of visitors in the summer.

In recent years several occurences of extensive phytoplankton growth, part of which escalated to real algal blooms, have been recorded in the southern part of the bay. Among the undesirable effects of this phenomenon there is clogging of the filters during collection of water for treatment and unpleasant influences on taste and odor of the treated water, which cannot be removed with reasonable effort. Up to now there have been no real dangers like toxic substances released by algae.

However, the occurrence of algal blooms is an alarming indicator of instabilities in the ecosystem and its complex equilibrium mechanisms. Among the long-term effects of such disturbances is the reduction of the variety of species in the ecosystem. From the study of smaller water basins with complete eutrophication (i.e. the state of excessive accumulation of nutrients) the dangerous consequences are known.

It is important to develop and apply preventive measures for both immediate and long-term inhibition of algal blooms, as well as to react to the actual occurences immediately (e.g. by cutting off the pumps in the vicinity). Knowledge-based computer support for these tasks requires models of the involved processes as a basis for analysis, monitoring, and prediction of causes and effects. The following section describes the most important ones among these processes.

3 Algal Bloom - Causes and Effects

Among the factors that are certain to influence phytoplankton growth, light conditions, water temperature and the availability of nutrients are the most important ones (see [Palmer 62], [Guerrin *et al.* 94]).

The following influences can be determined (compare figure 2):

- Phyto-growth: the most important nutrients are nitrogen and phosphorus. The best available forms are ammonia (NH_4), nitrate (NO_3) and ortho-phosphate (PO_4). The consumption of nutrients by phytoplankton is dependent on temperature and on the N:P ratio. A value around 9:1 yields optimal growth. If significantly different from that value, the less available nutrient acts as the so-called “limiting factor” for the phytoplankton growth.
- Nitrification: The forms of the nutrients are changed by the activity of microorganisms, for example *nitrosomona* (which change ammonia to nitrite) and *nitro-bacter* (which transform nitrite into nitrate), the process as a whole being called nitrification. The process is highly temperature dependent.
- Advection and dispersion: The actual concentration of these substances in a specific location is influenced by the hydrodynamics, which transport matter by advection (i.e. by means of the flow of water) and turbulative dispersion (i.e. assimilation of spatial concentration differences by means of turbulative mixing).
- Photosynthesis: The algal blooms in turn change the carbon-dioxide concentration by extensive photosynthesis, possibly even leading to CO_2 depletion.
- Acidity reaction: The decrease in CO_2 causes the pH to rise.
- Dissociation: This will cause the ammoniacal equilibrium ($NH_4 + OH \leftrightarrow NH_3 + H_2O$) to shift towards free ammonia (NH_3).
- Intoxication: There is the danger of toxic effects on fish and aquatic fauna by high concentrations of free ammonia.
- Predation and Respiration: The fish represent another important community in the aquatic ecosystem, interacting both directly with the phytoplankton and influencing the chemical system.

Presently we ignore the light conditions (which limit the algal growth to a layer near the surface of the water body and are affected by the algal bloom through a sharp increase in turbidity) and the oxygen equilibrium (influenced by respiration and photosynthesis as well).

4 Modeling

4.1 Qualitative Modeling

As indicated by this verbal description, the expert knowledge in the hydro-ecological domain and particularly in the case of algal blooms is imprecise and of qualitative nature. For instance, the energy balance and the exact influence coefficients are usually not known, but rather general tendencies and limiting cases are used to reason about ecosystems.

To adequately represent this knowledge, a qualitative modeling approach is chosen, since

- the conclusions that can be drawn without precise quantitative knowledge are obtained with tractable effort and in a transparent way, so that intuitive explications about the inferences leading to the conclusion can be given to the user.
- the communication with the user is facilitated by a modeling language containing intuitive conceptual entities (e.g. processes), instead of requiring familiarity with complex formalisms and coding.
- even where quantitative calculations are necessary, they should be integrated into a conceptual and qualitative framework to check for their applicability in advance and to continuously verify their validity during the calculations, which both are inferences beyond the scope of classical mathematical methods.
- the model should predict the possible behaviors of the ecosystem in terms of alarming and dangerous events related to bloom occurrence, i.e. the expected output are statements of possibilities rather than quantities that still have to be interpreted qualitatively. Our approach provides models enabling the direct inference of such assertions.
- to account for the lack of exact knowledge about some relationships and interactions, we can use qualitative expressions of proportionality and influence, thus avoiding unreliable tentative descriptions and gaining the property of completeness (relative to the model), since the inferences will be valid for all instances of mathematical relationships. For example the relation I+(A,B) states that B positively influences A, i.e. the derivative of A will increase with B (all other influences assumed constant). The inferences that can be made without knowing the precise form of the dependency between A and B (typically a partial differential equation) will therefore be valid for all covered dependencies.

We refer to the process ontology of [Forbus 84], describing the system with a set of state variables and processes acting on them. A modeling assumption is that all changes of the state variables are caused by the effects of processes. An implementation of such a process-oriented modeling language is the Qualitative Process Compiler QPC ([Crawford *et al.* 90]). The model specification consists of a domain ontology, a process library, a scenario description and initial conditions. The ontology introduces the classes of entities (like the chemical constituents taken into account), the quantity types (like concentrations and values), relations and predicates. The process library consists of generic process descriptions (such as Figure 1). The scenario description specifies the entities present in the specific situation to be analyzed and states some properties in terms of relations and predicates. The initial conditions assert initial parameter values.

From this specification, a constraint model is obtained which can be executed by the qualitative simulation system QSIM ([Kuipers 86], [Kuipers 94]):

Process:	**Nitrification**
Individuals:	location (type: compartments)
	ammonia (type: ammonia-instance)
	temperature (type: temperature-instance)
Conditions:	Bacteria-present (location)
Operating-conditions:	Concentration (ammonia, location) > 0
Quantity-types:	rate
Relations:	rate = Value (temperature, location)
	* Concentration (ammonia, location)
	I+ (Concentration (ammonia, location), rate)
	I- (Concentration (nitrate, location), rate)

Figure 1: A sample process description: the nitrification process

The processes will be instantiated for each set of entities specified in the *individuals* slot, satisfying additionally *conditions*. In each system state the *operating-conditions* determine whether the process is active, i.e. its relations and influences are in effect. The *quantity-types* are introduced as variables local to the process, used in the calculations of the effects in the *relations* slot. The influences I+, I- have been described above.

Figure 2 presents a graph of the influences of the modeled generic processes, with black arrows indicating positive influences and gray ones indicating negative influences.

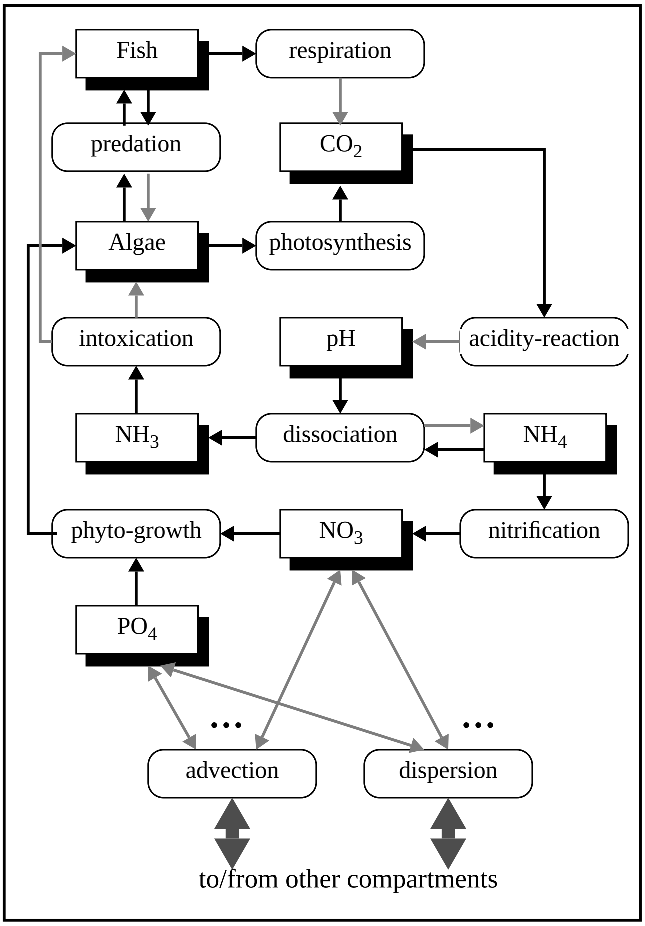

Figure 2: The basic process influence graph

4.2 Spatial Distinctions

Algal blooms are spatially limited phenomena, and in reasoning about them one of the important questions is “Where will they occur?”, instead of simply predicting the possibility of a bloom anywhere in the modeled system

Furthermore, a number of chemical as well as physical parameters show a distinctive spatial distribution, thus providing qualitatively different conditions for processes taking place at a local scale like chemical reactions or nutrient consumption. Therefore, reasoning about the involved processes will have to take spatial distinctions into account. These will be influenced by the advective flow of matter and the dispersive assimilation of constituent concentrations.

We adopt the notion of “hydrodynamic regimes”, a terminology used by experts while reasoning informally, referring to regions with (almost) homogeneous flow characteristics, which result in similar parameter values and consequently similar conditions for local processes. We introduce a more general class of entities called compartments for a spatial partitioning of the system. We use them similarly as in compartmental modeling ([Ironi, Stefanelli 94]).

In the first step, only the surface area of the aquatic ecosystem is decomposed into compartments, whose parameter values will be treated as homogeneous. We ignore the third dimension and leave a partitioning into different horizontal layers for further development, concentrating implicitly on a layer with typically good conditions for algal growth. For the time being, the decomposition into compartments is determined by the modeler and fixed during simulation.

We model the exchange of any constituent, such as nitrate (and similarly the assimilation of physical parameter values) between neighboring compartments by generic processes “advection” (see figure 3) and “dispersion”. Advective transport of constituents is caused by the (directed) flow of water, whose speed becomes a variable for each pair of compartments in the neighboring relation. Also, the turbulative dispersion, resulting in the decrease of concentration gradients is calculated by combining the turbulence parameter of each compartment.

Process: **Advection**
Individuals: constituent (type: constituents)
source (type: compartments)
destination (type: compartments)
Conditions: Distributed (constituent)
Neighbor (source, destination)
Operating-conditions: Flow (source, destination) > 0
Quantity-types: loss-absolute, loss-relative, gain-relative
Relations: loss-absolute = Flow (source, destination)
* Concentration (constituent, source)
loss-relative = loss-absolute / Volume (source)
; gain-absolute = loss-absolute
gain-relative = loss-absolute
/ Volume (destination)
I+ (Concentration (constituent, source),
loss-relative)
I- (Concentration (constituent, destination),
gain-relative)

Figure 3: Example process advection

One of the important settings for the simulation is the Praia do Lami, which is tentatively modeled as a single compartment. Two additional compartments represent the upstream and the downstream part of the river. Similar constructions are used to give boundary conditions like the incoming pollution from the smaller affluents.

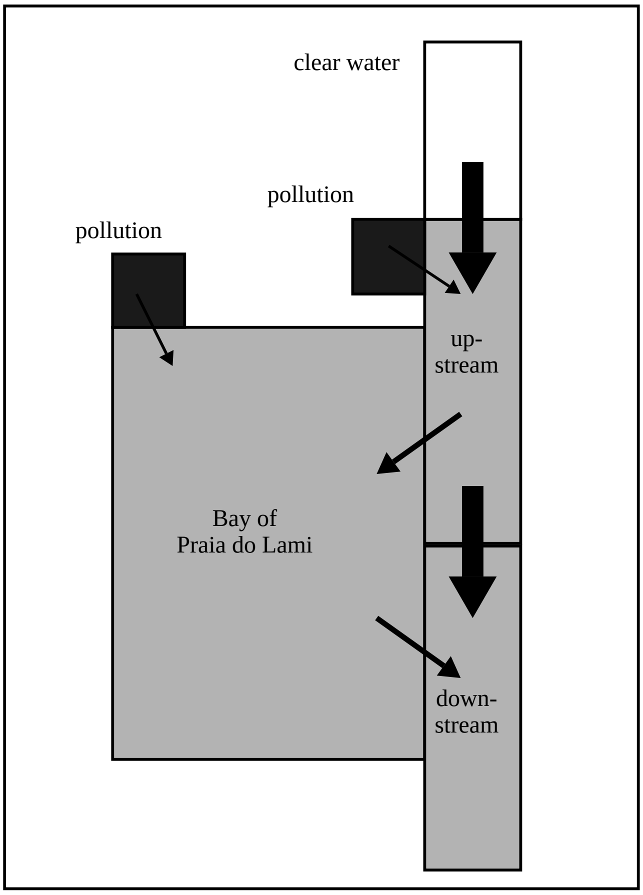

Figure 4: A simple compartmental structure of the scenario of Praia do Lami

4.3 Time-Scales

In a complex ecosystem like the Rio Guaba, there exist processes acting at very different time-scales. Chemical reactions for example will take place and eventually establish an equilibrium between the participating reactants very much faster than a fish population number will react to changed conditions.

If the processes relevant for some reasoning task span a wide range of time-scales, this imposes difficulties. Qualitative simulation may produce spurious solutions, which could be pruned by information about the relative time-scale. For instance, after increasing the concentration of free ammonia, the chemical reaction ionizing free ammonia, resulting in a diminished danger of intoxication of a fish population by free ammonia, comes into effect. The predictive engine could generate a behavior under the assumption that additional free ammonia is absorbed by fish more rapidly than transformed by the chemical mechanisms, thus maybe leading to severe intoxication of the animals, even though in reality the ammonia would be neutralized within parts of a second.

One possibility to avoid these problems is the technique of time-scale abstraction (as described in [Kuipers 87], [Kuipers 94], [Struss 93]). For a specific time-scale of interest all very much faster processes will be transformed into functional relationships, much slower ones will be treated as constant. As a second desirable effect, the model actually used for prediction will be much smaller and the reasoning will significantly gain efficiency.

The intention is to abstract the fastest equilibrium mechanisms (like chemical reactions) into functional relationships to facilitate long-term forecasts. By equilibrium mechanism, we denote a group of processes acting together to establish equilibrium. Since the exchange processes between the compartments are located at an intermediate time-scale (between chemical reactions and population developments), they can be used as a kind of reference time-scale. There is a correlation between the spatial partitioning and the different time-scales: The mechanisms with short response times (in relation to the global distribution mechanisms) can be seen as acting *inside* the single compartments, establishing a kind of dynamic equilibrium there, which will be shifted by the medium-term processes, the effects of which appear as if acting *between* the compartments. In a second step the system could derive a global equilibrium established by the exchange processes and the transformation acting on a comparable time-scale. The processes very much slower could either influence the compartment parameters, e.g. flow speeds, or even *change the spatial decomposition*.

We will give an example for a medium-term equilibrium for the nitrate concentration in a simplified scenario consisting of the upstream, the downstream and the bay compartment, without external pollution. We choose the processes of advection (with two instances of interest, the $\text{advection}_{\text{up,bay}}$ and the $\text{advection}_{\text{bay,down}}$) and the above mentioned nitrification of ammonia accomplished by bacteria (*nitro-bacter* and *nitrosomona*) approximately in the time-scale of hours, therefore matching the speed, with which concentration changes propagate through the river by means of advection.

The relevant influences on the state variable $\text{nitrate}_{\text{bay}}$ are the rate of the local nitrification process, which we denote rate_{bay}, and the loss ($\text{loss-rel}_{\text{bay,down}}$), respectively gain ($\text{gain}_{\text{up,bay}}$), by means of advection.

The equilibrium equation is obtained by setting the sum of the influences zero. Under the additional assumption, that influences combine linearly and the two advection processes even equally, we obtain:

$$\text{gain-rel}_{\text{up,bay}} + k \cdot \text{rate}_{\text{bay}} - \text{loss-rel}_{\text{bay,down}} = 0$$

which expands to

$$\frac{\text{flow}_{\text{up,bay}} \cdot \text{nitrate}_{\text{up}}}{\text{volume}_{\text{bay}}} + k \cdot \text{temp}_{\text{bay}} \cdot \text{ammonia}_{\text{bay}} = \frac{\text{flow}_{\text{bay,down}} \cdot \text{nitrate}_{\text{bay}}}{\text{volume}_{\text{bay}}}$$

and under the (reasonable) assumption that the sum of the inflows equals the sum of the outflows of a compartment, in our case $\text{flow}_{\text{up,bay}} = \text{flow}_{\text{bay,down}} =: \text{flow}_{\text{bay}}$, the equation can be solved for

$$\text{nitrate}_{\text{bay}} = \text{nitrate}_{\text{up}} + k \cdot \frac{\text{volume}_{\text{bay}}}{\text{flow}_{\text{bay}}} \cdot \text{temp}_{\text{bay}} \cdot \text{ammonia}_{\text{bay}}$$

therefore determining the nitrate concentration in the bay by the nitrate concentration upstream and the ammonia concentration in the bay, which can be calculated by a similar equilibrium equation. The obtained equilibrium can be coded into constraints replacing the original dynamic modeling of the relationship. For long-term forecasts the model will become smaller and more manageable in that way.

5 Open Issues and Future Work

As of now, the algal bloom model has been implemented in QPC and will be evaluated and further developed based on the qualitative simulation results. Further calibration and validation will be carried out using the data available from DMAE. We do not expect the purely qualitative model to unambiguously predict or rule out the occurrence of algal bloom. For a detailed analysis it has to be extended by a semi-quantitative level, for instance based on interval mathematics.

The current model is based on some simplifications that will have to be dropped in the future. First, some influences are presently ignored, for instance the oxygen equilibrium. Second, the compartments introduced so far do not reflect distinctions along the vertical axis. However, the influence of light, air temperature and wind primarily affects the surface layer, giving rise to important variations with depth, and, hence, a three-dimensional compartmental structure will be more adequate.

Third, the current compartmental structure is fixed. Ultimately, a dynamic structure will be regarded, e.g. when changing conditions lead to significant parameter differences within a single compartment, the simulator should be able to split the compartment later. This raises the issue of how to determine the partitioning into compartments. The rationale behind this decomposition of a continuous system is to identify areas of strong, or fast, interactions, that interact with each other only in a weak, or slow, fashion. In our domain, this is closely related to direction and speed of flow, but other influences can be present. Rather than being defined beforehand, an appropriate compartmental structure could be detected automatically as a result of executing the qualitative model and/or interpretation of data.

Also, the transformation of models under time-scale abstraction should be performed automatically taking into account the respective purpose of using the model.

The modeling efforts described are a first step in the development of the ΣIGMA system ([Guerrin *et al.* 95]), a proposed framework for environmental decision support systems, that aims at integrating model building, situation assessment, simulation, diagnosis, and action planning. Sharing models across the various tasks will impose further requirements on the structure and content of the models.

Acknowledgments

This work has been supported in part by the Brazilian research council (CNPq) and the German ministery of research and technology (project QUALIDADE). We benefitted from discussions with Paulo Marcos Amaral Alves, Elenara Correa Lersch and Maria Mercedes Bendati from the research department of DMAE.

References

[Crawford *et al.* 90]
Crawford J., Farquhar A., Kuipers B., *QPC: A Compiler from Physical Models into Qualitative Differential Equations.* In: Proceedings of AAAI-90.

[Forbus 84]
Forbus K., *Qualitative Process Theory.* In: Artificial Intelligence 24(1984): 85-168

[Guerrin *et al.* 94]
Guerrin F., Bousson K., Steyer J.-P., Trave-Massuyes L., *Qualitative reasoning methods for CELSS modeling.* In: Advances in Space Research, 14(11), 307-312, 1994.

[Guerrin *et al.* 95]
Guerrin F., Heller U., Roque W., Struss S., Σ*IGMA - um sistema integrado de gerenciamento do meio ambiente.* ECOS journal, DMAE Publications, Porto Alegre, 1995.

[Ironi, Stefanelli 94]
Ironi L., Stefanelli M., *QCMF: a tool for generating qualitative models from compartmental structures.* In: Working notes of QR-94.

[Kuipers 86]
Kuipers B., *Qualitative Simulation.* In: Artificial Intelligence 29(1986):289-338.

[Kuipers 87]
Kuipers B., *Abstraction by Time-Scale in Qualitative Simulation.* In: Proceedings of AAAI-87.

[Kuipers 94]
Kuipers B., *Qualitative Reasoning - Modeling and Simulation with Incomplete Knowledge.* MIT Press, 1994.

[Palmer 62]
Palmer C., *Algae in water supplies.* U.S. Department of Health, Education and Welfare, Public Health Service, Washington, 1962.

[Struss 93]
Struss P., *On Temporal Abstraction in Qualitative Physics.* In: Working notes of QR-93.