CONTRIBUTIONS to SCIENCE, 1 (4): 451-462 (2000)
Institut d'Estudis Catalans, Barcelona

Wastewater treatment improvement through an intelligent integrated supervisory system

M. Poch[1]*, I.R. Roda[1], J. Comas[1], J. Baeza[2], J. Lafuente[2], M. Sànchez-Marrè[3] and U. Cortés[3]

1 Laboratori d'Enginyeria Química i Ambiental (LEQUIA), Institut de Medi Ambient. Universitat de Girona

2 Grup d'Enginyeria Ambiental, Departament d'Enginyeria Química. Universitat Autònoma de Barcelona

3 Secció Intel·ligència Artificial, Departament de Llenguatges i Sistemes Informàtics. Universitat Politècnica de Catalunya

Abstract

This paper shows the result of years of work by a cooperative research group including chemical engineers, environmental scientists and computer scientists. This research has been focused on the development and implementation of new techniques for the optimisation of complex process management, mainly related to wastewater treatment plants (WWTP). The experience obtained indicates that the best approach is a Supervisory System that combines and integrates classical control of WWTP (automatic controller for maintaining a fixed dissolved oxygen level in the aeration tank, use of mathematical models to describe the process...) with the application of knowledge-based systems (mainly expert systems and case-based systems). The first part is an introduction to wastewater treatment processes and an explanation of the complexity of the management and control of such complex processes. The next section illustrates the architecture of the supervisory system and the work carried out to develop and build the expert system, the case-based system and the simulation model for implementation in a real plant (the Granollers WWTP). Finally, some results of the field validation phase of the Supervisory System when dealing with real situations in the plant are described.

Resum

Aquest article mostra el resultat de la col·laboració portada a terme durant els darrers anys entre grups d'enginyeria química, enginyeria ambiental i intel·ligència artificial. El treball se centra en el desenvolupament de tècniques per a la millora i supervisió de processos complexos, especialment del tractament biològic d'aigües residuals. L'experiència demostra que la millor opció requereix desenvolupar un sistema supervisor que combini i integri tècniques de control clàssic (controlador automàtic del nivell d'oxigen dissolt en el reactor biològic, ús de models descriptius del procés, etc.) amb sistemes basats en el coneixement (concretament sistemes experts i sistemes basats en casos). El present article descriu la complexitat de la gestió del procés de tractament de les aigües residuals, l'arquitectura integrada que es proposa i el desenvolupament i la construcció de cadascun dels mòduls d'aquesta proposta per a la implementació real a l'estació depuradora d'aigües residuals de Granollers. Finalment, es detallen alguns resultats del procés de validació del seu funcionament enfront de situacions quotidianes de la planta.

Key words: Biological wastewater treatment, activated sludge, knowledge-based systems, supervision, environmental decision support systems.

The survival of the human species and our quality of life depend upon our ability to manage the Earth's natural resources on scales ranging from local to global. This requires an assessment of the extent of these resources, including an understanding of their variation in time and space and of what causes these variations. The concentration of population at specific locations and the industrialization of our society (both caused by human activity) are responsible for non-sustainable exploitation of natural resources and for the breakdown of equilibrium in different natural ecosystems of our planet. It has resulted in environmental pollution, which negatively affects the quality of water, air, soil and therefore, animal, vegetal and human life. The increasing degradation of the environment has forced society to consider changes in human behaviour in order to ensure the essential conditions for life on Earth. This consideration has encouraged re-

* Author for correspondence: Manel Poch, Laboratori d'Enginyeria Química i Ambiental, Dept. Enginyeria Química, Agrària i Tecnologia Agro-Alimentària, Universitat de Girona. Campus Montilivi, s/n. 17071 Girona, Catalonia (Spain). Tel. 34 972 418 161. Fax: 34 972 418 150. Email: manel@lequia.udg.es.

search and great effort has been put into understanding, preventing and correcting environmental degradation.

In this sense, the treatment of water and wastewater has become one of the most important environmental issues. Wastewater treatment is fundamental for keeping water natural resources (rivers, lakes and seas) as high quality as possible. More and more restrictive social regulations have appeared due to these environmental reasons and the correct management of wastewater treatment facilities has become a very important topic during the last 20 years. For example, European directive 91/271 establishes that every city or village with a population of more than 2000 inhabitants must treat its wastewater before the year 2005, at least for the Suspended Solids (SS) and the Biological Oxygen Demand (BOD) contained in the wastewater. This criterion attempts to provide a regulated water effluent with low contaminant load in order to cause minimum environmental impact (including energy use) on the quality of the receiving water ecosystem. This concept has been enlarged with the new water Framework Directive 2000/60/EEC. In Catalonia, the rapid implementation of Directive 91/271 by the *Generalitat de Catalunya*, through its *Plà de Sanejament*, has led to the building of several Wastewater Treatment Plants (WWTPs) during the last few years.

By February 2000, there were 225 WWTPs operating in Catalonia (plus another 102 under construction or in an upgrading phase), being their management cost over 60 million Euros during the year 1999. However, more important than economics are the maintenance of quality criteria established by environmental regulations for treated effluent. In this sense, optimal WWTP management is concerned with trying to obtain good wastewater treatment efficiency and maintain maximum process stability while avoiding operational problems.

A typical wastewater treatment plant usually includes a primary treatment and a secondary treatment to remove organic matter and suspended solids from wastewater. Primary treatment is designed to physically remove solid material from the incoming wastewater. Coarse particles are removed by screens or reduced in size by grinding devices. Inorganic solids are removed in grit channels and many of the organic suspended solids are removed by sedimentation. Overall, the primary treatment removes almost one-half of the suspended solids in the raw wastewater. The wastewater flowing to the secondary treatment is called the primary effluent [14]. Secondary treatment usually consists of a biological conversion of dissolved and colloidal organic compounds into stabilized, low-energy compounds and new biomass cells, caused by a very diversified group of microorganisms, in the presence of oxygen. This mixture of microorganisms (living biomass) together with inorganic as well as organic particles contained in the suspended solids constitutes what is known as activated sludge. This mixture is kept moving in wastewater by stirring done by aerators, turbines or rotators, which simultaneously supply the required oxygen for the biological reactions. Some of the organic particles can be degraded by subjecting them to hydrolysis whereas others are non-degradable (inert). This mixture of microorganisms and particles has the ability to bioflocculate, that is, to form an aggregation called activated sludge floc if there exist a balance in population between floc-formers and filamentous bacteria. The activated sludge floc gives to the sludge the capacity to settle and separate from treated water in the clarifier. A biological reactor followed by a secondary settler or clarifier constitutes the activated sludge process, which is the most well known process of secondary treatment because it is also the most widely used (see Figure 1).

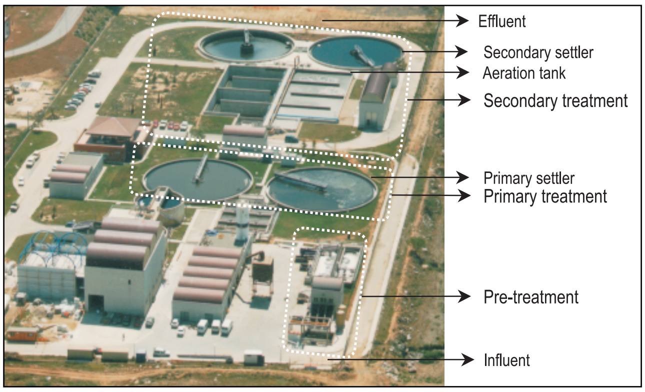

Figure 1. Primary and secondary treatment of a typical WWTP.

Like other environmental and biotechnological processes, WWTP are complex systems, involving many interactions between physical, chemical and biological processes, *e.g.* chemical or biological reactions, kinetics, catalysis, transport phenomena, separations, etc..The successful management of these systems requires multi-disciplinary approaches and expertise from different social and scientific fields. Some of the special and problematic features of environmental processes are:

- Intrinsic instability: most of the chemical and physical properties as well as the population of microorganisms (both in total quantity and number of species) involved in environmental processes do not remain constant over time.
- Many facts and principles underlying environmental domain cannot be characterized precisely only in terms of a mathematical theory or a deterministic model with clearly understood properties.
- Uncertainty and imprecision of data or approximate knowledge and vagueness: these processes generate a considerable amount of qualitative information. Moreover, on-line data is not sufficient to monitor and diagnose the process successfully since the use of on-line analysers is still rare and they are often unreliable. It is necessary to be able to access analytical information. Most of this information is acquired with global variables that cannot be obtained on-line but only with a delay of some hours or days.
- Huge quantity of data/information: the application of current computer technology to the control and super-

vision of these environmental systems has led to a significant increase in the amount of data acquired.
- Heterogeneity and scale: because the media in which environmental processes take place are not homogeneous and cannot easily be characterised by measurable parameters, data are often heterogeneous. Focusing on wastewater treatment processes, four different types of information can be identified in WWTPs: online quantitative data, off-line quantitative data, qualitative observations of plant operators and microscopic observations.

Due to the complexity of wastewater treatment process control, even the most advanced conventional *hard* control systems have encountered limitations when dealing with problem situations that require qualitative information and heuristic reasoning for their resolution (e.g. presence and interrelations among different microorganisms, substrate and operational conditions in filamentous bulking problems, foaming...). Indeed, in order to describe these qualitative phenomena or to evaluate circumstances that might call for a change in the control action, some kind of linguistic representation built on the concepts and methods of human reasoning, such as intelligent systems, is necessary. And this is the reason why human operators have, until now, constituted the final step in closed-loop plant control [18]. A deeper approach is necessary to overcome the limited capabilities of conventional automatic control techniques when dealing with *abnormal* situations in complex systems, and to provide the level and quality of control necessary to consistently meet environmental specifications.

For these reasons, the field of intelligent control systems began to look promising a few years ago in terms of solution to these problems. A reasonable, distributed proposal outlines the scope for the integration of tools like pattern recognition, knowledge-based systems, fuzzy logic, neural networks, or inductive decision trees, which handle the particular characteristics of complex processes (*e.g.*, environmental problems), with numerical and conventional computational techniques (statistical methods, advanced and robust control algorithms and system identification techniques).

In this paper we present the results of cooperative research conducted over the last years by the Environmental and Chemical Engineering Laboratory of the *Universitat de Girona* in close association with the Chemical Engineering Department of the *Universitat Autònoma de Barcelona* (UAB) and the Knowledge Engineering and Machine Learning Group (KEML) in the Software Department of the *Universitat Politècnica de Catalunya* (UPC) to develop an intelli-

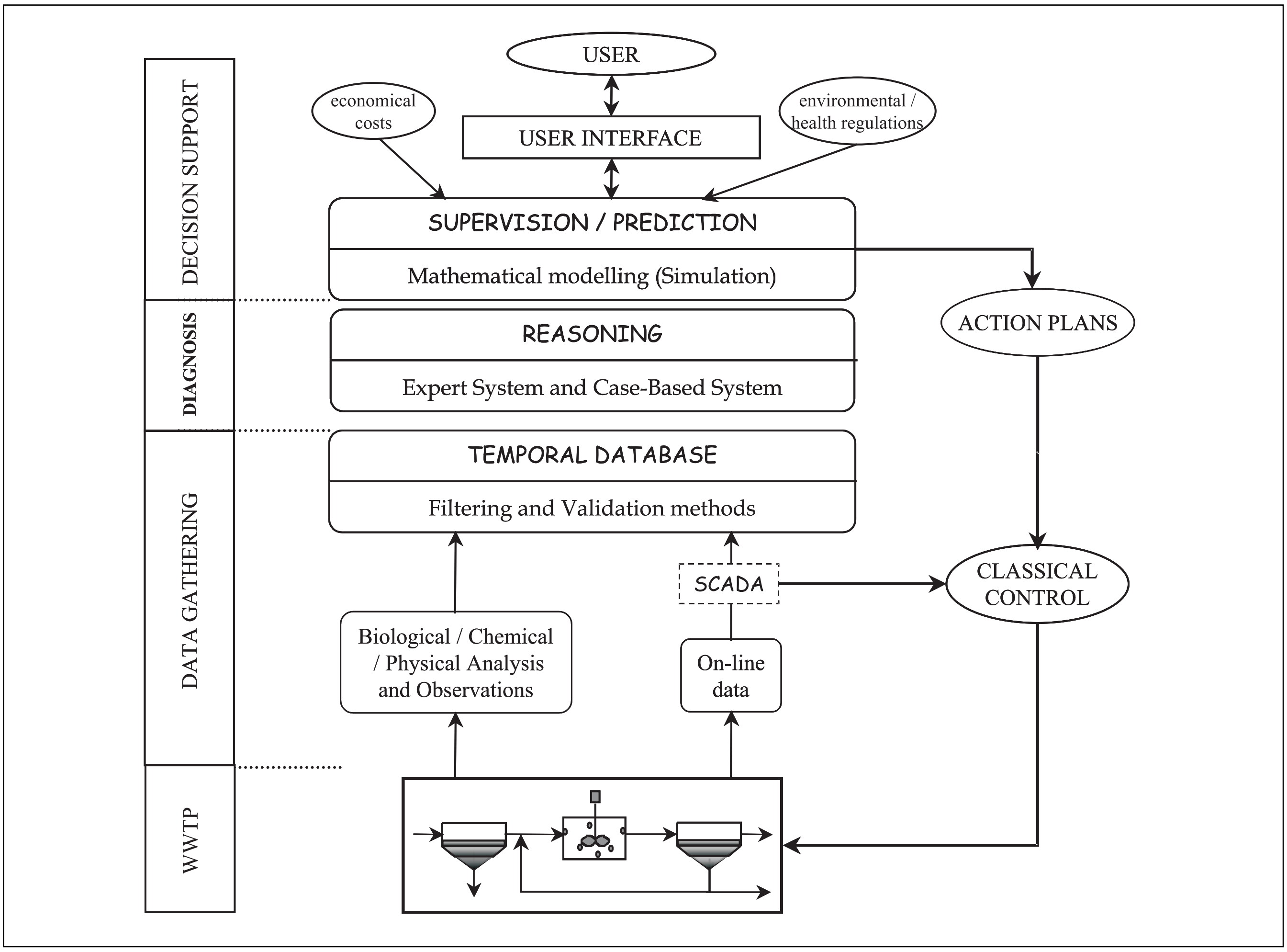

Figure 2. Integrated modules of the computing approach.

gent supervisory system for an activated sludge process. This Supervisory System integrates classical control systems with two knowledge-based systems (Expert System -ES- and Case-Based System -CBS-) and has already been implemented in the Granollers WWTP. The development, implementation and evaluation of the Supervisory System for the Granollers WWTP can be found in [4]. This real implementation has been called BIOMASS, that is, Besòs Intelligent Operation & Management for the Activated Sludge System.

The multi level-architecture

The architecture developed provides an integrated framework for easy access to three modules: data gathering (including database management); advanced tools of reasoning (diagnosis module); and the decision support module (simulation models to implement predictive and supervisory tasks in the plant). This multi-level architecture guarantees the useful and successful supervision of wastewater treatment processes ([21] and [24]). Figure 2 shows the structure of this computing approach to supervise the Granollers WWTP. This integrated architecture can be used to supervise any environmental process as an Environmental Decision Support System (EDSS) [5].

The first module is the level that supports the data gathering and updating processes. This level is composed of the acquisition systems for the on-line data (coming from sensors and equipment) and the off-line data (biological, chemical and physical analyses of water and sludge quality and other qualitative observations of the process, i.e., presence of bubbles on the settler surface...). Moreover, this level implements data filtering, validation and management processes over the temporal evolving (real-time) database where on-line, off-line and data calculated by the system are stored.

The second module of this architecture includes two artificial intelligence techniques (expert system and case-based reasoning); and numeric models, overcoming the caveats and limitations in the use of each single technique, which constitutes the reasoning module for situation assessment. This second level entails cooperation between knowledge-based control and automatic control for the supervision of this complex process. This level implements reasoning tasks to diagnose its state and behaviour and to propose the action to maintain or return the process to its normal operation.

The third level of the Supervisory System (the decision support module) implements a supervisory and predictive task over the WWTP. The supervisory task of this module is to seek a consensus on the diagnosis and actions to be taken proposed by the different reasoning tools. Meanwhile, the predictive task evaluates the possible alternatives of action by means of a dynamic model, which predicts the future behaviour of the plant and, finally, infers and suggests the most suitable action strategy to be considered. The decision-maker should consider the different alternatives in relation to socio-economic conditions and applicable legislative frameworks. This module also increases the interaction of the users with the computer system throughout an interactive, graphic user-machine interface (the user may query the system for justifications and explanations about suggested decisions, to consult certain values, etc.). A commercial shell for developing real-time knowledge-based systems with a user-friendly interface (G2, [11]) was selected as the suitable platform on which the supervisory system could be built. The next sections describe the ES, CBS and simulation model development and their implementation for the Granollers WWTP.

Expert System

An Expert System (ES) is defined as an interactive computer program that attempts to emulate the reasoning process of experts in a given domain over which the expert makes decisions. The ES has two main modules: the Knowledge Base (KB) and the inference engine. The knowledge base includes the overall knowledge of the process as a collection of facts, methods and heuristics, which are usually codified by means of production rules. The inference engine is the software that controls the reasoning operation of the ES by chaining the knowledge contained in the knowledge base in the best possible way.

The acquisition of the knowledge included in the knowledge base is the core and also the bottleneck of the ES development. It involves eliciting, analysing and interpreting the knowledge that experts use to solve a particular problem. This knowledge should be represented in an easy, structured way (*e.g.*, in tables, graphs, frames or decision trees).

Knowledge included in the KB can be obtained from several sources. These sources can be divided into two types: documented (based on existing literature about the topic and on the WWTP database) and undocumented (that is, experiences or expertise from experts on the process). Information can be identified and collected using any of the human senses (i.e. through interviews with the experts or reading books, journals, flow diagrams, etc.) or with the help of machines (the use of machine learning tools to acquire knowledge from historical database). Therefore, the different knowledge acquisition methods can be classified as manual or automatic (Figure 3).

In the knowledge base development of this study, different methods from both types have been used. Conventional knowledge acquisition methods (literature review, interviews, etc.) were used first. To overcome the limitations of conventional methods, they were complemented with the use of different automatic knowledge acquisition methods. These latter methods can be either supervised (mainly inductive learning techniques, CN2, C4.5 and *k*-NN [3]) or unsupervised (essentially Linneo$^+$ [26], which obtains information directly from the database [23]). Figure 3 illustrates

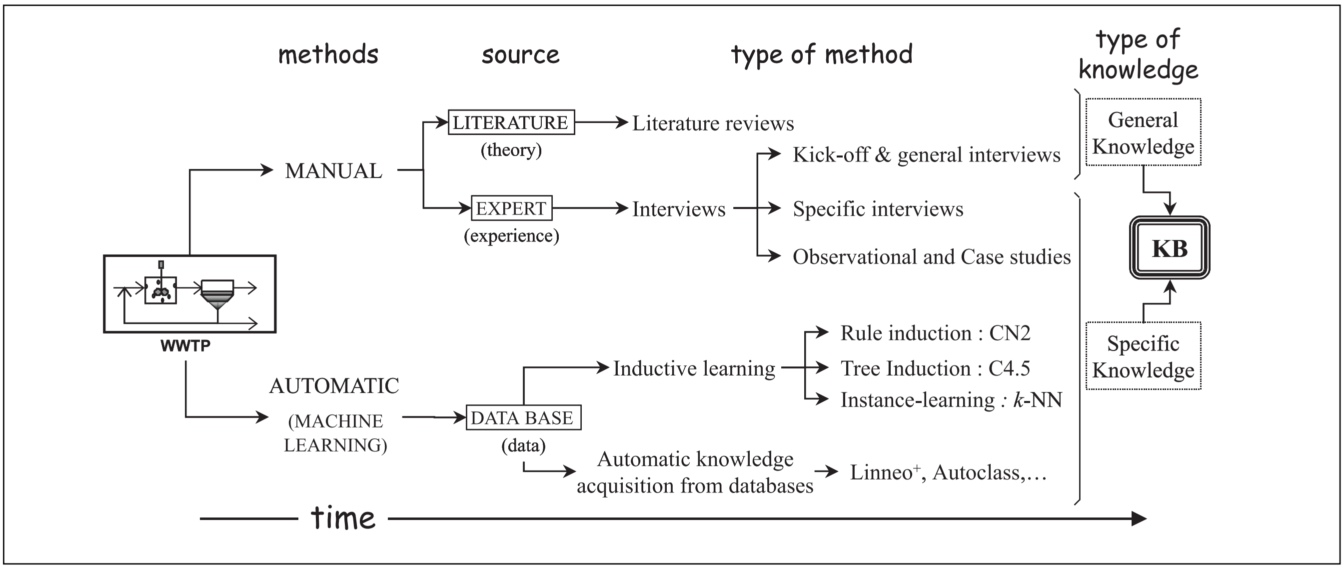

Figure 3. Classification of the sources and methods used to acquire knowledge from WWTP.

the main sources and methods that were used to acquire both kinds of knowledge on the wastewater treatment processes.

Among the different possibilities to represent the whole knowledge elicited (tables, decision trees or knowledge diagrams and frames), decision trees were selected as the most suitable representation. All symptoms, facts, procedures and relationships used for problem diagnosis can be cast into a set of decision trees. These trees consist of hierarchical, top-down descriptions of the linkages and interactions among any kind of knowledge utilized to describe facts and reasoning strategies for problem solving (objects, events, performance and meta-knowledge). The translation of the knowledge contained in a branch of decision trees into a production rule is direct. The trees designed include diagnosis, cause identification, and action strategies for a wide range of WWTP troubleshooting, avoiding contradictions and redundancies. These logic trees serve as a record of the expert's step-by-step information processing and decision-making activity. Some branches are specific and contain some peculiarities of the plant, while others are more general and can be applied to any plant. As an example of the developed trees, Figure 4 shows the tree for diagnosing and solving filamentous bulking problems. The inference can start from any kind of data, both quantitative (chemical or biological), related to the water and sludge characteristics, and qualitative referring to sludge settleability and *in-situ* observations (supernatant in the settling test, presence of foams on the clarifier surface...). Whenever a symptom of filamentous bulking is presented, the expert system activates an intermediate alarm, signalling the risk of the problem occurring. The diagnosis rules then evaluate the filamentous bulking problem according to the activity of microfauna, the settleability of the sludge (Sludge Volume Index, SVI) and the presence of filamentous bacteria. Once the expert system has concluded a situation, it tries to detect its specific cause. With the situation and cause diagnosed, the expert system sends its conclusions to the third level of BIOMASS, which conducts the processes of supervision and prediction and proposes an action plan.

Table 1 shows the list of all the situations contemplated. The list covers primary and secondary treatment, distinguishing between the non-biological and the biological origins of the problems. The latter origin causes a decrease in the biological reactor performance or dysfunctions in the secondary settler.

Table 1. List of decision trees developed for detecting and solving WWTP problems

	Secondary treatment	
Primary treatment problems	*Non-biological Origin*	*Biological Origin*
Old sludge	White foams	Filamentous bulking
Septic sludge	Overloading (and Organic shock)	Foaming (*Actynomicetes*)
Sludge removal systems breakdown	Nitrogen shock	Deflocculation (pinpoint)
Clogged pumps or pipes	Conductivity shock	Deflocculation (disperse growth)
Low efficiency of grit removal	Storms	Slime viscous bulking
Primary high sludge density	Hydraulic shock	Toxic shock
Inadequate sludge purges	Underloading	Nitrification/denitrification
Hydraulics shock	Aeration problems	(include rising sludge)
High solids loading	Clarifier problems	
Other mechanical problems	Mechanical and electrical problems	
	Transition state to some of these problems	

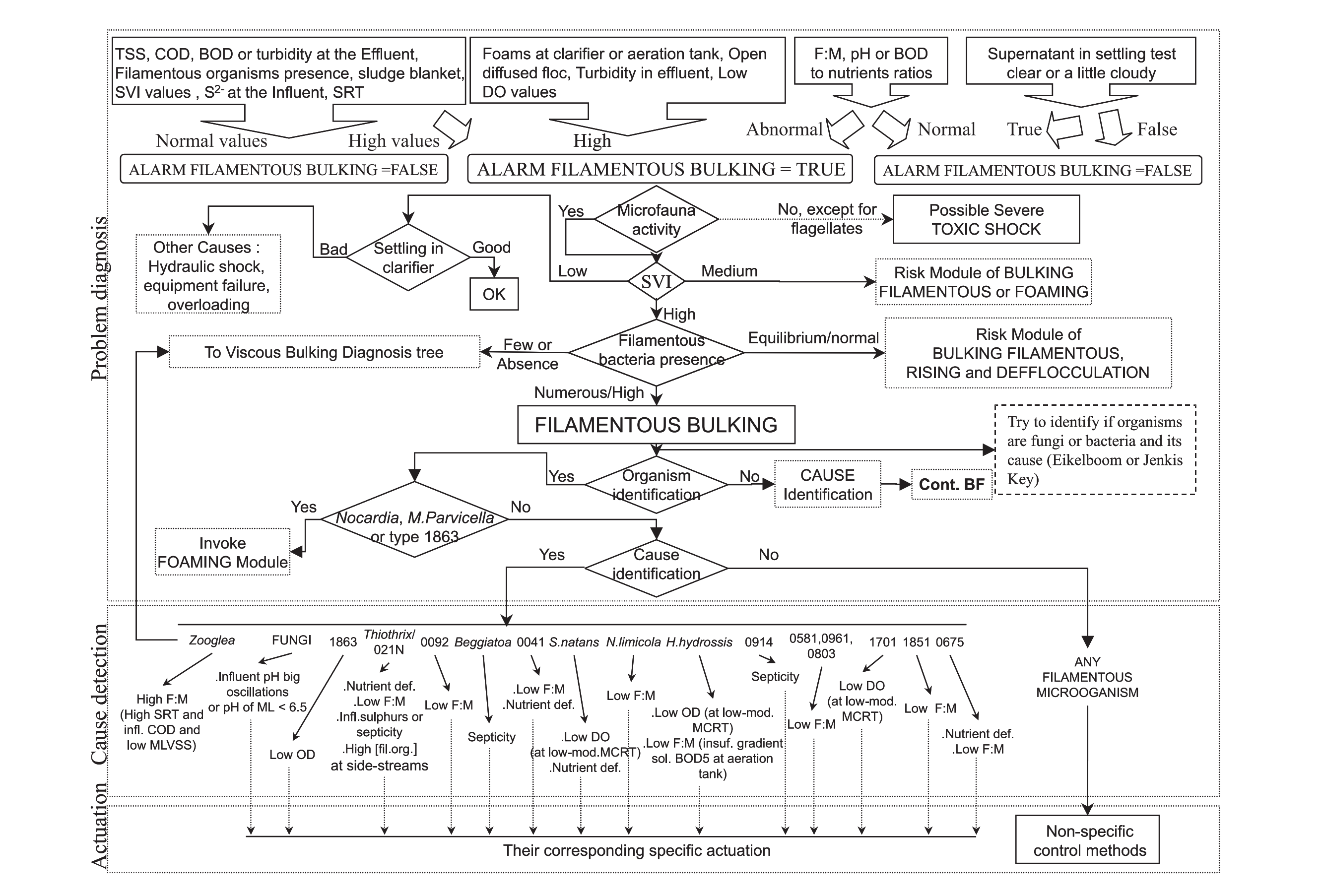

Figure 4. Decision tree for bulking filamentous problems.

The use of expert systems offers a number of advantages that overcome the limitations of other techniques: ES facilitate the inclusion and retention of heuristic knowledge from experts and allow qualitative information processing; knowledge is represented in an easily understandable form (rules); a well-validated ES offers potentially optimal answers because action plans are systematised for each problematic situation; in addition, ES make possible the acquisition of a large general knowledge base, with flexible use for any WWTP management. Finally, ES facilitate objective acquisition of specific knowledge throughout the use of machine learning techniques. The first results in ES development were published in [19] and [29]. Expert systems also show some limitations: most of the knowledge acquired is general knowledge for managing any WWTP (coming from a review of the literature), in which there is a lack of specificity, mainly in the repertory of actions proposed. In addition, this insufficient specific knowledge comes from interviews with plant managers and workers (thus involving bias, discrepancies and imprecision) and from database study and classification, which is often incomplete and almost never contains qualitative information. People tend to remember their past endeavours as being successful, regardless of whether they actually were or not. Complex problems require many (hundreds of) rules, involving long development time and may create problems in both using the system and maintaining it. And, perhaps most importantly, the knowledge base is static. Once developed, it is not an easy task, at least for the expert or final user, either to modify rules or to adapt the knowledge base to new specifications, and the system is unable to learn from new experiences. This last fact could provoke the systematic repetition of errors in diagnosis and proposed actions. All these limitations indicate that ES should be complemented with other approximations to manage complex processes optimally.

Case-Based System

The proposed Case-Based System (CBS) methodology permits the use of past experiences to solve new problems that arise in the process. It is based on the idea that the second time we solve a problem it is usually easier than the first time because we remember and repeat the previous solution or recall our mistakes and try to avoid them. The basic idea is to adapt solutions that were used on previous particular problems affecting process performance and use them for solving similar new problems with less effort than with other methods that start from scratch. A case is described as a conceptualised piece of knowledge representing an experience that teaches a fundamental lesson about how to achieve the reasoner's goals [16]. The set of specific cases is stored in a structured memory in a case-base (the case-library) and initialised with a set of typical cases in the plant. The CBS development includes the case definition, the case-library definition, and the selection of the initial seed.

The case is the codified description of each specific state or experience of the WWTP studied. Codification should be in storable form in order to be easily retrieved in the future. As an example, table 2 shows a real case for the Granollers WWTP. Each case must contain:

- An identifier or label. In this study, we have used the date of the day preceded by the word «Case-».
- The description of the situation, based on the values of the relevant variables that characterize the wastewater treatment process itself (corresponding set of observations related to that situation).
- The diagnosis implicit in each situation, indicating the expert evaluation of the «state» of the process.
- Action plan proposed or recommended to keep the process under control.
- The solution-result, with the evaluation of the application of the proposed action, specifying whether the result has been a success or a failure.
- The similarity or distance between the case retrieved from the library and the current case.

The CBS requires a library of cases to broadly cover the set of problems that may arise from the process. These cases are indexed in memory so that they are retrievable when their experiences can be used advantageously to contribute to achieving the goals of the process. Both successful cases and failures must be included in the library. The structure and organization of this case library is different from a normal database, where full match among variables is required. Case library definition in CBS is a key point because it has a great effect on the efficiency of the system both on response time and on success in finding a suitable stored case to match the new situation.

It is recommendable to initialize the library using a set of common situations (or cases) obtained from technical books or provided by experts on the process. Thus, the CBS is ready from the very start to propose solutions to problems similar to those considered in the initial «seed». Otherwise, the CBS would have to increase its knowledge from each new experience, increasing significantly the «set up» time necessary for the CBS to be used successful. The initial seed at Granollers included 74 real cases from the historical database, which covered a broad range of the main problems in the process and normal situations.

The library is updated with new cases as the knowledge about the process progresses; so the CBS evolves into a better reasoner and system accuracy benefits from these new acquisitions. However, because the library can become overcrowded with large amounts of information, it is crucial to include only the most relevant cases.

Table 2. Example of a real case stored in the case-library of the Granollers WWTP

:identifier	((	Case-29-11-99		))
:description of the situation	((	Influent flow rate (Flow-I)	19380	)
	(	Influent-COD (COD-I)	727	)
	(	Influent-Suspended Solids (SS-I)	254	)
	(	Influent-Total kjeldhal Nitrogen (TKN-I)	110	)
	(	Primary effluent-COD (COD-P)	693	)
	(	Primary effluent-SS (COD-P)	140	)
	(	Secondary effluent-COD (COD-E)	61	)
	(	Secondary effluent-Suspended Solids (SS-E)	15	)
	(	Secondary effluent -Total Nitrogen (TN-E)	80	)
	(	Biomass Concentration (MLSS-AS)	4271	)
	(	Sludge Volumetric Index (SVI-AS)	111	)
	(	Waste Activated sludge Rate (WAS)	663	)
	(	Recycle Activated sludge Rate (RAS)	66 %	)
	(	Dissolved Oxygen_line1 (DO1)	2.5	)
	(	Dissolved Oxygen_line2 (DO2)	2.3	)
	(	Sludge Residence Time (SRT)	6	)
	(	Food to Microorganism ratio (F/M)	0.18	)
	(	Filamentous_organism_Predominant (*Filam*)	*Microthrix Parvicella*	*)*
	(	V30-setttling test observations (V30-settling test)	Good but foams on supernatant	))
:diagnosis		*Foaming* caused *by Microthrix Parvicella*	)	
:action plan	((	1. Cause Identification: Low F/M ratio		)
	(	2. Physical aeration tank foam and clarifier foam removal		)
	(	3. High wasting activated sludge flow rate		)
	(	4. Low Recycle activated sludge flow rate to facilitate good compaction		)
	(	Check also for *Nocardia* trends during a week		)
:evaluation	((	Success. In five days, Nocardia population starts to decrease		))
:similarity	((	94.1 %		))

When a new case occurs, CBS start a new cyclic process (figure 5), which consists of the following steps:

- Gathering and processing data from the process in order to define the current case.
- Searching for and retrieving from the case library the case that best fits the current case. This step is accomplished by using a suitable algorithm, which is based on a numerical similarity comparison between the cases.
- Adapting the solution proposed by the retrieved case if it does not perfectly match with the current case. To carry out this step a simple parameter adjustment algorithm is usually applied, under the assumption that the distance between the case retrieved from the library and the current case is small enough so that only a linear interpolation adjustment of the parameter involved in the solution is needed.
- Applying the adapted solution to the process, and evaluating its consequences. There are different ways to evaluate these results: by asking the operator directly, by simulating the proposed actions, or by directly getting feedback on the results of the proposed solution.
- Learning each new experience to update the knowledge contained in the case library. When the evaluation has been successful, the case is incorporated into the library so that the CBS can benefit from it in the future when similar cases arise. This stage is known as learning from success, and, to avoid the introduction of redundant information in the library, it must be considered only when this new case offers complementary knowledge. Alternatively, the proposed solution may be a failure. In this case the learning process is known as learning from failure, and the system should record it to prevent continued mistakes in trying to solve future problems. From the beginning of the CBS implementation in BIOMASS and during the five-month field validation phase, the case-library was enlarged by more than 100 complete new cases, registering diagnosis and action plan carried out. More details about our CBS approach for complex process supervision can be found in [20] and, specially for WWTP, in [22], [25], [27] and [28].

Simulation Model

According to the Fishwick definition, to model is to abstract from reality a formal description of a dynamic system [17]. The characterisation of the behaviour of a system is reached by simulations. Thus, simulation explains the process of conducting experiments with this model for the purpose of enabling decision makers to either: (1) understand the behaviour of the system, (2) evaluate various strategies for the operation of the system, or (3) predict possible future conditions [17]. *Mechanistic* models describe whole processes taking place in the biological organic and/or nutrient removal processes, including substrate and biomass evolution. To accomplish this objective, different kinds of substrates and biomass are established. The relationships (stoichiometric parameters) between all these state variables are obtained by applying the mass balances to the process. The rates proposed by the International Water Association (IWA) *Task Group on Mathematical Modelling for Design and Operation of Biological Wastewater Treatment* to model the organic matter, nitrogen and phosphorous biological removal are based on Monod kinetics or modifications of them. Usually these models are represented in a matrix fashion (see Table

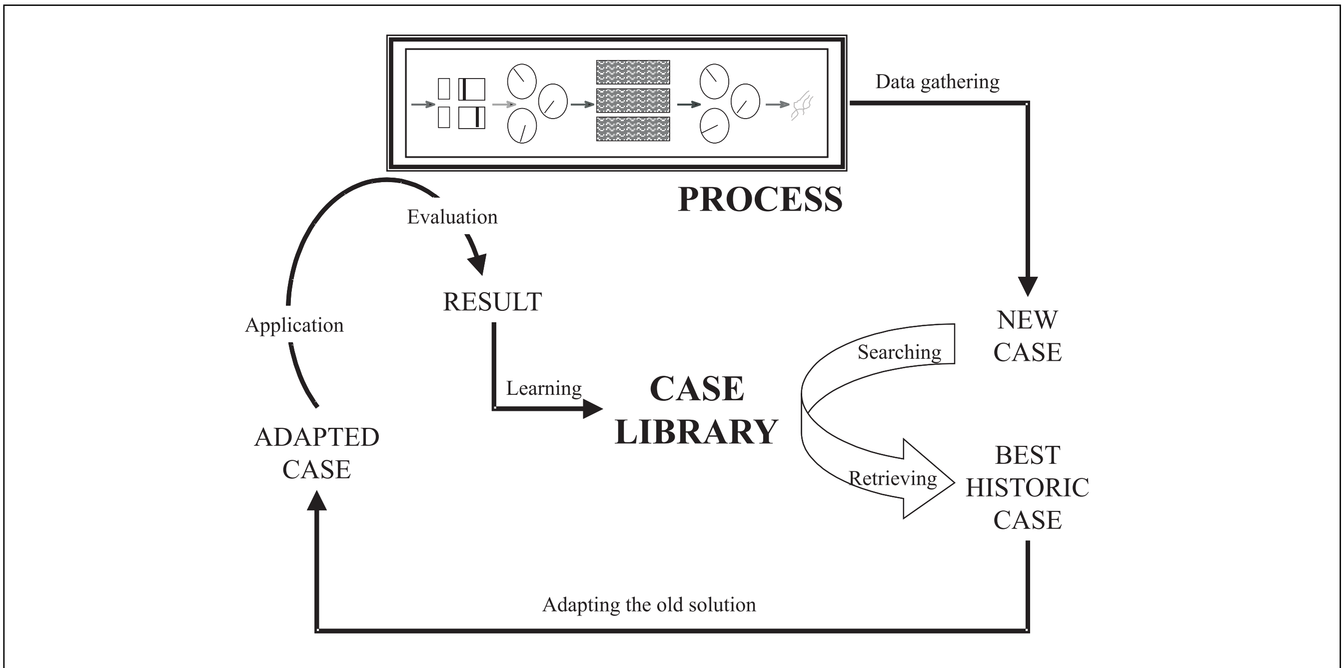

Figure 5. Case-Based System working cycle.

3), which should include all the components and processes used to define the model. Many years ago the IWA group proposed Activated Sludge Model nr. 1 (ASM nr. 1, [12]), which has been internationally accepted as the model of reference for carbon and nitrogen removal. These models require a large number of kinetic and stoichiometric parameters, which should be previously adjusted, by means of a calibration phase with experimental data. The validity of the model developed will largely depend on the reliability of the kinetic and stoichiometric parameters selected.

Table 3. Matrix representation of models

Processes	State Variables X	State Variables S	Velocity Rates
Growth	1 *	$-\frac{1}{Y}$	$\frac{\mu_{max} \cdot S}{K_s + S} X$
+ Dead	-1 *	-	$K_d \cdot X$
Stoichiometric parameters	Biomass	Substrate	**Kinetic parameters**
Y, yield			μ_{max}, K_s, K_d

The first simulation software using mechanistic models for nutrient removal was developed in the late 80s [6]. Nowadays, there exist several commercial simulators for wastewater treatment processes. Among them, the GPS-X software, developed by [15], includes all the models developed by the IWA group and some modifications of them.

A mechanistic model of the Granollers wastewater treatment plant was developed using the GPS-X commercial software [15]. The biological reactor was modelled as four Continuous Stirred Tank Reactors, while primary and secondary settlers use a model that consists of a one-dimensional tank with 10 layers of solids flux without biological reaction. A calibration process has been carried out to adjust the kinetic, stoichiometric and settling parameters of the ASM2 model [13] of this plant. The standard values for these parameters given by the GPS-X were used in the first step of the calibration procedure being changed manually along the calibration process. A mechanistic model enables to simulate several *off-line* scenarios with different operational conditions, changes in the influent characteristics (underloading, overloading, storms...), and alternative actions proposed by the Supervisory System. In spite of these capabilities, these models present limitations when dealing with problematic situations of biological origin (filamentous bulking, foaming, rising...) as well as with situations for which they have not been calibrated. In this sense, the utilization of soft-computing techniques to build a non-mechanistic model to simulate the behaviour of the plant in any situation is also being studied.

In addition, the group has also explored the field of soft computing techniques, in particular, by experimenting with neural networks and fuzzy approaches. The final aim is to develop a non-mechanistic or black-box model to predict the evolution of activated sludge processes in short-term under any situation and to integrate it into BIOMASS. Specifically, system identification and real-time pattern recognition by neural networks were studied to estimate key parameters in the activated sludge process ([7], [8], [9] and [10]). In research that is still going on, heterogeneous time-delay neural networks are being experimented to study the influence of qualitative variables coming from microscopic examinations and subjective remarks of the operators on the prediction of the sludge settleability. This model, in which inputs are a mixture of continuous (crisp, rough or fuzzy) and discrete values, performs effectively even when dealing with missing values ([1] and [2]).

The supervisory cycle

The different tasks of the third level of BIOMASS are performed cyclically, using a supervisory cycle (see Figure 6). Each cycle is composed of six steps: data gathering and update; diagnosis; supervision; prediction; user-validation and action phase; and evaluation phase.

Every time the supervisory cycle is launched, the first task to be carried out is *data gathering* and *updating* the current data for the inference process, in other words, gathering the most recent, both quantitative and qualitative data from the evolving database. According to the expert, there are minimal essential variables - basic information - that must be updated in order to make a reliable diagnosis of the current state of the process. In the Granollers WWTP, these are the influent flow rate and the Chemical Oxygen Demand of the biological influent.

Once the information has been collected, it is sent to the *diagnosis* module where the knowledge-based systems (ES and CBS) are executed concurrently without any kind of interaction between them. The current state of the process will be diagnosed through a reasoning task carried out according to both the expert rules and the most similar cases retrieved. If a problem is detected or suspected, the diagnosis module will also try to identify the specific cause. The solution of the most similar case is modified to adapt it to the new situation. The conclusions of diagnosis phase are sent to the decision support module. This upper module infers a global situation of the WWTP and suggests a proper action plan as a result of the *supervision* and *prediction* tasks integrating the expert recommendations sent by the ES and the experience retrieved by the CBS, while evaluating any possible conflict. The final result of BIOMASS is sent through the computer interface to the operator who will finally decide on the action to be taken (*user-validation* and *action*). The expert can use the dynamic model implemented in the GPS-X shell to support the selection of an action plan by simulating the possible consequences of applying different alternatives. Finally, the *evaluation* of the results of the application of the action plan to solve the problem in the process allows the system *to close* the CBS cycle, in other words, to learn from failure or successful experiences and to upgrade the case-library. Thus, if the current case is quite different from all historical cases from the case-library, it should be stored as a

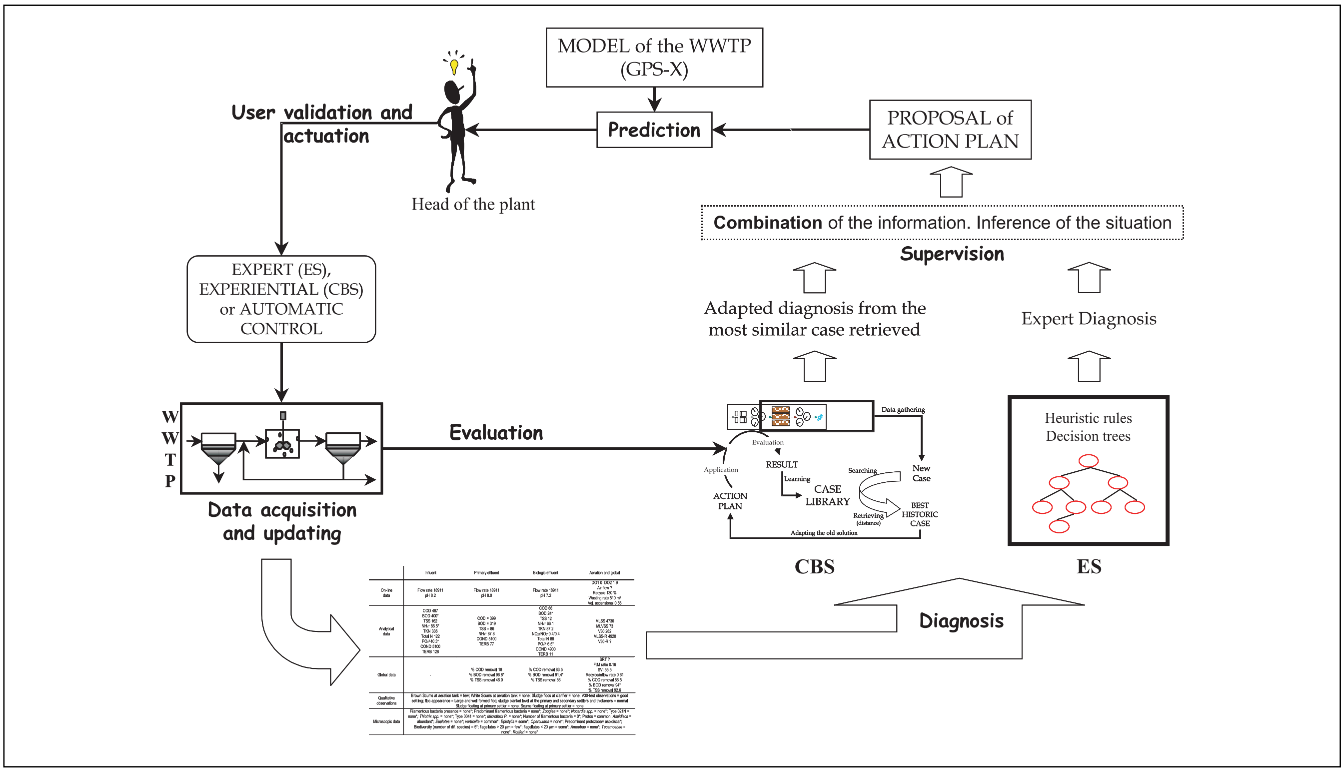

Figure 6. Supervisory cycle.

new experience with the diagnosis and comments of the process state, the action carried out and the evaluation of its application to the plant. These features can be detected by the Supervisory System itself (unless a manual operation is carried out), but it is essential to provide the confirmation by the plant manager who will have the opportunity to change misleading or add missing information. On the other hand, the Supervisory System can also extend the knowledge base by acquiring new knowledge from new sources (new experts or new automatic data classification).

Experimental System

The wastewater treatment plant selected to develop and apply our proposed Supervisory System prototype is located in Granollers, in the Besòs river basin (Catalonia, NE of Spain). This plant initially included preliminary and physical-chemical treatment for organic matter and suspended solids removal (built in 1992). In April 1998, the plant was expanded to include biological treatment, and physical-chemical treatment was completely replaced. Therefore, nowadays, this facility provides preliminary, primary and secondary treatment to remove the organic matter, suspended solids and, under some conditions, nitrogen contained in the raw water of about 130,000 inhabitant-equivalents. The raw influent comes from a sewer that collects the urban and industrial wastewater together. A current plan of the Granollers WWTP is shown in Figure 1.

The Granollers WWTP has several particular characteristics that increase the potential advantages of the development and application of an intelligent supervisory system to control and supervise the wastewater treatment process. Among these characteristics, we emphasize the following:

- Availability of a significant amount of historical records describing plant operation. These records include either quantitative information (e.g., analytic determinations of sludge and water quality at different locations in the plant, and on-line signals from different sensors) and qualitative information (e.g., microscopic observations of mixed liquor twice or three times a week, biological foam presence, filamentous bulking sludge which interfere with the settleability, sludge floating in clarifiers and, in general, any abnormality).
- This plant has a high level of automation centralised in a computer that collects on-line data and controls most of the plant operations.
- The Granollers WWTP is a highly variable system . Continuous change occurs in hourly loading because the plant is located in an area of the Besòs river basin with an important contribution of industrial activities.
- A wide range of different situations taking place throughout the year (storms, overloading, nitrification in hot periods, uncontrolled industrial spills...), causing significant changes in the influent characteristics, which affect standard process operation.
- High level of specialization of plant experts who have been working in the plant from the beginning of its operation. They are perfectly acquainted with all sorts of details that make up the heuristic knowledge of the plant.

Validation of the Supervisory System

Field testing was considered to be the most effective validity test. The main objective of field validation was to test the use of the overall Supervisory System *in situ* with actual real cases. We were attempting to test the system within its real environment and to identify needs for further modifications. The System performance was tested in its actual operating environment working as a real-time decision support system for more than 10 months. During this period of exhaustive validation of the Supervisory System, it was able to successfully identify 123 different problem situations, suggesting suitable action strategies. Nowadays, the Supervisory System is used as a complementary tool of diagnosis for the usual management of the activated sludge process.

Conclusions

An integrated intelligent Supervisory System for the supervision of WWTP management has been developed, implemented and tested in a real plant (Granollers WWTP). The system integrates an expert system and a case-based system with classical control in a three-level architecture: data gathering, diagnosis module and decision support module. We think that the research presented in this paper can be considered as a good example of a successful application of a multidisciplinary approach to tackle complex problems, starting from a basic research and finishing to its implementation at an industrial scale, in order to contribute to the amelioration of our environment.

Acknowledgements

This research received support from the CICyT, CIRIT, *Agència Catalana de l'Aigua* (ACA) of the *Generalitat de Catalunya* and *Consorci per la Defensa de la Conca del Riu Besòs*.

References

[1] Belanche, Ll., Valdés, J.J., Comas, J., R.Roda, I. and Poch, M., Towards a Model of Input-Output Behaviour of Wastewater Treatment Plants using Soft Computing Techniques, *Environmental Modelling & Software*, 14(5), pp. 409-419, 1999.

[2] Belanche, Ll., Valdés, J.J., Comas, J., R.Roda, I. and Poch, M., Prediction of the Bulking Phenomenon in Wastewater Treatment Plants, *Artificial Intelligence in Engineering*, 14(4), pp. 307-317, 2000.

[3] Comas, J., Dzeroski, S., Gibert, K., R.Roda, I. and Sánchez-Marrè, M., Knowledge Discovery by means of Inductive Methods in Wastewater Treatment Plant Data, *AI communications*, 14(1), pp. 45-62, 2001.

[4] Comas, J., *Development, Implementation and Evaluation of an Activated Sludge Supervisory System for the Granollers WWTP*, PhD Thesis, Universitat de Girona, 2000.

[5] Cortés, U., Sànchez-Marrè, M., Cecaronni, L., R.Roda, I. and Poch, M., Environmental Decision Support Systems, *Applied Intelligence*, 13(1), pp. 77-91, 2000.

[6] EFOR Aps., clo Krüger Systems, Gladsaxevej 363, DK-2860 Soborg, Denmark.

[7] Fu, C.S. and Poch, M., Application of multilayered neural network to system identification, *International Journal of Systems Science*, 24, pp. 1601-1609, 1993.

[8] Fu, C.S. and Poch, M., System identification and real-time pattern recognition by neural networks for an activated sludge process, *Environment International*, 21, pp.57-69, 1995.

[9] Fu, C.S. and Poch, M., Fuzzy modeling and pattern recognition for dynamic processes and its application, *Chemical Engineering Science*, 50, pp. 3715-3725, 1995.

[10] Fu, C. S. and Poch, M., Fuzzy model and decision for COD control of an activated sludge process, *Fuzzy Sets and Systems*, 93, pp.281-292, 1998.

[11] *Gensym G2 Reference Manual Version 5.0*, Gensym Corporation, Cambridge, Massachusetts, 1997.

[12] Henze, M., Grady, C.P.L.Jr., Gujer, W., Marais, G.vR. and Matsuo, T., *Activated Sludge Model nr. 1, IAWPRC Scientific and Technical Reports nr. 1*, IAWPRC, London, 1987.

[13] Henze, M., Gujer, W., Mino, T., Matsuo, T., Wentzel, M.C. and Marais, G.v R., *Activated Sludge Model No. 2, IAWPRC Scientific and Technical Reports nr. 3*, IAWPRC, London, 1995.

[14] Henze, M., Harremoës, P., La Cour Jansen, J. and Arvin, E., *Wastewater Treatment. Biological and Chemical Processes*, 2nd ed., Springer-Verlag, Berlin, Heidelberg, 1997.

[15] Hydromantis. *GPS-X Technical Reference*, Hydromantis, Inc., Ontario, 1995.

[16] Kolodner, J., *Case-Based Reasoning*, Morgan Kaufmann, 1993.

[17] Lein, J.K., *Environmental Decision Making. An Information Technology Approach,* Blackwell Science, Ltd. , Malden (MA) USA, 1997.

[18] Manesis, S.A., Sapidis, D.J. and King, R.E., Intelligent Control of Wastewater Treatment Plants, *Artificial Intelligence in Engineering*, 12, pp.275-281, 1998.

[19] R.Roda, I., Comas, J., Sànchez-Marrè, M., Cortés, U., Lafuente, J. and Poch, M., Expert system development for a real waste water treatment plant, *Chemical Industry and Environment III*, Edited by R. Zarzycki and Z. Malecki, ISBN 83-87198-68-4, Kraków (Poland), pp. 653-660, 1999.

[20] R.Roda, I., Sànchez-Marrè, M., Cortés, U., Lafuente, J. and Poch, M., Case-Based Reasoning Systems: a Suitable Tool for the Supervision of Complex Processes, *Chemical Engineering Progress*, 95(6), pp. 39-47, 1999.

[21] R.Roda, I., Sànchez-Marrè, M., Comas, J., Cortés, U., Lafuente, J. and Poch, M., «An intelligent integration to improve the control and supervision of large WWTP», 8 IAWQ Conference on *Design, Operation and Economics of Large Wastewater Treatment Plants*, Budapest, pp. 566-571, 1999.

[22] R.Roda, I., Sànchez-Marrè, M., Comas, J., Cortés, U. and Poch, M., Development of a case-based system for the supervision of an activated sludge process, *Environmental Technology*, 22(4), pp. 477-486, 2001.

[23] R.Roda, I., Comas, J., Poch, M., Sànchez-Marrè, M. and Cortés, U., Automatic Knowledge Acquisition from Complex Processes for the development of Knowledge Based Systems, accepted for publication to *Industrial & Engineering Chemical Research* (march 2001).

[24] Sànchez-Marrè, M., Cortés, U., R.Roda, I., Lafuente, J. and Poch, M., DAI-DEPUR: A distributed architecture for wastewater treatment plants supervision, *Artificial Intelligence in Engineering*, 10, pp. 275-285, 1996.

[25] Sànchez-Marrè, M., Cortés, U., R.Roda, I., Poch, M. and Lafuente, J., Learning and adaptation in wastewater treatment plants through case-based reasoning, *Microcomputers in Civil Engineering*, 12 (4), pp. 251-266, 1997.

[26] Sànchez, M., Cortés, U., Béjar, J., de Gracia, J., Lafuente, J., and Poch, M., Concept formation in WWTP by means of classification techniques: a compared study, *Applied Intelligence*, 7, pp.147-166, 1997.

[27] Sànchez-Marrè, M., Cortés, U., R.Roda, I. and Poch, M., Sustainable Case Learning for Continuous Domains, *Environmental Modelling & Software*, 14, pp. 349-357, 1999.

[28] Sànchez-Marrè, M., R.Roda, I., Comas, J., Cortés, U. and Poch, M. Using Meta-cases to Improve Accuracy in Hierarchical Case Retrieval, *Computación y Sistemas*, vol.4, nr. 1, pp.53-63, 2000.

[29] Serra, P., Sànchez-Marrè, M., Lafuente, J., Cortés, U. and Poch, M., ISCWAP: A knowledge-based system for supervising activated sludge processes, *Computers and Chemical Engineering*, 21(2) pp. 211-221, 1997.

About the authors

The Laboratori d'Enginyeria Química i Ambiental *(LEQUIA, http://lequia.udg.es) is a research group at the* Universitat de Girona*, which was set up ten years ago as a dynamic framework in which different academic staff could combine their research areas, from different points of view, all of them focused on environmental aspects. Recently, the LEQUIA has become a member of the* Xarxa de Centres de Suport a la Innovació Tecnològica *from CIDEM (*Generalitat de Catalunya*). LEQUIA has wide experience in the field of wastewater treatment plant control and supervision, mainly in the application of computing techniques to improve management of real plants. Though we started with a classical approach, applying our previous experience in the control of lab bioreactors, we soon realised that classical control methods show some limitations when dealing with complex environmental processes. Based on this idea, LEQUIA has been working in close association with several research groups in Environmental Engineering, Chemical Engineering and Artificial Intelligence, mainly with the KEML group, concluding that it is necessary to integrate an array of specific supervisory intelligent systems and numerical computations for detailed engineering when trying to optimally manage complex environmental problems. In this sense, our research covers different approaches such as classical modelling, knowledge-based systems (mainly expert systems and case-based reasoning) and soft computing.*

The Environmental Engineering Group (http://eq3.uab.es /depuradoras/) is part of the Chemical Engineering Department of the Universitat Autònoma de Barcelona. *This group started its activities 20 years ago, and during this time it has been working in several fields, such as the design and implementation of anaerobic processes, detoxification processes, nutrient removal processes (as nitrification-denitrification and enhanced biological phosphorus removal) and wastewater treatment plant automatic control. This last topic includes development and implementation of real-time knowledge-based expert systems for different full-scale wastewater treatment systems, done in co-operation with the LEQUIA and KEML groups.*

The Knowledge Engineering and Machine Learning group (KEML, http://www.lsi.upc.es/~webia/gr-SBC/ KEMLG.html) is a research group of the Artificial Intelligence Section of the Software Department at the Universitat Politècnica de Catalunya *(UPC). The group has been active in the Artificial Intelligence field since 1988. Main research areas of KEML group are Knowledge-Based Systems, Data Mining, Knowledge Engineering and Management, Case-Based Reasoning, Machine Learning, Autonomous Agents and Integrated AI architectures. KEML has a special interest on the application of AI techniques to Environmental Decision and Design Support Systems, and has been working intensely with several research groups, especially with LEQUIA.*