PERGAMON

Water Research 37 (2003) 2377–2387

WATER
RESEARCH

www.elsevier.com/locate/watres

A knowledge-based approach to the deflocculation problem: integrating on-line, off-line, and heuristic information

J. Comas[a,]*, I. Rodríguez-Roda[a], M. Sànchez-Marrè[b], U. Cortés[b], A. Freixó[c], J. Arráez[c], M. Poch[a]

[a] *Chemical and Environmental Engineering Laboratory, University of Girona, Campus Montilivi s/n, Girona E-17071, Spain*
[b] *Software Department, Technical University of Catalonia, C/Jordi Girona 1-3, Barcelona E-08034, Spain*
[c] *Consorci per a la Defensa de la Conca del Riu Besòs, Camí Ral, s/n. Granollers 08400, Spain*

Received 31 October 2001; accepted 30 December 2002

Abstract

A knowledge-based approach for the supervision of the deflocculation problem in activated sludge processes was considered and successfully applied to a full-scale plant. To do that, a methodology that integrates on-line, off-line and heuristic information has been proposed. This methodology consists of three steps: (i) development of a decision tree (which involves knowledge acquisition and representation); (ii) implementation into a rule-based system; and (iii) validation. The set of symptoms most useful in diagnosing the deflocculation problem has been identified, the different branches to diagnose pin-point floc and dispersed growth have been built (using generic and specific knowledge), and all this knowledge has been codified into an object-oriented shell. The results obtained in the application of this knowledge-based approach to the Granollers WWTP (which treats about 130,000 inhabitants-equivalents) showed that the system was able to identify correctly the problem with reasonable accuracy. Our positive experience building this system suggests that this approach is a practical and valuable element to include in an intelligent supervisory system combining numerical and reasoning techniques.

Keywords: Deflocculation; Wastewater treatment; Activated sludge; Operational problems; Knowledge-based systems

1. Introduction

The operational problems of the activated sludge process that have a biological cause are among the most serious and most difficult to detect and solve in wastewater treatment plants (WWTP). They can decrease the biological removal efficiency, and hinder the compactation and settling processes. Filamentous

---

Abbreviations: BOD, biological oxygen demand ($g\,O_2/m^3$); COD, chemical oxygen demand ($g\,O_2/m^3$); Cond, conductivity (μS/cm); DO, dissolved oxygen ($g\,O_2/m^3$); EPS, extracellular polymeric substances; F/M ratio, food to microorganism ratio (g BOD/g VSS d); HRT, hydraulic residence time (d); KB, knowledge base; KBS, knowledge-based system; MLSS, mixed liquor suspended solids (g TSS/m^3); MLVSS, mixed liquor volatile suspended solids (g VSS/m^3); NH_4^+, ammonia (g N/m^3); NO_2^-, nitrite (g N/m^3); NO_3^-, nitrate (g N/m^3); P, phosphorus (g P/m^3); PLC, programmable logic controller; RAS flow rate, return activated sludge flow rate (m^3/d); RBS, rule-based system; SCADA, supervisory control and data acquisition; SRT, sludge residence time (d); SVI, sludge volume index (ml/g); T, temperature (°C); TKN, total Kjeldhal nitrogen (g N/m^3); TSS, total suspended solids (g TSS/m^3); Turb, turbidity (NTU); V30, volume of 30-min settled sludge (ml/l); WAS flow rate, waste activated sludge flow rate (m^3/d); VSS, volatile suspended solids (g VSS/m^3); WWTP, wastewater treatment plant

*Corresponding author.

E-mail addresses: quim@lequia.udg.es (J. Comas), ignasi@lequia.udg.es (I. Rodríguez-Roda), miquel@lsi.upc.es (M. Sànchez-Marrè), ia@lsi.upc.es (U. Cortés), freixora@diba.es (A. Freixó), arraezej@diba.es (J. Arráez), manel@lequia.udg.es (M. Poch).

0043-1354/03/$ - see front matter © 2003 Elsevier Science Ltd. All rights reserved.
doi:10.1016/S0043-1354(03)00018-6

bulking and deflocculation are the most common problems that cause activated sludge settling problems and a deterioration of effluent quality [1–3]. Of these, deflocculation is the least well-known. Microscopical observations reveal that the poor settling conditions characteristic of the deflocculation problem are not due to the excessive presence of filamentous bacteria (bulking), but poor floc properties [4].

Deflocculation refers to a dysfunction of the activated sludge process characterised by the formation of a very small sludge floc, or the absence of floc formation. Two types of deflocculation problems can be distinguished: pin-point floc and dispersed growth (some authors make no distinction between dispersed growth and deflocculation) [5]. The improper formation of the activated sludge flocs causes bad separation between the microorganisms and the treated water, which results in effluent turbidity and, consequently, lower clarifier efficiency.

Diagnosis of deflocculation involves the use of multiple and heterogeneous data sources. Detection involves on-line data (e.g., dissolved oxygen, DO), off-line numerical measurements (e.g., sludge volume index, SVI), and most importantly, the evaluation of qualitative features such as the presence of turbidity, the absence of bubbles in the settling test, and specially, microscopic observations of the floc characteristics. Moreover, the actions needed to solve the problem depend largely on the characteristics of the treatment plant.

Currently, most WWTPs are equipped with automatic control and data acquisition systems capable of monitoring plant conditions by retrieving numerical data from database management systems, and of controlling the essential operational and control equipment through a PLC network [6,7]. Besides this on-line data, the analytical results of the water and sludge quality and specially, the microbiological information of activated sludge or qualitative observations about the process state, supply a type of information that, once processed and interpreted by an expert, is fundamental to increase the reliability of the WWTP control and management [8]. Incorporating all these heterogeneous sources of knowledge, and most importantly, reasoning with the whole information collected is beyond the scope of classical control systems. A method based on human reasoning can help in this task.

Knowledge-based systems (KBS) are computational tools that enable the integration of numerical data and heuristic knowledge and mimic the human-decision making processes to solve complex problems. These systems are capable of using the knowledge of human experts, acquired through years of experience, to diagnose the state of a process, and to propose solutions to new problems. The potential use of KBS to support the operation of WWTP came into picture in the early 1980s. In the 1990s several interesting proposals were developed (e.g., [9–12]). These applications never really succeeded because they were too complex, and the available knowledge could not be captured in reliable models and advisory systems [35]. In fact, a rigorous evaluation of these proposals could never be done since they were not installed in real facilities, but performed under hypothetical simulated problem testing, presented as simplified case studies, or supervising specific experiments carried out in pilot plants. A new generation of KBS is recently being proposed, based on hybrid architectures that combine classical expert systems with soft computing (neural networks, fuzzy logic) and complementary knowledge-based techniques to deal with specific knowledge of the process (e.g., case-based reasoning). Some of these systems, which include real-time intelligent data interpretation and validation (also handling different scale qualitative information), have been successfully implemented and evaluated in real facilities (e.g., [13–16]).

The development of a KBS involves the crucial steps of knowledge acquisition, knowledge representation and knowledge implementation. The knowledge acquisition can easily become a bottleneck in building a good quality system. In order to build a complete knowledge base (KB), several knowledge sources and methods can be used. Methods to identify and collect this information include those involving the human senses (i.e. carrying out interviews with the experts or reading specialised literature), and those involving the use of machines (automatic learning tools) to acquire knowledge from a database [17–19]). All the knowledge acquired during the knowledge acquisition process can be then synthesized in a suitable and understandable form of representation such as a decision tree.

This paper illustrates the development, implementation, and application of a KBS for the diagnosis and solution of the deflocculation problem in WWTPs. Although it focuses on one specific problem it presents a methodology of general applicability to any problem.

The goals of the paper are threefold. Firstly, the methodology developed for building the knowledge-based approach to the deflocculation problem to assist in the operation of a full-scale WWTP is discussed in detail. Secondly, the deflocculation problem decision tree resulting from the use of this methodology is examined and discussed. Finally, a validation of the KBS is effected by checking the performance of the system during a four-month trial in real-life situations at the Granollers WWTP.

2. Experimental system

2.1. Plant description

A KBS to support the diagnosis and solution of the deflocculation problem was developed for the Granollers

WWTP located in the Besòs river basin (Catalonia, NE Spain). This facility provides preliminary, primary and secondary treatment to remove the organic matter, suspended solids and nitrogen contained in the raw water of about 130,000 inhabitant-equivalents. The configuration of the secondary treatment is based on the modified Ludzack-Ettinger configuration for biological nitrogen removal. The raw influent comes from a sewer that collects together urban and industrial wastewater.

2.2. Description of the database

Plant and laboratory operators carry out a daily characterisation of the water and sludge quality and process state variables, including both quantitative and qualitative information, at different sample points (influent, primary settler, aeration tank, secondary settler, and effluent). Quantitative data can be divided into on-line data provided by sensors and directly acquired by the SCADA system, off-line data provided by the analytical determinations from different sample points, and combinations of quantitative data that allow the calculation of global process state variables (e.g., sludge residence time, hydraulic residence time (HRT), sludge volume index, food to microorganism ratio, and COD, BOD, and SS removal rates of primary, secondary, and overall treatment). The qualitative data include microscopic examinations of the activated sludge (identification and counting of the different species of ciliates, flagellates, amoebae and metazoans, identification and counting of the filamentous bacteria and floc characterisation, i.e., morphology and size of the sludge floc and filament effects on its structure), and daily in situ macroscopic observations about plant performance, quality of biomass, and settling characteristics. Crucially for this application, qualitative information was recorded in a standardised format to ensure that it would be easily and effectively managed, retrieved and interpreted.

3. Methodology

The methodology proposed in this paper for a knowledge-based approach to the deflocculation problem consists of three steps: (i) development of a decision tree, (ii) implementation into a KBS, and (iii) validation.

3.1. Development of a decision tree for the deflocculation problem

Decision trees represent the causal chains of interactions from symptoms to problems, causes and solutions. To build a decision tree, it is first necessary to state explicitly the kind and number of problems that are to be addressed by the tree. In this paper, the problems of dispersed growth and pin-point floc were considered. With the intention of not missing any indispensable information for the detection and troubleshooting of the deflocculation problem, the development of this decision tree was accomplished in two key steps (Fig. 1): (1) knowledge acquisition, and (2) knowledge representation.

3.1.1. Knowledge acquisition

Knowledge acquisition involved searching and compiling information contained in diverse sources of knowledge, and identifying the reasoning mechanisms followed by experts to detect and solve the problem:

(i). The first task involved identifying, striving for multiplicity, the warning symptoms that indicate a potential deflocculation situation (an intermediate alarm). This was accomplished through an analysis of the historical database (on-line and off-line data) and a series of interviews with experts. A set of 334 consecutive days and 78 variables was selected from the historical database as a representative period containing the most common situations that occur in the process (including deflocculation). It is important to learn about any incidence that could have affected the information registered during this period (i.e., changes in plant configuration or mechanical failures). Then, the raw data contained in this 334 × 78 matrix was filtered to avoid type-errors and to remove incorrect signals and poorly informative variables (redundancies). A method of automatic knowledge acquisition was then applied to the database to extract objectively the most relevant information. This methodology, consisting of a four-step process, i.e., data handling, automatic classification, interpretation, and codification [20], allowed us to narrow down the selection of the most relevant variables for the identification of each of the operational states of the plant, and the locations where they were observed in the plant. Once the variables displaying deflocculation symptoms were identified, their qualitative warning values and ranges (high, low, normal or trends—increasing, decreasing or stable) were established with the help of a statistical analysis based on histogram representations and a fuzzyfication process to avoid crisp confines.

(ii) The identification of the reasoning mechanisms to diagnose deflocculation was complemented with interviews with experts in biological wastewater treatment. Interviews are essential because they provide heuristic knowledge that makes the general reasoning mechanisms pointed out by the literature

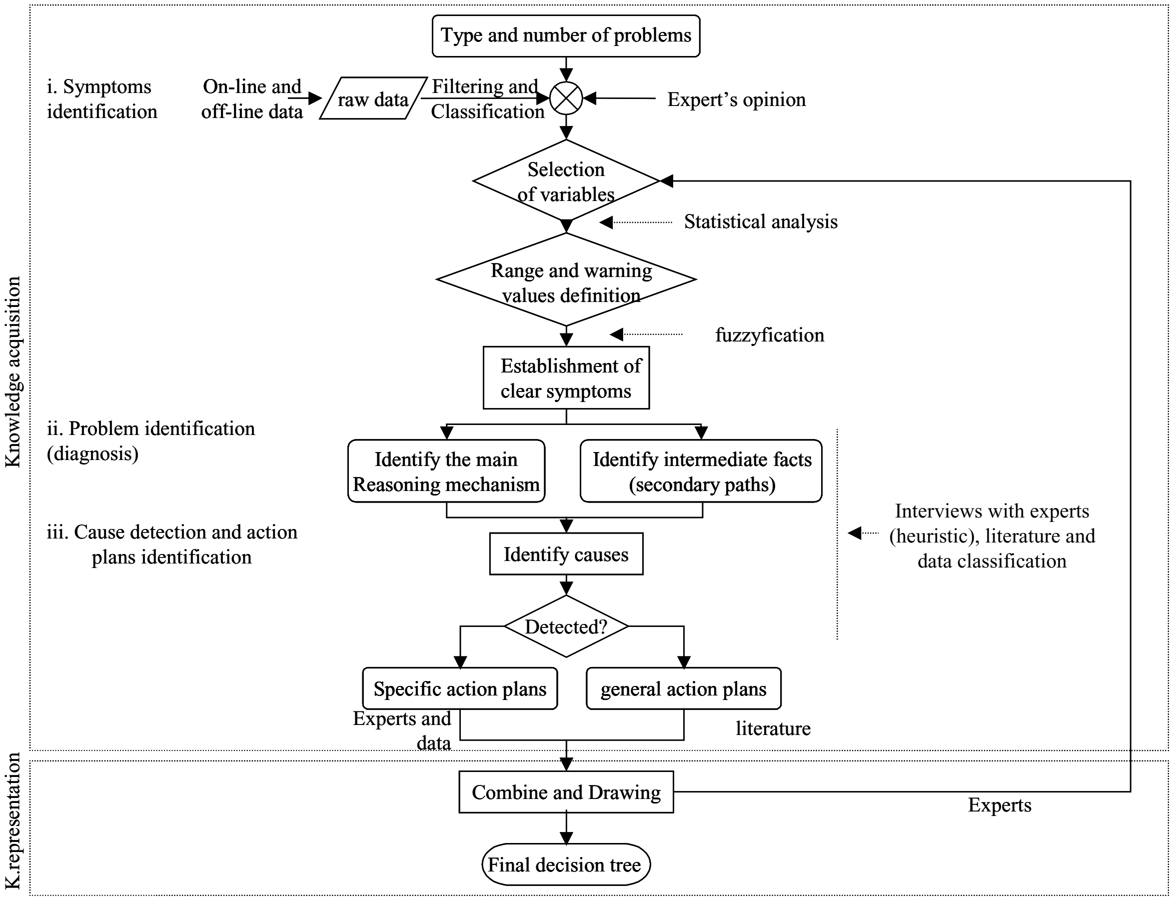

Fig. 1. Flow chart for the development of a decision tree.

more specific to the plant and problem under consideration.

(iii) Finally, cause detection and action plans were identified and integrated. Deflocculation conditions have multiple causes, all of which need to be identified so that specific action plans can be recommended. The specific action plans are the particular strategies that have proven efficient and have been validated by the manager of the studied plant. These strategies are gathered during the course of interviews with the experts of the plant and through the automatic data classification process. Specific action plans are preferred whenever the cause is well detected. However, if the cause is not well determined or if specific knowledge does not cover totally the situation, it may be necessary to resort to the application of a general action plan.

3.1.2. Knowledge representation

The symptoms, facts, relationships and methods used by the experts in their reasoning strategy, and identified during the previous phase, were organised and documented as a decision tree. Decision trees consist of hierarchical, top–down descriptions of the linkages and interactions among any pieces of knowledge used for problem solving. To ensure an optimal representation, the development of decision trees involved two components:

- The identification of the sources of information used by the experts to detect the problem and the identification of the solutions to the problem being discussed. This *descriptive knowledge* is represented as the nodes or leafs of the branches of the tree.
- The reasoning procedures used in the problem-solving processes. This *procedural knowledge* is translated into the branches of the decision tree.

Once the tree was documented and adapted to the Granollers WWTP, the experts could easily revise the correctness of each branch through an iterative process.

3.2. Implementation of the trees

Since the decision trees can be simply converted to production rules by traversing each branch from the root to the leaves, the most suitable knowledge-based tool to implement these trees is a rule-based system (RBS). The collection of IF–THEN rules and sequences of actions (procedures or plans) extracted from the deflocculation decision tree constitute part of the KB of the RBS.

The RBS presented here was developed with the G2 object-oriented programming shell [21], which already includes an inference engine, i.e., the computational mechanism that chains this knowledge, and allows for the development of a user-friendly front end. In addition, G2 supports on-line and off-line data acquisition from different data sources and provides easy connectivity with conventional SCADAs. For the complete characterisation of the Granollers WWTP, 679 classes of hierarchical objects and about 1000 parameters and variables were defined [22]. A menu-based interface provides a simple way for users to communicate with the KBS, allowing the user to know the current state and the tendency of the process. Nonetheless, the system has been built so as to minimise input from the user. Whenever the inference engine requires new data to reach a conclusion, it tries to get an updated value by itself, or to evolve the reasoning through other branches or rules using the available information.

3.3. Validation

The objective of the validation phase was to guarantee the performance of the KBS module related to deflocculation, while checking for its adequacy, accuracy, and compliance with user-required specifications. The validation was performed in two stages: *laboratory testing* and *field testing* [23,24]

- *Laboratory testing* involved the execution of several series of *off-line* experiments to validate the correctness, consistency and usability of the KBS. These *off-line* experiments were first performed with real data selected from the historical database. Then, a set of cases representing situations not collected in the database was prepared to learn how the expert system faced them.
- *Field testing* involved letting the KBS face real situations while working as a real-time support system at the Granollers WWTP. The objective was to test the system within its real environment and to identify needs for further modifications. To do that, it was necessary to integrate the KBS into the supervisory system that manages the plant [25,16].

The flow of information through the system was exhaustively followed and compared against the reasoning mechanisms of experts to detect weak reasoning and to ensure that the system could deal with real qualitative variables and missing information. When necessary, the KB was refined, adjusted or corrected to cover unexpected errors, and extended to deal with new situations [22].

4. Results

4.1. Symptoms identification

The set of symptoms that were found to be most useful in diagnosing the deflocculation problem included

- off-line data from analytical determinations (high values of COD, BOD and TSS at the effluent, and an SVI value not high),
- calculated values derived from off-line and on-line data (F:M ratio high or increasing),
- and qualitative in-situ observations of plant performance (turbidity in the plant effluent, turbidity in the 30 min-settling test, and presence of floating sludge in the secondary clarifier surface).

In the Granollers WWTP, in particular, deflocculation appears to be associated with certain qualitative features of the V30-settling characteristics test (i.e., fast settling with supernatant moderately cloudy). All these signals constitute the antecedents of the decision tree and, when reached, launch an intermediate alarm of deflocculation (Fig. 2).

4.2. Problem diagnosis

Three main branches were obtained to diagnose the deflocculation process (Fig. 2):

(a) The central branch was obtained mainly from the literature. If information from the settling test is available, the KBS checks the turbidity values and the presence of bubbles. If bubbles are observed, the inference path is directed to search for possible denitrification problems in the secondary settler. If bubbles are not observed, or if this qualitative observation is not available, the decision tree guides the inference path to check for the SVI value. If SVI is high and the abundance of filamentous bacteria is high or increasing, the activated sludge could be suffering an episode of filamentous bulking or foaming, and the decision trees for these problems are invoked; otherwise, if the SVI is not high or decreasing, the system proceeds to the deflocculation diagnosis. The floc characterisation is evaluated next. If the floc is well

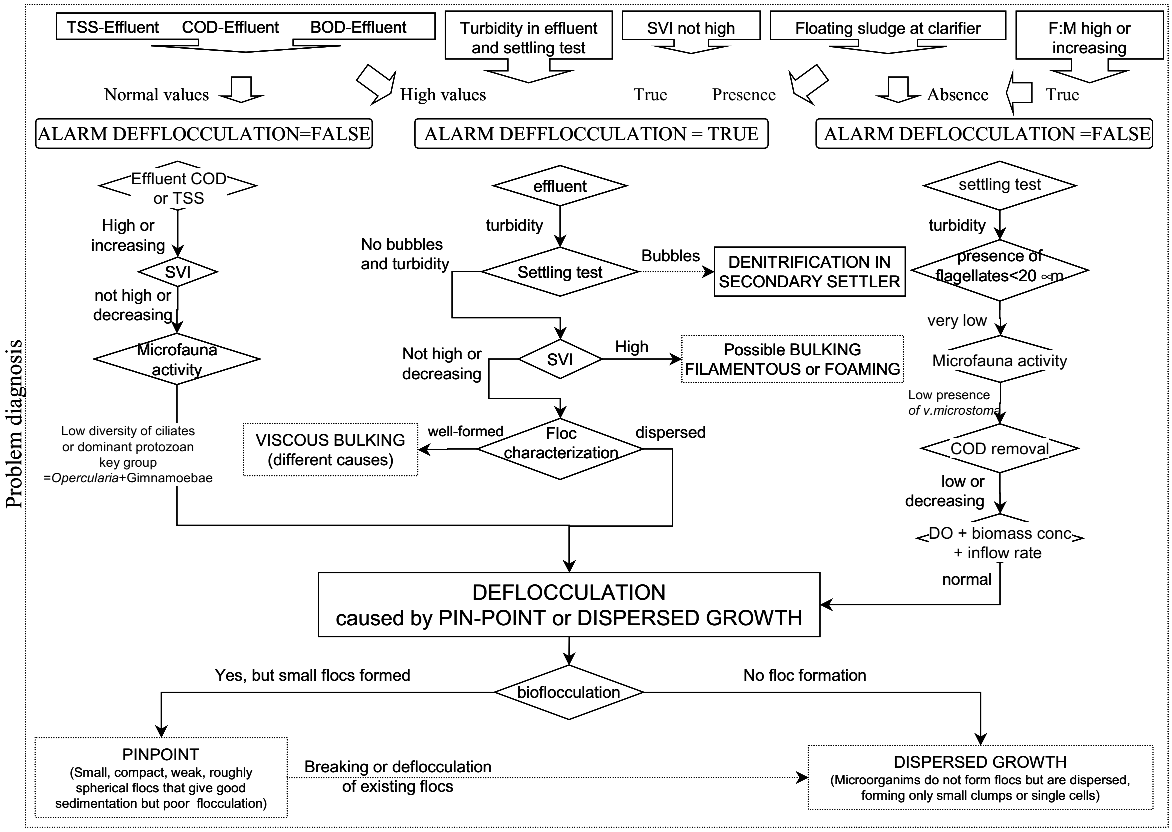

Fig. 2. Decision tree for diagnosing deflocculation problems.

formed, the viscous bulking decision tree is launched; if flocs are dispersed, a deflocculation situation is concluded.

(b) The left branch to diagnose deflocculation was obtained through the automatic classification process; therefore, it is completely specific to the Granollers WWTP. The most relevant variables of the cluster, obtained by automatic classification of the historical database and labelled by the expert as deflocculation, are: high or increasing values of the effluent COD or effluent TSS; low, normal or decreasing values of SVI; and low diversity of ciliates or dominance of *Opercularia* and *Gimnamoebae* among other protozoan key groups [20].

(c) Finally, a third branch to achieve the deflocculation diagnosis was obtained mainly from interviews with experts in WWTP management. It checks the mixed liquor for low abundance of small flagellates and low presence of *Vorticella microstoma*, and the supernatant of the V30-settling test for the presence of turbidity. Finally, if the total COD removal efficiency is low or getting worse despite normal levels of DO, biomass concentration, and inflow rate, then deflocculation is diagnosed.

Once deflocculation has been diagnosed, it is necessary to distinguish between the two types of deflocculation; this is done by checking for the occurrence of bioflocculation (Fig. 2) [26,27]. If only small flocs or microflocs are formed, the so-called pin-point problem is diagnosed; otherwise, if microorganisms do not form consistent biological flocs, but are dispersed, forming only small clumps (smaller than 30 μm) or growing as single cells causing uniform turbidity and hindering settleability, then dispersed growth is diagnosed. Although these two types of deflocculation are similar, their origin and microscopic observations are different. Dispersed growth is caused by a lack of production of extracellular polymeric substances (EPS), which prevents the bioflocculation. Pin-point is caused by a total absence of filamentous bacteria, which prevents the growth of activated sludge flocs; only small, compact and weak flocs are produced (smaller than 75 μm). Usually pin-point flocs are accompanied by a fraction of large flocs with a high density giving an activated sludge with good sedimentation but very poor flocculation.

4.3. *Cause detection and actions plans*

4.3.1. *Dispersed growth*

When dispersed growth is diagnosed (Fig. 3), the activity of the protozoa in the activated sludge is evaluated under the microscope:

(a) If protozoa are present but inactive, then a poisoning of the biomass is the most likely cause. By looking at the dominant group of protozoa, it is possible to determine whether the toxic shock was severe, or recent and weak [28]. Dominance of flagellates suggests a severe toxic shock load has occurred (possibly industrial dumping with metals or flocculation inhibitor); conversely, dominance by amoebae indicates that a recent toxic shock has occurred (possibly an industrial by-pass). Toxic shock actions involve the identification and elimination of the toxic source. If an industrial source is identified, a pre-treatment at the industrial plant is needed and sewer-use ordinance must be enforced. If the toxics are still in the system, the waste flow rate must be increased to remove the toxics with the excess sludge. Otherwise, the waste flow rate must be decreased and the DO level kept above 1.5 mg/L. If the number of microorganisms affected were large, it might become necessary to reseed with activated sludge from another plant. If a nutrient deficiency is identified as the cause, then an addition of nutrients will solve the problem. In any case, the execution of respirometric tests enables to check for the possible presence of toxics or inhibitors at the influent.

(b) If the observation of activated sludge shows an absence or inadequate presence of protozoa, the DO level is checked. If it is adequate according to the values established for the Granollers WWTP, the system concludes again that a severe toxic shock load has occurred and proposes the severe toxic shock actuations. If the DO level is high, the inferred cause is an over-aeration. Finally, oxygen deficit can cause old sludge (in the Granollers WWTP, SRT longer than 9 days). Old sludge actuations involve increasing sludge wasting to reduce SRT and adjusting the F:M ratio and increasing aeration to overcome the temporary oxygen deficit [29].

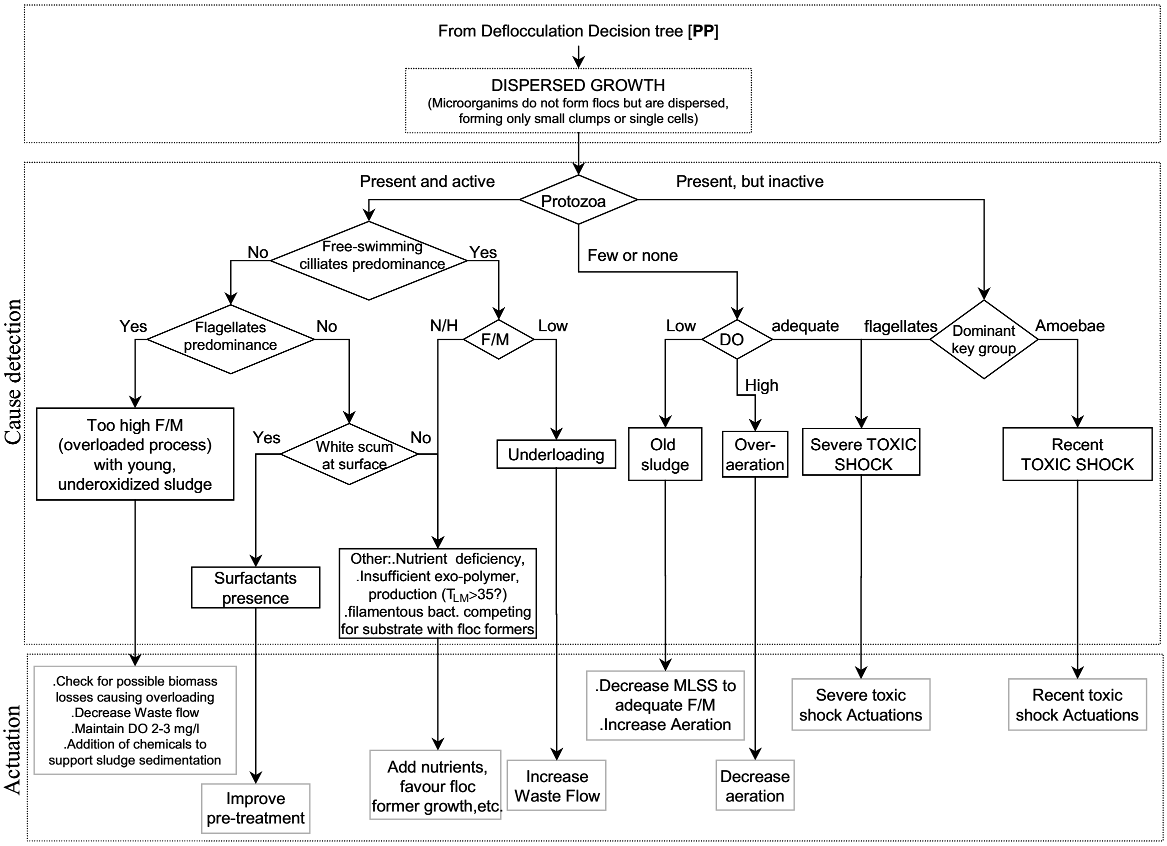

Fig. 3. Decision tree for solving disperse-growth problems.

(c) If active protozoa are present but free-swimming ciliates dominate, the F:M ratio is checked. If the F:M ratio is low, then the process is underloaded (old, over-oxidised sludge). The recommended actuation is to increase wasting rates. If the F:M ratio is normal or high, other facts must be checked: nutrient deficiency, temperature, presence of slowly biodegradable surfactants, and competition between filamentous bacteria and floc-forming bacteria for readily biodegradable substrate.

(d) When the observation of activated sludge shows normal presence of active protozoa and predominance of flagellates, the most likely cause of the dispersed-growth is organic overloading in the biological reactor (with young, under-oxidised sludge). Aside from process start-ups, the most usual causes of organic overloading are massive losses of solids to the effluent or the presence of readily biodegradable substrate. The organic overloading actuations entail adjusting an adequate F:M ratio (decreasing sludge wasting rates and avoiding undesirable losses of solids), maintaining proper DO levels or adding chemicals to improve the sludge settleability [8]. If batch wasting is effected, wasting when the BOD loading is increasing must be avoided. BOD load from side-streams should be included in F:M calculations.

(e) Finally, when active protozoa are present, but flagellates are not dominant, the system checks for the presence of white scum at the surface of the aeration reactor. If it is present, the presence of surfactants and tensoactive agents is inferred as the cause of dispersed-growth; if it is not, the cause could be nutrient deficiency, abnormal temperature of the mixed liquor, presence of slowly biodegradable surfactants, or competition between filamentous bacteria and floc-forming bacteria.

4.3.2. *Pin-point floc*

To distinguish among the different causes of pin-point floc formation (Fig. 4), the KBS starts by checking the DO level in the aeration basin:

(a) If the DO level is high, the pin-point floc is likely due to the break-up of large flocs due to excessive turbulence.

(b) If the DO level is low, the system checks for anaerobic conditions [30]. Recommended actions include checking for the aeration systems and, depending on DO readings, increasing or decreasing the airflow rate or stirring.

(c) If the DO level is normal, the HRT of the clarifier is checked. In Granollers, if the HRT is longer than 2 days, the system infers anaerobic conditions. The recommended actions are calculating the sludge residence time in the clarifier and increasing the recycle flow rate.

(d) Finally, if the HRT of the clarifier is low or normal and the SRT is high, pin-point floc formation is attributed to an old and overoxidised sludge due to extremely low-substrate concentrations (extremely low F:M ratio) in the biological reactor or wastewater with a high portion of poorly biodegradable organic substances, leading to endogenous metabolism. This produces a floating ash-like material in the surface of the V30-settling test, which does not release bubbles nor settles when stirring (*ashing*). In this case, microscopic observations show a high abundance of rotifers, and the floc is dark and compact. If the suspended particles release bubbles and settle when stirred, it is an indication that denitrification is occurring. The KBS mandates an increase in the wasting rate to adjust the F:M ratio and reduce SRT. Before concluding definitely that an *ashing* problem exists, excessive grease in the activated sludge (above 15% of MLSS by weight) must be ruled out. If the problem is excessive grease, an improvement of upstream in-plant grease capture and the identification of grease discharges would be necessary.

4.4. *Validation*

The KBS has been performing real-time support to process operation in the Granollers WWTP since September 1999. During the period of validation (four months), plant operators detected 123 problem situations, of which 9 were deflocculation problems (7%). The KBS correctly diagnosed as deflocculation 8 of the 9 cases (89% accuracy), and misdiagnosed a deflocculation problem in 2 out of 112 cases (98% accuracy). From its implementation (September 1999) until January 2001, 17 cases of deflocculation were detected by the KBS; 70% of these situations were identified on the same day that the problem manifested noticeably itself, and 30% were identified at a transition state.

The following paragraphs present two real case studies where a deflocculation problem was identified and different causes were inferred. They illustrate the flow of information and the evolution of knowledge along different branches of the same tree.

In the first case (October 16th, 1999), underloading, bulking, foaming, tendency to foaming, high influent TKN loading and deflocculation were detected in the plant. Deflocculation diagnosis was effected after checking that flagellates and *V. microstoma* abundance was low, the supernatant of the V30-settling test was turbid and total COD removal efficiency was low even though DO, biomass concentration, and inflow rate were normal. Microscopic observations of activated

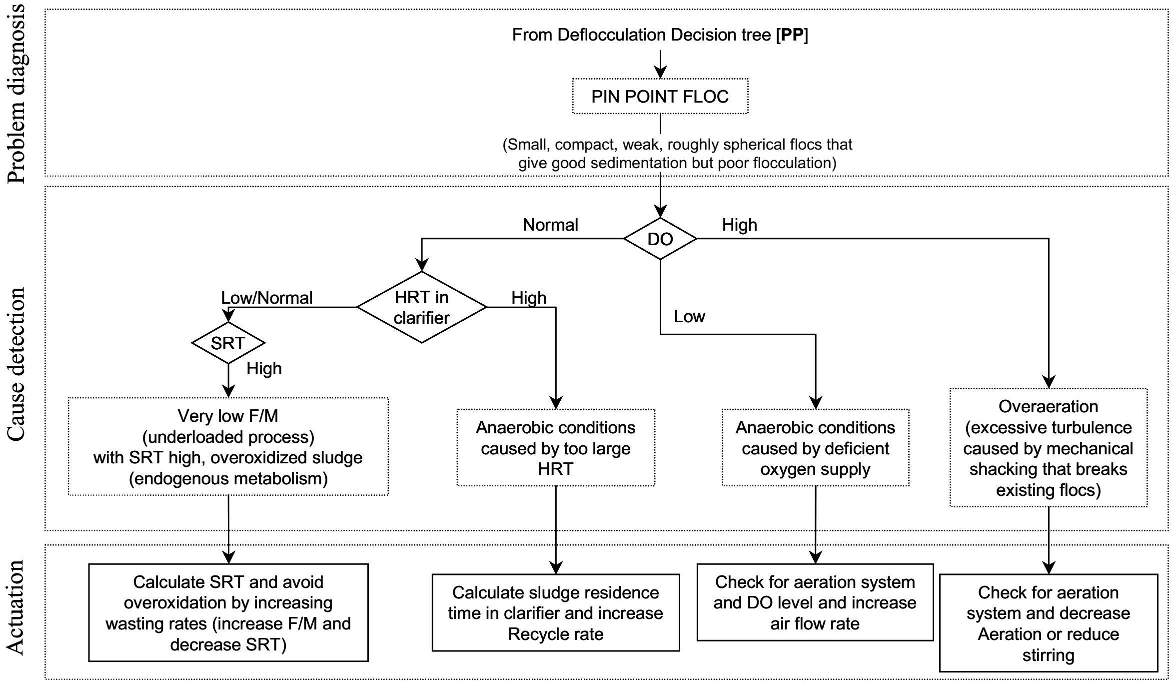

Fig. 4. Decision tree for solving pin-point problems.

sludge confirmed this situation. The deflocculation of the activated sludge floc was probably caused by the inhibitory or toxic action of some substance entered with the influent, but this could not be confirmed. Moreover, the high presence of filamentous bacteria around the floc (both *Microthrix parvicella* and type 021N) also affected adversely floc formation, causing large quantities of inter-floc bridges and floc breaking. The KBS recommended increasing the F:M ratio to favour the growth of floc-former bacteria over filamentous bacteria, and to eliminate the possible remainders of toxic substances while maintaining a proper DO level. After two days of large purges, the filamentous population started to decrease, while flagellates and *V. microstoma* recovered common rates and activated sludge improved settleability.

In the second case (December 26th, 2000), the KBS detected underloading, filamentous bulking, high influent nitrogen loads, foaming and deflocculation. In this case, a different cause for the deflocculation problem was pointed: an excessively underloaded process due to old sludge. Old sludge caused dispersed growth of the activated sludge floc, preventing good bioflocculation (SVI was not high, sludge floc was formed but was small and dispersed, the plant effluent and V30-settling test presented turbidity, and bubbles were not present in the settling test). Operational controls were fixed to favour a good formation of activated sludge floc: high wasting rates (during the 24 h) and minimum recycle rates (only one pump) were maintained for several days. On December 29th, the activated sludge floc already presented better flocculation and compactation, foams of *M. parvicella* had significantly decreased, and there was no floating sludge in the clarifiers.

Misdiagnoses of deflocculation problems by the KBS were generally due to errors in the diagnosis or in the cause-detection rules. Some problems were undetected because premises of certain diagnosis rules were too restricting, because of inadequate limits of antecedents, or simply because of the absence of a diagnosis path for a given problem. Nevertheless, misdiagnosed and non-detected cases proved useful to find new errors in the branches of the decision tree.

5. Discussion

A classical control system cannot, by itself, identify problem situations in activated sludge processes when detection requires on-line, off-line and heuristic information. A correct action would not be possible. In contrast, a KBS performs a systematic examination of all the available information by establishing and interpreting relationships among the different variables even much more easily than the operators in normal control, managing data from sensors, lab results, microscopic observations, trends of variables, etc. Thus, a KBS is able to detect and diagnose rapidly and efficiently complex problems, such as problems with a microbiological origin, of which deflocculation is a good example. Our experience building the KBS for the deflocculation problem illustrates how a KBS can reach a correct diagnosis or detect transition states with reasonable accuracy. In the case of the deflocculation problem, for instance, a classical control system would be able to detect turbidity at the effluent, if the plant was equipped with a turbidimeter or a solids sensor, and would be able to actuate automatically, for example, by modifying the speed of the clarifier bridge. However, a classical control system would not be able, by itself, to detect if this effluent turbidity is due to overloading of the clarifiers, to undesired denitrification in the clarifiers, or to bad formation of activated sludge flocs. Similarly, the proliferation of filamentous bacteria or deflocculation of existing flocs cannot be detected by a conventional control system based only on numerical algorithms.

Most importantly, the examination of the temporal trends in certain significant variables enables the KBS to detect some problems before severe episodes occur in the plant. Hence, the system helps the user to become acquainted with the tendencies of the process by detecting transition states and avoiding persistent causes that may lead to severe operational problems. This is a desirable feature of the application of KBS to WWTP operation, as it may help to prevent a severe modification of the activated sludge population when problems with microbiological origin are impending. If a severe situation of deflocculation or filamentous proliferation develops, recovering the proper activated sludge floc or re-directing the process to its normal state may take a long time.

The capability of the knowledge-based supervision to offer explanations about the deductive processes followed and justifications about the conclusions reached increases its reliability. The KBS may activate, deactivate, or just modify the classical control strategy (e.g., change the values of the set-points in order to adequate the control action to the specific situation). Furthermore, it can suggest additional actions based on knowledge, or, if necessary, carry out additional reasoning processes to look for other operational problems. Despite their potential, KBSs such as the one developed and tested in the Granollers WWTP, still present some shortcomings. Allowing final users to automatically update the KB (e.g., by adding, deleting or modifying rules) still proves difficult. In addition, the knowledge contained in the KBS is mostly general and static, rather that specific and adaptable. Accordingly, a KBS is not advisable as the only supervisory tool. Jeppsson et al. [7] suggest that supervisory, process-wide and integrated control supervised by overlaying rule-based support systems should be the main focus of instrumentation, control and automation in wastewater treatment in the future. Indeed, in agreement with Stephanopoulos and Han [31], we recommend the development of *intelligent decision supervisory systems*, a framework integrating numerical computations and KBSs to provide tools aimed at mission-critical problems, such as deflocculation. In such a framework, the classical management techniques (simulation, control, design, optimisation) would be allocated to numerical computations, while delegating the logical analyses and reasoning to KBSs ([25,32–34,16]).

6. Conclusions

This paper presents a knowledge-based approach to the management of the deflocculation problem. The main contribution is the methodology proposed and applied, which is based on the development of decision trees, their implementation in a RBS, and their full-scale validation. Although applied here to the deflocculation problem, this methodology is applicable to any other operational problem of WWTPs causing poor settling problems and/or poor efficiency removal. Expert opinion has proved fundamental at each step of the KBS development. In the field validation, the KBS showed good potential, with high accuracy in the detection of deflocculation situations ($\geqslant 89\%$), which could not have been easily identified with conventional control schemes.

Acknowledgements

This research has been partially supported by the Spanish CICyT projects TIC2000-1011 and DPI2000-0665-C02-01. We acknowledge the support of the European Union projects IST-1999-176101 and EVK1-CT-2000-00081. The authors also wish to acknowledge the co-operation of the *Consorci per a la Defensa de la Conca del Riu Besòs*, owner of the Granollers WWTP.

References

[1] WEF (Water Environment Federation). Operation of Municipal Wastewater Treatment Plants. Manual of Practice, ASCE Manual and Report on Engineering Practice, 5th ed., Vol. 2. Water Environment Federation, Alexandria, USA, 1996.

[2] Jenkins D, Richard MG, Daigger GT. Manual on the causes and control of activated sludge bulking and foaming. Chelsea: Lewis Publishers, 1993.

[3] Wanner J, Ruzicková I, Jetmarová P, Krhutková O, Paraniaková L. A national survey of activated sludge separation problems in the Czech Republic: filaments, floc characteristics and activated sludge metabolic properties. Water Sci Technol 1998;37(4/5):271–9.

[4] Kjellerup BV, Keiding K, Nielsen PH. Monitoring and troubleshooting of non-filamentous settling and dewatering problems in an industrial activated sludge treatment plant. Water Sci Technol 2001;44(2/3):155–62.

[5] Pipes WO. Bulking, deflocculation and pinpoint floc. J WPCF 1979;51:62–70.

[6] Ingildsen P, Jeppsson U, Olsson G. Dissolved oxygen controller based on on-line measurements of ammonia combining feed-forward and feedback. Water Sci Technol 2002;45(4/5):453–60.

[7] Jeppsson U, Alex J, Pons MN, Spanjers H, Vanrolleghem PA. Status and future trends of ICA in Wastewater treatment—A European perspective. Water Sci Technol 2002;45(4/5):485–94.

[8] Müller E, Kriebitzsch K, Snaider J, Wagner M, Wilderer PA, Wuertz S. Community Structure of micro- and macroflocs in Pin-point sludge and the influence of sludge age and potassium addition on microfloc formation, 3rd IWA conference on Microorganisms in Activated Sludge and Biofilm Processes, Rome, Italy, 13–15 June 2001.

[9] Chan WT, Koe LCC. A knowledge based framework for the diagnosis of sludge bulking in the activated sludge process. Water Sci. Technol 1991;23:847–55.

[10] Ladiges G, Kayser R. On-line and off-line expert systems for the operation of wastewater treatment plants. Water Sci Technol 1993;28(11-12):315–23.

[11] Ozgur NH, Stenstrom MK. Development of a knowledge-based expert system for process control of nitrification in the activated sludge process. J Environ Eng ASCE 1994; 120:87–107.

[12] Bergh SG, Olsson G. Knowledge based diagnosis of solids-liquid separation problems. Water Sci Technol 1996;33(2): 219–26.

[13] Wen C, Vassiliadis CA. Applying hybrid artificial intelligence techniques in wastewater treatment. Eng Appl Artif Intell 1998;11:685–705.

[14] Puñal A, Rodríguez J, Franco A, Carrasco EF, Roca E, Lema JM. Advanced monitoring and control of anaerobic wastewater treatment plants: diagnosis and supervision by a fuzzy-based expert system. Water Sci Technol 2001; 43(7):191–8.

[15] Baeza JA, Gabriel D, Lafuente J. Improving the nitrogen removal efficiency of an A2/O based WWTP by using an on-line knowledge based expert system. Water Res 2002; 36(8):2109–23.

[16] Rodríguez-Roda I, Sànchez-Marrè M, Comas J, Baeza J, Colprim J, Lafuente J, Cortés U, Poch M. A hybrid supervisory system to support wastewater treatment plant operation. Water Sci Technol 2002;45(4/5):289–97.

[17] Quinlan J. C4.5: programs for machine learning. San Mateo, CA: Morgan Kaufmann, 1993.

[18] Sànchez-Marrè M, Cortés U, Béjar J, de Gracia J, Lafuente J, Poch M. Concept formation in WWTP by means of classification techniques: a compared study. Appl Intell 1997;7(2):147–66.

[19] Comas J, Dzeroski S, Gibert K, R-Roda I, Sánchez-Marrè M. Knowledge discovery by means of inductive methods in wastewater treatment plant data. AI Commun 2001;14(1): 45–62.

[20] R-Roda I, Comas J, Poch M, Sànchez-Marrè M, Cortés U. Automatic knowledge acquisition from complex processes for the development of knowledge based systems. Ind Eng Chem Res 2001;40(15):3353–60.

[21] Gensym. G2Classic5.1Rev.7, Gensym Corporation, Cambridge, USA, 2000.

[22] Comas J. Development, Implementation and Validation of an Activated Sludge Supervisory System for the Granollers WWTP, Ph.D. thesis, University of Girona, 2000 (in English, accessible via http://www. tdcat. cbuc.es).

[23] Juristo N. Structuring the knowledge-based systems evaluation process, Intelligent Information Systems. IIS'97, IEEE Computer Society Press, 8–10 December 1997; 266–70.

[24] Borenstein D. Towards a practical method to validate decision support systems. Decision Support Systems 1998;23(3):227–39.

[25] R-Roda I, Sànchez-Marrè M, Comas J, Cortés U, Lafuente J, Poch M. An intelligent integration to improve the control and supervision of large WWTP. Proceedings of the 8th IAWQ Conference on Design, Operation and Economics of Large Wastewater Treatment Plants, Budapest, Hungary, 1999. p. 566–70.

[26] Urbain V, Block JC, Manem J. Bioflocculation in activated sludge: an analytic approach. Water Res 1993;27: 829–38.

[27] Liao BQ, Allen DG, Droppo IG, Leppard GG, Liss SN. Surface properties of sludge and their role in bioflocculation and settleability. Water Res 2001;35(2):339–50.

[28] Madoni P. A sludge biotic index (SBI) for the evaluation of the biological performance of activated sludge plants based on the Microfauna analysis. Water Res 1994;28(1): 67–75.

[29] Wilén BM, Balmér P. The effect of dissolved oxygen concentration on the structure, size and size distribution of activated sludge flocs. Water Res 1999;33(2): 391–400.

[30] Wilén BM, Keiding K, Nielsen PH. Anaerobic deflocculation and aerobic reflocculation of activated sludge. Water Res 2000;34(16):3933–42.

[31] Stephanopoulos G, Han C. Intelligent Systems in process Engineering: a review. Comp Chem Eng 1996;20(6/7): 743–91.

[32] Sànchez-Marrè M, Cortés U, R-Roda I, Lafuente J, Poch M. DAI-DEPUR: a distributed architecture for wastewater treatment plants supervision. Artif Intell Eng 1996; 10(3):275–85.

[33] Cortés U, Sànchez-Marrè M, Cecaronni L, R-Roda I, Poch M. Environmental decision support systems. Appl Intell 2000;13(1):77–91.

[34] Poch M, Comas J, Rodríguez-Roda I, Sànchez-Marrè M, Cortés U. Ten years of experience in designing and building real environmental decision support systems. In: Rizzoli AA, Jakeman AJ, editors. What have we learnt? Integrated Assessment and Decision Support. 2002; Vol. 2. p. 22–33. ISBN: 88-900787-0-7.

[35] Olsson G, Aspegren H, Nielsen MK. Operation and control of wastewater treatment — a Scandinavian perspective over 20 years. Water Sci Technol 1998;37(12):1–13.