Universitat de Girona

A DYNAMIC KNOWLEDGE-BASED DECISION SUPPORT SYSTEM TO HANDLE SOLIDS SEPARATION PROBLEMS IN ACTIVATED SLUDGE SYSTEMS: DEVELOPMENT AND VALIDATION

Montserrat MARTÍNEZ PUENTES

ISBN: 84-689-7705-5
Dipòsit legal: GI-381-2006
http://hdl.handle.net/10803/7792

ADVERTIMENT. L'accés als continguts d'aquesta tesi doctoral i la seva utilització ha de respectar els drets de la persona autora. Pot ser utilitzada per a consulta o estudi personal, així com en activitats o materials d'investigació i docència en els termes establerts a l'art. 32 del Text Refós de la Llei de Propietat Intel·lectual (RDL 1/1996). Per altres utilitzacions es requereix l'autorització prèvia i expressa de la persona autora. En qualsevol cas, en la utilització dels seus continguts caldrà indicar de forma clara el nom i cognoms de la persona autora i el títol de la tesi doctoral. No s'autoritza la seva reproducció o altres formes d'explotació efectuades amb finalitats de lucre ni la seva comunicació pública des d'un lloc aliè al servei TDX. Tampoc s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX (framing). Aquesta reserva de drets afecta tant als continguts de la tesi com als seus resums i índexs.

ADVERTENCIA. El acceso a los contenidos de esta tesis doctoral y su utilización debe respetar los derechos de la persona autora. Puede ser utilizada para consulta o estudio personal, así como en actividades o materiales de investigación y docencia en los términos establecidos en el art. 32 del Texto Refundido de la Ley de Propiedad Intelectual (RDL 1/1996). Para otros usos se requiere la autorización previa y expresa de la persona autora. En cualquier caso, en la utilización de sus contenidos se deberá indicar de forma clara el nombre y apellidos de la persona autora y el título de la tesis doctoral. No se autoriza su reproducción u otras formas de explotación efectuadas con fines lucrativos ni su comunicación pública desde un sitio ajeno al servicio TDR. Tampoco se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR (framing). Esta reserva de derechos afecta tanto al contenido de la tesis como a sus resúmenes e índices.

WARNING. Access to the contents of this doctoral thesis and its use must respect the rights of the author. It can be used for reference or private study, as well as research and learning activities or materials in the terms established by the 32nd article of the Spanish Consolidated Copyright Act (RDL 1/1996). Express and previous authorization of the author is required for any other uses. In any case, when using its content, full name of the author and title of the thesis must be clearly indicated. Reproduction or other forms of for profit use or public communication from outside TDX service is not allowed. Presentation of its content in a window or frame external to TDX (framing) is not authorized either. These rights affect both the content of the thesis and its abstracts and indexes.

Universitat de Girona

Departament d’Enginyeria Química, Agrària i Tecnologia Agroalimentària

Laboratori d’Enginyeria Química i Ambiental

A Dynamic Knowledge-Based Decision Support System to Handle Solids Separation Problems in Activated Sludge Systems: Development and Validation

Montse Martínez Puentes

PhD Thesis - Girona, December 2005

JOAQUIM COMAS MATAS i IGNASI RODRÍGUEZ-RODA LAYRET,
Professors del Departament d'Enginyeria Química, Agrària i Tecnologia Agroalimentària de la Universitat de Girona

CERTIFIQUEN

Que la llicenciada Montse Martínez Puentes ha dut a terme, sota la seva direcció, el treball que, amb el títol *A Dynamic Knowledge-Based Decision Support System to Handle Solids Separation Problems in Activated Sludge Systems: Development and Validation*, presenta en aquesta memòria, la qual constitueix la seva Tesi per a optar al Grau de Doctora per la Universitat de Girona.

I perquè en prengueu coneixement i tingui els efectes que correspongui, presentem davant la Facultat de Ciències de la Universitat de Girona l'esmentada Tesi, signant aquesta certificació a

Girona, 19 de desembre del 2005

Joaquim Comas Matas　　　　Ignasi Rodríguez-Roda Layret

Agraïments

Arribats a aquest punt, i després d'escriure més de 300 pàgines en anglès sembla que això dels agraïments ha de representar la tasca més senzilla realitzada al llarg d'aquests 5 anys de doctorat. Doncs bé, per a mi no és gens senzill això d'explicar en unes línies tota la gratitud que sento vers aquelles persones que han fet possible que la tesi de la Montse sobre el bulking vegi la llum. Espero doncs que tots aquells que llegeixin aquests agraïments s'hi vegin reflectits, si no és així, gràcies a TU també per dedicar uns moments a llegir part de la tesi...

Voldria doncs en primer lloc agrair als meus directors de tesi el seu suport incondicional. Ells han estat el meu Decision Support System diari i han estat presents sempre que els he necessitat. Gràcies Ignasi per la teva sinceritat i bones idees que han anat il·luminant aquest llarg camí (sobretot en la foscor del final). Gràcies igualment a tu Quim per les teves paraules plenes de suport, per trobar sempre un moment per mí i per la teva visió pràctica de les coses. També volia agrair al Manel Poch l'oportunitat que em va oferir ara fa ja 5 anys d'iniciar-me en l'apassionant món de la depuració d'aigües.

Al Ministeri de Ciència i Tecnologia, que mitjançant la concessió d'una beca FPI (CICYT DPI2000-0665-C02-01) han finançat aquests 5 anys de recerca.

Als meus pares per facilitar-me l'accés als estudis i per recolzar totes les meves decisions. A la meva germana, al Carles i a l'alegria de la família el Gabriel, que amb tan sols una paraula o un somriure pot aixecar l'ànim a qualsevol. Molt especialment també vull dedicar part d'aquesta tesi al Josep, per estar sempre amb mi i aguantar-me en els bons i mals moments. Sense tu això tampoc hauria estat possible. També volia agrair als pares i germanes del Josep el seu recolzament en aquesta recta final. Us estimo molt a tots.

Al motor d'inferència d'aquest projecte: el Miquel Sànchez i l'Ulises Cortés de la UPC. Gràcies per facilitar-me l'accés a l'apassionant món de la IA i pels vostres savis consells. Moltes gràcies a tu també Carles, per les teves context-sensitive idees.

A la base de coneixement composada per grans professionals: Àngel Freixò, David Güell, Sandra Rovira, Francesc Granés, Ricard Tomàs, Gabriela Poch, Elvira Cèsar i un llarg etcètera. Gràcies per fer-me partícip de la vostra gran experiència i per transmetre'm tants coneixements.

A tota la gent de la depuradora de Girona, moltes gràcies per donar suport a aquest projecte i per col·laborar-hi tant activament: Gabriela, Dolors, Estel, Jaume, Lluís... sense vosaltres el sistema s'hauria quedat orfe.

També volia agrair molt especialment al Joan Lluís Riera pel suport lingüístic ofert durant la redacció de la tesi.

I would also express my sincere gratitude to professor Krishna Pagilla for his kind support during my stay in the Illinois Institute of Technology in Chicago. I am also grateful to Slav Hermanowicz for facilitating my stay in the University of California, Berkeley, where I had the great pleasure to meet Professor David Jenkins.
Moltes gràcies també a la penya berkeleyiana que van fer de la meva estada a Berkeley una experiència inoblidable: Imma, Magda, Xavi, Aitor, Sabrina, Jordi, Roberto... moltes gràcies per deixar-me formar part de la vostra petita comunitat a l'altra banda de l'Atlàntic, ah, i per ensenyar-me a esquiar!!!

A les meves amigues de tota la vida: Òria, Laura, Mercè, Noelia... Tot i que no sabeu molt bé que coi faig a la uni gràcies per pensar sempre que faig coses interessants.

A les grans amigues de la universitat: Lídia, Raquel, Anna, Núria, Alba, Eva, Rosi, Berta... moltes gràcies a vosaltres també per saber mantindre aquesta amistat i per compartir tants bons moments.

Als meus companys infatigables de doctorat a la Mireia i al Francesc. Us estimo molt i us estic molt agraïda per tots els moments compartits durant aquests 10 anys. Gràcies per proporcionar-me una segona residència a Girona, per les partides al Catan en els moments més crítics, per les taronges de València, per la perspectiva de vacances davant el mar... gràcies per tot de veritat. Us trobaré molt a faltar.

De la mateixa manera volia agrair a la resta de companys de la guarderia el brindar-me l'oportunitat de compartir tants bons moments amb vosaltres i fer que els 150 km diaris valguessin la pena. Lluís, Helio, Anna, Sebas, Ramon, Marta, Carla, Xavi, Clàudia, Peter, Gemma, Hèctor, Jordis, Oriol, Christian, Teia, Elvira, Núria... També a la resta de staff del Lequia: Jesús, Miquel, Maria, Marilós, Jaume... gràcies a tots per tot.

També voldria agrair als meus companys de trajecte: Montse, Marc, Manel, Tamara per fer que les hores mortes al tren fossin més suportables.

Finalment, voldria acabar la darrera part d'aquests agraïments amb una frase regalada per 2 grans amics: "*We do not remember days... we remember moments*" que guanya molt de sentit al finalitzar aquesta darrera reflexió. Els records d'aquests 5 darrers anys estan plens de bons moments que no oblidaré mai i m'acompanyaran sigui quin sigui el següent pas del meu camí. Fent un balanç doncs abans de passar aquesta pàgina de la meva vida tant important tan sols puc dir:

....GRÀCIES A TOTS...

Resum

El sistema de fangs activats és el tipus de tractament biològic més àmpliament utilitzat arreu del món per la depuració de les aigües residuals. El seu funcionament depèn de la correcta operació dels seus dos

components principals, el reactor biològic, on la majoria de contaminants presents a l'aigua residual (especialment matèria orgànica i nutrients) són degradats mitjançant l'acció d'una comunitat de microorganismes molt diversa, i el decantador secundari, on finalment aquests microorganismes (altrament anomenats biomassa) són físicament separats de l'aigua depurada mitjançant la seva sedimentació. Quan aquesta darrera fase de sedimentació no es realitza correctament, la biomassa no decantada s'escapa amb l'efluent (en forma de sòlids en suspensió i matèria orgànica) causant un impacte sobre el medi receptor. Aquests tipus de problemes de sedimentació, anomenats **problemes de separació de sòlids**, són actualment una de les principals causes d'ineficiència en l'operació dels sistemes de fangs activats arreu del món. Els següents problemes es poden incloure en aquest grup: bulking filamentós, bulking viscós, escumes biològiques, creixement dispers, flòcul pin-point i desnitrificació incontrolada.

L'origen dels problemes de separació de sòlids generalment es troba en un desequilibri entre les principals comunitats de microorganismes implicades en la sedimentació de la biomassa: els bacteris formadors de flòcul i els bacteris filamentosos. Degut a aquest **origen microbiològic**, la seva identificació i control no és una tasca fàcil pels caps de planta. Algunes de les característiques que fan que els problemes de separació siguin complexos de solucionar són: (1) hi ha una gran quantitat de variables, tant quantitatives com qualitatives que s'han de mesurar i monitoritzar per tal d'identificar la causa del desequilibri i poder actuar conseqüentment; (2) actualment no existeix cap model matemàtic fiable que ajudi a interpretar i/o predir de forma genèrica el desenvolupament dels problemes de separació; (3) el control clàssic no és efectiu en prevenir i solucionar aquest tipus de situacions, ja que el control integrat de totes les variables implicades, no és encara possible. A més, en la majoria dels casos, a part de l'eina de control es requereix un raonament expert addicional; (4) l'origen microbiològic fa que la seva dinàmica sigui molt lenta, tant durant la fase d'aparició com durant la seva desaparició i (5) degut a la seva complexitat i dinamisme, les estratègies de control proposades han d'estar ben estructurades en un complet pla d'actuacions, a seguir mentre el problema persisteixi en el procés. Actualment, les principals eines de control que els caps de planta segueixen per afrontar aquest tipus de problemes, per una banda es basen en els diferents manuals i articles de recerca destinats a la solució dels problemes de separació i, per altra banda, i encara més important, en coneixement heurístic i en la pròpia experiència en el control del procés.

Els **Sistemes de Suport a la Presa de Decisions basats en el coneixement** (KBDSS) són un grup d'eines informàtiques caracteritzades per la seva capacitat de representar coneixement heurístic i tractar grans quantitats de dades (qualitatives, quantitatives, incertes i inexactes). Aquestes qualitats han fet que la seva aplicació s'hagi estès a diferents dominis ambientals complexes, com ara la depuració d'aigües residuals. El relatiu èxit d'aquestes aplicacions, ens ha fet pensar en la seva potencial aplicació com a eina de suport en el control dels problemes de separació de sòlids. D'aquesta manera, l'**objectiu** de la present tesi és el desenvolupament i validació d'un KBDSS específicament dissenyat per donar suport als caps de planta en el control dels problemes de separació de sòlids d'orígen microbiològic en els sistemes de fangs activats. Per aconseguir aquest objectiu principal, el KBDSS ha de presentar les següents **característiques**: (1) la implementació del sistema ha de ser ***viable*** i ***realista*** per garantir el seu correcte funcionament; (2) el raonament del sistema ha de ser ***dinàmic*** i ***evolutiu*** per adaptar-se a les necessitats del domini al qual es vol aplicar i (3) el raonament del sistema ha de ser ***intel·ligent***, basat en els principis del raonament humà que inclou coneixement expert i basat en l'experiència.

En primer lloc, a fi de garantir la viabilitat a l'hora d'implementar el KBDSS proposat, a part de realitzar una revisió teòrica de les actuals tècniques existents pel control dels problemes de separació -***secció 1.5***-, s'ha realitzat un estudi exhaustiu, a petita escala (Catalunya), de la realitat dels problemes de separació microbiològics, tant pel que fa a la magnitud de la seva problemàtica com a les possibilitats reals de seguiment i control que actualment disposen les EDARs. Aquest estudi –*Capítol 3*– ha determinat tant les variables més utilitzades per a la diagnosi i monitorització dels problemes i els mètodes de control més viables, com la detecció de les principals limitacions que el sistema hauria de resoldre. De les conclusions derivades de l'estudi, s'ha pogut establir l'enfoc idoni per desenvolupar un KBDSS **viable, realista i eficient**.

Per altra banda, els resultats d'anteriors aplicacions han demostrat que la principal limitació en el desenvolupament de KBDSSs és l'estructura de la **base de coneixement** (KB), on es representa tot el coneixement adquirit sobre el domini, juntament amb els processos de raonament a seguir. En el nostre cas, tenint en compte la **dinàmica** del domini, aquestes limitacions es podrien veure incrementades si el disseny i desenvolupament del sistema no fos òptim. S'ha considerat doncs, indispensable, una fase prèvia de disseny conceptual, destinada al plantejament inicial del KBDSS (incloent l'estructura i interrelació entre les bases de coneixement i el procés de raonament a utilitzar). En aquest sentit, s'ha proposat el **Domino Model** com a eina per dissenyar conceptualment el sistema. Durant el seu estudi experimental –*capítol 4*–, l'ús del Domino Model ha aportat: (1) una visualització intuïtiva de l'organització i interrelació de les diferents bases de coneixement i processos de raonament. Degut a la seva disposició cíclica, aquestes interrelacions no són estàtiques sinó dinàmiques; (2) la possibilitat d'estructurar els plans d'acció en una sèrie de tasques ordenades i (3) la possibilitat d'utilitzar qualsevol tècnica de representació del coneixement per acomplir els objectius establerts. En el nostre cas els objectius a acomplir són: (i) diagnosi o predicció dels problemes de separació d'origen microbiològic dels sistemes de fangs activats; (ii) identificació de les possibles causes i (iii) recomanació i seguiment d'un pla d'actuacions complet per solucionar els problemes identificats que tingui en compte l'evolució del procés. D'aquesta manera, el desenvolupament del KBDSS dinàmic proposat s'ha estructurat en 3 cicles diferents.

Finalment, el darrer objectiu principal o característica que el KBDSS ha de complir, és el seguiment d'un **raonament intel·ligent** que li permeti complir amb les tasques pel qual ha estat desenvolupat. En general, es podria dir que el raonament seguit pel sistema ha d'emular el raonament humà, el qual generalment utilitza tant coneixement expert (adquirit de la bibliografia existent), com derivat de la seva pròpia experiència a l'hora de resoldre problemes de separació de sòlids. Per tal d'introduir tots dos tipus de raonament, l'ús d'un Sistema Expert (basat en coneixement expert) i l'ús d'un Sistema de Raonament Basat en Casos (basat en l'experiència) han estat escollits i integrats com els principals sistemes intel·ligents encarregats de dur a terme el raonament del KBDSS. Als *capítols 5 i 6* respectivament, es presenten el desenvolupament del Sistema Expert dinàmic (ES) i del Sistema de Raonament Basat en Casos temporal, anomenat Sistema de Raonament Basat en Episodis (EBRS). Tots dos han estat desenvolupats a partir del model conceptual proposat al capítol 4. Al final de cada capítol es mostra, a mode de discussió, una sèrie de conclusions preliminars derivades del nou enfoc dinàmic, temporal i intel·ligent aplicat a ambdues eines.

A continuació, al *capítol 7*, es presenten detalls de la implementació del sistema global (KBDSS) en l'entorn G2, destacant la seva estructura modular i els avantatges derivats de la integració d'ambdues eines. Seguidament, al *capítol 8*, es mostren els resultats obtinguts durant els 11 mesos de validació del sistema, on aspectes com la precisió, capacitat i utilitat del sistema han estat validats tant experimentalment (prèviament a la implementació) com a partir de la seva implementació real a l'EDAR de Girona. Finalment, al *capítol 9* s'enumeren les principals conclusions derivades de la present tesi.

Resumen

El sistema de fangos activados es el tipo de tratamiento biológico más ampliamente utilizado en todo el mundo para la depuración de aguas residuales. Su funcionamiento depende de la correcta operación de sus dos componentes principales, el reactor biológico donde la mayoría de contaminantes presentes en el agua residual (especialmente materia orgánica y nutrientes) son degradados mediante la acción de una comunidad de microorganismos muy diversa, y el decantador secundario, donde finalmente estos microorganismos son físicamente separados del agua depurada mediante su propia sedimentación. Cuando esta segunda fase de sedimentación no se realiza correctamente, la biomasa no decantada se escapa por el efluente, causando un impacto sobre el medio receptor y afectando la eficiencia del proceso. Este tipo de problemas de sedimentación, llamados **problemas de separación de sólidos**, es actualmente una de las principales causas de ineficiencia en la operación de los sistemas de fangos activados de todo el mundo. Los siguientes problemas se pueden incluir en este grupo: bulking filamentoso, bulking viscoso, espumas biológicas, crecimiento disperso, flóculo pin-point y desnitrificación incontrolada.

El origen de los problemas de separación de sólidos generalmente se encuentra en un desequilibrio entre las principales comunidades de microorganismos implicadas en la sedimentación de la biomasa: las bacterias formadoras de flóculo y las bacterias filamentosas. Debido a este **origen microbiológico**, su identificación y control no resulta ser una tarea fácil para los jefes de planta. Algunas de las características que otorgan complejidad durante el control de los problemas de separación son: (1) la gran cantidad de variables, cuantitativas y cualitativas, que han de ser analizadas y monitorizadas con el fin de identificar la causa del desequilibrio y poder actuar consecuentemente; (2) actualmente no existe ningún modelo matemático fiable que pueda ayudar a interpretar y/o predecir de forma genérica el desarrollo de estos problemas; (3) el control clásico no es efectivo durante la predicción y solución de éste tipo de situaciones, ya que el control integral de todas las variables implicadas, actualmente no es posible; (4) el origen microbiológico hace que su dinámica sea muy lenta, tanto durante la fase de aparición como durante su desaparición y (5) las estrategias de control propuestas han de estar bien estructuradas en un completo plan de actuaciones a seguir mientras el problema persista. Actualmente, las principales herramientas de control disponibles son, por un lado, diferentes manuales y artículos de investigación destinados a la resolución de los problemas de separación (conocimiento heurístico) y por otro lado y si cabe aún más importante, el cocimiento heurístico y la valiosa experiencia en el control del propio proceso.

Los **Sistemas de Soporte a la Toma de Decisiones** (KBDSS) son un grupo de herramientas informáticas caracterizadas por su capacidad de representar conocimiento heurístico y manejar grandes cantidades de datos (cualitativos, cuantitativos, inciertos e inexactos). Estas cualidades han hecho que su aplicación se haya extendido a diferentes dominios ambientales complejos como por ejemplo la depuración de aguas residuales. El relativo éxito de estas aplicaciones nos ha llevado a pensar en su potencial aplicación como herramienta de soporte en el control de los problemas de separación de sólidos. De esta manera, el **objetivo** de la presente tesis es el desarrollo y validación de un KBDSS específicamente diseñado para dar soporte a los jefes de planta durante el control de los problemas de separación de sólidos de orígen microbiológico en sistemas de fangos activados. Para conseguir este objetivo principal, el sistema ha de presentar las siguientes **características**: (1) la implementación del sistema ha de ser ***viable*** y ***realista*** para garantizar su correcto funcionamiento. Su uso ha de suponer al usuario más ventajas que inconvenientes; (2) el razonamiento seguido por el sistema ha de ser ***dinámico*** y ***evolutivo*** para adaptarse a las necesidades del dominio y (3) el razonamiento ha de ser ***inteligente***, basado en los principios del razonamiento humano, que incluyen conocimiento experto y basado en la experiencia.

En primer lugar, a fin de garantizar la viabilidad durante la implementación del KBDSS, a parte de una revisión teórica de las técnicas existentes para el control de los problemas de separación -***sección 1.5***-, se ha realizado un estudio exhaustivo, a pequeña escala (Cataluña), de la realidad de los problemas de separación incluyendo tanto la magnitud de su problemática como las posibilidades reales de seguimiento y control disponibles en las EDARs. Dicho estudio –*Capítulo 3*– ha determinado tanto las variables más utilizadas para la diagnosis y monitorización de los problemas y los métodos de control más viables, como las principales limitaciones que el sistema tendría que resolver. De las conclusiones derivadas del estudio, se ha podido establecer el enfoque idóneo para desarrollar un KBDSS **viable, realista y eficiente.**

Por otro lado, los resultados de anteriores aplicaciones han demostrado que la principal limitación en el desarrollo de KBDSS es la estructura de la **base de conocimiento** (KB), donde se representa todo el conocimiento adquirido sobre el dominio, juntamente con los procesos de razonamiento a seguir. En nuestro caso, teniendo en cuenta la **dinámica** del dominio, estas limitaciones se podrían ver incrementadas si el diseño y desarrollo del sistema no fuera óptimo. Se ha considerado entonces indispensable una fase previa de diseño conceptual, destinada a la planificación inicial del KBDSS (incluyendo la estructura e interrelación entre las bases de conocimiento y el proceso de razonamiento). El **Domino Model** ha sido estudiado como posible herramienta para diseñar conceptualmente nuestro sistema ***-capítulo 4-***. El uso del modelo ha aportado: (1) una visualización intuitiva de la organización e interrelación de las diferentes bases de conocimiento y procesos de razonamiento. Debido a su disposición cíclica, estas interrelaciones no son estáticas sino dinámicas; (2) la posibilidad de estructurar los planes de acción en una serie de tareas ordenadas y (3) la posibilidad de utilizar cualquier técnica de representación del conocimiento para cumplir os objetivos establecidos. En nuestro caso, los objetivos son: (i) diagnosis o predicción de los problemas de separación de origen microbiológico en los sistemas de fangos activados; (ii) identificación de posibles causas y (iii) recomendación y seguimiento de un plan de acción completo para solucionar los problemas identificados que tenga en cuenta la evolución del proceso. El desarrollo del KBDSS dinámico propuesto se ha estructurado en 3 ciclos diferentes.

Finalmente, el último objetivo principal o característica que el KBDSS ha de cumplir es el seguimiento de un **razonamiento inteligente** que le permita cumplir las tareas por las cuales ha sido desarrollado. En general, se podría decir que el razonamiento seguido por el sistema ha de imitar el razonamiento humano, el cual, generalmente utiliza tanto conocimiento experto (adquirido de bibliografía), como derivado de su propia experiencia en la resolución de problemas. Con el fin de utilizar los dos tipos de razonamiento, el uso de un Sistema Experto (basado en conocimiento heurístico) y el uso de un Sistema de Razonamiento Basado en Casos (basado en la experiencia) han sido escogidos e integrados como los principales sistemas inteligentes encargados de cumplir los objetivos del KBDSS. En los ***capítulos 5 y 6***, se presenta el desarrollo del Sistema Experto dinámico (ES) y del Sistema de Razonamiento Basado en Episodios (EBRS), respectivamente. Los dos han estado desarrollados a partir del modelo conceptual propuesto en el capítulo 4. Al final de cada capítulo se muestra, a modo de discusión, las conclusiones preliminares derivadas del nuevo enfoque dinámico, temporal e inteligente aplicado a ambas herramientas.

A continuación, en el ***capítulo7***, se comentan algunos detalles sobre la implementación del KBDSS, en el entorno G2, destacando su estructura modular y las ventajas derivadas de la integración de las 2 herramientas. Seguidamente, en el ***capítulo 8*** se muestran los resultados obtenidos durante los 11 meses de validación del sistema, donde aspectos como la precisión, capacidad y utilidad del sistema han sido validados tanto experimentalmente (previamente a la implementación) como a partir de su implementación real en la EDAR de Girona. Finalmente en el ***capítulo 9*** se enumeran las principales conclusiones derivadas de la presente tesis.

Abstract

The activated sludge system is the most widely used technology for biological wastewater treatment in the world. Its successful performance relies on the correct operation of two main units, the bioreactor where the organic matter and nutrients are partially removed by the microorganisms and the secondary settler where the biomass is separated from the treated water. When settleability deteriorates, the inefficient separation of biomass can affect the quality of the activated sludge effluent, implying on most occasions both a deterioration of the water quality (in terms of suspended solids and organic matter), with the corresponding impact on the receiving ecosystem and affecting the process efficiency. These kind of undesired situations, known as activated sludge solids separation problems, are one of the main causes of inefficiency in activated sludge systems. They include: filamentous bulking, non-filamentous bulking, biological foaming, dispersed growth, pin-point floc and rising sludge.

The usual origin of solids separation problems is (except from rising sludge) an imbalance between the different microbiological communities responsible for the biomass settleability: the floc-forming bacteria and the filamentous bacteria. Due to this **microbiological origin**, their identification and control is a tough task for plant operators. Some of the characteristics that makes their control so complex are: (1) there is a wide number of variables, including quantitative and qualitative variables, that must be measured and monitored to identify the cause of such imbalance and to act consequently; (2) no reliable mathematical model capable of simulating or predicting the development of solids separation problems has been developed yet; (3) classical control is not effective enough in preventing and solving these type of situations, given that the integrated control of all the relevant variables and parameters is still not possible. Moreover, on most occasions, an expert reasoning is additionally required; (4) the microbiological origin of solids separation problems makes its dynamics very slow; (5) due to its complexity and dynamism, the suggested control strategies must follow a well-structured and complete action plan to be applied meanwhile the problem persists in the process. Currently, the main control tools available for plant operators to solve this type of problems are, on the one hand, different handbooks and research papers dedicated to the solution of solids separation problems and on the other hand, and even more important, the heuristics and experience of plant operators in solving this type of situations.

Knowledge-Based Decision Support Systems (KBDSS) are a group of tools from the Artificial Intelligence domain characterized by their capability to represent heuristic knowledge and to work with large amounts of data (qualitative, quantitative, uncertain and inexact). These capabilities have increased their application in several complex environmental domains, such as wastewater treatment. The exit of such applications has made us consider them as a potential support tool to handle solids separation problems. Hence, the main **objective** of the present thesis is to develop and validate a KBDSS specially designed to support plant operators to handle solids separation problems of microbiological origin occurring in activated sludge systems. In order to achieve this objective, the developed KBDSS must accomplish with the following characteristics: (1) the implementation of the system must be ***viable*** and ***realistic*** in order to ensure its proper operation; (2) the reasoning process followed by the system must be ***dynamic*** and ***evolutive*** in order to match the necessities of the domain and (3) the reasoning must be also ***intelligent***, based on the principles of human reasoning, and including both expert and experiential knowledge.

First of all, in order to guarantee the feasibility of the system, apart from carrying out a deep study on the control methods currently available in literature to handle solids separation problems -***section 1.5***-, a thorough study, at a local scale (Catalonia) has been performed to investigate the real magnitude of the problems in Catalonia and to analyse the real possibilities of monitoring and control such undesired situations. This study -***chapter 3***- has contributed to the determination of the most common parameters generally used to diagnose, monitor and control these problems as well as to the detection of the existing limitations that the suggested KBDSS should overcome. Thus, the main conclusions derived from the present study have allowed us to set up the basic approach to develop a **viable**, **realistic** and **efficient** KBDSS.

On the other hand, the results obtained from past applications of KBDSS has demonstrated that the main bottleneck in developing KBDSS is the structure of the **knowledge base** (KB), where all the knowledge acquired from the domain is represented, together with the necessary reasoning processes. In our approach, the additional complexity and the corresponding necessities imposed by the **dynamic** nature of the domain exacerbate the limitations in developing a feasible system. In this case, a previous conceptualisation phase was considered in which a conceptual design of the system was set up. In this sense, the **Domino Model** was suggested as a tool to conceptually design the system. During its experimental study -***chapter 4***- the use of the Domino Model has provided: (1) an intuitive visualization and a direct characterization and organization of the different knowledge bases and reasoning processes. The cyclic structure has facilitated dynamic interrelations between the different knowledge bases; (2) the possibility of structuring the suggested control plans into a well-organized and scheduled set of actions and (3) the use of any suitable technique of knowledge representation that must permit the system to accomplish the pre-defined objectives including: (i) diagnosis and prediction of solids separation problems of microbiological origin occurring in activated sludge systems; (ii) identification of the possible causes and (3) proposal and monitoring of complete control strategies to restore the process. According to this, the development of the system has been structured in 3 different cycles.

Finally, in order to efficiently fulfill its main tasks, the last main objective or characteristic that the KBDSS must accomplish is the use of **intelligent reasoning**. In general, it has been established that the reasoning process followed by the system should emulate human reasoning, which usually use both expert knowledge (acquired from the existing literature) and knowledge acquired from his/her own experience in solving solids separation problems. In order to use both types of knowledge, an Expert system (based on expert knowledge) and a Case-Based Reasoning System (based on experiential knowledge) has been chosen and integrated as the main intelligent tools to carry out the goals of the KBDSS. In ***chapter 5*** the development of the dynamic ES is presented. In chapter 6, a new temporal approach for classical CBRS is depicted, the Episode-Based Reasoning System (EBRS). Both systems have been developed by using the conceptual model suggested in chapter 4. At the end of each chapter, a list of the preliminary conclusions derived from the new dynamic, temporal and intelligent approach of both techniques is presented.

Next, in ***chapter 7*** some details of the KBDSS implementation in the G2 environment are presented, including the advantages resulting from the suggested modular structure and from the integration of both tools. After that, in ***chapter 8*** the results obtained during the 11 months of validation are depicted, including the results regarding the accuracy, adequacy, usefulness and usability of the system, which have been validated both experimentally (before the implementation) and as a result of the system's implementation in the Girona WWTP. Finally, in ***chapter 9***, the main conclusions derived from the present thesis are enumerated.

The results of cooperative research between the Laboratory of Chemical and Environmental Engineering (LEQUIA) of the *Universitat de Girona* (UdG) in close association with the Knowledge Engineering and Machine Learning Group (KEMLg) at the Software Department of the *Universitat Politècnica de Catalunya* (UPC) represent the work carried out during the last 15 years up to the present thesis.

Though this research group began with a classical approach, *e.g.* statistical modelling or feedback control (Robusté, 1990), soon they realised that classical control techniques show some limitations and that more powerful strategies were necessary to deal with such complex environmental processes as WWTP. Special characteristics of these processes made this research group think in the possibility of applying knowledge based techniques, originally developed within the Artificial Intelligence field, as a new proper step for the WWTP management and control. LEQUIA and KEMLg began with standard expert systems (Serra, 1993), and soon evolved to the integration of Case-based Reasoning module to consider learning processes with new experiences (Sànchez-Marrè, 1996). In spite of knowledge-based systems suppose a significant improvement, the application of a single technique also presented some pitfalls since only enabled to overcome some of the limitations of classical control techniques. The cooperative group became conscious that complex problems should be better addressed from a multidisciplinary way, integrating different techniques from artificial intelligence and conventional methods to solve the WWTP operational problems precisely. Indeed, as a result of the experience obtained thanks to several years of cooperative research, they have developed an intelligent integrated supervisory system to manage optimally complex wastewater treatment systems (R-Roda, 1998; Comas, 2000; Cortés, 2003). The new approach presented in this Thesis represents a step further in the supervision of complex wastewater treatment systems, specifically in the supervision of the solids separation problems of microbiological origin occurring in activated sludge systems.

The PhD research work carried by Montse Martínez and presented in this document has enabled to produce the following publications:

- **Martínez, M.**, *Feasibility testing of a conceptual model to develop knowledge based decision support systems to manage filamentous bulking in a WWTP*. Master Thesis, University of Girona (2002).

- Cortés, U., **Martínez, M.**, Comas, J., Sànchez-Marrè, M., Poch, M and Rodríguez-Roda, I., A conceptual model to facilitate knowledge sharing in waste water treatment plants, *AI Com.*, **16(3)**, 193-207 (2003).

- **Martínez, M.**, Cortés, U., Bonmatí, A., Poch, M. and Rodríguez-Roda, I., Conceptualisation of a decision support system to solve separation problems, *Proc. of the 9th IWA Specialised Conference on Design, Operation and Economics of Large Wastewater Treatment Plants*, pp. 131-134, September 1-4, Praha, Czech Republic (2003).

- Nuñez, H., Sànchez-Marré, M., Cortés, U., Comas, J., **Martínez, M.**, Rodríguez-Roda, I. and Poch, M., A comparative study on the use of similarity measures in case-based reasoning to improve the classification of environmental systems situations, *Environ. Modell. Softw.*, **19(9)**, 809-819 (2004).

- Sànchez-Marrè, M., **Martínez, M.**, Rodríguez-Roda, I., Alemany, J. and Cortés, C., Using CBR to improve intelligent supervision and management of wastewater treatment plants: the atl_EDAR system, Proc. of the Industrial day in *7th European Conference on Case-based Reasoning*, 79-91., September, Madrid (2004).

- **Martínez, M.**, Rodríguez-Roda, I., Poch, M., Cortés, C. and Comas, J., Dynamic reasoning to solve complex problems in activated sludge processes: a step further in Decision Support Systems, *Water Sci. Technol*, (in press) (2005).

- **Martínez, M.**, Sànchez-Marrè, M., Comas, J. and Rodríguez-Roda, I. Case-Based Reasoning, a promising tool to face solids separation problems in the activated sludge process. *Water Sci. Technol*, (in press) (2005).

- **Martínez, M.**, Mérida-Campos, C., Sànchez-Marrè, M., Comas J. and Rodríguez-Roda, I. Improving the efficiency of Case-Based Reasoning to deal with activated sludge solids separation problems. *Environ. Technol.* (in press) (2005).

- Sànchez-Marrè, M., Cortés, U., **Martínez, M.**, Comas, J. and Rodríguez-Roda, I., An approach for temporal Case-Based Reasoning: Episode-Based Reasoning, *Procc. of the 6th International Conference, ICCBR 2005* pages 465-476, Chicago,IL,USA, August (2005).

On the other hand, two short research stays were carried out during the research period of Montse Martínez. One at the Department of Chemical and Environmental Engineering in the Illinois Institute of Technology (IIT, Chicago, USA) under the supervision of Professor Krishna Pagilla (March-May 2002) and, the other one, at the Department of Civil and Environmental Engineering in the University of California, Berkeley (California, USA) under the supervision of Professor Slaw Hermanowicz (February-April 2004).

Index

Acronyms, Abbreviations and Nomenclature

AI	Artificial Intelligence
ASM1	Activated Sludge Model number 1
ASM2	Activated Sludge Model number 2
ASM3	Activated Sludge Model number 3
BNR	Biological Nutrient Removal
BOD	Biological Oxygen Demand
C	Carbon
Ca	Calcium
CBR	Case-Based Reasoning
CBRS	Case-Based Reasoning System
Cl_2	Chlorine
COC	Chlorine Organic Compounds
COD	Chemical Oxygen Demand
CO_2	Carbon dioxide
DGGE	Denaturing Gradient Gel Electrophoresis
DNA	Deoxyribonucleic Acid
DO	Dissolved Oxygen
DSS	Decision Support System
DSVI	Diluted Sludge Volume Index
EBPR	Enhanced Biological Phosphorus Removal
EBRS	Episode-Based Reasoning System
EDAR	Estació Depuradora d'Aigües Residuals
EDSS	Environmental Decision Support System
ES	Expert System
EPS	Extracellular Polymeric Substances
FAME	Fatty Acid Methyl Ester
Fe	Iron
FISH	Fluorescent *In Situ* Hybridization
F/M	Food to Microorganisms ratio
GSI	G2 Standard Interface
H	Hydrogen
H_2O_2	Hydrogen peroxide
HRT	Hydraulic Retention Time
ICA	Instrumentation, Control and Automation
IWA	International Water Association
KB	Knowledge Base
KBDSS	Knowledge-Based Decision Support System
KBS	Knowledge-Based System
Ks	Saturation Constant
LOOCV	Leave-One-Out Cross-Validation
MAR	Microautoradiography
MCRT	Mean Cell Residence Time
Mg	Magnesium
MLSS	Mixed Liquor Suspended Solids
MLVSS	Mixed Liquor Volatile Suspended Solids
N	Nitrogen
Na	Sodium
NALO	Nocardia amarae-like organism
N_2	Nitrogen gas
N_2O	Di-nitrogen oxide
NN	Neural Networks
NO_2	Nitrite
NO_3	Nitrate
N&P	Nitrogen and phosphorus
O	Oxygen
OUR	Oxygen Uptake Rate
O_3	Ozone
P	Phosphorus
PAO	Polyphosphate Accumulating Organism
PCA	Principal Components Analysis
p.e.	Population equivalent
PI	Proportional Integral
PLC	Programmable Logic Controller
Q1	First quartile
Q3	Third quartile
RAS	Recycle Activated Sludge
RBCOD	Readily Biodegradable COD
RNA	Ribonucleic Acid
RTC	Real-Time Control
S	Sulphur
SBI	Sludge Biotic Index
SBR	Sequencing Batch Reactor
SCADA	Supervisory Control And Data Acquisition
SI	Scum Index
SRT	Sludge Residence/Retention Time
SSVI	Stirred Sludge Volume Index
sp	Specie
SVI	Sludge Volume Index
T	Temperature
TEFL	Total Extended Filamentous Length
TKN	Total Kjeldhal Nitrogen
TN	Total Nitrogen
TP	Total Phosphorous
TSS	Total Suspended Solids
Turb	Turbidity
UCT	University of Cape Town
UV	Ultraviolet
V	Volume
VFA	Volatile Fatty Acids
VIP process	A modification of the UCT process
V30	Volume of activated sludge settled in 30 min.
WAS	Waste Activated Sludge
WEF	Water Environment Federation
WWTP	WasteWater Treatment Plant
X	Biomass Concentration (=MLVSS)
ZSV	Zone settling velocity
μ_{max}	Maximum specific growth rate

Species of filamentous bacteria

H. hydrossis	*Haliscomenobacter hydrossis*
M. parvicella	*Microthrix parvicella*
N. limicola	*Nostocoida limicola*
S. natans	*Sphaerotilus natans*

CHAPTER 1

INTRODUCTION

In recent years, wastewater treatment plants (WWTP) have undergone important improvements as a result of the present, more restrictive regulations and the demands of the general public, which is, to a greater extent, more aware of environmental problems. Wastewater may be defined as a combination of the liquid or water-carried wastes removed from residences, institutions and commercial and industrial establishments, together with groundwater, surface water, and storm water (Metcalf and Eddy, 2003). Its typical composition, which makes its treatment so necessary, mainly includes: carbohydrates, fats, amino acids, non-volatile acids, detergents, uric acid, creatine, amino sugars, amides, inorganic suspended solids, nutrients (particularly nitrogen and phosphorus), toxic compounds, heavy metals and numerous pathogenic organisms such as bacteria, viruses and intestinal parasites.

Early systems were developed to cope with wastewater from small populations, dealt only with domestic wastes (fill-and-draw reactors, biofilters, etc.), and were slow in operation and rather inefficient. Later on, in the early 1900s Ardern and Lockett developed the activated sludge system which, since its development, has undergone many changes in its operational features that have made it the most widely used technology for biological wastewater treatment. Conventional activated sludge systems were originally designed to remove carbonaceous organic compounds and ammonia, which may be toxic to fish, but due to the current and more restrictive regulations many systems are now upgraded to also attempt to microbiologically remove other nitrogen and phosphorous compounds. In Catalonia, the adoption of the European Directive 91/271 establishes a limit for organic matter of 25 mg/L as BOD_5, a limit of 35 mg/L for suspended solids and, in terms of nutrients, and whenever wastewater is discharged to a sensitive zone (i.e., with an eutrophication risk), the limit is 10 mg/L for total nitrogen and 1 mg/L for total phosphorus.

The contemporary activated sludge process reflects its 90 years development in all aspects. From its very beginning, its evolution has combined to a large extent natural sciences and technology with empirical knowledge (Wanner, 1994). But despite all the advances in trying to improve the operation and efficiency of activated sludge systems, many problems still develop in its operation that adversely affect process efficiency and thus effluent quality. These problems, which originate either in the engineering, hydraulic or microbiological components of the process, include hydraulic or organic overloads, inhibition of nitrification and/or denitrification, bulking sludge or scum formation. It seems clear that in order to understand the nature of these problems, it is necessary to collect a great deal of knowledge from different sources. To apply this knowledge to practical control measures, a combined approach is therefore necessary. It is evident also that, despite the numerous manuals written to help solve common operational problems and despite much research, some process problems continue to be a nuisance in operating wastewater treatment plants. One example of these insidious problems are the activated sludge solids separation problems, a group of microbiological problems whose complexity makes it extremely difficult to find a good general solution to evade them permanently.

The main objective of this thesis is to go one step further from classical control up to the applicaton of some of the most promising techniques of Artificial Intelligence, testing their feasibility in resolving the complex problems that arise in WWTP operation. Can Artificial Intelligence tools become the solution for the complex nuisances in activated sludge systems? Are they the key approach to the control of solids separation problems? These and other questions are addressed in the present thesis.

In the present chapter, a general overview of the activated sludge system, its microbiology and the description of the different types of solids separation problems occurring in the process is presented, together with a deep review regarding the control methods currently available to solve these situations.

1.1 The Activated Sludge System

The activated sludge system is the most widely used technology for biological wastewater treatment in Europe and the rest of the world. Since its development by Ardern and Lockett in 1914, it has gained increasing importance in the treatment of both municipal and industrial wastewaters. This is mainly a consequence of the adaptability of the process to variations in wastewater composition, of its high rates of organic matter removal and of its ability to reduce nutrients, in particular nitrogen (N) and phosphorus (P), to low levels without the use of additional chemicals. In addition, it is probably the most closely studied process, both in terms of the description of the processes involved and in its characterization.

A typical wastewater treatment process consists of primary and secondary treatments to remove both the organic compounds and the suspended solids contained in the raw influent. The primary treatment usually starts with the removal of floatable substances and gross solids such as large objects, rags and grit in order to avoid maintenance or operational problems arising during the process. Immediately afterwards, the primary settler is where approximately one-half of the suspended solids and a portion of the organic matter are removed by sedimentation. After the primary treatment, the essential phase of the wastewater process begins. The secondary or biological treatment can vary according to the physiological conditions involved (aerobic, anaerobic, anoxic), or to the biomass configuration (suspended growth or attached growth). However, the most widely used secondary treatment is the suspended growth biological treatment process: the activated sludge process.

The activated sludge process generally consists of at least one aeration tank or aeration phase where the organic wastes are generally oxidized to form carbon dioxide and water, followed by a sedimentation tank or settling stage where the suspended biomass is separated from the treated water (Figure 1.1). Another relevant characteristic is that the activated sludge is a suspended growth process that maintains a high population of microorganisms (biomass) in aeration tanks by means of solids recycling from the sedimentation tank (WEF, 2002). As the primary effluent is introduced into the aeration tank, it is mixed with return activated sludge to form the mixed liquor, which typically contains between 1500-3500 mg/L of suspended solids. This practice helps to maintain a large number of microorganisms that effectively oxidize organic compounds in a relatively short time.

The aeration tank or biological reactor is where organic carbon, ammonium and phosphate are removed from the wastewater by the microorganisms, which convert these compounds into new cell biomass and products of metabolism. Simplifying the process, it can be considered that microorganisms, while kept under the correct environmental conditions, use the oxygen provided and present in the process to consume the substrate (biodegradable organic particles) present in the wastewater. Biomass is then separated from the treated wastewater in the sedimentation tank for recycling or wasting to solids-handling processes. The sedimentation tank, also known as secondary settler or clarifier, is located downstream from the aeration tank, and apart from performing the separation of the solids from its suspending medium (water), it also removes floating foam and scum produced in and discharged from the aeration tank. The settled sludge makes up the settled solids or sludge blanket of the secondary clarifier. Solids in the sludge blanket may be returned to the aeration tank to treat more wastewater, a fraction known as return activated sludge (RAS), or may be removed (wasted) from the activated sludge process for further treatment and disposal, composing the waste activated sludge (WAS). A good separation (settling) and compaction (thickening) of activated sludge in the secondary clarifier is a necessary condition to guarantee a good effluent quality from the activated sludge process.

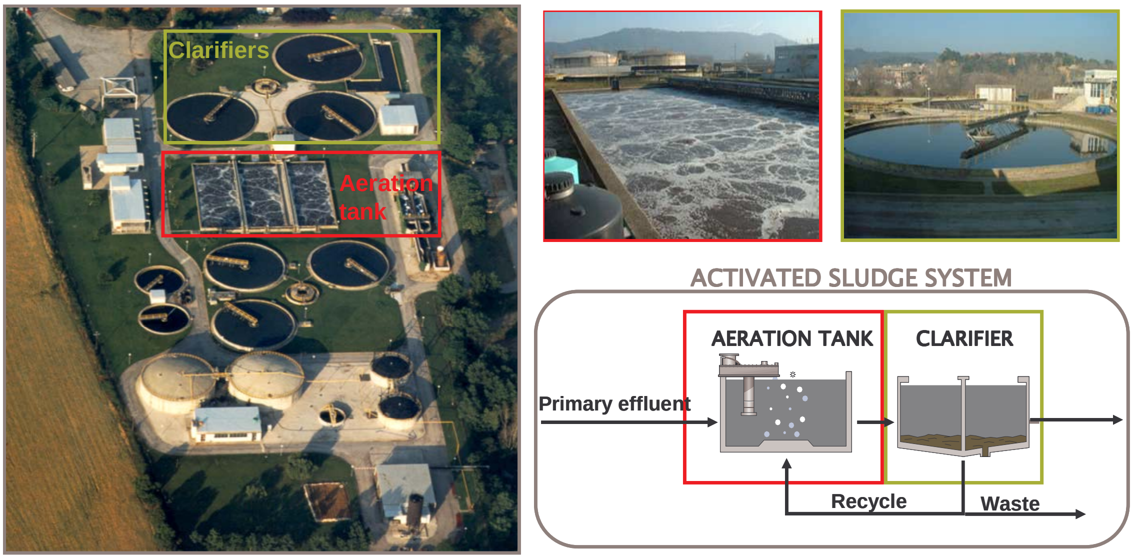


Figure 1.1 Aerial picture and diagram of a typical activated sludge system

One of the main advantages of the activated sludge system is the considerable flexibility and adaptability of its design to the different objectives that need to be accomplished. During the 20th century and up to now, the early activated configurations have evolved into many variations mainly depending on the treatment goals (organic matter removal, nitrification, denitrification or phosphorus removal) or site and operational constraints, but always trying to find the best configuration to treat the incoming wastewater more cheaply and more efficiently in a smaller space, to improve the treatment of particular types of wastes, or to meet increasingly stringent discharge criteria. The main configurations for the activated sludge system include (Metcalf and Eddy, 2003; Grady *et al.*, 1999):

- **Complete mix.** In this type of system, the composition of the mixed liquor is the same throughout the reactor volume. The influent wastewater and the RAS immediately and completely mix with the reactor contents so that the concentration of a given component is the same throughout the reactor. In order to achieve the completely mix, an intense mixing system or a multiple feed are provided. Aeration is supplied uniformly throughout the reactor by mechanical aeration equipment or diffused aeration. The biggest advantage of completely mixed reactors is its capability to attenuate wide swings in effluent concentration. They have, however, several disadvantages. In particular, they are often plagued by filamentous bulking problems due to the high sludge residence time (SRT) required to achieve a good effluent quality. For this reason, this kind of configuration is generally reserved for low-rate applications or for treating industrial or industrial-municipal wastewater where large variations in load are anticipated.

- **Plug-flow**. In this kind of system, the particles entering the process move uniformly along the reactor length without dispersing in the fluid. Consequently, the concentration of substrate and the oxygen uptake decrease along the reactor length while the production of biomass increases. Since organic load and oxygen uptake rates are very high in the influent, there is an implicit necessity for high oxygen transfer rates at the inlet. Theoretically, the plug-flow reactors deliver the highest removal rate per unit volume. They are also generally thought to encourage less filamentous bacterial growth (Wanner, 1994) provided that sufficient dissolved oxygen is present at the inlet and hence, produce better settling sludge than completely mixed systems (Poole, 1984).

- **Step feed**. This configuration is a variation of the conventional plug-flow where the influent wastewater is introduced into the reactor at 2 or more points, usually along the first 50-75% of its length. This maintains similar oxygen consumption rates along the reactor.

- **Contact stabilization**. This system consists of two reactors (a contact reactor and a regeneration reactor) together with the sedimentation tank. In this process, wastewater is in contact with the microorganisms for a much shorter time (1 to 2 hours as compared to 6 to 8 hours for the conventional treatment). Activated sludge in the contactor absorbs particulate substrate, which is enmeshed in activated sludge flocs. At the same time, individual cells of some activated sludge bacteria quickly accumulate soluble substrates. In the regenerator, the activated sludge is aerated and the enmeshed particulate substrate is hydrolyzed and oxidized together with internal accumulated substrates and storage products. This system is commonly used in WWTPs treating small flows and high sludge residence time. The reaeration of RAS means a reduction of oxygen requirement of 30-40% compared to a plug-flow reactor. Although the configuration is slightly less efficient than the conventional activated sludge process, it is more stable when subjected to large variations in flow or BOD loading.

- **Oxidation ditch**. In this type of configuration the wastewater is pumped around an oval or circular pathway by means of brushes, rotors, or other mechanical aeration devices and pumping equipment. The pumping devices are located at one or more points along the flow circuit, and the velocity of water (0.2-0.37m/s) impedes the sedimentation of the suspended solids. In this kind of system, different operational zones can be established within the same reactor, in order to accomplish different reaction phases, such as those required for nutrient removal.

- **High-purity oxygen.** These systems incorporate an enclosed and staged reactor. In the first stage, influent wastewater, RAS and oxygen (typically 98% pure) are added. The headspace gas and mixed liquor flow concurrently from stage to stage, where mixing is provided for solids suspension and oxygen dissolution. The use of high-purity oxygen increases the volumetric oxygen transfer rates, allowing smaller bioreactor volumes and lower hydraulic and sludge retention times.

- **Sequencing batch reactor**. This is a fill-and-draw activated sludge system where wastewater is added to a single reactor which operates in a batch treatment mode repeating a cycle (sequence) continuously. All the steps concerning filling, reaction, clarification and drawing are achieved in a single batch reactor. The use of SBR systems precedes the application of continuous flow activated sludge technology and nowadays successfully competes with the conventional systems.

- **Multistage treatment systems**. Usually consist of two stage activated sludge processes, each consisting of a biological reactor and a settler. During the first stage, the influent is only partially treated (in general, partial removal of organic matter), while in the second stage organic matter removal is completed and nitrogen is also eliminated by providing longer sludge age. If the first stage is heavily loaded, the system is known as the AB-process. The main advantage of this process is the reduction of the total reaction volumes required.

In all these configurations, when the objective of treatment includes nutrient removal (usually nitrogen and phosphorus), the reactor configuration needs to be modified by means of combining different physiological conditions (aerobic, anaerobic or anoxic) to increase the process efficiency. Some examples of configurations for nitrogen removal include: the Wuhrmann, the Ludzack-Ettinger, the modified Ludzack-Ettinger, the Bardenpho and the Bio-denitro processes. On the other hand, whenever both nitrogen and phosphorus need to be removed (also known as Enhanced Biological Phosphorus Removal - EBPR), the following configurations are commonly used: the Phoredox, the A/O, the Johannesburg, the UCT and modified UCT, the VIP or the Biodenipho processes and the alternating aerobic and anaerobic processes (AAA) where aerobic and anaerobic conditions are alternated (Metcalf and Eddy, 2003).

Despite the wide possibilities in activated sludge configuration, a common characteristic of all these kinds of systems is the presence of a group of microorganisms in suspension or in suspended growth also known as biomass or sludge. Because the biomass is aerated, the microorganisms become very active in the degradation and removal of wastes. Therefore, the term “activated sludge” is used to describe the process in which microbiological solids are active in the treatment of wastewater (Gerardi, 2002). The ability of the activated sludge process to degrade all these compounds is therefore achieved mainly through the growth and maintenance of a large and diverse group of microorganisms: the activated sludge microbiological community.

1.2 The Activated Sludge Microbiology

The activated sludge is a microbiological enrichment culture consisting of a mixed, and largely uncontrolled, consortium of micro- and macroorganisms that remove (metabolize) wastewater inorganics and organics or transform them into environmentally acceptable forms (Richard, 1989). The microorganisms present within the process are of the same kind as those which exist in natural systems, but the conditions in the aeration tank and in the clarifier determine the selection of specific species which built a large and diverse microbiological community. Therefore, the growth in numbers and diversity of microoganims occurs over time by increasing the mean residence time of microorganisms in the system, the so-called MCRT (Mean Cell Residence Time) or sludge age. While the number and concentration of microorganisms is very low in the influent and none of the different groups and species are predominant at the beginning, the limiting conditions in which the activated sludge is cultivated lead to a strong competition between individual groups of microorganisms where only the best-adapted microbes prevail. The selection of the most efficient species will improve wastewater treatment, yielding low concentrations of organic compounds (carbon and energy source) and inorganic nutrients in the effluent.

Since the influencing conditions such as organic load, oxygen concentration, toxics presence, pH, water temperature or even the sludge age are not constant in wastewater treatment plants, the dominant species change over time reflecting all the effects that the activated sludge system is exposed to (Wanner, 1994). In this complex and diverse ecosystem, bacteria, which usually are about 95% of the total microbial population, play a key role (Jenkins *et al.*, 2003). Viruses, protozoa, fungi, metazoa and algae compose the remaining 5%.

1.2.1 Bacteria in the activated sludge

Bacteria form the most numerous as well as the most important group of activated sludge microorganisms due to their vital role in metabolizing the wide diversity of organic compounds and, in most occasions, also some nutrients present in the raw influent. The bacteria are present in the aeration tank in floc particles in billions per gram and in the bulk solution in millions per millilitre although, due to their small size, they are usually difficult to observe through an optic microscope. Most bacteria are heterotrophic aerobic or facultative, gram negative, mobile and with a spherical or cylindrical shape (Horan, 1990). According to their type of metabolism, the following groups can be distinguished:

- <u>Chemoheterotrophs or chemoorganotrophs</u>: most of the bacteria present in the activated sludge belong to this group, which is the main responsible for complete removal of organic substances from wastewaters. These bacteria obtain energy from carbonaceous organic matter in influent wastewater and use it for the synthesis of new cells. At the same time, they release energy via the conversion of organic matter into compounds such as carbon dioxide and water. The subgroup of **oxic organotrophic bacteria** need oxygen to degrade the complex (particulate) organic substrates such as lipids, acids, alcohols, sugars or proteinaceous materials. Examples of this subgroup of bacteria are the genera *Bacillus*, *Pseudomonas*, *Micrococcus*, *Alcaligenes*, *Moraxella*, *Flavobacterium* or *Zoogloea*. The **fermentative bacteria** is another important subgroup whose main contribution to wastewater treatment is the conversion of organic compounds to volatile fatty acids (especially acetic acid) and low molecular alcohols during anaerobic conditions, a key process within enhanced biological phosphorus removal activated sludge systems. The genera *Aeromonas*, *Pasteurella*, *Alcaligenes*, *Acetobacter* or *Clostridium* are some examples of fermentative bacteria. The **polyphosphate accumulating bacteria**, part of the important group of the polyphosphate accumulating organisms (PAOs) present in the EBPR systems, is a complex group responsible for the storage and subsequent removal of phosphorus-containing organic substrates. *Acinetobacter* is one of the main genuses involved in this process. Finally, the important subgroup of **denitrifiers**, also classified as anoxic organotrophic bacteria, is a wide group within the activated sludge characterized by its ability to use nitrate as the final electron acceptor in its biochemical reactions. It is estimated that 82-97% of microorganisms in activated sludges from systems with an anoxic zone are able to denitrify (Grabinska-Loniewska, 1991). The genera *Achromobacter*, *Alcaligenes*, *Arthrobacter*, *Bacillus*, *Flavobacterium*, *Hypomicrobium*, *Moraxella* and *Pseudomonas* are typical organotrophic denitrifiers.

- <u>Chemoautotrophs or chemolithotrophs</u>: this group uses inorganic compounds such as carbon dioxide or carbonate as a carbon source for cell growth. The subgroup of **nitrifiers** or aerobic chemolithotrophic bacteria is the most significant group due to its role in nitrogen removal. These bacteria obtain their energy by oxidizing ammonia nitrogen to nitrate nitrogen in a two-stage conversion process known as nitrification. Because very little energy is derived from these oxidization reactions, and because energy is required to convert carbon dioxide to cellular carbon, nitrifying bacteria represent a small percentage of the total population of microorganisms in activated sludge. In addition, autotrophic nitrifying bacteria have a slower rate of reproduction than heterotrophic, carbon-removing bacteria. The main genera are *Nitrosomonas*, *Nitrococcus*, *Nitrospira* and *Nitrosocystis*, responsible for the oxidation of ammonia to nitrite, and *Nitrobacter*, *Nitrospina* and *Nitrococcus*, responsible for the oxidation of nitrite to nitrate. Other important subgroups within this category are the **sulphide/sulphur oxidizing bacteria**, the **iron bacteria** and the **hydrogen bacteria**.

- Photoautotrophic and photoheterotrophic bacteria: these two categories are characterized by the use of light as their energy source instead of the oxidisable chemical compounds. This group though is not very common in activated sludge. The main subgroups include **green bacteria**, **cyanobacteria** and some subgroups of the **proteobacteria**.

Apart from this general classification and regarding bacteria physiology, it is important to notice that although most bacteria usually live as individual cells or forming microcolonies, other types of aggregation are also common in activated sludge. One example of aggregation is the genera **Zoogloea** (Figure 1.2a), which can be defined as an aggregate or colony of different species of bacteria embedded in a polymeric substance, that usually imparts the aggregation a characteristic “finger” shape. On the other hand, the **filamentous bacteria** (Figure 1.2b), represent another example of aggregation where the different cells are enveloped inside a trichome. Growth in the form of filaments represents an ideal form mainly for two reasons: first, all the cells embedded within the trichome are exposed to the same concentration of substrate as in the bulk liquid; secondly, the trichomes create large structures which are not easily washed out from a continuous flow system, especially when the filaments form bundles, or are attached to flocs or solid surfaces. Hence the filamentous growth form is advantageous in environments with low concentrations of substrates and rather high turbulence and flow of cultivation medium (Wanner, 1994). Most filamentous bacteria are chemoheterotrophs, although little is yet known of their metabolic contributions to the performance of plants and their utilization and degradation of organic compounds (Seviour and Blackall, 1999).

Filamentous bacteria will always be present in small numbers in healthy activated sludge systems which operate normally (Jenkins *et al.*, 2003 and Wanner, 1994). Their overabundance though usually results in a problem with sludge settleability. The role of filamentous bacteria in sludge settleability will be discussed in detail in later sections of this thesis.

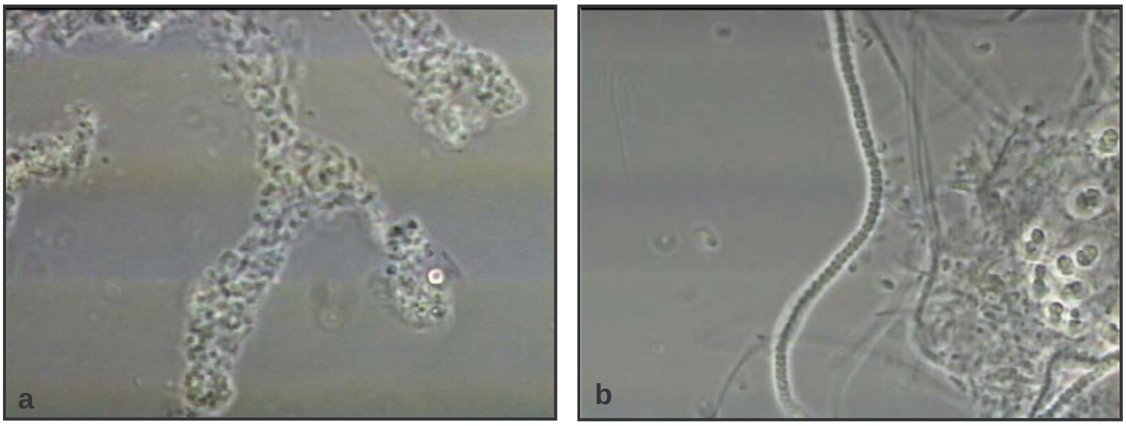


Figure 1.2 Phase contrast micrographs of two examples of bacteria aggregation: (a) Zoogloeal aggregate; (b) filamentous aggregate (*N. limicola*). (Magnification x 1000).

1.2.2 Protozoa in the activated sludge

Protozoa play a secondary, but important role in the purification of aerobic wastewater. The presence of particular types of protozoa is related to effluent quality and plant performance, so their presence and composition tells us how the sludge is changing with the organic load and other operational changes. They can be considered good indicators of the state of the process, not only because of their high sensibility to the operational conditions but also because they are easily detected under the microscope because (their size varies from 1 µm up to a few millimetres).

Protozoa are unicellular eucaryotic microorganisms that feed mostly by active predation on suspended particulate material, including bacterial cells, so they play an important role on BOD removal, free-bacteria and pathogens elimination and on the general renovation of the microbial community. The protozoa in activated sludge treatment systems fall into four major classes: amoebae, flagellates, ciliates (free-swimming, crawling, and stalked) and suctoria:

- Amoebae or Rhizopods: this kind of microorganisms lack cilia or flagella to move, their motion is caused by cytoplasmatic streaming and the production of pseudopodia (ameboid movement) (Seviour and Blackall, 1999). They ingest their prey by phagocytosis and they usually multiply by simple cell division. They range in size from 10 to 200 µm. Amoebae are divided into two different types, testate and naked (see Figure 1.3a). Testate amoebae have shells (or tests); naked amoeba do not. The shells can be proteinaceous or siliceous.

  Amoebae are frequently present in the raw influent, but are scarce in the aeration basin. Amoebae can only multiply when there is an abundance of nutrients in the aeration tank. Therefore, the presence of large numbers of amoebae in the aeration basin usually indicates that some sort of shock loading has occurred and there is a lot of food available. Their high tolerance to very low amounts of dissoloved oxygen (DO) may also indicate oxygen deficiency within the aeration tank.

- Flagellates: flagellate protozoa (Figure 1.3b) possess small numbers of flagella through which they move. Flagellates range in size from 5 to 20 µm. Many flagellates are able to feed autotrophically as well as heterotrophically. Like bacteria, this group of microorganisms feeds on organic nutrients in the sewage and consequently, as the nutrient levels decline, they find it increasingly difficult to compete with the bacteria for soluble food and their numbers begin to decrease. Thus, flagellates (as well as rhizopods) are typical only of start-up periods (low sludge age), when large amounts of soluble food are available (high F/M or high BOD), prior to the stabilization of a bacterial population. They quickly dominate over the amoebae because they are more efficient feeders. The presence of large amounts of flagellates in the later stages of the activated sludge development usually indicates that either the wastewater still contains a large amount of soluble organic nutrients or that the system is recovering from a toxic discharge. They can also indicate conditions of low dissolved oxygen.

- Ciliates: the major function of the ciliate protozoa is predatory activity. While bacteria and flagellates compete for dissolved nutrients, ciliates compete with other ciliates and rotifers for bacteria and other protozoa. The presence of ciliates generally indicates a healthy sludge, because they dominate after the floc has been formed and after most of the organic nutrients have been removed (medium sludge age). According to their physical location within the biomass, the different species can be classified into three main subgroups (Madoni, 1994a): (1) **free-**

swimming ciliates, such as *Litonotus* spp. or *Paramecium* spp., whose main characteristic is their free motion between flocs in the bulk liquid; (2) **crawling ciliates** (Figure 1.3c), such as *Aspidica* spp. or *Euplotes* spp., that crawl over and through the flocs; and (3) attached, **sessile (or stalked) ciliates** (Figure 1.3d), such as *Vorticella* spp., *Opercularia* spp., *Carchesium* spp., *Epistylis* spp. or *Vaginicola* spp., which are fixed to flocs through holdfasts, and extend into the bulk liquid, growing alone or in large colonies which can even form whole flocs. The colonial forms of stalked ciliates usually occur at higher MCRTs. Sessile ciliates are found in large numbers when the bacterial population and dissolved oxygen concentration of the treatment process are high, the wastewater environment is stable and a mature floc structure has developed. Thus, stalked ciliates usually indicate a stable wastewater environment and a healthy biomass.

- Suctoria: they are probably closely related to the ciliates, although their cilia have evolved into fine tubes or tentacles. They use these to spear their prey, such as free swimming ciliates and flagellates. They produce ciliate larvae and are likely to be sessile, feeding largely upon ciliate and flagellate.

In general, protozoa perform a positive role in the activated sludge process. On the one hand they add weight to the floc particles and improve solids settleability by cropping bacteria from the bulk solution and releasing sticky secretions that coat fine solids. On the other hand, they help to recycle nitrogen and phosphorus by cropping and digesting bacteria and excreting nitrogen-containing and phosphorus-containing wastes to the bulk solution (Gerardi, 2002).

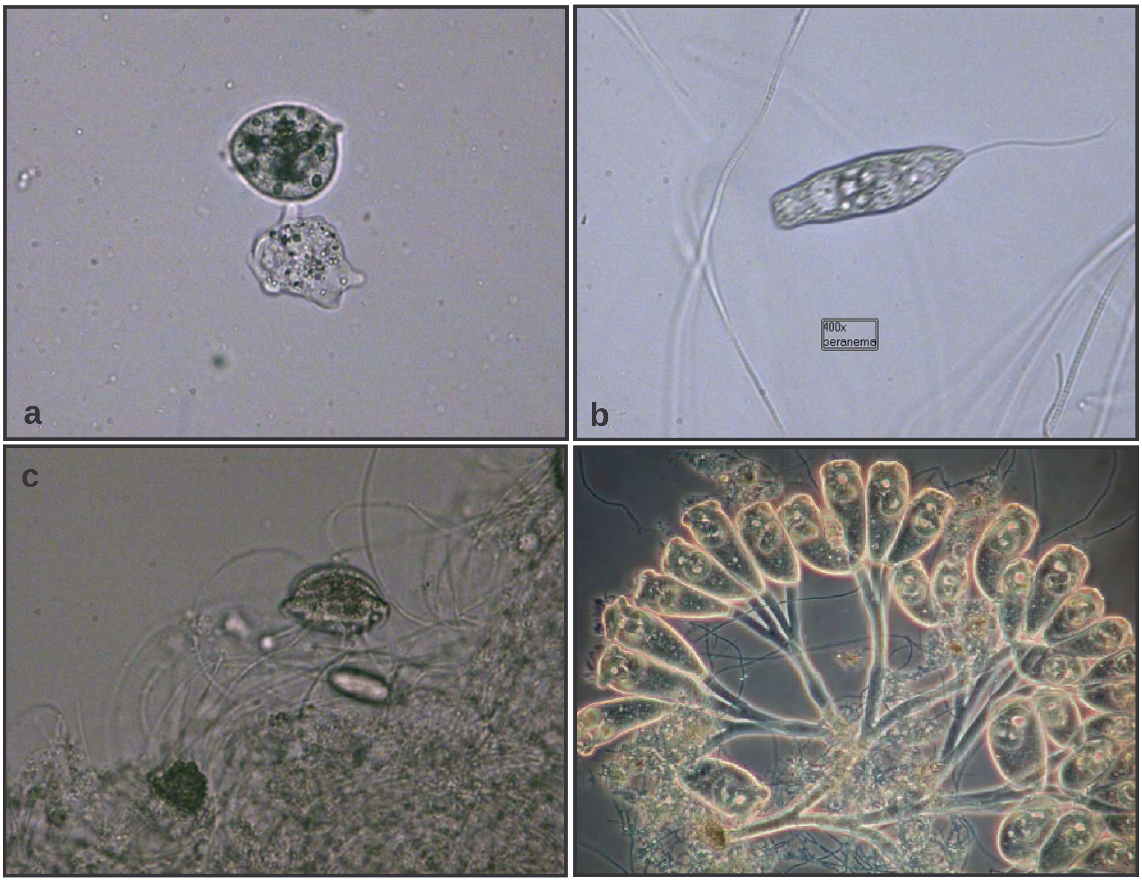


Figure 1.3 Phase contrast micrographs of four examples of protozoa: (a) testate and naked amoebae; (b) flagellates (*Peranema* sp.); (c) crawling ciliate (*Aspidica* spp) and (d) sessile ciliates (*Epistyilis* spp.) (Magnification x 400).

1.2.3 Metazoa in the activated sludge

Metazoa are pluricellular eukaryotic microorganisms. Their cells are grouped forming tissues and specialized organs and their size is quite big varying from 40μm up to 3mm. They occupy a niche similar to attached ciliates, so they are mostly found in well-established activated sludges with good flocculation. They probably play important roles in plant operations as predators, consuming bacterial cells, but their true significance is not as well understood as that of protozoa (Woombs and Laybourn-Parry, 1987). The most common metazoa in the activated sludge can be divided into four subgroups: rotifers, gastrotrichs, nematodes and oligochaete worms.

- Rotifers: These are the most common metazoa in the activated sludge. Rotifers (Figures 1.4a and 1.4b), range in size from 40 to 500 μm. The mouth opening of the rotifer is surrounded by two bands of cilia. The beating of these cilia creates water currents for locomotion and food gathering. Rotifers move by swimming freely or by crawling. The principal role of rotifers is the removal of bacteria and the development of flocs. Rotifers are able to consume both bacteria and particulate matter and contribute to the removal of effluent turbidity by removing non-flocculated bacteria. Mucus secreted by these organisms at either their mouth opening or their foot helps in floc formation. Moreover, they graze on the floc structure, which results in increased oxygen penetration. Rotifers require a longer time to become established in the treatment process, so their presence indicates stabilization of organic wastes, and thus a high sludge age and a low organic load. Like protozoa, these microorganisms are strict aerobes and are more sensitive to toxic conditions and high organic loads than bacteria. Therefore, their presence generally indicates a good, stable sludge with plenty of oxygen.

Figure 1.4 Phase contrast micrographs of four examples of metazoa: (a and b) rotifers; (c) nematodes and (d) oligochaete worm (*Aelosoma* spp.) (a and c magnification x 400 and b and d magnification x 100)

- Gastrotrichs: A group of metazoa similar to the ciliates. Their body is full of bristles and cilia but they are usually bigger than ciliates (100-500 μm). They are very active, sliding and twisting up continuously. As the rest of metazoa, gastrotrichs have developed organs (digestive system, musculature and nervous system, reproductive and excretory organs). They feed on protozoa, algae and detritus present in the wastewater. This group of microorganisms is not very common in activated sludge processes so their presence is not especially significant.

- Nematodes: As the rest of metazoa, nematodes are often observed in systems operating at high sludge ages. These microorganisms secrete a sticky substance in order to anchor themselves to a substrate (media) or floc particles so that they can feed without interference by currents or turbulence. Nematodes (Figure 1.4c) range in size from 0.5 to 3.0 mm in length and from 0.02 to 0.05 mm in width. The nematodes present in WWTPs are usually intestinal parasites, so their presence is indicative of faecal contamination.

- Oligochaete worms: They are occasional visitors of the activated sludge. The most common worm observed in wastewater is *Aelosoma* spp. (Figure 1.4d), a bristle worm with a long segmented body full of pink spots and bristles. They are generally very long (up to 3.0 mm in length) and have many internal organs. They usually indicate systems operating at very high sludge ages (or low food-to-microorganisms ratio). They crop the bacterial population and recycle nutrients. They can also indicate the presence of a high nitrate concentration.

1.2.4 Other important microorganisms in the activated sludge

Although bacteria, protozoa and metazoa are the most important and common groups found in activated sludge processes, sporadically other groups such as viruses, algae or fungi can also occur (Seviour and Blackall, 1999).

- Viruses and bacteriophages: Animal viruses, including the human immunodeficiency virus (Ansari *et al.*, 1992) may pose a serious potential health hazard. Some, like rotaviruses, often survive the treatment process (Lewis *et al.*, 1986).

- Algae: Microorganisms sporadically observed within the microbiological community of activated sludge containing vegetal substances such as cellulose and starch and photosensitive pigments such as chlorophyll. These organisms do not grow in activated sludge because light does not penetrate significantly into the mixed liquor. Algae in activated sludge usually originate from another treatment unit in the process stream and are fed or recycled into the activated sludge process.

- Fungi: They are not important members of the microfauna, except under certain conditions. Since they are not good competitors due to its μ_{max} lower than the μ_{max} of bacteria, its occurrence may indicate a low presence of bacteria as a consequence of some extreme conditions such as low pH. Fungi have been found to cause bulking in plants with a low pH (Jenkins *et al.*, 2003).

1.2.5 The importance of microfauna as indicator of plant performance

Some of the microorganisms observed in activated sludge are capable of developing under a wide range of different conditions. Others, however, are strictly dependent on environmental variables, including physical, chemical and biological conditions. These microorganisms, especially protozoa and metazoa, are therefore of maximum importance as indicators of plant performance. According to the results of several studies based on protozoan and metazoan ecology (Curds, 1982; Mudrack and Kunst, 1986; Al-Shahwani and Horan, 1991; Esteban *et al.*, 1991; Madoni, 1994b, Salvadó *et al.*, 1995 and Jenkins *et al.*, 2003), a number of general conclusions and relations can be established:

⇒ Flagellates, amoebae, and small free-swimming ciliates require high prey densities (>10^6 to 10^7 bacteria/L) because their feeding mechanisms are inefficient. These groups appear during plant start-up and are indicative of low sludge age and high organic load conditions.

⇒ Attached ciliates, rotifers and other invertebrates develop at lower prey densities because of their attachment to the activated sludge floc and their ability to feed by ciliary action. These organisms are indicators of high sludge age and low organic load conditions.

⇒ Individual protozoan species can also be indicators of process performance. For example, the shelled amoebae *Arcella* and *Vaginicola* spp. are associated particularly with nitrifying plants, while sessile ciliates like *Opercularia spp*. and *Vorticella microstoma*, small flagellates and the swimming bacteriovore ciliates are said to indicate poor effluent quality (Madoni, 1994b).

Satisfactory activated sludge performance occurs when there is a balance among free-swimming and attached ciliates and rotifers and when few or no flagellates are present. According to Madoni (1994b), an efficiently operating plant would also present a high population density of protozoa of more than 10^3 protozoa per mL with dominant crawling and sessile forms and a well diversified community where no single species or group of species is overwhelming. Whenever this is not the case, the predominant group or groups may give some clues about the performance of the process.

Regular microscopic observation of activated sludge can tell the trained plant operator the physiological state of the biocenosis. This kind of visual analysis can lead to a set of reliable results covering a wide number of possible factors. These factors would be more difficult to determine by classical chemical analysis, or at least it would take more time to get results. Therefore, it can be concluded that a periodic microscopic examination (with a daily recommended periodicity) of activated sludge may provide a rapid, simple and convenient method for indicating sudden changes in performance (Curds, 1982). It will contribute also to understand how the process is evolving and even to predict some frequent problems occurring in the facility, helping operators to correct them before they get worse.

One example of an alteration detected by biomass monitoring is toxicity. Ciliates and rotifers in activated sludge are generally the first to be impacted by toxic materials and thus can be used as *in situ* biomonitoring tests for toxicants and other adverse stresses on the process. Signs of toxicity or stress usually include the slowing or cessation of cilia movement and an overabundance of flagellates and small free-swimming ciliates. Stresses other than toxicity that induce these responses include low DO, pH outside the range of 6.5 to 8.5 and high temperatures (Jenkins *et al.*, 2003).

1.3 Floc Formation

Apart from the crucial role of microfauna in wastewater treatment, the quality of the effluent from activated sludge treatment plants is highly dependent on the correct separation between the treated wastewater and the biomass, which takes place in the secondary settlers. For a successful separation, the microorganisms must form **flocs** that settle and compact well without leaving a high concentration of suspended solids in the supernatant.

Activated sludge flocs are made up of two types of components: a biological component consisting of a wide variety of bacteria, fungi, protozoa and some metazoa; and a non-biological component, made up of inorganic and organic particles from the incoming wastewater and extracellular polymeric substances, which play a key role in the bioflocculation of the activated sludge. The basis of the floc seems to be a number of heterotrophic bacteria, called **floc-formers,** which include such genera as *Pseudomonas*, *Achromobacter, Flavobacterium*, *Alcaligenes*, *Arthrobacter*, *Citromonas*, and *Zoogloea* (Jenkins *et al.*, 2003). These and many other chemoheterotrophic bacteria are able to convert organic substrates to specific extracellular polymeric substances (EPS) termed also **glycocalyx** (Costerton *et al.*, 1981). Glycocalyx is mainly composed of carbohydrates, humic substances and proteins, of which proteins are believed to be most important for the flocculation (Higgins and Novak, 1997). As any organic polymer, glycocalyx increases water viscosity and enables individual cells to stick together or attach to solids surfaces, forming larger aggregates that settle more easily due to their increasing weight. The bridging of cells is based on a decrease in surface electrostatic charge as in the flocculation by synthetic polyelectrolytes (Wanner, 1994).

The **filamentous backbone theory** of activated sludge flocs (Sezgin *et al.*, 1978) assumes that the structure of the floc is formed at two levels (Figure 1.5). These have been termed the “microstructure” and the “macrostructure”.

- The **microstructure** is provided by processes of microbial adhesion, aggregation and bioflocculation. It is the basis for floc formation because, without the ability of one microorganism to stick to another, the large aggregates of microorganisms that exist in activated sludge would never form. The bioflocculation results from the interaction between the EPS, which function as polyelectrolytes. This glycocalyx forms a felt-like envelope around individual cells or groups of cells. Researchers have noticed that for an activated sludge in which only “floc-forming” bacteria are present (flocs that have only the so-called microstructure), the flocs are usually small (up to about 75 µm in dimension), spherical and compact, but mechanically rather weak.

- The **macrostructure** is provided by filamentous microorganisms. When an activated sludge culture contains filamentous organisms, large floc sizes are possible because the filamentous organisms form a backbone within the floc, to which the floc-formers are firmly attached by their EPS. This backbone provides the floc with strength so that it can hold together in the turbulent environment of the aeration basin and resist the shearing action. Large flocs containing filamentous organisms become therefore irregularly shaped, rather than approximately spherical (as they do when filamentous organisms are absent or present only in small numbers). With increasing sludge age, an increase in size of the floc particles provides a larger number and a greater diversity of microfauna for the treatment of a larger quantity and a larger variety of wastes.

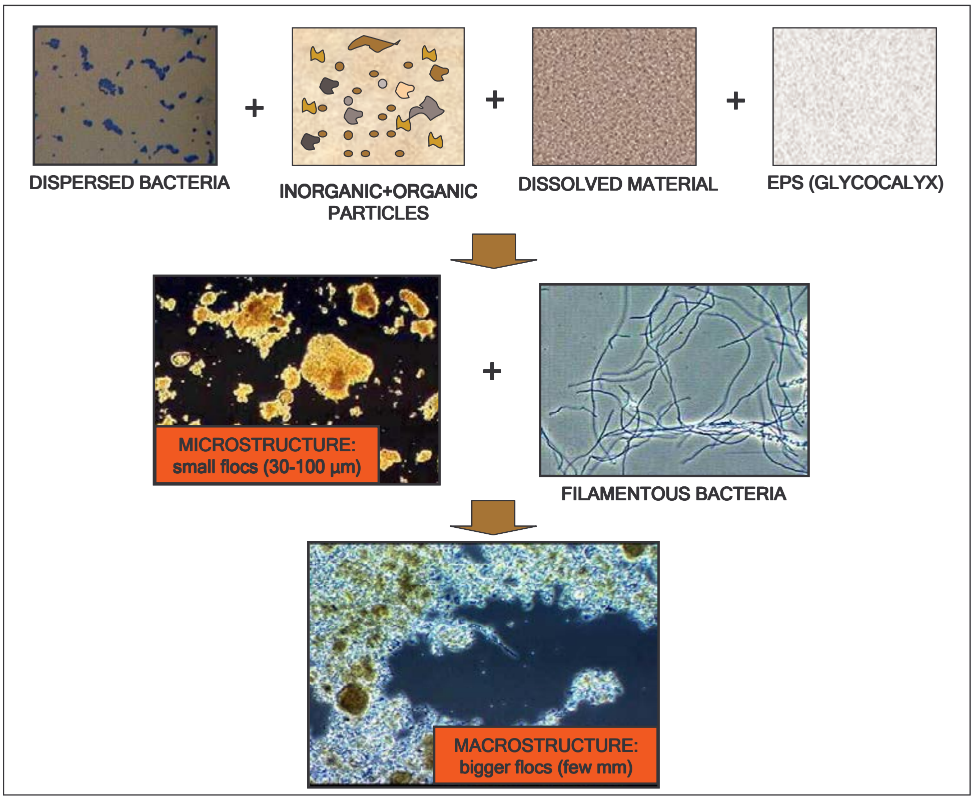


Figure 1.5 Microstructure and macrostructure of the activated sludge floc

According to the filamentous backbone theory, the ideal floc is the result of a balanced growth of filamentous and floc-forming organisms, which implies the existence of big and compact flocs that settle easily, resulting in a clear effluent (Figure 1.6a). Whenever an imbalance between the floc-formers and the filamentous bacteria exists, problems will occur. A total absence of filamentous microorganisms can lead to totally dispersed flocs (Figure 1.6b). This type of sludge settles rapidly but can produce a turbid supernatant and a low sludge volume index (SVI). It is the larger compact flocs that settle rapidly; the smaller aggregates, sheared off from larger flocs, settle slowly and create the turbid supernatant. On the other hand, if there is an excessive proliferation of filamentous bacteria (Figure 1.6c), the floc structure will lead to excessive growth of the floc size, which gives rise to compaction problems and prevents the settling of these spongy flocs. As a consequence, the SVI will increase and the biomass will probably spill out when the sludge level surpasses the secondary settler capacity.

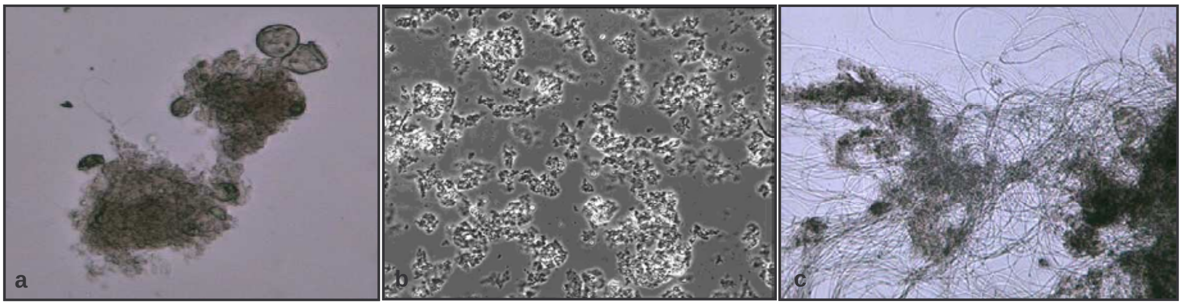


Figure 1.6 Phase contrast micrographs of different floc structures: (a) ideal floc, (b) deflocculation and (c) overabundance of filaments

Although the theory of Sezgin *et al.* (1978) is widely accepted, some studies (e.g., Chudoba *et al.*, 1973 and Wanner and Grau, 1989) have reported many well-floculating and settling activated sludge systems that had no filamentous backbones. Recent studies (Wilén *et al.*, 2004 and Liao *et al.*, 2001) support the theory that the strength of activated sludge flocs is, just like in the case of other biological aggregates, dependent on the interparticle forces between the different floc constituents: microorganisms, EPS, and organic and inorganic components. Chemical interactions help bridging of EPS by means of divalent and trivalent cations, and the active aerobic metabolism of bacteria is also essential to maintain strong aggregates. This implies that apart from filament abundance, the aerobic microbial activity and the surface characteristics of sludge, such as charge and hidrophobicity, need also to be considered whenever sludge settleability is being assessed.

The correct formation of flocs determines important qualities of activated sludge due to its direct correlation with sludge settleability, sludge thickening and dewatering and clarity of the final effluent. Once again, a complete observation of the sludge, including macroscopic analysis such as the determination and observation of the sludge settleability (V30 test) and an exhaustive microscopic analysis to examine typical characteristics of the floc such as size, density, strength, porosity, shape or presence of filaments are strictly necessary to avoid undesired settleability problems.

1.4 The Activated Sludge Solids Separation Problems

The Activated Sludge Solids Separation Problems is the general name or classification given to a group of problems occurring within the activated sludge process. Their main and common characteristic or consequence is the loss of solids through the WWTP effluent due to the inefficient separation between the biomass and the treated wastewater during the secondary settling. The physical and microbiological causes for each one of these problems have been fairly well defined in numerous studies and publications (see, for example, Jenkins *et al.*, 2003, Wanner, 1994, Casey *et al.*, 1995, Eikelboom and van Buijsen, 1983 and Gerardi, 2002).

These phenomena can also be divided into four groups: those that are clearly caused by excessive quantities of filamentous organisms like **filamentous bulking** and **biological foaming**; those that are associated with small, non-settleable particles due to deflocculation of flocs such as **dispersed growth** or the formation of **pin-point floc**; the **viscous bulking** caused by an excessive production of EPS, and finally **rising sludge**. Despite having no connection with poor floc formation, rising sludge has also been considered as a solids separation problem because of its negative impact on the effluent quality due to the loss of solids that it causes.

1.4.1 Filamentous Bulking

Filamentous bulking or bulking sludge is probably the most frequent solids separation problem occurring in WWTPs all over the world. The term bulking defines a process in which the activated sludge tends to be bulky, i.e. its density decreases as a consequence of the overabundance of filamentous bacteria which hinders the sedimentation of the sludge during secondary settling. Whenever the settling velocities of the sludge (which directly depends on the outflow rate and the surface area of clarifiers) are within the acceptable range for efficient separation in secondary settling tanks, the effluent will not be affected, since all the microflocs can be enmeshed and trapped in the filamentous network resulting in a clear effluent.

However, when the filamentous bulking is severe (Figure 1.7), and the sedimentation zone of the secondary settling tank is full of poorly compacted sludge, an overflow of sludge blanket may occur (Wanner, 1993). This problem can be reflected also in the sludge settleability measurement (V30), and consequently in the SVI value. As a rule of thumb, the probability of having a bulking sludge problem becomes significant whenever the SVI takes values above 150-200 mL/g.

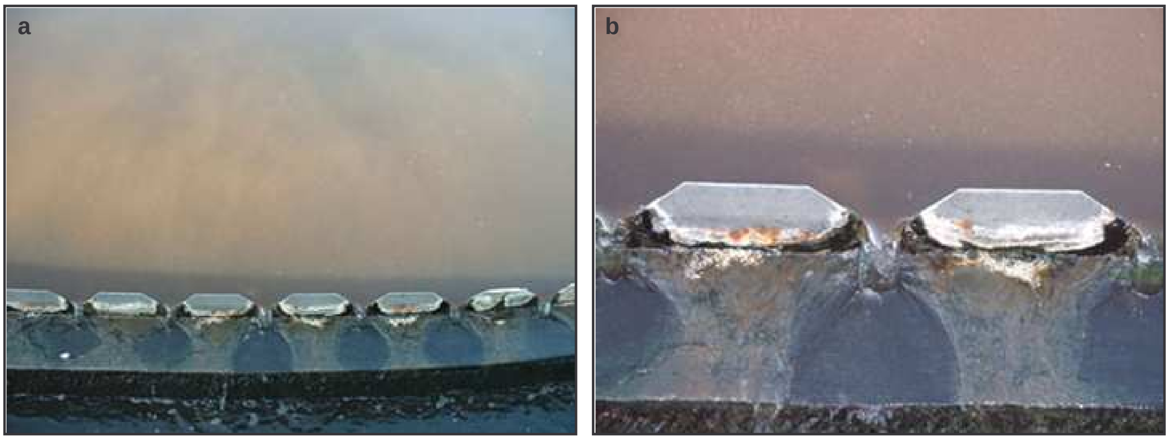


Figure 1.7 Example of filamentous bulking in a secondary settler showing (a) the bulky appearance of the sludge and (b) the loss of solids with the effluent

Filamentous bulking can thus be defined as a typical solids separation problem due to its main implication on the loss of solids through the effluent. As a consequence of this poor activated sludge compaction and loss of solids, the following outcomes can appear: (1) the recipient water ecosystem is polluted with organic matter with a high COD[1] content, creating problems of oxygen depletion or future contamination due to the long term release of N, P and organic matter due to degradation; (2) bacterial contamination; (3) the process efficiency is affected, given that that part of the sludge that should have been used in the treatment process is removed and (4) the concentration of the return and waste sludge is very poor with negative consequences both on the control of sludge concentration in reaction basins and on the sludge dewaterability process during sludge-handling operations. In addition, in countries where facilities are penalized for effluents containing excessive solids levels, bulking can also be very costly (Seviour and Blackall, 1999).

The bulking phenomenon was first reported in the 1920s and early 1930s in countries where the aerobic activated sludge system was becoming the established treatment method. Notably these reports were by Scott (1928) in England, Ruchoft and Watkins (1928) in the United States, and Smit (1934) in Holland. According to the ATV Working Group (1989), until the 1950s bulking was not a serious problem in municipal plants; it was encountered mainly in plants treating a high proportion of industrial wastes. However, by the 1950s, bulking was becoming increasingly problematic in municipal activated sludge plants. In the 1970s and early 1980s, surveys of several hundred wastewater treatment plants in Europe showed that bulking occurred in about 50% of the plants (Eikelboom, 1975 and Wagner, 1982). Nowadays, the need for N and P removal from municipal wastewaters to limit eutrophication has resulted in even the earliest applications of the activated sludge system being redesigned for biological N removal and, later, for P removal too. The environmental conditions encountered by the biomass in the N and P removal systems are very different from those in conventional fully aerobic systems. Nevertheless, the N and P

---

[1] COD (Chemical Oxygen Demand). Indirect measure of the amount of organic compounds present in wastewater.

systems do not appear to be exempt from filamentous bulking problems; indeed, it appears that these systems tend to produce generally poorer settling sludge than their more conventional short sludge age, fully aerobic counterparts.

Bulking due to filamentous organisms seems to have thus a long history of occurrence and at present constitutes a major problem in activated sludge systems. Bulking is encountered over virtually the entire spectrum of systems, from short and long sludge age aerobic systems operating in plug-flow, in-series and completely mixed regimes, to long sludge age anoxic-aerobic N removal and anaerobic-anoxic-aerobic N and P removal systems (Casey *et al.*, 1995). Nowadays though, the use of Membrane Bioreactors (MBRs) seems to be a promising type of configuration to avoid solids separation problems.

1.4.1.1 Filamentous microorganisms and bulking

Although filamentous bacteria overabundance is the main origin of filamentous bulking, it has to be noticed that their excessive abundance need not always cause bulking, given that not all filaments have the same effects on the settling properties of the flocs. It usually depends on the filament type, shape, morphology and on their differing abilities to grow out from the floc (Wanner, 1994). Filamentous bacteria usually interfere with the compacting and settling of the activated sludge either by producing a very diffuse floc structure (Figure 1.8a) or by growing in profusion beyond the confines of the activated sludge floc into the bulk medium and bridging between flocs (Figure 1.8b). The way this interferes with compacting and settling depends on the filamentous organism causing the problem (Jenkins *et al.*, 2003). Thus, *M. parvicella*, *N. limicola* I and II, Type 0092 or Type 0675 usually cause diffuse floc structures while *N. limicola* III, *S. natans*, *Thiothrix spp*. or Type 021N originate bridges between flocs.

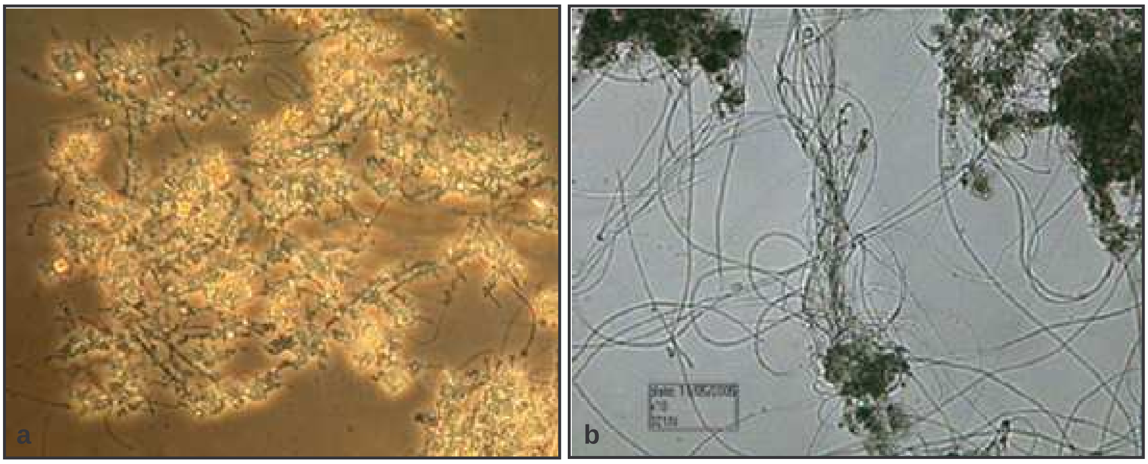


Figure 1.8 Phase contrast micrographs of different filaments interferences: (a) *Nocardia* spp causing diffuse floc (magnification x 400) and (b) Type 021N originating floc-bridging (magnification x 100)

Older publications dealing with the bulking problem always mentioned "*Sphaerotilus*" as the microorganism responsible for the bulking phenomenon. Nowadays though, thanks to the identification methods developed by Eikelboom (1975), and the subsequent development of dichotomous keys for *in situ* identification of filamentous microorganisms (Eikelboom and van Buijsen, 1983 and Jenkins *et al.*, 2003) it is possible to identify the many different filamentous microorganisms that can lead to this problem and to relate their presence to the existence of specific characteristics of the wastewater.

In general, it is rare that the activated sludge sample from a plant with bulking contains only one filamentous organism type. Usually three or more types are seen in a municipal WWTP, although often there is one or more filamentous organism types present in significant (dominant) quantities along with several others present in smaller amounts (secondary). The correlation with the cause of bulking is made with the dominant filamentous organism types (Jenkins *et al.*, 2003). In this way, it has been found that almost 30 different kinds of bacteria can play a role in a bulking situation, and that approximately 10-12 types of them account for the great majority of all bulking episodes, as shown in Table 1.1.

The existing results, obtained from several investigations in different countries, show that the microorganisms lists of the United States and Europe are similar in that the top six to eight most frequently occurring filaments are from four of the six causative categories (see 1.4.1.2 Causes of filamentous bulking). However, both lists differ substantially from those for South Africa and Australia, where six of the eight filaments are all low F/M ones. At the time of the US and European surveys (early 1980s), the plants were principally aerobic with relatively short sludge ages (high F/M ratios). In contrast, South African and Australian plants have long sludge ages (20-30 days) and usually incorporate anoxic-aerobic or anaerobic-anoxic-aerobic zones for biological N and N&P removal. Thus, the significant differences in environmental factors clearly give rise to considerably different filament populations. However, as nowadays Europe and the USA increasingly implement biological nutrient removal (BNR) plants, it can be expected that the problematic filaments will belong to the low F/M group (Ekama *et al.*, 1997).

Table 1.1 Comparison of dominant filamentous organisms in bulking activated sludge from several geographical areas (adapted from Jenkins *et al.*, 2003)

Filamentous organism	Ranking in order of prevalence								
	USA	The Netherlands	Germany	South Africa	Australia	Denmark	Italy	Czech Republic	Japan
NALO	1	-	7	2	7	-	2	6	4
Type 1701	2	5	8	-	14	-	9	13	-
Type 021N	3	2	1	-	13	3	6	7	1
Type 0041	4	6	3	6	2	2	3	5	6
Thiothrix spp.	5	19	-	-	15	-	7	8	-
S. natans	6	7	4	-	17	-	-	-	6
M. parvicella	7	1	2	3	1	1	1	1	3
Type 0092	8	4	-	1	3	3	4	3	-
H. hydrossis	9	3	6	-	4	-	-	11	6
Type 0675	10	-	-	5	2	-	3	11	6
Type 0803	11	9	10	8	10	-	-	4	6
N. limicola	12	11	7	-	5	-	5	2	6
Type 1851	13	12	-	4	9	6	-	-	2
Type 0961	14	10	9	-	11	-	8	9	4
Type 0581	15	8	-	-	-	-	-	-	-
Beggiatoa spp.	16	18	-	-	18	-	-	-	-
Fungi	17	15	-	-	6	-	-	-	-
Type 0914	18	-	-	2	8	5	10	10	-
Type 1863	-	-	-	-	16	-	-	-	-

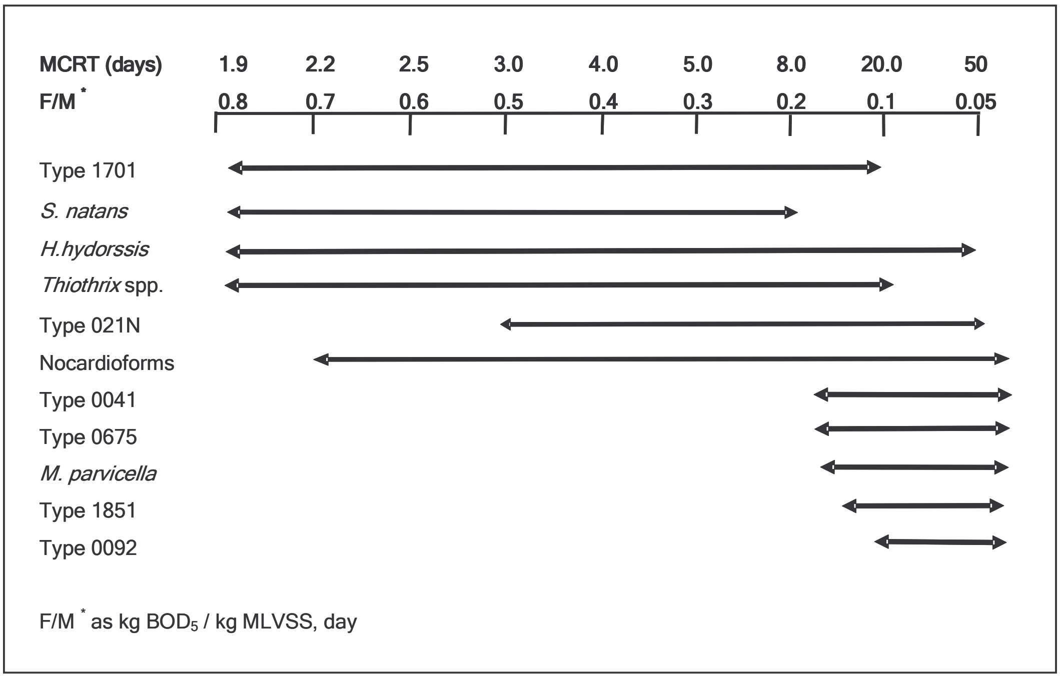
1.4.1.2 Causes of filamentous bulking

Filamentous bacteria are distinguished from floc-forming bacteria not only by their morphology but also by their physiology and growth kinetics. Filamentous bacteria can grow at a different rate than floc-forming bacteria depending on the environmental and nutritional conditions. According to the results of numerous comparative studies mentioned above, the different factors that can favour filamentous bacteria and thus provoke filamentous bulking in activated sludge processes can be divided into two main categories: operational conditions and wastewater composition.

Operational conditions include those operational parameters that plant operators can generally modify to accomplish specific treatment goals. The operational conditions directly affect the balance between floc-formers and filamentous microorganisms, becoming on most occasions the source of bulking problems. These conditions are the following:

- <u>Sludge age.</u> The sludge age, measured as Mean Cell Residence Time (MCRT), can be defined as the mean residence time of microorganisms within the activated sludge system. The correlation between sludge age and the presence of filamentous bacteria is based on the filamentous backbone theory (see 1.3 Floc Formation), which states that generally a high sludge age is associated with the growth of a higher number of filaments than a low sludge age. Although some filamentous microorganisms occur over a fairly wide range of MCRT values (*e.g.*, *S. Natans*, type 1701, *Thiothrix* spp., and *Nocardia* spp.), in general, most filamentous bacteria, as shown in Table 1.2, (*e.g.* types 0092, 1851, 0675, 0041, and *M. Parvicella*), are associated with high MCRTs (>10 days) (Richard, 1989).

Table 1.2 Relationship of MCRT (and F/M) to the occurrence of specific filamentous organisms in activated sludge (adapted from Jenkins *et al.*, 2003)



F/M * as kg BOD_5 / kg MLVSS, day

- <u>F/M ratio</u>. The food-to-microorganisms ratio indicates the proportion of food (measured as kg of BOD) per kg of microorganisms (measured as MLVSS). The influence of the F/M ratio on the presence of bulking can be inversely related to sludge age (Table 1.2), given that most WWTPs operating at high sludge age usually operate also at low F/M ratios due to the higher number of microorganisms present in the activated sludge (higher MLVSS). According to some experiments, such as the one carried out by the UCT group (Chudoba, 1985), it has been found that the relative abundance of filamentous and non-filamentous (floc-forming) organisms is related to their relative growth rates when exposed to varying concentrations of substrate (Figure 1.9).

At low substrate concentrations (low F/M ratios) the non-filamentous, floc-forming microorganisms have a high μ_{max} but a low affinity for the substrate (high K_s), whereas the filamentous forms are slow-growing organisms that have a low μ_{max} but a high affinity for the substrate (low K_s). Furthermore, the filamentous bacteria have a greater surface area to volume ratio than the floc-forming bacteria and so are better able to absorb a substance when it is only present at a low concentration. Low F/M ratio is therefore a major cause for the increase of different filamentous bacteria. In any case, no explicit definition for low F/M (or equivalent load factor) has been determined mainly due to its dependence on the definitions or measures whereby it can be expressed (COD, BOD, MLSS or MLVSS) and on wastewater characteristics (i.e. raw or settled wastewater). For this reason, when defining the low F/M group of filamentous bacteria, it is preferred to relate its cause to a more consistent and fundamental parameter: the sludge age (Casey *et al.*, 1995). As a consequence, from here on, low F/M or high sludge age will indistinctly refer to the same cause.

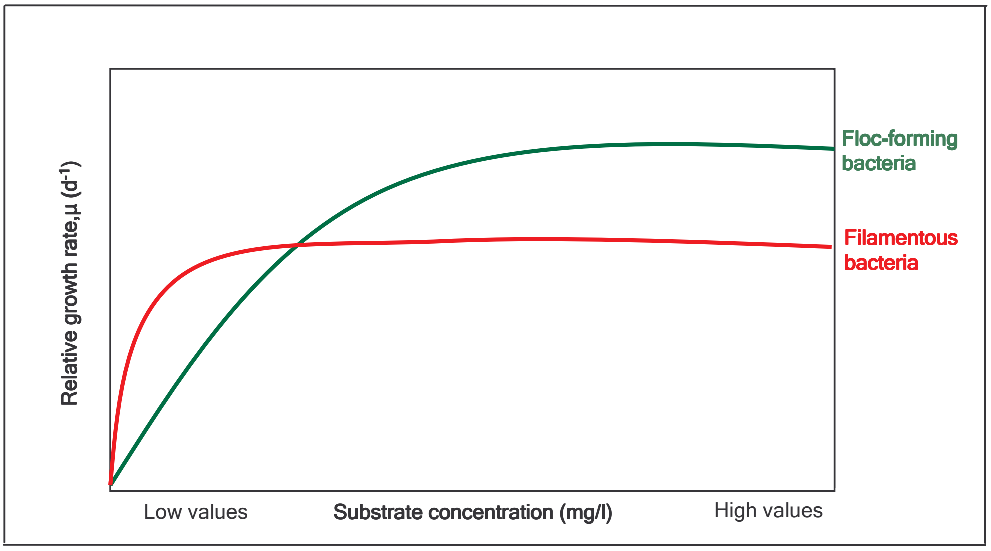


Figure 1.9 Relative growth rates of floc-forming and filamentous bacteria in relation to substrate concentration.

Also related to the filamentous microorganisms developed at low F/M ratios, recent studies (Kappeler and Brodmann, 1995 and Musvoto *et al.*, 1999) have demonstrated that the incomplete denitrification occurring in low F/M plants can favour low F/M filamentous microorganisms due to the inhibition of the heterotrophic facultative floc-forming microorganisms in the aerobic zones. This inhibition starts as soon as nitrification is accomplished within the aerobic reactor, but due to limited concentrations of dissolved oxygen and low concentrations of readily biodegradable substrate, partial denitrification occurs within the aerobic reactor. Nitric oxide, a denitrification intermediate generated within the reactor and accumulated by the heterotrophic facultative aerobic floc-forming microorganisms, inhibits some of the enzymes which are responsible for oxygen utilisation and growth of these bacteria during aerobic conditions. Low F/M filamentous bacteria are not able to reduce nitrite, hence their enzymes, which are responsible for aerobic growth, are not inhibited by nitric oxide. Musvoto *et al*. (1999) experimentally reported that in N and N&P removal systems, where the sludge is subjected to sequential anoxic-aerobic conditions, concentrations greater than about 2.0 mg NO_3^--N/L or 1.0 mgNO_2^--N/L within the anoxic zone effluent can favour low F/M filamentous microorganisms. In the specific case of *M. parvicella*, though, it has been reported (Tsai *et al*., 2003) that this microorganism prefers ammonia (during aerobic conditions) to nitrate or nitrite (during anoxic conditions) as a N source because partial denitrification provides the filament less energy than the oxidation of ammonia.

- Dissolved Oxygen Concentration (DO). The growth of certain filamentous bacteria (*e.g.*, *S. natans*, type 1701 or *H. hydrossis*), is favoured by relatively low dissolved oxygen levels in the aeration tank. It has to be noticed though that the required aeration basin DO concentration to prevent “low DO” bulking is not a constant, but a function of the F/M rate (Palm *et al*., 1980). The higher the F/M ratio, the higher DO is required. In general, it is stated that aeration tanks should be operated with a minimum of 2.0 mg O_2/L, to avoid a predominance of these specific filamentous microorganisms (Chudoba, 1985) and to allow the correct diffusion of oxygen into the floc interiors. This desired DO concentration should be maintained at the point of greatest oxygen demand, for example, at the head of a plug-flow system. It has been demonstrated that in plug-flow systems, bulking only occurs with F/M ratios higher than 0.5 kg BOD/kg MLVSS,d (ATV Working Group, 1989). Some industrial waste systems and high rate domestic plants operating at higher F/M may therefore need DO values higher than 2.0 mg/L, due to oxygen diffusion limitations.

  Although the average dissolved oxygen level in a plant may appear to be sufficient to avoid bulking, the sludge may cycle through short or longer periods of low oxygen by, for example, overshooting set-points on aeration control equipment, periodic high loads to the plant, dead spots in the aeration tank or extended periods in the unaerated clarifier. On the other hand, it has been found (Lakay *et al*., 1999) that filaments growing at low F/M ratios such as *M. parvicella*, *H. hydrossis*, and types 0092, 0041 and 0675 proliferate whenever anoxic and aerobic periods are alternated within the reactor.

- Aeration Basin Configuration. The growth of many, but not all, filamentous organisms in activated sludge (*e.g.*, type 021N, *Thiothrix* spp., *S. natans*, *N. limicola*, Type 1701, *H. hydrossis*, Type 1851 or *Nocardia* spp.) is generally encouraged by the use of uniformly aerated, completely mixed, continuously fed aeration basins. Low F/M filamentous bacteria are specially favoured by alternating anoxic-aerobic conditions (Musvoto *et al*., 1999). On the other hand, when the aeration basin is compartmentalized to include an initial high F/M feed zone (acting as a “selector”), and especially when this feed zone is anoxic or anaerobic, the growth of these filamentous organisms can be suppressed (see 1.5.2 Control of solids separation problems). The same effect is seen when a true feed-batch mode of operation, the sequencing batch reactor (SBR), is utilized with a non-aerated feed and an initial reaction period (Jenkins *et al*., 2003). Finally, other configurations that can favour floc formation are: WWTPs that do not include primary settlement, due to the high presence of settleable particles entering the reactor which facilitates the formation of flocs with high density, or plants containing Bio-P bacteria (where P removal is performed) which also increase the weight of the flocs (Eikelboom *et al*., 1998).

Apart from the different operational conditions (MCRT, F/M, DO and aeration basin configuration), usually manipulated by plant operators, that can favour the growth of filamentous bacteria, there exist a wide group of causes related to wastewater characteristics that can also lead to the development of filamentous bulking. Since "real" activated sludge plants do not operate under steady state conditions but usually under constantly changing volumetric and organic loading rates, changes in wastewater characteristics are the norm in most municipal WWTPs. When the plant is treating industrial effluent, these fluctuations are even more pronounced. This instability in the process can mask the cause of filamentous bulking and may also contribute to the problem. The causes of filamentous bulking that are related to **wastewater characteristics** are the following:

- Nutrient Deficiency. Microorganisms require six main macronutrients: C, O, H, N, S and P mostly for the synthesis of carbohydrates, lipids, proteins and nucleic acids. Na, Ca, Mg and Fe are also needed by microorganisms to accomplish different functions such as transport or stabilisation of membranes so they remain as cations within the cell (Prescott *et al.*, 1990). Among these ten macronutrients, nitrogen and phosphorus are usually the growth limiting nutrients when they are not present in sufficient amounts in influent wastewater. This is a common problem with industrial wastes due to its large content of soluble BOD that degrade quickly and demand relatively large quantities of nutrients, but it is generally not a problem in plants treating domestic or municipal wastes. In general, a C(BOD$_5$)/N/P ratio of 100/5/1 is required in wastewater for complete BOD removal (Grau, 1991). Nutrient deficiency can also be controlled by checking that at least 1.0 mg/L total inorganic nitrogen (TIN = ammonia + nitrite + nitrate) and 0.5 - 1.0 mg/L ortho-phosphate remains in the effluent at all times (depending on the effluent permissions of each country). The remaining macronutrients are usually not deficient in wastewater because they are required in very low concentrations (generally < 1 mg/L).

  The growth of *Thiothrix* spp. and type 021N may be associated with nitrogen deficiencies while *S. natans*, *H. hydrossis* or *N. limicola* III may be related to phosphorus deficiencies. Other nutrients such as iron or sulphur have been reported as limiting to activated sludge, but this is not common. The growth of some filamentous bacteria under nutrient deficiencies is correlated again with the theory used to explain the growth of the low-F/M filaments. During nutrient deficiency, filamentous bacteria may require a smaller quantity of nutrients than floc-forming bacteria due to their low K_s. On the other hand, their longer surface area also allows them to obtain nutrients in adequate amounts when their concentration becomes limited in the bulk solution.

- pH. The optimum pH in the aeration tank is between 7 and 7.5. PH values of < 6.0 may favour the growth of fungi and cause filamentous bulking (Jenkins *et al.*, 2003). Although some nitrifying activated sludge systems treating low alkalinity wastewater can exhibit punctual low pH conditions, fungal bulking has never been observed in these circumstances but in facilities where the low pH is already present within the incoming wastewater. Fungi may be observed occasionally in the mixed liquor of coupled fixed-film/activated sludge systems, especially if there is no intermediate clarifier between the two systems (Jenkins *et al.*, 2003).

- Sulphide Concentration (Septicity). Influent wastewater septicity is usually indicated by odours (H_2S or the "rotten egg" smell) and a dark colour imparted to the wastewater by precipitated ferric sulphide. Sulphides may reach the aeration basins from the influent, meaning that they have been already originated in sewers due to long retention time, or they can be produced in the aeration basin itself by sulphate reduction, when the concentration of dissolved oxygen is insufficient and a fraction of the sludge flocs permanently becomes anaerobic. Several filamentous organisms can utilize sulphide as a source of energy, oxidizing it to sulphur and then depositing the sulphur as intracellular granules. The sulphur granules can be seen within the cell under the microscope. These are *Thiothrix* spp.*,* Type 021N, *Beggiatoa* spp., and Type 0914. Some of these organisms (*Thiothrix* spp. and Type 021N) use both sulphide (autotrophically oxidized to sulphates under oxic conditions) and low molecular weight organic acids (e.g., acetic acid) as an advantageous source of energy. Thus, their growth is strongly encouraged by septicity. In general, a sulphide concentration > 1-2 mg/L and an organic acid concentration > 100 mg/L can favour filament growth (Richard, 1989 and Echeverria *et al.*, 1992).

- Nature of organic substrate. The nature of the organic substrate, including whether it is soluble or particulate or slowly biodegradable can also be related to the development of filamentous bacteria. Readily biodegradable organic substrates such as low molecular weight fatty acids and simple sugars composed of glucose, maltose, and lactose, support the growth of some filamentous bacteria (Chudoba, 1985). *S. Natans*, *Thiothrix spp.*, *H. Hydrossis*, *N. Nimicola* II, *N. Nimicola* III, Type 021N, Type 1701 and Type 1851, appear to be favoured by simple sugars while Type 021N, Type 0411, *Thiothrix spp.*, Type 0914, *N. Nimicola* II and *N. Nimicola* III grow on wastewaters containing volatile fatty acids. On the other hand, *M. parvicella*, Type 0041, Type 0675 and Type 0092, are apparently able to grow on slowly biodegradable substrates (Jenkins *et al.*, 2003).

- Temperature. It can be generally stated that filamentous organisms (like most microorganisms in activated sludge) grow more rapidly as the temperature increases within the typical range in activated sludge aeration basins (8-25 ºC). *M. parvicella*, however, is an exception since it has been demonstrated (Knoop and Kunst, 1998) that this microorganism seems to proliferate at low temperatures (usually below 12-15 ºC). This can be explained by the lower solubility of fat at lower temperature, which makes this material more available to *M. parvicella* than to less competitive floc-forming bacteria (Richard, 1989; Wanner, 1994 and Eikelboom *et al.*, 1998).

Table 1.3 correlates the most frequent filamentous bacteria and their related possible causes.

Table 1.3 Summary of conditions associated with filamentous organisms growth in activated sludge systems (adapted from Jenkins *et al.*, 2003)

POSSIBLE CAUSES	FILAMENTOUS ORGANISM
Low DO concentration	Type 1701, *S. natans*, *H. hydrossis*
Low F/M (high MCRT)	*M. parvicella*, *H. hydrossis*, Type 021N, Type 0041, Type 0675, Type 0092, Type 0581, Type 0961, Type 0803 and Type 1851
Septicity	*Thiothrix spp.*, *Beggiatoa spp.* and Type 021N
Nutrient deficiency	*Thiothrix spp.*, *S. natans*, Type 021N, Type 0041, Type 0675, *N. limicola* III and *H. hydrossis*
Readily biodegradable substrate	*Thiothrix spp.*, *S. natans*, Type 021N, Type 0914, *N. limicola* II, *N. limicola* III and *H. hydrossis*
Slowly biodegradable substrate	*M. parvicella*, Type 0041, Type 0092 and Type 0675
Low pH	Fungi

1.4.2 Biological Foaming

Apart from filamentous bulking, activated sludge systems suffer from another common solids separation problem: Biological Foaming, a variant of filamentous bulking where the overabundance of some specific species of filamentous bacteria results in the occurrence of foams in the surface of both the biological reactor and the secondary settler. Although other kinds of foaming can occur, such as those caused by biodegradable or non-biodegradable surfactants (typical of plant start-up), the type of foam discussed here is specifically caused by filamentous bacteria.

The surveys of filamentous microorganisms in biological foams indicate that the diversity of filaments responsible for foaming problems is less than in the case of filamentous bulking. In general, three filamentous organisms (or groups) have been identified as the typical filaments causing activated sludge foaming: (1) nocardioforms (general term given to the group of actinomycetes similar to *Nocardia* sp.),; (2) *Microthrix parvicella* (the most common) and (3) Type 1863 (less common) (Jenkins *et al.*, 2003 and Richard, 1989). These three types of filamentous organisms can be distinguished from one another through microscopic examination. The common characteristic that makes these microorganisms to cause foam is the hydrophobic (water repellent), waxy nature of their cell wall, which in contact with the air bubbles of aeration basins make the flocs to float. On the other hand, foam-forming filamentous microorganisms also produce extracellular materials like lipids, lipopeptides, proteins and carbohydrates which have the properties of surface active agents (biosurfactants), which enable the hydrophobic cells of foam-formers to be frothed up to form a floating foam (Wanner, 1994). The result is the production of stable, viscous foam on the biological reactor surface that can shift over onto the clarifier surface (Figure 1.10). The colour of the foam can vary from greyish-white in foams produced by Type 1863 to brown or dark-brown whenever nocardioforms or *M. parvicella* are the causative filament. In many cases foaming due to *M. parvicella* or nocardioforms results in bulking by the same species (Eikelboom, 1994), so it is through these two microoganims that the two phenomena are often linked.

Apart from these three filamentous bacteria commonly found in foams, several surveys (Seviour *et al.*, 1994, Blackbeard *et al.*, 1986, Pujol *et al.*, 1991 and Eikelboom, 1991) have shown that some other species such as Type 0041, Type 0675, Type 0092 or *Nostocoida limicola* are usually observed in foams and thus their potential role as foam-forming bacteria need further research.

Figure 1.10 Example of foams in activated sludge systems. (a) Foam caused by nocardioforms on the aeration tank, (b) foam caused by *M. parvicella* on the aeration tank and (c) on clarifiers

Foaming can range from being a nuisance to being a serious problem. The typical consequences include:

⇒ The foam concentrated on the surface of clarifiers can escape with the effluent, affecting water quality (increasing both TSS and BOD_5 concentration) and violating permit limits.

⇒ If the foam layer is deep it can overflow onto walkways and surrounding areas creating hazardous slippery areas.

⇒ Foam can cause severe problems on the mechanisms of superficial aeration where, due to the reduction of oxygen transfer efficiency, the electricity consumption increases.

⇒ The foam layer creates a microhabitat with some specific characteristics that are difficult to control by operators. For example, the sludge age within the foam is much higher than the sludge age within the mixed liquor. On the other hand, since the foam may contain up to 40% of the total solids in the plant the operator may also have difficulties in keeping track of concentration of solids in aeration basins (Richard, 1989). For example, if the amount of waste activated sludge is kept the same as before foaming, a significant decrease of the biomass retention time in the mixed liquor may occur, which is especially dangerous in nutrient-removal plants.

⇒ In cold weather foam can freeze and in warm weather it can dry up making their removal a difficult task. During warmer climates foam can also putrefy rapidly and become odorous.

⇒ Since some organisms found in foam are opportunistic pathogens, the aerosols derived from foam-producing organisms are considered a potential health hazard (Blackall *et al.*, 1988).

⇒ In case foaming sludge fed into anaerobic digesters, it may also cause them to foam causing blockage of gas-mixing devices, inversion of solids profiles, gas binding of sludge recirculation pumps, etc. (van Niekerk *et al.*, 1987).

The operational troubles resulting from foam formation in activated sludge plants were first observed and reported at the Jones Island East WWTP in Milwaukee, USA, in 1969. The phenomenon was called the “Milwaukee mystery”. From then until recent years, biological foaming has remained mysterious due to the lack of knowledge concerning their possible causes and the consequent lack of control methods to specifically face these situations. Nowadays, the problem is widespread throughout the world and, although biological foaming is mostly reported in nutrient removal plants, conventional aerobic activated sludge systems are not resistant to foaming either. In South Africa, for example, a survey done in the mid-80s revealed that about 40% of activated sludge plants suffered from biological foaming (Blackbeard *et al.*, 1986). In the United States, 66% of surveyed plants experienced some kind of foaming at one time or another (Pitt and Jenkins, 1990). In Australia (Seviour *et al.*, 1994), 68% of the surveyed plants had foaming problems and finally, in France almost 20% of all the studied plants and 87% of the extended aeration plants also presented foaming problems at one time or another (Pujol *et al.*, 1991).

The results of all these surveys and some other studies based on filamentous microorganisms’ predominance carried out around the world (see results of Table 1.1) showed that nocardioforms and *M. parvicella* were some of the most commonly observed filamentous organisms in activated sludge systems. On the other hand, the results reported from all these surveys also show that at present there exists no reliable active tool for efficiently preventing and controlling biological foaming.

1.4.2.1 Causes of biological foaming

Apart from the hydrophobic nature of the filament surface and the presence of gas bubbles (usually generated by aeration), the existence of fats, greases, oils or surfactants in the influent is one of the most common causes that enhance the amount of foam produced by an activated sludge system. The presence of these substances increases hydrophobity and favours foam stability (Jenkins *et al.*, 2003). In general, then, systems that lack primary clarification (where most greases and oils are removed) appear to suffer more foaming problems. On the other hand, surface trapping of the scums in activated sludge systems also contributes to foaming development. Trapping can either be caused by physical barriers such as surface baffles, or by hydrodynamic effects such as the rotational momentum in carrousel plants and the surface pump-back action of surface aerators (Pitman, 1996).

In spite of these general causes, there is still much work to do to find the potential causes that promote each one of the foam-forming species. Studies based on filamentous bacteria physiology have found that the occurrence of foam-forming bacteria can be related to some of the following causes:

- **Nocardioforms**. *Nocardia* growth is associated with warmer temperatures, grease, oil and fat present in treated wastes (which increase the hydrophobic nature of cells). Although its presence is favoured by long MCRTs (generally greater than 9 days), once *Nocardia* is present in the activated sludge system it can develop at MCRTs as low as 2 days. Concerning the substrate, nocardioforms can use a wide variety of organic substrates as carbon sources. These include readily degradable compounds such as sugars and low molecular weight fatty acids and slowly biodegradable, high molecular weight compounds such as polysaccharides, proteins, pesticides, and aromatic compounds (Lemmer and Kroppenstedt, 1984). On the other hand, results of different studies (Richard, 1989 and Jenkins *et al.*, 2003) have demonstrated that this filament can grow within a wide range of F/M ratios, from 0.08 to 0.7 kgBOD/kg MLVSS, d so it can not be correlated with low or high F/M conditions. In general, foam production can occur with both diffused and mechanical aeration but it is more pronounced with diffused aeration and with higher air flow rates.

  In general, it can be concluded that plants prone to *Nocardia* foaming often: (1) receive oil and grease wastes, (2) have poor or no primary scum removal, and (3) recycle scum rather than remove it from the plant.

- ***M. parvicella*.** This is a common filamentous bacteria that can cause both filamentous bulking and foaming. It is mostly found in nutrient removal systems since its growth is favoured by low F/M ratios (in general, F/M ratios ≤ 0.1 kg BOD/kg MLVSS,d (Knoop and Kunst, 1998)), or long sludge ages (greater than 8 days). As occurs in episodes of bulking, *M. parvicella* can also outcompete floc-forming bacteria whenever high nitrate and/or nitrite concentrations are present in the anoxic reactor effluent (see 1.5.1.2. Filamentous bulking causes). Apart from these specific causes, the proliferation of these bacteria is also influenced by low temperatures (usually < 12 ºC) (Knoop and Kunst, 1998), probably due to the reduction of lipids solubility which become more available for this microorganism. Hence, *M. parvicella* is mostly found during colder seasons such as winter and spring. It has been also observed that its optimum pH range is 7.7-8.0. Regarding DO concentration, it has been demonstrated (Ekama *et al.*, 1995), that although *M. parvicella* can proliferate over a wide range of oxygen partial pressures including aerobic, anoxic or anaerobic conditions, intermittent aeration could be defined as the condition that most favours its development. Finally, in general *M. parvicella* can not use common readily biodegradable substrates, preferring long chain acids in esterified forms such as oleic acid. The presence of septic conditions, which favour the breakdown of grease and fat to organic acids, or the presence of mechanical stresses, which can also release long chain fatty acids from destroyed cell membranes, are other examples of factors that can favour *M. parvicella*.

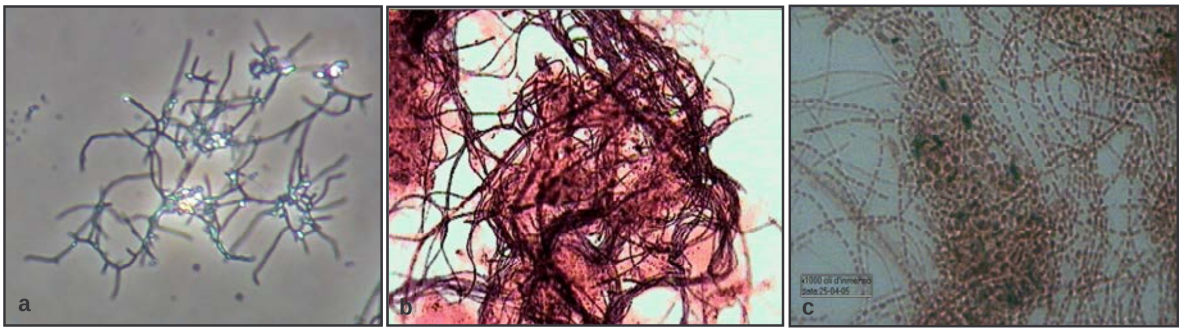


Figure 1.11 Phase contrast micrographs of different foam-forming filaments: (a) Nocardioforms (magnification x1000) ; (b) *M.parvicella* (Gram positive stained) and (c) Type 1863 (Gram negative stained) (magnification x1000)

- **Type 1863.** Although Type 1863 is not very common in activated sludge systems and as a consequence it has not been deeply studied, its proliferation is commonly related to high F/M ratios in plants working at low sludge ages (generally < 5 days). Related also to high F/M presence, its growth is favoured by very low DO concentrations (usually < 1.0 mg 0_2/L).

1.4.3 Non-Filamentous Bulking

Apart from filamentous bulking, there exists another microbiological problem in activated sludge processes that can transform suspended solids into a bulky solution hindering its correct settleability in clarifiers. This problem is known as non-filamentous bulking, viscous bulking, slime bulking or zoogloeal bulking and it is caused by an excessive production of extracellular polymers (glycocalyx) that may or may not be associated with the growth of a group of floc-forming bacteria: *Zoogloea* spp. (Jenkins *et al.*, 2003). The production of biopolymers is characteristic of most floc-forming microorganism, but under normal conditions the amount of byopolimers generated is just enough for the formation of firm flocs. Bacteria from genus *Zoogloea* spp. are accustomed to slime colonies (see Figure 1.2a), which is why the term zoogloeal bulking is often used to describe the settling problems caused by an excessive presence of EPS (Wanner, 1994).

Even though zoogloeal microorganisms constitute the major part of well-settling sludge flocs, their excessive growth can cause an increase of slime production that may affect the compaction and settleability of the sludge in several ways. On the one hand the produced slime is insoluble in water and less dense than water. Second, the dispersed and flocculent microbial cells are surrounded by large amounts of water-retentive polymers which produce a viscous sludge that is difficult to settle and to become compact. On the other hand, the slime stored outside the bacterial cells can entrap air bubbles and gases contributing to a more buoyant sludge and, under mixing action or aeration, the production and accumulation of billowy white foam on the surface of the aeration tank (Gerardi, 2002). As in filamentous bulking, when sludge settleability becomes poor an overflow of sludge blanket may occur, increasing suspended solids and BOD within the effluent and diluting return and waste flows.

The excessive presence of viscous biopolymers in the sludge can easily be detected microscopically with live samples reverse stained with India ink (Jenkins *et al.*, 2003). The principle of this reverse staining is that since the carbon microparticles from the India ink can not penetrate in the viscous polymer, whenever large amount of EPS are present in the sample most of the floc does not get stained.

1.4.3.1 Causes of non-filamentous bulking

Since the occurrence of non-filamentous bulking is not as common as filamentous bulking or foaming, the related references to the study of its main causes and possible control methods are not frequent in the literature. In general though, the most commonly referred causes are:

- Nutrient deficiency. During nutrient deficiency (specially N and P deficiency), the bacteria within the floc particles can not degrade some of the soluble COD and this is stored within the floc particles as an insoluble polysaccharide or slime that can be solubilized and degraded later when nutrients become available to the bacteria cells. As in filamentous bulking, the concentration of nutrients required to ensure complete BOD removal corresponds to a C(BOD_5)/N/P ratio of 100/5/1 (Grau, 1991).

- High F/M and high RBCOD. Highly loaded plants, particularly those with significant amounts of simple or easily degraded carbohydrates, volatile fatty acids or in general readily biodegradable COD in the wastewater are prone to viscous bulking. It is common thus to find zoogloeal bulking in those activated sludge systems designed to improve settling properties by using a concentration gradient (e.g., activated sludge systems with selectors or plug-flow systems) due to the high F/M ratio obtained within the first compartments. As when nutrient deficiency occurs, the origin is the big amounts of soluble COD that can not be degraded and is stored within the floc.

1.4.4 Pin-point Floc

Pin-point floc can be defined as a solid separation problem originated when, during floc formation, only floc-forming bacteria are present (Figure 1.12a). The absence of filaments hinders the formation of macrostructure and consequently activated sludge flocs come to be small (up to 75 µm), spherical and compact. This type of sludge settles rapidly producing a low SVI and a cloudy or turbid supernatant (Figure 1.12b). The larger, more compact flocs settle rapidly; the smaller aggregates that may have been sheared off from larger flocs settle slowly and create a turbid supernatant. This loss of solids, generally known as deflocculation, apart from decreasing sludge dewaterability, induces an increase in TSS and COD in the effluent. Even small amounts of released particles have great implications for the effluent quality. For example, if only 2% of the flocs deflocculated, the suspended solids concentration in the supernatant would be 80 mg/L (Wilén *et al.*, 2000), which is far beyond the effluent limit in most countries (35 mg/L in Catalonia).

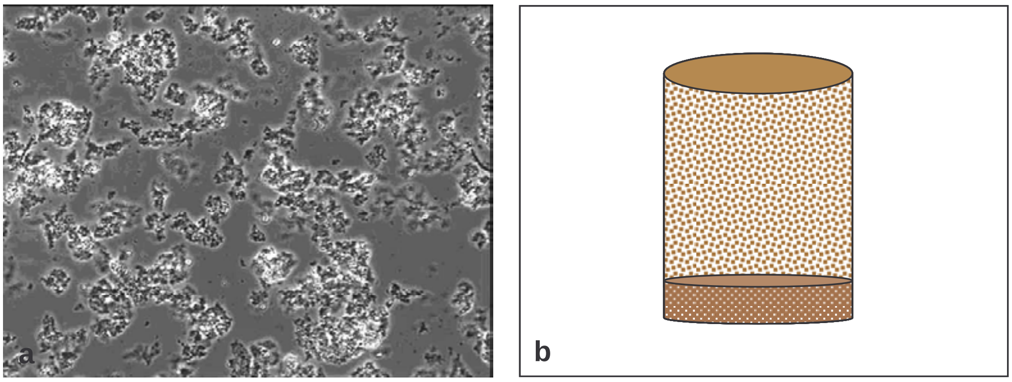


Figure 1.12 Pin-point floc, (a) microscopic appearance of flocs; (b) drawing representing a V30 test of pin-point floc

1.4.4.1 Causes of pin-point floc

Factors that can interrupt floc formation and cause pin-point flocs can be listed as follows:

- Extremely high values for sludge age (overoxidised sludge). High values of sludge age imply an overoxidation of sludge and an exposition of flocs to low concentrations of exogenous substrates (low F/M ratio). The extracellular polymeric substances serve, under these conditions, as carbon and energy sources (endogenous metabolism), which leads to the destruction of the polymeric matrices of activated sludge flocs (Wanner, 1994). Commonly, this overoxidised sludge produces floating ash-like particles on the surface of the secondary settler, generally referred to as *ashing*. The ash may be composed of particles of dead cells, as well as normal sludge particles or greases. The difference with rising sludge is that these floating flocs do not show gas bubbles.

- Low sludge age. Activated sludge with a very low sludge age (< 3days) can also lead to pin-point floc whenever the SRT is too low to allow filaments to develop and conform the macrostructure of flocs. At low sludge age, many weak and buoyant particles are produced, which are easily sheared and often float out of the secondary clarifier. In general, a low sludge age usually occurs during the start-up of an activated sludge process. There are other operational conditions though, that can mimic a low sludge age, including excess sludge wasting, hydraulic washout, recovery from toxicity and a slug discharge of soluble BOD (Gerardi, 2002).

- Excessive shearing. Excessive turbulence, usually caused by overaeration, in the biological reactor may hinder floc formation and result in pin-point floc being carried over with the clarifier effluent. Common sources of excessive turbulence are surface aerators, high rates of coarse aeration, high speed pumps, and high rates of RAS. These sources of excessive turbulence place considerable physical stress on the floc particles and the community of stalk ciliated protozoa. Regarding the aeration system, all kinds of aerators causing excessive turbulence such as turbines or surface aerators should be avoided. If the plant uses fine-bubble diffusers or pure oxygen, high DO concentrations can be safely maintained. Otherwise, DO values > 4.0 ppm could damage floc structure.

- Presence of toxics or surfactants. The presence of toxics within the system such as heavy metals, or overdoses of oxidants such as chlorine can interrupt floc formation and originate deflocculation.

The existence of surfactants, present in soaps or detergents, may also result in deflocculation. When surfactants come in contact with particles, they act as dispersing agents and the floc particles become weak, releasing fine solids to the bulk solution.

- Oxygen limitation. Oxygen deficiency or anaerobic conditions can also be a cause of deflocculation. It has been demonstrated (Eikelboom and van Buijsen, 1983 and Starkey and Karr, 1984) that during anaerobic conditions the growth of aerobic floc-forming microorganisms and the production of EPS can be inhibited. Nevetheless, the same conditions can sometimes cause quite opposite troubles connected with an enormous production of extracellular polymers or an overabundance of some filamentous bacteria, so this cause, like most of the causes related to floc formation or deflocculation, need further research.

1.4.5 Dispersed Growth

Activated sludges suffering from dispersed growth are composed of only small flocs due to the absence or disruption of the exopolymer bridging, which hinders the ability of microorganisms to stick to each other (Figure 1.13a). Bacteria are suspended in the liquid portion of the mixed liquor as individual cells or small clumps with a diameter of 10-20 μm (Wanner, 1994), and although they are still growing rapidly they have not begun to flocculate (Figure 1.13b). Since particles density is too low, gravity settleability is difficult to occur and no zone settling develops in clarifiers. This has two effects on the activated sludge process: (1) the separation efficiency of secondary settlers is very low, and the final effluent is turbid (high TSS and high COD) and (2) because of poor sludge compaction, the concentration on the return sludge is very poor and consequently the control of MLSS concentration is difficult to achieve.

In general, dispersed growth is not a very typical problem for conventional or nutrient-removal activated sludge plants, given that it is usually observed at extremely low values of sludge age, SRT=1-3 days.

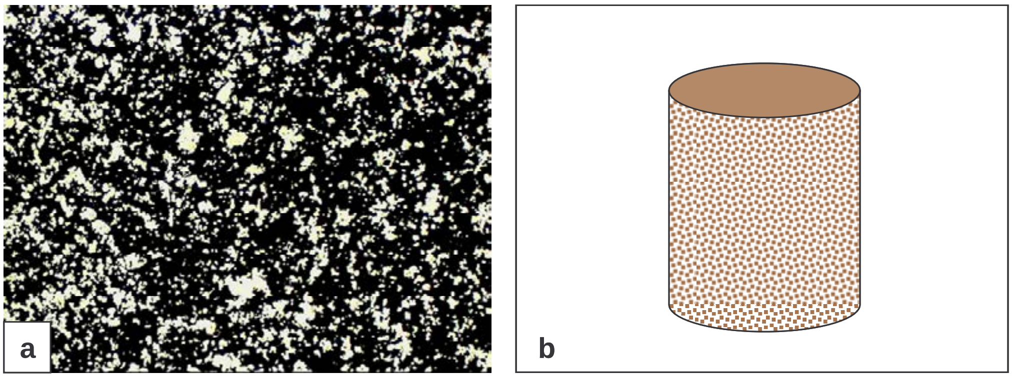


Figure 1.13 Dispersed growth, (a) microscopic appearance of flocs; (b) drawing representing a V30 test with dispersed growth

1.4.5.1 Causes of dispersed growth

There are many (but poorly understood) causes that can hinder the production of exopolymers and result in dispersed growth. Most of these factors also cause pin-point floc and are the following:

- Low sludge age. A significant amount of dispersed growth is present at the start-up of an activated sludge process (SRT < 3 days). At low sludge age, a lot of food is available and, although bacteria are very active and multiply rapidly, the SRT is not high enough to form bigger aggregates. The microbial structures contain considerable bound water, resulting in a low specific gravity.

- Presence of toxics or surfactants. The presence of toxics within the system such as heavy metals and cyanide, or overdoses of oxidants such as chlorine, can interrupt floc formation and develop dispersed growth due to deflocculation. The existence of surfactants present in soaps or detergents may also result in dispersed growth.

- High concentration of readily degradable substrates. Under high organic loadings, especially during high RBCOD concentrations, the cells of organotrophic floc-forming bacteria are not metabolically forced to synthesize EPS as storage products so their production may be reduced.

- Excessive shearing. Excessive turbulence (overaeration) in the biological reactor caused by aggressive aeration or excessive turbulence may also hinder floc formation and result in dispersed growth being carried over with the clarifier effluent.

- High monovalent cations to divalent cations ratio. Floc particles contain billions of bacteria and EPS that possess a net anionic charge. High concentrations of monovalent cations (e.g., K^+ and Na^+) relative to the concentrations of divalent cations (e.g. Ca^{2+} and Mg^{2+}) can affect floc formation (Sanin and Vesilind, 2000). The argument is that if less positive charges are present, then the whole negative charge of flocs will not be neutralized and flocs will have difficulties to settle. On the other hand, changes in surface charges can also be induced by pH variations.

1.4.6 Rising sludge

Rising sludge is the last setback occurring during the activated sludge process considered as a solids separation problem. In this case, although the origin is not an imbalance between the floc-forming and the filamentous microorganisms, the source is also correlated with some groups of the microfauna, and the consequence on the process is still a loss of solids through the effluent.

The phenomenon of rising sludge, also known as “blanket rising” or “clumping” was firstly reported in the literature in the early 1940s (Sawyer and Bradney, 1945) can be defined as a process of denitrification which takes place within the secondary settler, instead of the reactor. During denitrification, nitrites and nitrates present in the wastewater are converted to nitrogen gas. Therefore, for denitrification to take place, nitrification must have occurred first within the aerobic reactor, unless industrial wastes are contributing nitrates. Apart from this requirement, the following conditions must also be present within the sludge blanket to cause denitrification:

1. High nitrate concentration (typically more than 5 mg/L)
2. A source of organic matter or residual BOD_5 (typically more than 10 mg/L)

3. Low DO concentration (less than 0.5 mg/L) within the secondary settler. This can be caused by a slow removal of sludge from the clarifier.
4. Presence of denitrifying bacteria, which use nitrate ions (NO_3^-) or nitrite ions (NO_2^-) to degrade the soluble BOD present in clarifiers.

As nitrogen gas is formed in the sludge layer due to denitrification, much of it is trapped in the sludge mass. Apart from N_2, N_2O and CO_2 are also produced during denitrification and can also contribute to rising sludge. If enough gas is formed, the sludge mass becomes buoyant and rises or floats to the surface, resulting in poor compaction of solids, loss of settleability and solids washout through the effluent (with the corresponding increase both in TSS and COD concentrations). Rising sludge can be differentiated from bulking sludge by noting the presence of small gas bubbles attached to the floating solids (Figure 1.14a). During the settleability test of V30, the presence of bubbles can also be observed. In general, during the first 30 minutes the sludge shows good settling characteristics, but from 30 to 60 minutes after the start of the settleability test, bubbles of gas may be found rising to the surface of the cylinder or attached to the surface of the rising solids (Figure 1.14b). Often, when the rising solids reach the surface they break apart and numerous bubbles of gas are released.

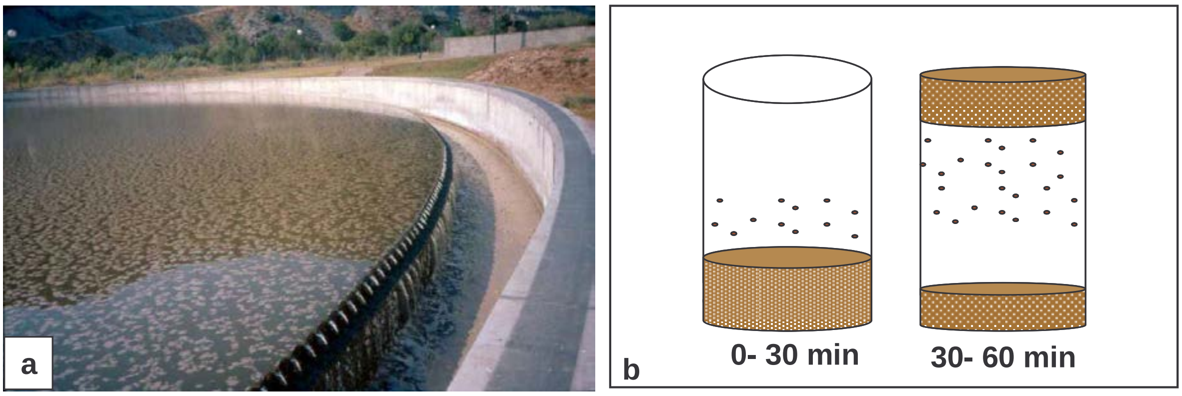


Figure 1.14 Rising sludge, (a) sludge rising on clarifiers; (b) drawing representing a V30 test with rising sludge

1.4.6.1 Causes of rising sludge

In order to determine the possible causes of rising sludge, it is useful to distinguish between WWTPs that usually perform complete nitrification and denitrification, and plants that are not required to nitrify. In plants that usually remove nitrogen, rising sludge can be originated by an incomplete denitrification in biological reactors due to some problem during the anoxic stage: insufficient anoxic retention time, insufficient organic matter, etc. On the other hand, WWTPs that are not required to nitrify can nonetheless undergo periods of nitrification induced by favourable conditions such as warmer temperatures, long sludge ages, high oxygen concentrations, etc. Whenever desired or undesired nitrification occurs but denitrification is not completed within the reactor itself, there exist some causes that can lead to undesired denitrification in clarifiers and consequently cause rising sludge:

- Long SRT in clarifiers. When the sludge is being held too long in the clarifier, two main consequences which favour denitrification can occur. First of all, long sludge ages facilitate the growth of denitrifying bacteria, which require quite a long time to develop. Secondly, due to the high SRT, the DO present in clarifiers is quickly consumed by microorganisms and the environment becomes anoxic.

- Warmer temperatures. When there is a higher wastewater temperature than normal, it results in a higher rate of microorganism activity that causes the process to nitrify at a higher F/M ratio. A higher rate of microorganism activity will also more quickly deplete the DO in the settled sludge and, consequently, poses a greater potential for denitrification.

- Presence of organic matter. Together with high SRT and anoxic conditions, the presence of some organic matter in the influent of the clarifier is also required for denitrification to occur.

1.5 Monitoring and Control of Activated Sludge Solids Separation Problems: State-of-the-art

The present section reviews the methods currently available to monitor and control the occurrence of solids separation problems in activated sludge systems.

1.5.1 Monitoring of solids separation problems

In the process of monitoring activated sludge solids separation problems, several parameters can be used as symptoms of setback in sludge settleability, or even as determinants of the possible causes that may lead to these undesired situations. Most of these parameters are usually obtained by means of laboratory analysis because, although the number of sensors is increasing everyday and more and more parameters can be obtained on-line (dissolved oxygen, pH, suspended solids, nutrients, etc.), most facilities still resort to laboratory analysis to get most of the data, mainly because of the high cost of using such sophisticated techniques. Nevertheless, laboratory analysis are not immediate, meaning that plant operators must wait for long periods of time, varying from few minutes to several days, to obtain the results that might help them to characterize the process.

1.5.1.1 Sludge settleability

A variety of laboratory and macroscopic tests exist to help plant operators to determine the settling characteristics of secondary solids and to obtain basic information on the extent of solids separation problems. Unfortunately, most current methods do not provide early warnings of problematic situations. The most used tests are the following:

- **30-minute settling test (V30).** The V30 or mixed liquor settleability test is one of the most useful tools for understanding what is happening to the sludge settleability. It can be defined as the volume occupied by the sludge when one litre of sample from a reactor is allowed to settle for 30 minutes. Although this parameter depends on the suspended solids, the major benefit of the test is that the results, which help to demonstrate how mixed liquor solids may behave in secondary clarifiers, are available in 30 minutes.

It needs to be mentioned, however, that instead of using a settleometer (a specially marked wide-mouth beaker) to perform the V30 test, some operators use a 1-litre graduated cylinder, which decreases the accuracy of the test due to the sidewall effects of the narrow graduate which can interfere with the settling (WEF, 2002).

- **Sludge Volume Index (SVI).** The sludge volume index is the most commonly used test for assessing the settleability of activated sludge. It was described by Mohlman (1934), hence it is sometimes termed as the Mohlman Index. SVI is described as the volume (in mL) occupied by 1 g of settled sludge in a 1-litre unstirred measuring cylinder after 30 minutes of settling. In general, it is considered that the SVI must not surpass the value of 150 mL/g to ensure proper settleability. A high SVI (>150 mL/g) indicates bulking conditions, whereas a SVI below 70 mL/g indicates the predominance of small flocs. However, it is difficult to determine the exact value at which sludge is lost to the final effluent, because the secondary settler efficiency not only depends on the SVI but also on other factors, such as the dimensions of the settling tank, the MLSS or the characteristics of the effluent. Optimum SVI therefore must be determined for each plant experimentally. The SVI is given by the following formula:

$$SVI = \frac{V_{30} \times 1000}{MLSS} = \frac{mL}{g}$$

Despite its high usefulness, this parameter has attracted severe adverse criticism as a measure for assessing sludge settleability since its value depends on the sludge concentration. Thus, a sludge with bad settleability (e.g. V30=1000 mL) but high concentration of biomass (e.g.: MLSS= 8 g/L) will have a SVI= 125 mL/g, which would incorrectly indicate a good settleability (Ekama *et al.*, 1997). The variation of SVI values with the suspended solids concentration and the sludge volume makes it difficult to compare the SVI values between different plants. Two sludges with the same SVI may have different sedimentation and dewatering properties.

In order to avoid these misinterpreted results, it has been proposed to modify the SVI test and use the Diluted Sludge Volume Index (DSVI) instead, which yields more useful information (Lee *et al.*, 1983 and Ekama and Marais, 1984).

- **DSVI.** The Diluted Sludge Volume Index (Stobbe, 1964 and Lee *et al.*, 1983) is an alternative to SVI and is similar to it, except that in this case, the mixed liquor is diluted with process effluent so that the settled volume in the graduated cylinder (V30) has a value between 150 and 250 mL. The dilution was experimentally determined by the authors in order to avoid the direct dependency with the sludge concentration. This insensitivity to the sludge concentration makes it possible to compare sludge settleability in different activated sludge plants.

Due to its higher accuracy, DSVI has gradually increased in popularity to become the method of choice (instead of SVI) in the United States, South Africa and progressively also in Europe.

- **SSVI.** The Stirred SVI or $SSVI_{3.5}$ (White, 1975) is also an alternative to SVI where the V30 is calculated for a standard initial MLSS concentration of 3.5 g/L. Slow stirring is used to minimize wall effects on the settled sludge volume. It tries to reproduce the non-ideal situation in settling tanks whereas the SVI is measured under complete quiescence.

- **Sludge Blanket Depth.** The sludge blanket is a measurement of the solids depth in the secondary clarifier. The measurement may be performed using a sonic meter, density meter or Secchi disk, and it is measured to ensure proper operation of the collector and sludge return mechanisms. It measures the depth of the settled sludge (normally in meters). A large depth (usually higher than 90 cm) indicates a poor settleability or a failure in retaining the MLSS in the secondary settlers, which can lead to a decrease in the effluent quality.

- **Zone settling velocity (ZSV).** The ZSV can be defined as the present technique which best simulates real conditions during settling in secondary settling tanks (Wanner, 1994). The zone settling velocity represents the maximum rate of sedimentation and can be expressed as the slope of the linear part of the sedimentation curve obtained from the plot of the sludge-supernatant interface against time. The sedimentation curve can be obtained by measuring the velocity of movement of a sludge-supernatant interface motion in high calibrated cylinders. As with SVI, ZSV depends on sludge concentration, so it is recommended to calculate these velocities only for the range of biomass concentrations common in the reactors of activated sludge systems: 3-5 kg/m^3.

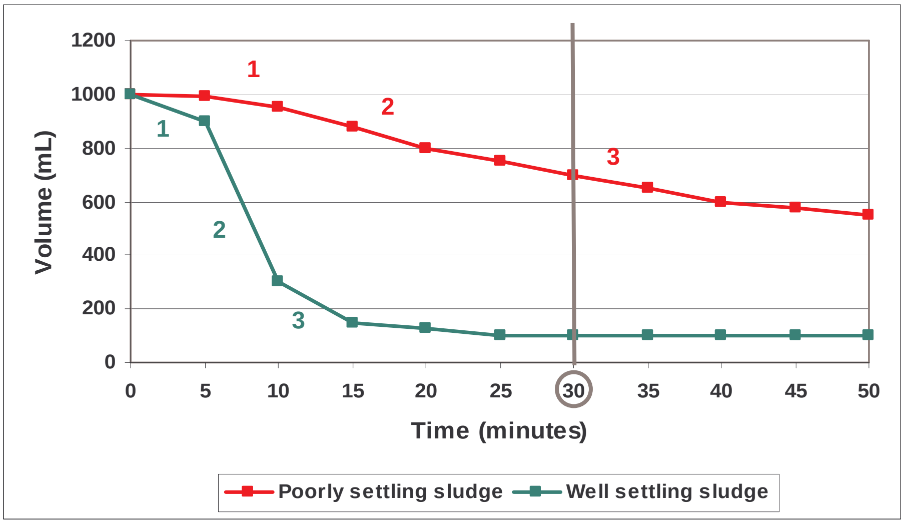


Figure 1.15 Activated sludge settling curves obtained from real data at the Girona WWTP: (red) well settling sludge and (green) poorly settling sludge; 1: reflocculation; 2: zone settling; 3: transition and compaction (adapted from Wanner, 1994)

As depicted in Figure 1.15, 4 different phases can usually be distinguished in the settling curve. Phase 1 corresponds to the **reflocculation** stage where little sedimentation is observed. In general, the higher the number of filaments present in flocs, the longer the period of time needed for aggregation and thus the longer the reflocculation phase. Phase 2 is where a sharp interface between sludge and supernatant is formed. In this phase, the settling velocity is expressed as a slope of the settling curve, which is constant and maximum. This is termed **zone settling velocity**. Finally the last phase defines a process of transition and compaction of the floc where the settling velocity starts to slow down due to floc stabilization.

There are many attempts in the literature to find a correlation between the SVI and ZSV values. For practical needs, Table 1.4 was estimated in order to classify different activated sludges (Wanner 1994).

Table 1.4 Classification of activated sludges according to SVI and ZSV values (Wanner 1994)

Type of Sludge	SVI (mL/g)	ZSV at X= 3g/L (m/h)
Well settling	< 100	> 3
Light	100-200	2-3
Bulking	> 200	< 1.2

- **On-line sensors.** In order to overcome the imprecision derived from conventional settleability tests such as SVI, which have often been criticised due to the low frequency with which it is measured (generally once a day) or because of their vague approximation to the reality of full-scale settlers (Dick and Vesilind, 1969), on-line sensors have been developed to automatically record sludge settling characteristics and some of them are commercially available. These devices allow high frequency monitoring of the settling properties by taking sludge samples about every hour (De Clerq *et al.*, 2003). Vanderhasselt *et al.* (1999), for example, developed and tested in a real plant, an on-line settlometer where the height of the sludge water interface is monitored by a moving scanner and the sedimentation curve is automatically represented. From this curve, the zone settling velocity is obtained as the maximum downward slope. This sensor was used both to pinpoint recurring operational settling problems and to identify possible improvements to the operation of secondary clarifiers, e.g. through better controlled polymer addition. Other examples include: Simon *et al.*, (2005) who developed an in-situ sludge volume probe and Vanrolleghem *et al.* (2005) who recently tested other possible on-line sensors such as an Acoustic Doppler Current Profiler for characterization of the hydrodynamics in the settler, a laser backscattering method or a non-invasive radiotracer for the study of the settling properties in clarifiers.

1.5.1.2 Foaming Tests

Although the number of plants suffering from foaming problems is quite high, there are not too many available methods to measure and evaluate these plant nuisances. The most used tests are the following:

- **Scum Index (SI).** This is a procedure developed by Pretorius and Laubscher (1987). It is calculated by means of the following equation:

$$SI(\%) = \frac{mass\ of\ biomass\ in\ foam}{total\ mass\ of\ biomass} \times 100$$

The portion of biomass in the foam is estimated by means of fractionary flotation performed with a standard aeration rate of 10 m^3/m^3,h. The flotation is repeated with the sediments after separating the scum from the settleable biomass several times until all foam-forming microorganisms are transferred to the scum. The scum index test can be used for the prediction of operational problems expected as a result of the presence of foam-forming microorganisms in activated sludge, as shown in Table 1.5.

Table 1.5 Qualitative interpretation of the results obtained by the Scum Index calculation (adapted from Pretorius and Laubscher, 1987)

SI (%)	Extent of problems
0-0.5	Negligible
0.5-6	Small
6-10	Medium
10-15	Serious
>15	Catastrophic

In Australia, Blackall *et al*. (1991) developed a variation of this foaming test where the 50 mL sample within a glass cylinder is aerated with air diffused through a sintered glass disc with an aeration intensity of 200 mL/min. The generated foam is evaluated and rated according to a scale from 0, indicative of no foam formation, up to 7, where the formed foams are classified as dense and very stable.

- **Foam coverage.** The foam coverage method was tested in England by Kocianova *et al*. (1992). In this test, the foam cover of the surface of aeration basins and secondary settling tanks is assessed visually as a percentage of surface cover. Readings should be taken at the same time each day, when loading conditions and operating procedures are similar.

- **Determination of surface hydrophobicity of solids**. This sophisticated technique was developed by Rosenberg and Doyle (1990) to measure surface hydrophobicity in activated sludge plants with common foaming problems. The test analyses the adherence of the microorganisms present in the sample to hydrocarbons. The assay involves vortexing a suspension of sludge with a known volume of non-polar hydrocarbon such as hexadecane or xylene, allowing phase separation to occur, and measuring the absorbance of the aqueous phase with spectrophotometer. The result, expressed as percent of hydrophobicity, is calculated as the increment of absorbance between the initial and the final solution.

1.5.1.3 Microscopic analysis

The evaluation of settling and foaming properties of the activated sludge can be considered as an important measurement of the settling characteristics of secondary solids. However, they only show the symptoms of solids separation problems, without revealing why the secondary solids settle well or poorly. In contrast, routine microscopic observations provide valuable monitoring information about the condition of the microbial population and floc formation within the activated sludge process. The operator does not need to be able to identify all the organisms in the activated sludge but, by using the experience gained from previous observations, make judgments as to what is normal for his/her facility. Specific information usually gathered includes:

a) Changes in floc size and density

b) Presence of protozoa and metazoa which provides qualitative information about the performance of the biological process

c) Presence and effect of filamentous organisms on floc structure

This observation of the flocs and filamentous organisms provides a measurement that reflects not just the current conditions, but to some degree the conditions that have existed over at least the previous 1 or 2 MCRTs. Changes in these characteristics can provide an indication of changes in the wastewater characteristics or the occurrence of any operational problem.

It is important that a wastewater treatment plant had a good phase microscope suitable for the identification of the main characteristics mentioned above. The microscope should have objective lenses at 100, 200 or 400 and 1000x.

a) Floc Characteristics

The common characteristics often measured in flocs are: size (classification of flocs in small: ≤ 150 µm, medium: 150-500 µm or large: ≥ 500 µm); shape (round or irregular), density (compact or diffuse); porosity; strength; surface charge; settling velocity; coherence and specific surface area. It needs to be mentioned that it is difficult to measure most of the characteristics of activated sludge flocs since they are heterogeneous and the range of differences is very broad. Difficulties are also encountered in sampling to avoid destroying or otherwise physically altering the flocs.

Although in conventional WWTPs floc characteristics are usually measured at 100x magnification, various methods have been developed and used to measure the different characteristics of activated sludge flocs:

- Direct microscopic observation with an eyepiece micrometer (Sadalgekar *et al.*, 1988; Barbusinki and Koscielnak, 1995)
- Photographs of individual flocs (Magara and Nambu, 1976; Li and Ganzarczyck, 1987)
- Image analysis, which can be improved by the use of Confocal Laser Scanning Microscopes (Li and Ganzarczyck, 1991; Andreadakis, 1993; Grijspeerdt and Verstraete, 1997; Jenné *et al.*, 2004)

b) Presence of protozoa and other microorganisms

Microscopic observation of protozoa and other higher life forms in activated sludge is a common and widespread practice. In a very general way, the types of these organisms present can be related to plant performance and effluent quality. They are useful for toxicity assessment but they are of little or no value for determining the properties of activated sludge that influence its behaviour in solids separation processes (Jenkins *et al.*, 2003). In such a way, the presence of specific groups or species is generally used as an approach of specific wastewater characteristics or operational conditions which can be related as possible causes of the different solids separation problems.

The observation of microfauna usually includes the determination of the following parameters:

- **Determination of the predominant group.** By determining the predominant group of the microfauna, general correlations to different process characteristics such as F/M or SRT can be determined (see 1.2.5 The importance of microfauna as indicator of plant performance). This generic evaluation is usually performed by firstly using 100x and increasing magnification up to 200x, 400x or 1000x whenever most specific determinations (e.g., at species level) are needed.

- **Counting of groups or individual species.** The counting of specific groups or species is usually performed by placing a drop of known volume (generally 0.05 mL) and counting the number of microorganisms present in different fields of view. In general, 4-5 separate preparations are prepared and each one is divided into different fields of view. The total fields of view usually range from 10 to 20. The observation is performed at 100x and the average number of microorganisms obtained per field of view is finally multiplied by the total number of fields of view and by the cover slip area.

- **Sludge Biotic Index (SBI).** This is an index developed by Madoni (1994b) to describe and monitor the activated sludge plant performance by means of the number and types of protozoa and metazoa present in a sample of mixed liquor. The different microfaunal groups are enumerated. These groups include the small and large flagellates, ciliates, testate amoebae, rotifers and nematodes. Once the count is completed, the SBI index is determined based on the density of individual populations, the population diversity and the number of small flagellates. In general, the greater the abundance and diversity of crawling and sessile ciliate protozoa, and the lower the numbers of free-swimming ciliates and flagellates, the healthier the mixed liquor is. The calculation of this index is optimal, due to its objectivity. Its main disadvantage is that skilled users are necessary to identify the organisms to genus and sometimes species level.

c) Presence of filaments

When filamentous microorganisms are the cause of separation problems, it is necessary to know the extent of their presence in biocenoses of activated sludges. Several practical methods of quantifying filamentous microorganisms have been developed and multiple modern techniques have been researched an applied to improve these quantifications.

- **Total Extended Filament Length (TEFL).** It is known that different filamentous organisms affect sludge settleability differently. Some filaments are short and do not protrude very far beyond the floc into the bulk liquid and therefore, even in significant numbers, have relatively little influence on sludge settleability (Jenkins *et al.*, 2003). Other filaments are long and coiled and exist mainly in the bulk liquid; these have a major impact on sludge settleability. Accordingly, the best way to quantify the occurrence of filamentous microorganisms and correlate this value with sludge settleability is by directly counting the total length of the existing filaments. Sezgin *et al.* (1978) developed for these purposes a counting technique to measure the total length of filaments protruding from the flocs as well as filaments floating freely in the bulk liquid. Since the microscopic measurement of the total length of all filaments in the sample would be too laborious and time-consuming, the filaments are divided into seven groups (in µm): (1) 0-10, (2) 10-25, (3) 25-50, (4) 50-100, (5) 100-200, (6) 200-400 and (7) 400-800. The measure is done at 100x magnification and the numbers of filaments within these intervals are counted.. The length of filaments greater than 800 µm is then measured individually. Filaments with branches are counted as two filaments. Finally, the resulting total extended filament length is expressed as measured length (µm) per unit mass (g or mL of MLSS). If the sample is diluted before the measurement, it will have to be multiplied by the dilution factor. A value higher than 30 km/g, according to Lee *et al*. (1983) or just higher than 10 km/g MLSS, according to Gabb *et al*. (1989), can be considered as an indicator of a bulking problem.

Matsui and Yamamoto (1984) facilitated the microscopic counting by using a video camera and a colour TV monitor.

- **Simplified filament counting technique.** The counting technique was further simplified by Beebe *et al* (1982), especially for the quantification of filaments with an irregular shape such as the branched *Nocardia* spp. During the procedure of measurement, a diluted mixed liquor sample of a known volume is placed on a glass slide and covered completely with a cover slip. The covered area is divided into fields. The eyepiece of the microscope is equipped with a simple hairline. By moving the objective from one edge of the cover slip to the other, consecutive fields are observed, and the number of cases in which a filamentous microorganism intersects the hairline is recorded. Then the number of intersections in all the examined fields is summed up, multiplied by the number of fields on the slide, and divided by the volume or concentration of the sample. The final result can be expressed as a filament count per volume or biomass unit. This method is suitable for evaluating some remedial actions in a particular plant against filamentous microorganisms.

  Pitt and Jenkins (1990) modified the counting method for use with irregular and branched *Nocardia* filaments by using Gram-stained samples. The methodology is the same as above, but only the Gram positive branched filaments greater than 1 µm in length are counted. The final result of this modified filament counting technique is expressed as the number of intersections per gram of biomass. Pitt and Jenkins (1990), tried to correlate these values with foam production, suggesting that values higher than 1 x 10^6 intersections/g VSS (volatile suspended solids) can be expected for foaming episodes caused by *Nocardia*.

- **Abundance of filaments.** For a routine examination of the extent of filamentous growth, a rapid and simple method was elaborated by Jenkins *et al*. (1986): the subjective scoring of filament abundance. This is a counting technique based on microscopic observation of a wet mount under phase contrast at 100x magnification. The abundance of filamentous microorganisms is classified according to the scale given in Table 1.6, which ranges from *none* to *excessive*. Some examples of filament abundance are shown in Figure 1.16. In general, an abundance rating for all filaments present together in the sample, and an abundance rating for each filamentous organism type are determined. Abundance of filaments scored as very common or greater are considered to account for solids separation problems (Jenkins *et al*., 2003).

Table 1.6 Subjective Scoring of Filament Abundance (Jenkins *et al*., 1986)

Numerical Value	Abundance	Explanation
0	None	
1	Few	Filaments present, but only observed in an occasional floc
2	Some	Filaments commonly observed, but not present in all flocs
3	Common	Filaments observed in all flocs, but at low density (1-5 filaments/floc)
4	Very Common	Filaments observed in all flocs at medium density (5-20 fil./floc)
5	Abundant	Filaments observed in all flocs at high density (>20 fil./floc)
6	Excessive	Filaments present in all flocs, with more filaments than flocs and/or filaments growing in high abundance in bulk solution

Although this method is not as exact and reproducible as the TEFL method, if the subjective scoring is performed by the same individual for a long time, it provides a true picture of the biocenosis of a given activated sludge. Using this method, an experienced observer can reveal the tendency to bulking much earlier than the settling properties start to deteriorate, so that appropriate remedial actions can be taken in advance (Wanner, 1994).

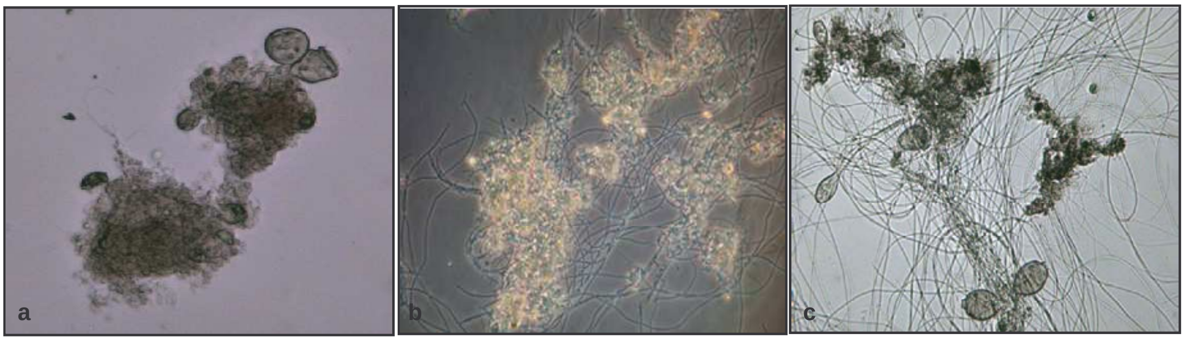


Figure 1.16 Different abundance of filaments: (a) few; (b) very common and (c) abundant

As in the determination of floc characteristics, **Image Analysis** has also been applied in microbiological systems to automatically discriminate between flocs and filaments (Cenens *et al.*, 2002), and to determine morphological parameters of filaments such as total length or filaments number (da Motta *et al.*, 2002 and Jenné *et al.*, 2004). They also found various correlations between some of the features of filaments determined automatically, and usual measurements of sludge settleability such as SVI or settling velocity, in order to provide tools for early detection of bulking and foaming.

- **Identification of filaments**. In the case of solids separation problems, it is also useful to know the type of filamentous microorganisms present in the sludge and the extent of their growth. It is a well-recognized fact that different types of filamentous microorganisms cause problems of different intensity. For example *N. limicola* or Type 0092 do not cause episodes of bulking as severe as the ones caused by Type 021N or *S. natans*. On the other hand, correct identification of filamentous microorganisms can help plant managers to estimate the causes of the presence of filamentous species in the biocenosis, which is also crucial for finding the proper methods of controlling their growth.

  The conventional, standard methods for identifying bacteria and other microorganisms are based on their isolation from mixed cultures and on subsequent exhaustive tests of the morphological, biochemical and physiological features that allow determination of the exact taxonomic position of the filamentous species. The use of these conventional methods, though, has demonstrated that these techniques are not very suitable for practical purposes of identification mainly because, apart from being expensive, are tedious and time consuming. In addition, the isolation of filamentous bacteria is problematic due to their slow growth and the difficulties in obtaining pure cultures from activated sludge samples (Wanner, 1994). Hence, only a few of these filamentous bacteria have been obtained in pure cultures and consequently described taxonomically.

  Eikelboom (1975) made one of the first attempts at classification of filamentous bacteria by dividing the 29 different morphological types he recognized into several groups based on their microscopic properties and staining reactions. He also developed an identification key for distinguishing the individual observed filamentous microorganisms. In this key, Eikelboom identified the filamentous microorganisms as *types* and not genera or species. The taxonomic position of these types is uncertain and out of the 29 categories only a few represent undoubted taxonomical species. However, this fact in no case contradicts the enormous advantages of this classification system in which the identification of types, according to its author, is fast and can be performed even by trained non-microbiologists. The Eikelboom system of types has been applied in many manuals on microscopic examination of activated sludges aimed at bulking and foaming control. Some examples are: Eikelboom and van Buijsen (1983); Jenkins *et al.* (2003); Richard (1989); Pujol *et al*. (1990) and Wanner (1994). Most of these identification tools are based on dichotomous keys, very easy to follow, which classify microorganisms according to the following features:

- Staining reactions (Gram stain and Neisser stain)
- Filament shape
- Size and shape of cells
- Branching
- Filament motility
- Presence of epiphytic bacteria on filament surfaces
- Filament width and length
- Presence of granules or sulphur deposits
- Presence of sheath
- Location

Despite their wide applications, conventional detection techniques based on morphological characteristics of filaments have been criticised mainly because of their possible misinterpretation. This criticism is based on the fact that: (1) single species may exhibit polymorphism or different species may look the same; (2) only indicative quantification is possible, given that filaments that reside inside the flocs can remain unnoticed, and (3) the method is subjective to the extent that it depends on the training level of the personnel (van der Waarde *et al.*, 2002).

During the 1990s, the development of molecular identification techniques based on new biological molecular tools, have made it possible both to determine the composition of microbial communities and to detect and quantify the presence of individual microbial species more accurately. An especially promising genetic method for detecting filamentous microorganisms is the **Fluorescent *In Situ* Hybridization (FISH)** technique. In this technique, a gene probe (a fragment of genetic material approximately 15 to 30 bases in length) complementary to a unique portion of one of the RNA molecules in a cell's ribosomes (usually the 16S rRNA) is prepared. The probe is then reacted with a fluorescent molecule that, when illuminated with ultraviolet (UV) light, can be visualized. When the probe is added to a sample, it will bind only to the 16S rRNA to which it is complementary. The extent and location of the probe binding can be detected by observing the sample through a microscope with UV illumination.

Despite the promising results obtained in several applications of FISH in activated sludge wastewater treatment systems such as Soddel *et al*. (1993); Hernandez *et al*. (1994); Wagner *et al*. (1994); De los Reyes *et al*. (1997); Snaidr *et al*. (1997); and van der Waarde *et al*. (2002), the application of oligonucleotide probes is limited by the number of probes available. Out of the 80 different morphotypes of filamentous microorganisms that have been identified, only less than 20 species can be currently identified with specific gene probes by FISH (Martins *et al.*, 2004). Thus, more research effort has to be undertaken in the development of more specific gene probes to ensure the proper phylogenetic classification of most species. Moreover, the application of FISH is constrained by the necessity of well-trained operators, which implies both money and time expenses that are still too great for its routine use in real plant monitoring.

Other promising techniques recently developed which can help in the domain of microbial communities identification are: **Denaturing Gradient Gel Electrophoresis (DGGE)**, a DNA/RNA fingerprinting methods that has shown to be a powerful technique to demonstrate the diversity of a particular wastewater community, and especially to compare samples differing only in the effect of one or more variables such as time or operating conditions (Wilderer *et al.*, 2002); **Microautoradiography (MAR)**, which has been used in combination with FISH to determine specific uptake of organic and inorganic radiolabeled substrates by activated sludge microbial communities. This method can be used, for example, to verify that certain organisms are active during specific conditions (Lee *et al.*, 1999). Other applications include the study of substrate uptake by different filamentous bacteria (Andreasen and Nielsen, 1997); the **Fatty Acid Methyl Ester (FAME)** technology, which has also been used for monitoring microbial communities, such as that of *Nocardia* spp., in activated sludge processes. The characteristic of this technique is that the analysis is performed for a whole community of microorganisms based on the examination of the relative abundances of the various fatty acids associated with the organisms present in the sample (White *et al.*, 1997).

As with the application of FISH, these sophisticated techniques to identify different species or communities of microorganisms are still applied within the research domain, and their application to regular, full-scale monitoring is not yet feasible.

1.5.1.4 Analytical data and operational parameters

Although microscopic and settleability tests widely contribute to the detection of solids separation problems, this information will not be of practical use for plant operators as long as the operational factors responsible for the occurrence of these specific problems remain unknown. The study of biological and ecological properties of the different microorganisms within the activated sludge system makes it easy to define the conditions which can promote their growth (Wanner, 1994 and Jenkins *et al.*, 2003). The analysis of operational parameters and analytical data from the process is therefore crucial to identify those operational conditions or wastewater characteristics which have an effect on the types of microorganisms present on the system and thus on the settleability of the sludge.

Some analytical data from the process can be used as indicative or symptomatic of solids separation problems. These include:

- **Effluent Total Suspended Solids (TSS)**. Since most solids separation problems, by definition, result in a loss of solids with the effluent, the measurement of the TSS present at the effluent is of major importance as a monitoring parameter.

- **Effluent COD**. Effluent organic load, measured as BOD_5 (Biological Oxygen Demand) or COD (Chemical Oxygen Demand) is also measured to both monitor treatment efficiency and loss of solids. COD is generally preferred because results are obtained faster (in a few hours) than for BOD_5 (after 5 days of digestion).

- **Turbidity.** This is another parameter that determines the presence of suspended matter within the effluent. It can be measured with a nephelometer or can be appraised visually by making transparency measurements directly in the secondary clarifier with a Secchi disc.

The analytical data from the process that can be measured to determine the possible causes of solids separation problems generally include parameters measured at the primary effluent (influent of the biological reactor) that, based on literature, are potentially related to sources of solids separation problems. These parameters include:

- **Readily biodegradable substrate**. This can be either determined from a biological response method (OUR) or estimated by a physical separation technique (Metcalf and Eddy, 2003). The physical separation technique is the most used technique to determine the readily biodegradable COD (RBCOD), usually by means of a floc/filtration method (Mamais *et al.*, 1993), which uses $ZnSO_4$ or $Zn(OH)_2$ to induce floc formation.

- **Toxics presence.** It can be approached by studying the characteristics of the biomass (see 1.2.5) or by specifically analysing the concentration of the suspected toxics: heavy metals, chlorine, etc.

- **Sulphide concentration.** It is usually measured by using the methylene blue method or by using a hatch kit.

- **Nutrients concentration**. Usually calculated by means of its relationship to organic matter: COD/Total Nitrogen and COD/Total Phosphorus are the ratios usually calculated. The existence of on-line sensors capable of measuring BOD, COD, Total Nitrogen and Total Phosphorus can be very helpful in determining these parameters on-line. Their general use though is still quite limited.

- **Temperature and pH**. These two parameters are usually measured on-line.

There exist also some operating data from the activated sludge process that, apart from being very useful for establishing operational control strategies, they can also indicate some characteristics of the process operation that can be linked with solids separation problems causes. The most used operating data are the following:

- **Dissolved Oxygen concentration (DO)**. This is generally measured by on-line sensors and often linked with automatic controllers.

- **Mixed Liquor Suspended Solids (MLSS)**. MLSS is the total amount of organic and mineral suspended solids, including microorganisms, in the mixed liquor. It is determined by filtering an aliquot of mixed liquor, drying the filter at 105 ºC, and determining the weight of solids in the sample.

- **Mixed Liquor Volatile Suspended Solids (MLVSS)**. The organic portion of the MLSS is represented by MLVSS, which comprises non-microbial organic matter, as well as dead and live microorganisms and cellular debris. MLVSS is determined after heating dried filtered samples at 600-650 ºC, and represents approximately 65-75% of MLSS.

- **Food-to-Microorganism Ratio (F/M)**. It indicates the organic load introduced into the activated sludge system and is expressed in kilograms of BOD per kilogram of MLVSS per day, although using COD (and sometimes MLSS) is usually preferred because the results are available sooner.

$$\frac{F}{M} = \frac{Inflow\ rate\ (m^3/d) \times BOD_{5\ primary\ effluent}\ (g/m^3)}{Vreactor\ (m^3) \times MLVSS\ (g/m^3)} = \frac{kg\ BOD_5}{kg\ MLVSS, d}$$

The F/M ratio is generally controlled by the rate of activated sludge wasting. The higher the wasting rate, the higher the F/M ratio. For conventional aeration tanks, the F/M ratio ranges from 0.2 - 0.5 kg COD/kg MLSS *d, but it can be higher (> 1.5) for activated sludge using high-purity oxygen. A low F/M ratio means that the microorganisms in the aeration tank are starved, leading to a more efficient wastewater treatment.

- **Sludge age**. The sludge age, or sludge residence time (SRT), is the mean residence time of microorganisms in the system. In its calculation it is usually assumed that the biomass separator (i.e., final clarifier) works perfectly, so that no biomass is lost in the final effluent. The sludge residence time values may be in the order of days. This parameter is given by the following formula:

$$SRT = \frac{(V_{reactor} \times MLSS)}{(WAS\ rate \times MLSS_{rec.}) + (Outflow\ rate \times SST_{effluent})} = days^{-1}$$

The common range of sludge age in a conventional activated sludge plant is between 3 and 15 days. For extended aeration activated sludge plants the range is between 15 and 30 days. Generally during the winter months, higher sludge ages are required to maintain a sufficient biological mass. In the summer time, biological activity increases and lower sludge ages normally produce a higher quality effluent.

- **Mean Cell Residence Time (MCRT)**. This parameter is a refinement of the sludge age and takes into consideration the total solids in the secondary or biological system (solids in the reactor plus solids in the secondary settler). Nevertheless, the mass of solids in the secondary clarifiers is difficult to determine due to solids distribution patterns that mainly depend on the settleability of the sludge. The most generalised assumption is to consider that the concentration of suspended solids in the secondary settler is the same as in the reactor. The MCRT formula is the following:

$$MCRT = \frac{(V_{reactor} + V_{clarifier}) \times MLSS}{(WAS\ rate \times MLSS_{rec.}) + (Outflow\ rate\ \times TSS_{effluent})} = days^{-1}$$

- **Hydraulic retention time**. The hydraulic retention time (HRT) is the average time spent by the influent liquid in the aeration tank of the activated sludge process. It is represented as:

$$HRT = \frac{V_{reactor}}{primary\ effluent\ rate} = hours^{-1}$$

- **Mixed liquor respiration rate.** The mixed liquor respiration rate, also referred to as Oxygen Uptake Rate (OUR), measures the rate at which oxygen is removed from the wastewater, thus providing an indication of the activity of activated sludge biomass. This analysis of microorganisms' activity can help predict settling characteristics and sludge quality, and can assist in verifying changes in mode of operation and in troubleshooting presence of toxics, excessive or insufficient MCRT or F/M, filamentous bacteria overgrowth, etc. (WEF, 2002).

1.5.2 Control of solids separation problems

According to Jenkins *et al*. (2003), the beginning of any solids separation problem solving has to involve microscope examination of the activated sludge in order to reveal the origin of the problem. If the problem is caused by filaments, identification of the causative filamentous bacteria yields a direction or approach to take for a remedy. After that, the analysis of the plant operating conditions and the wastewater characteristics may help to delimit the possible causes. If the probable cause is identified, it may be possible to apply a **specific control method** to suppress those specific causes that promote the solids separation problems.

In general, since wastewater fractions can hardly be manipulated in practice, at least operating conditions and reactor design should be optimized in order to obtain a better performance of the plant. If successful, these methods will provide a permanent solution to the particular situation. Some plants with solids separation problems, however, may require major design or operational changes that can take a long time to implement (e.g., additional aeration capacity, changes in aeration basin configuration, industrial waste control, etc.). In addition, it also needs to be considered that, once changes have been made to specifically thwart the possible factors encouraging separation problems, sludge settleability may improve only slowly, given that the activated sludge microbial population changes at a rate proportional to the MCRT. For example, in a completely mixed system, a period equal to approximately 3 MCRT is required for a bulking activated sludge to return to a non-filamentous condition.

On the other hand, when a serious solids separation problem occurs, usually rapid, **non-specific control methods** are applied. These methods deal with the consequences (symptoms) of the separation problems, i.e., improve sludge settleability, but do not remove the conditions that caused those undesired situations. As a consequence, the effects of these methods are just temporary and usually, once the control method is stopped, the still-present conditions will originate again the same problem forcing the action to be repeated after a certain period of time.

The best approach for controlling activated sludge solids separation problems and ensuring long-term, problem-free plant operation is using appropriate design and operational methods. However, many of these techniques have been developed only relatively recently, and its application implies some modifications on the existing operational strategies or plant design that usually prevent them from being chosen as a possible control method. Hence, currently many existing activated sludge plants still face solids separation problems. So, which criteria do plant operators follow to choose the proper control method to restore the process? The proper selection and use of control methods may allow the operator to achieve the following objective: the optimum degree of treatment (according to the regulations) at the lowest cost.

The present section reviews the existing control methods (both specific and non-specific) that can be applied in activated sludge systems, defining their main advantages and disadvantages, based on several manuals of control such as: Wanner (1994), Gerardi (2002), WEF (2002) or Jenkins *et al.* (2003). A general conclusion about the most effective control methods can not be established since, despite extensive research to develop and test different control methods, most studies show that generally the success or failure of control actions varies from one WWTP to another. For example, while the chlorination of RAS was the most effective method in controlling foaming in 48% of the surveyed plants in the USA (Pitt and Jenkins, 1990), in Australia Blackall *et al.* (1991), found that the reduction of sludge age was generally more effective than chlorination (57% of plants versus 20% of plants, respectively). The same occurs when the efficiency of different configurations of selectors was tested. While the use of anaerobic selectors was effective in controlling *Nocardia* spp. in some studies, it did not show any improvement on foams in other applications.

Tables 1.7 and 1.8 review the current control methods, both specific and non-specific, that are usually applied to control solids separation problems in activated sludge systems, pointing out both the advantages and disadvantages of their application.

Table 1.7 Specific control methods for activated sludge solids separation problems

CAUSES	SPECIFIC CONTROL METHODS
Nutrient deficiency	-Addition of the required nutrients to accomplish the ratio COD/N/P=100/5/1. Specifically formulated nutrient and micronutrient commercial solutions are available to correct nutrient deficiency. N sources: anhydrous ammonia, urea or ammoniac salts; P sources: phosphoric acid and sodium phosphates; N&P sources: ammonium phosphate. Recycle streams within the treatment plant should be tested for the presence of readily available nutrients that can be used to correct the nutrient deficiency. Nutrients should never be added in quantities that would cause a final effluent violation of N or P limits. Advantages: (1) This is a very effective method against certain types of microorganisms, such as Type 021N or *Thiothrix* spp; (2) Cheap commercial compounds are available that can be added to restructure the required nutrients. Disadvantages: (1) An excess of nutrients can contribute to the polluting load of the effluent. As a consequence, the effluent quality needs to be checked frequently.
Oxygen deficiency	-Increase of air flow rate or increase of DO set-point. Advantages: (1) After a certain period of time, it is an effective method against *S. natans*, *H. hydrossis* or Type 1701. Disadvantages: (1) Requires a long time to be effective (3 MCRT); (2) Expensive method due to the high energy required; (3) Can cause onset of nitrification in systems with high MCRT; (4) The ability to rise DO concentration may be limited by the available aeration capacity.
Low F/M ratio or high MCRT	-Increase of F/M ratio or reduction of sludge age by decreasing MLSS (increasing WAS or decreasing recirculation). It can be applied whenever the populations of filaments or the nitrifying/denitrifying bacteria need to be controlled. Also whenever the sludge age needs to be decreased to control the overoxidation of flocs. It has been demonstrated that the washout of filamentous bacteria can be achieved only with extremely short biomass retention times, typically less than 3 days. -MCRT control by means of automatic sludge wasting and continuous TSS analyses has demonstrated to allow for much less variation in discrete MCRT values. Advantages: (1) Cheap and fast method (except when concerning high sludge production); (2) Very effective in the control of filaments especially encouraged by high SRT such as *Nocardia* spp. Disadvantages: (1) Can cause the loss of nitrification; (2) Increase in waste sludge production; (3) At extremely short MCRT, deflocculation problems can appear.

Low F/M ratio and/or Readily biodegradable substrate	-Changes in operation that effectively increases the substrate concentration available. These methods include: system configuration modifications so as to incorporate alternating or sequential feed-starve conditions into the system (e.g. intermittent or batch feeding); multi-reactor or plug flow conditions. -Use of biological selectors where the incoming wastewater and the return activated sludge are mixed prior to entering the aeration basin. The goal is to provide a short term (15-30 min. of contact time between RAS and influent) to encourage both a high substrate condition and a high RBCOD uptake rate which, in general, favours certain floc-formers but discourages filaments (kinetic selection). High efficiency is reached when the selector is sectionalized (at least 3 compartments are recommended with a HRT about 5 minutes in each compartment). The terminal electron acceptor can also be manipulated (metabolic selection) to favour floc-forming and to discourage filaments. A combination of these factors has led to the development of three types of selectors: aerobic, anoxic (O_2 not present, NO_3^- present) and anaerobic (O_2 and NO_3^- not present). A combination of different kinds of selectors can also be used. In general, the use of anoxic and anaerobic selectors discourages the growth of filaments, given that most (but not all) filamentous microorganisms are obligate aerobes. It can also discourage the growth of *Zoogloea* spp. The SRT within the anoxic or anaerobic selector is of 1-2 hours. Feast-and-famine selectors provide copious quantities of DO and substrate for a period of time and then no DO or substrate along another period. Nowadays, selectors are usually designed into new installations but they can also be retrofitted to older plants. Advantages: (1) Very effective method against: *S. natans*, Type 1701, Type 021N, *Thiothrix*, *N. limicola, H. hydrossis*, Type 1851 and *Nocardia* (just anoxic selectors effective); (2) It is a permanent method against low F/M filamentous bulking and foaming, because the conditions that favour the development of this type of microorganisms are changed. Disadvantages: (1) Very expensive method because sometimes a change in the plant configuration is needed; (2) Ineffective against: *M. parvicella*, Types 0041, 0675 or 0092 (which can take up substrate in anoxic and anaerobic zones) as well as against filamentous bacteria caused by low pH, septicity or low nutrients; (3) Anoxic selectors are very difficult to operate because the readily biodegradable substrate is not constant and thus, the COD/N ratio is usually changing.
High F/M ratio or low MCRT	-Decrease the F/M ratio or increase sludge age by raising the MLSS concentration (decrease or stop WAS or increase recirculation). Advantages: This is a general cost-effective method. Disadvantages: (1) If the MLSS concentration is increased too much the solids handling capacity of the clarifiers may be exceeded; (2) It can induce an increase of carbonaceous oxygen requirements due to the increased endogenous respiration; (3) It can favour nitrification.
Low pH	-Adjustment of the pH by adding a base such as caustic soda ($NaOH$), lime ($Ca(OH)_2$), sodium bicarbonate ($NaHCO_3$) or magnesium hydroxide ($Mg(OH)_2$ Advantages: If identified, the source of acidity can be stopped and future problems can be avoided. Disadvantages: (1)Titration tests are required to adjust the proper dose of the chemical to be added; (2) Adding a base to increase pH is expensive.

Presence of septicity or sulphides	-Preaeration, chemical oxidation with chlorine, hydrogen peroxide or potassium permanganate; chemical precipitation with ferric chloride or use of ferric nitrate or sodium nitrate as an oxygen source must be applied to the influent suspected to contain sulphides. Advantages: If identified, the source of septicity can be stopped and future problems can be avoided. Disadvantages: (1) Titration tests are required to adjust the proper dose of the chemical to be added; (2) Most chemicals are quite expensive.
Excessive shearing	-Identification of the source of excessive turbulence in the activated sludge process. This can be checked by the collection and examination of samples before and after the suspect equipment or treatment tank. Once the source of excessive turbulence is identified, the responsible unit should be repaired, removed or replaced from service. If the source is excessive aeration in the reactor, the aeration flow or the set-point should be decreased. Advantages: If identified, the source of excessive shearing can be stopped and future problems can be avoided. Disadvantages: (1) Decreasing aeration can reduce treatment efficiency; (2) The replacement of aeration units can be expensive.
Presence of toxics or surfactants	-Identification and reduction of the source of toxicity. -If the toxic is still present in the mixed liquor, WAS rate should be increased for approximately 1 week to purge the system. After this action, it is recommended to add large quantities of healthy seed sludge from another plant to build up the biomass. -If the toxic is not present within the mixed liquor, it is recommended to increase MLVSS (by stopping or decreasing WAS) or to add bioaugmentation products to recover the activity of the microfauna. This increase in the number of microorganisms lowers the portion of the biomass affected by toxicity. Advantages: If identified, the source of toxicity can be stopped and future problems can be avoided. Disadvantages: The process of identifying the possible toxic and the elimination of the process can be quite long and the process efficiency can be affected.

Table 1.8 Non-specific control methods for activated sludge solids separation problems

NON-SPECIFIC CONTROL METHODS
Chlorination of RAS. Chlorine, usually in the form of hypoclorous acid (HOCl) is added to the activated sludge with the purpose of selectively damaging the filaments extending beyond the flocs into the bulk liquid. If the chlorine dosage is properly adjusted, the floc-formers will not be seriously affected by the toxicant because they will find protection inside the sludge flocs. For this reason, it is easier to damage filaments that project from the floc (for example: types 021N, 0961 or *S. natans*). Filaments integrated inside the floc are less sensible (*M. parvicella* or *Nocardia* spp.); hence they require higher doses to be affected. Apart from this, the hydrophobicity of cell wall showed by *Nocardia* and *M. parvicella*, which hinders the penetration of chlorine into the cells, has also been studied as a possible cause of inefficiency. Chlorine should be added at a well-mixed point, that is, directly in the aeration tank or to the return-activated sludge (RAS), where a high turbulence is present. If added to the RAS, it is recommended to apply it in such a way that the return sludge is in contact with the chlorine solution for about 1 minute before the sludge is mixed with the aerator influent. Most domestic WWTPs can manage with a chlorine exposure frequency of three or more times per day in the RAS line, although it is a function of the relative growth rates of filamentous and floc-forming organisms and the efficiency of chlorine in killing them. Success has been achieved at frequencies as low as once per day (2.5-3 day^{-1} is recommended). Chlorine dosage is also adjusted so that its concentration is lethal at the floc surface but it is sub-lethal within the floc. In order to determine the proper chlorine dose, it is suggested to perform adequate preliminary trials such as OUR determinations or Live/Dead staining to verify the effectiveness of chlorine on the microfauna. The required dose is measured conveniently on the basis of the sludge inventory in the plant (overall chlorine mass dose). Effective chlorine dosages are usually in the range of 1-10 g Cl_2/ kg MLVSS,d (2-4 should work). Initially, low doses are recommended; these are subsequently increased little by little until they are effective. In nitrification systems, dosage should be lower and strict control of effluent nitrates and nitrifying capacity is required. Signs of overchlorination are a turbid (milky) effluent, a significant increase in effluent TSS (Total Suspended Solids), a loss of the higher life forms (protozoa), and a reduction in BOD removal. In some occasions, reseeding with healthy activated sludge or bioaugmentation with freeze dried bacteria may be helpful after treatment. Microscopic examination of the activated sludge during chlorination is recommended to control chlorine application. Chlorine effects on filaments include cell deformity, empty and broken sheaths or filament lysis. The use of chlorine for this purpose is widespread in the USA where it has been applied since the 1930s. In Europe, chlorination is widely used in the United Kingdom while in Germany or other countries of central Europe it is recommended only as an emergency measure. Advantages: (1) Chlorine is relatively inexpensive and available on-site at most plants; (2) It has shown quite good results in temporarily controlling some filamentous microorganisms. Disadvantages: (1) Since chlorine does not selectively affect filaments, it may cause deterioration in other sludge microorganisms, especially in protozoa and nitrifying bacteria, affecting COD, nitrogen and phosphorus removal. An exhaustive monitoring of chlorination effects on the microfauna should be performed; (2) It is not very effective for controlling nocardioforms since these filaments are usually present inside the flocs; (3) It can make sludge settleability worse if the problem of settleability is non-filamentous e.g., zoogloeal bulking or poor floc development; (4) Different susceptibility to chlorine must be considered, since control of one organism may lead to eventual predominance of another; (5) Cl_2 can react with excess nitrate or ammonia and form monochloramines, which in extreme situations, apart from reducing the availability of chlorine to attack filamentous bacteria, can form trihalomethanes or organohalogens, becoming a pollutant for the receiving ecosystem .

Chlorination of foams. Chlorine can be applied as a chlorine solution by means of a fine spray (mist) directly to the aeration basin and/or secondary settler surface. It should be possible to enhance the effectiveness of the chlorine by installing fine bubble aeration to enhance the production of foams just before or within the chlorine surface spray system. In general, the solution applied is composed of 10-15% of sodium hypochlorite or powdered calcium hypochlorite. It is suggested to allow the foam and the solution to be in contact for 2-3 hours. In order to favour the collapse of foam, effluent water can also be scattered on the surface.

Advantages: (1) In general, it does not affect the rest of microfauna and thus it does not cause effluent quality deterioration or loss of treatment efficiency.

Disadvantages: (1) When using surface spray chlorination, it is important to avoid dispersal of the sprayed chlorine solution because of its hazardous and corrosive nature.

Oxidation by Hydrogen Peroxide. Hydrogen peroxide is used to control filamentous bacteria in a similar fashion to chlorine. The minimum effective H_2O_2 dose for filaments control is approximately 0.1 kg H_2O_2/kg MLSS,d. This method is recommended for nitrification plants.

Advantages: (1) H_2O_2 could also be used as a source of oxygen and ameliorate those episodes caused by low DO concentration; (2) The effect on nitrogen and phosphorus removal is lower than with chlorine.

Disadvantages: (1) Doses of H_2O_2 need to be higher than doses of chlorine to be effective; (2) H_2O_2 penetrates deeper into the flocs and affects floc-formers; (3) The cost of H_2O_2 is twice that of the chlorine.

Oxidation by Ozone. The use of ozone as an oxidant has a small but consistently positive effect on filamentous bacteria control.

Advantages: (1) The use of high doses (6 g O_3/kg MLSS,d) has been demonstrated to likely decrease effluent colour, turbidity and TSS concentrations; (2) Ammonia removal by nitrification and phosphate removal can be improved; (3) The risk of formation of halo-organics is suppressed.

Disadvantages: Ozonation is limited by the high cost of its application.

Addition of chemicals. Synthetic organic polymers, such as polyacrylamides and inorganic coagulants and flocculants such as ferric chloride, aluminium sulphate, ferrous sulphate or lime can also be added to improve sludge settleability. The choice is based on suitability for a particular waste, availability and cost of the coagulant, and sludge treatment and disposal considerations.

The function of these chemicals is to capture solids, reduce the surface area of extended filamentous microorganism, improve the density of floc particles and add weight to the sludge, which helps in settling small, sheared floc particles in the secondary clarifier. The most widely used additives are synthetic, cationic polymers, of high molecular weight, alone or in combination with an anionic polymer, that serve to overcome the physical effects of filaments on sludge settling. These are usually added to the MLSS as it leaves the aeration basin or to the secondary clarifier centerwell. It is wise to perform frequent jar testing to determine the most effective chemical and the exact dose to be applied. Doses are about 35 g Fe/m^3 or 10-15 g Ca/m^3 for at least 10 days.

Recent studies have reported the effectiveness of adding polyaluminium chloride (e.g. PAX-14) to control different filamentous microorganisms, especially to control *M. parvicella*. The mechanism of PAX-14 in controlling this microorganism is still unknown, although it is believed to affect the hydrophobic nature of the cell wall. The recommended dosage is 2-3 g Al/kg MLSS,d in the RAS for at least 3 weeks.

Addition of metal salts or polymers can also help control the creation of anoxic zones within the sludge blanket of clarifiers and avoid undesired denitrification, since the sludge may be thickened and removed more rapidly.

Advantages: (1) If the type of chemical and the dose are properly selected, the results on settleability improvements are immediate; (2) Synthetic organic polymers do not significantly increase the mass of solids; (3) Iron or aluminum coagulants, apart from increasing the sedimentability, can help in the removal of phosphate or sulphate (effective against Type 021N).

Disadvantages: (1) Very expensive methods. The use of polymers involves a higher cost with respect to RAS chlorination (chemical cost in the range of 3-15 $ / ML may be expected); (2) Produce an increase in waste sludge production; (3) The positive effects on the sludge have a relative short duration, and repeated additions are required; (4) An excessive addition of cationic polymers can produce foams.

Physical removal of foams: foam trapping or skimming, water sprays and selective flotation. Spraying the foam with effluent water through a bib sprayer or lawn sprinkler can dilute the foam and collapse it. As a consequence, the quantity of foams can decrease. Foam can also be vacuumed or raked from the surface of the aeration tank. In some plants, the use of a flotation unit, where the activated sludge with foam-forming filaments is manipulated to form foams, is used. The foams are then mechanically removed.

Surface trappers of scums such as zone dividing baffles or mixer shafts within the reactor need to be avoided. Finally, since foams can contain a high concentration of filaments, this foam should never be recycled to the inlet of the plant as it can cause reseeding of the process with foaming. Appropriate treatment and disposal of foams should be performed.

Advantages: (1) In general, physical removal of foams do not imply a high cost; (2) If surface trappers are properly detected and removed, the formation of foams can be prevented in future episodes.

Disadvantages: Water sprays have only limited application because of their inability to control moderate to severe foaming conditions.

Addition of antifoam compounds. Several defoaming agents can be used to decrease foams. An example of one of the most used compounds is the polyglycol-based defoaming agent.

Advantages: It is more effective than spraying effluent water since the defoaming agent physically and chemically affects the foam.

Disadvantages: (1) The efficiency is not very high probably because biological foams are much more stable than the foams against which the commercial antifoam agents were developed; (2) they are very expensive.

Addition of inert particles: sawdust. Some experiments have demonstrated that the continuous contact of filaments with small inert particles of sawdust added to the mixed liquor and kept in suspension with the flocs proved to be effective in controlling the development of filaments due to shear stress.

Advantages: If recycled as a sub-product of some processes it does not imply a high cost.

Disadvantages: It can cause a temporary increase in the turbidity of the final effluent.

Addition of other ballasting compounds: talc powder. The addition of talc powder into the aeration tank aims to weight the sludge and further reinforce the floc structure. Recently, a new additive consisting of a combination of synthetic polymer, talc powder and a quaternary ammonium compound has been developed to improve sedimentation in poorly settling sludge. Its use can decrease both the frequency of addition and the amount of traditional additive required

Advantages: (1) The application of talc usually results in an immediate sludge sedimentation improvement; (2) Ballasting agents have no adverse effects on floc-forming microorganisms.

Disadvantages: (1) Ballasting agents have no adverse effects on the filaments causing the bulking or foaming; (2) Their implementation requires repeated additions of relatively large amounts (up to 70-100% of the MLSS) of the ballasting agents to the sludge; (3) Addition of talc can cause foams; (4) The amount (in terms of dry solids) of the waste sludge is increased and the polyelectrolyte required for dewatering is also augmented; (5) The use of quaternary ammonium compounds is very expensive.

Biological control. Addition of mixtures of microoganims, sometimes containing extra nutrients and enzymes. Microorganisms, mainly bacteria, phages and actinomycetes, isolated from various sources are capable of lysing filamentous bacteria due to a lytic substance they produce. Other microorganisms, such as specific predatory ciliates, have demonstrated a potential in ingesting the trichomes of filamentous bacteria. Bioaugmentation products containing lipase enzymes can also be applied to degrade the lipids present in foams.

Advantages: Effective against certain types of filamentous microorganisms such as: *S. natans*, Type 021N or Type 1701.

Disadvantages: (1) Not effective against all types of filamentous microorganisms; (2) High dosages are needed, which implies a high cost; (3) Its application can impart toxicity to the effluent.

Manipulation of RAS and/or WAS. Increasing the RAS flow rate helps prevent failure of the clarifier due to the reduction of MLSS concentration in the clarifier before the solids are lost within the effluent. MLSS reduction can also be achieved by decreasing the mixed liquor solids inventory, which is obtained by increasing the sludge wasting rate (WAS).

Increasing the RAS and/or the WAS can also help controlling the creation of anoxic zones within the sludge blanket of clarifiers and avoid undesired denitrification.

Advantages: With the correct calculations, it can be an effective and non-expensive method, except from the consequent higher production of sludge caused by a higher WAS rate.

Disadvantages: (1) Biomass reduction in the system can reduce the removal efficiency and lead to dispersed growth of the activated sludge floc; (2) The sludge blanket level in the secondary settling can be decreased only for a limited period of time.

1.6 Information Technology Contribution to Decision Support

During the occurrence of solids separation problems, emphasis is placed on the plant operator's ability to quickly identify the problem, perform diagnosis, and initiate recovery actions. However, the reliability of human action is adversely affected at the time of crisis due to time stress and psychological factors. The availability of operational support capable of monitoring the status of the plant and quickly identifying the deviation from normal operation is expected to significantly improve the operator reliability. Reacting fast and properly to upsets of activated sludge wastewater facilities implies the optimal use of multiple and heterogeneous data sources.

In wastewater systems, inputs are unpredictable and suffer from drastic changes (both in quality and quantity), kinetics are complex, important variables of the process are difficult to measure in real-time, there are clear interactions between process units, and design of the facilities has been traditionally done in steady state (Vitasovic, 2001). Progress in Information Technology, including automatic acquisition, storage, manipulation, management, movement, control, display, switching, interchange, transmission, or reception of data or information has facilitated the management of complex processes, such as WWTPs. According to Olsson *et al*. (2003), the capability of computers to extract useful information, however, is rarely utilised beyond simple graphing. The use of Information Technology to encapsulate process knowledge, i.e. knowledge about how the process works and how to best operate, is extremely important in the management of complex processes. If process knowledge can be encapsulated, then not only is it retained but the computer can assist decision-making in plant operation.

During the present section, a general overview of the Information Technology contribution to decision support in WWTP operation is presented, including progress in data acquisition, instrumentation and process control; mathematical modelling and the application of some Artificial Intelligence techniques.

1.6.1 Data Acquisition, Instrumentation and Process Control

Progress in instrumentation, control and computer technology has enabled the measurement of an increasing number of variables as well as automatic control improvements. Presently, many wastewater facilities are equipped with automatic control and data acquisition systems capable of monitoring plant conditions by retrieving numerical data from database management systems (e.g., dissolved oxygen, pH, temperature, flows, etc.), and of controlling the essential operational and control equipment through a programmable logic controller (PLC) network. According to the results of a review performed by Jeppson *et al*. (2002) about the status and future trends of ICA (Instrumentation, Control and Automation) in wastewater treatment, it can be stated that most WWTPs in Europe (>10.000 p.e.) are equipped with SCADA (Supervisory Control And Data Acquisition) systems. The general function of an SCADA system is the retrieval of on-line data from the process, such as dissolved oxygen and flows, or the monitoring of equipment status. This allows plant operators to supervise the process. SCADAs usually incorporate a system of alarms which also alerts plant operators of possible failures or abnormal conditions within the process. According to Jeppson *et al*. (2002), though, the possibilities of SCADAs as control systems are usually underestimated and in general their use is limited to data acquisition and process monitoring.

Regarding the improvements in instrumentation, today it is possible to measure physical variables like flow rate, temperature, pH, redox, conductivity, turbidity and sludge blanket level. Concentrations of dissolved oxygen, sludge, ammonium, nitrate, phosphate, and organic matter can also be measured on-line. Yet, only few on-line data are used in closed loop control schemes (Olsson, 2005).

Research and progress in instrumentation, monitoring, modelling and computer technology has also enabled automatic process control improvement. Classical current control trends include instrumentation and control devices integration (distributed control), multivariate control, optimal control and hydraulic capacity control. Almost all applications, however, focus on controlling variables with fast response dynamics, such as flow rates or DO, which are by far the most common type of applied real-time control (RTC) based on on-line measurements (feedback). Other secondary process variables such as Sludge Residence Time (SRT), Mixed Liquor Suspended Solids (MLSS) and OUR also are controlled, but very few applications focus on controlling the real process objectives (effluent quality), which requires monitoring variables with slow response dynamics. Most recent applications try to control variables with medium response dynamics such as effluent COD and nutrient concentrations (especially ammonia and nitrate).

Nowadays new and sophisticated systems for ICA are often proposed in scientific journals but few are ever applied in real plants. The most fundamental barrier for more widespread acceptance is, apart from the required investment of money, the lack of flexible actuators, which makes these facilities unprepared for real-time control (Jeppson *et al*. 2002). According to Olsson (2005), there are still a lot of opportunities to further apply ICA but often there is a lack of training and education of plant operators to use the proper tools.

Regarding solids separation problems, the use of advanced on-line monitoring methods can partially support plant operators in controlling these undesired situations. For example, the on-line monitoring and control of the DO, or the SRT (Ekster and Jenkins, 1999) can prevent the process from developing specific conditions that directly encourage the development of undesired episodes of filamentous bulking or foaming. So, indirectly this type of automatic control can occasionally contribute to avoiding future problematic situations. Even so, in these cases, the automatic control by itself cannot prevent solid-liquid separation problems. A kind of further intelligent reasoning is always necessary to make certain decisions such as fixing the desired set-point.

1.6.2 Mathematical Modelling

Usually, the transition from data to information requires an appropriate application of algorithms on data and generally it needs to be supported by an appropriate model or theory of the application domain. Likewise, models may have to be calibrated, algorithms may have to be parameterised, etc. These tasks are very demanding and require massive knowledge processing and the availability of an appropriate model base (Cortés *et al*., 2000).

Mathematical modelling can be considered as a very useful tool in the study of complex ecosystems such as activated sludge systems, where a lot of parameters, processes and reactions are involved. In the last few decades, the significant increase in the knowledge of process fundamentals and the enormous development of calculation potentialities, including improvements in the characterisation of wastewater substrates, such as soluble and particulate, biodegradable and unbiodegradable, allowed the development of models that were able to give a more reliable mathematical representation of biological systems. Currently available activated sludge models include (Gearney *et al.*, 2004): the group of Activated Sludge Models developed by the International Water Association (IWA) including ASM 1, 2, 2D and 3 developed by Henze *et al.* (1987, 1995, 1999) and Gujer *et al.* (1999), and the metabolic model developed at the Delft University of Technology (TUDP Model) (van Veldhuizen *et al.*, 1999 and Brdjanovic *et al.*, 2000). These activated sludge models, together with the sedimentation tank model developed by Takács *et al.* (1991) have been widely used to study population dynamics, and to predict effluent loads or effluent quality parameters in activated sludge systems. One of the most valuable utilities of models is that once a well validated model is available, numerous variations of scenarios can be simulated in order to study, for example, the effect of different process conditions or the efficiency of some control strategies on the activated sludge system. This makes mathematical modelling an important tool for simulating or studying different situations at significantly lower cost and in a shorter time scale than the corresponding experimental studies.

Despite the wide research done in the development of activated sludge models, the characterisation of biomass has not gained such good improvements. This is mainly because of the complexity of defining the parameters related to morphology, physiology and kinetics of microorganisms (e.g. μ_{max}, K_s or k_d), as well as those related to all the possible interactions and processes present in such a community. These must be usually determined by using pure cultures, which increases the complexity of their definition. Nowadays though, the use of *in situ* molecular biological techniques makes its identification a little easier. Up till now, heterotrophs, nitrifiers and sometimes bio-P bacteria were the common bacterial populations considered in the existing applications. Therefore, no accurate model has been developed yet that is capable of simulating the growth of filamentous bacteria and the development of solids separation problems. Some of them are in the early phase of development with still no definitive results.

The first bulking sludge mathematical modelling was developed by Lau *et al.* (1984). It was used to predict the volume-average rate of a specific filamentous bacteria (*S. natans*) and a non-filamentous bacteria (*Citrobacter* sp.), according to the diffusion of soluble organic substrate, DO concentration, floc shapes and sizes and specific kinetic parameters of each bacteria. The limitation of the model to two specific bacteria makes it difficult to use it for the prediction of growth in other species. Then Chiesa and Irvine (1985) modelled the growth of 3 type of microorganisms (floc formers; slowly growing, starvation resistant filamentous microorganisms and rapidly growing, starvation susceptible filamentous microorganisms) to explain the growth and control of filamentous organisms using combined kinetic selection and accumulation/storage concept. Hermanowicz (1987) developed a model to describe the growth of floc-forming and filamentous bacteria according to the concentration of a single substrate and to DO. Later on, Kappeler and Gujer (1994) developed the AEROFIL model to study the effect of readily biodegradable COD on filamentous bacteria and Takács and Fleit (1995) modelled low DO and low F/M bulking. More applied models include the prediction of chlorine effect on microorganisms (e.g. Neethling *et al.*, 1987 and Caravelli *et al.*, 2003). Recent studies such as Cenens *et al.*, (2000_a, 2000_b), who developed a model which demonstrated the coexistence of filamentous bacteria for a wide range of dilution rates, concluded that the coexistence of floc formers and filaments could not be predicted in activated sludge systems with simple models based on the kinetic selection theory. The filamentous backbone theory had also to be considered. According to these authors, the kinetics, SRT and substrate feed concentration are key parameters in determining which type of bacteria will remain in the system.

According to Martins *et al*. (2004), other key factors to study solids separation problems are the interactions between bacterial morphology and bacterial physiology in gradient-governed microenvironments of activated sludge flocs. A promising modelling approach based on these interactions for biofilm processes is presented in Kreft *et al*., (2001) which probably represents a better approach for evaluating the competition between filamentous and non-filamentous bacteria.

Despite all the important advances in mathematical modelling, no reliable model capable of simulating the growth of filamentous bacteria or predicting other solids separations problems has been developed yet. Given that a well-formulated model must be based on well-known physical, chemical and/or biological processes, until the physical, chemical and biological interactions involving the solids separation problems is not properly understood, the development of mathematical models will not be reliable enough. Moreover, the results obtained in most of the referenced studies are difficult to be integrated in a general model capable of simulating or predicting the occurrence of solids separation problems, given that must results just relate the settleability of the sludge to the specific conditions under which these experiments were performed.

1.6.3 Use of Artificial Intelligence techniques

As mentioned before, the analytical testing results of the wastewater and sludge quality and, specially, the microbiological information of activated sludge (e.g., abundance of filaments) or qualitative observations about the process state (e.g., presence of bubbles in the settling test), supply a type of information that, once processed and interpreted by the operator, can help in monitoring and even preventing solids separation problems. However, it is not an easy task for plant operators and process engineers to acquire, integrate, and understand all this ever-increasing amount of collected data, which include both quantitative and qualitative information. Moreover, the actions they have to take to solve the problem depend largely on the specific characteristics of the treatment plant. A deeper approach is necessary to overcome the limited capabilities of mathematical modelling and conventional automatic control techniques when dealing with abnormal situations of complex systems, such as the solids separation problems, and to provide the level and quality of control necessary to always satisfy environmental specifications. The key is incorporating all these heterogeneous sources of data and knowledge, and, most importantly, reasoning to establish control actions based on the total information collected.

An automatic method based on the principles of human reasoning can help in this task. To solve these problems, **Intelligent Control** appeared as a promising field and had a significant impact in the process industry (Beltramini and Motard, 1988). The intelligent control term is applied to a control system that uses a knowledge-based approach (mainly expert systems and case-based reasoning) and/or soft computing techniques (mainly fuzzy logic and neural networks), as defined below. According to Stephanopoulos and Han (1996), intelligent controllers have the following features: (1) in addition to numerical algorithms, they also use logic, sequencing, reasoning, and/or heuristics; (2) are essentially non-linear controllers possessing autonomy which is broader than that of conventional controllers; (3) to carry out their expanded functionality, they rely on representational forms and decision-making procedures, which emulate human and/or biological systems.

Intelligent control techniques can be divided into knowledge-based and soft computing techniques. **Knowledge-based systems (KBS)** comprise several Artificial Intelligence (AI) techniques, which involve reasoning with some kind of knowledge (heuristic on expert systems, experience on case-based systems, etc.) to solve a problem. When applied to environmental issues, KBS receive different denominations such as Decision Support Systems (DSS), Knowledge-Based Decision Support Systems (KBDSS), Environmental Decision Support Systems (EDSS) or Multiple Objective Decision Support Systems (MODSS) (Cortés *et al.*, 2000). Among all the possible names, from here on we choose to use Knowledge-based Decision Support Systems.

Some of the problems with conventional control systems have been the focus, for the last few years, of much of the research effort in AI, especially in the development of KBDSS. KBDSSs have shown promising results due to their capabilities to represent heuristic reasoning and to work with large amounts of symbolic, uncertain and inexact data, as well as qualitative information which human operators comprehend best. They also permit implementation of human-like control strategies. Conventional or classical control methods cannot deal with these tasks. One of their main goals is therefore to assist decision-makers in choosing between alternative beliefs or actions by applying knowledge about the decision domain to arrive at recommendations for the various options (as when a doctor applies medical knowledge to decide on a diagnosis, or choose an appropriate treatment for a patient). In this case, a response may be a simple action, or a plan made up of sequences of actions over time. In more complex domains, these proposed plans may also require dynamic changes to intended actions if circumstances unfold unexpectedly, or if actions fail to achieve the intended goals.

Expert Systems (ES) are the core of this class of knowledge-based intelligent systems. Expert Systems (ES) are computer programs that attempt to emulate the reasoning process of experts making decisions to deal with a problem (Jackson, 1999). They use linguistic rules or conditional statements elicited from human experts to encode expertise (hence, they also are known as rule-based systems), which when chained in logical sequences can easily explore a situation and reach conclusions (e.g., IF F/M is decreasing and SVI is increasing THEN risk of filamentous bulking is true). The potential use of ES in the wastewater industry was introduced in the late 1970s and 1980s (e.g., Beck *et al.*, 1978; Johnston, 1985; Geselbracht *et al.*, 1986; Berthouex *et al.*, 1987). The first prototypes were based on simplified decision trees, which only included general knowledge available in textbooks, and applied a long and tedious algorithm, based on a series of questions and answers between the computer and the operator to infer the solution. In the 1990s, the literature describes more sophisticated and realistic ES in the field of wastewater treatment, applied to process condition diagnosis (e.g., Lapointe *et al.*, 1989; Gall and Patry, 1989; Krichten *et al.*, 1991; Bergh and Olsson, 1996), process design (e.g., Hudson *et al.*, 1997), decision aid (e.g., Maeda 1989; Patry and Chapman, 1989; Chan and Koe, 1991; Ohtsuki *et al.*, 1998), process selection (e.g., Okubo *et al.*, 1994; Yang and Kao, 1996), process optimization (Huang *et al.*, 1991), and operation and control (e.g., Ladiges and Kayser, 1993; Ozgur and Stenstrom, 1994; Zhu and Simpson, 1996; Serra *et al.*, 1997; Furukawa *et al.*, 1998; Ashraf Islam *et al.*, 1999; Vouros *et al.*, 2000; Puñal *et al.*, 2002; Baeza *et al.*, 2002;). However, not many of theses approaches have been found to be applied in the domain of solids separation problems (Parker and Parker, 1989; Barnett *et al.*, 1993; Ng *et al.*, 2000 and Comas *et al.*, 2003).

Olsson *et al.* (1998) pointed out though that these applications never really succeeded because they were too complex, and the available knowledge could not be captured in reliable models and advisory systems. In fact, a realistic and rigorous evaluation of these expert systems under day-to-day plant operational conditions could not be done since most of these applications 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.

The second widely used knowledge-based intelligent control technique is Case-Based Reasoning. **Case Based Reasoning Systems (CBRS)**, or simply Case-Based Systems, are computer programs that use past experiences to solve new problems that arise in the process. They automatically identify similarities of process conditions (cases) to previous conditions and reuse results and experience from previous particular situations that have affected the process performance to solve the current problem. Thus, by using this technique it is possible to solve brand-new situations similar to previous ones with less effort than with other methods, which start building up new solutions without the benefit of previous solved cases of similar nature (Kolodner, 1993; Aamodt and Plaza, 1994). Due to their optimal results and to their relative easy development, Case-Based Systems have been used recently in several fields such as fault diagnosis, medical detection, equipment selection (Kraslawski *et al.*, 1995), information retrieval from historical meteorological databases (Jones and Roydhouse, 1995), planning forest fire fighting (Avesani *et al.*, 2000), prediction for rangeland pest management advisories (Branting *et al.*, 1997) or design for process engineering (Surma and Brauschweig, 1996). CBS has also been proposed as a support tool in the activated sludge process (Krovvidy and Wee, 1993; R-Roda *et al.*, 1999; Sànchez-Marrè *et al.*, 1999). Regarding solids separation problems, CBRS have been scarcely applied as a decision support tool. Only one example of application can be found in Wiese *et al.* (2003).

Finally, the other important group within Intelligence Control is the group of **Soft-computing Techniques**. According to Zadeh (1965), soft computing differs from conventional (hard) computing in that, unlike hard computing, these techniques are tolerant of imprecision, uncertainty and partial truth. Indeed, the role model for soft computing is the human mind. Among all the principal constituents of soft computing, Fuzzy Logic and Neural Networks are the techniques generally applied in the supervision of wastewater processes.

Fuzzy logic control is in essence rule-based control but without the crispness of the ES controllers. It applies a kind of logic that recognizes more than simple true and false values. With fuzzy logic, propositions can be represented with degrees of truthfulness and falsehood. For example, the statement "there is foaming" might be 100% true if the reactor is fully covered, 80% true if there is lots of foam, 50% true if there is a little foam, and 0% true if there is no foam at all. In soft computing fuzzy logic control, knowledge of the equations describing the process dynamics is not essential. A fuzzy controller for the cement industry was indeed the first intelligent control system implemented and today such controllers are applied to control hundreds of industrial plants worldwide. Their application in the process industry has led to significant improvements in product quality, productivity, and energy conservation. In the domain of wastewater treatment, fuzzy control has been efficiently applied in real facilities to control different process parameters such as the aeration, RAS flow or the addition of chemicals. The following are some examples of application: Manesis *et al.* (1998); Rammacher and Hansen (2000); Bongards (2001); Chen *et al.* (2001) and Fiter *et al.* (2005a).

On the other hand, **Neural Networks (NN)** are soft computing systems made up of a number of simple, highly interconnected nodes, which process information by their dynamic state response to external inputs (Hecht-Nielsen, 1990 and Hertz *et al.*, 1991). NN are known to be very effective in capturing the non-linear relationships existing among variables in complex systems and to be resilient to noise. One limitation of NN technology though, is that they act as a black box, (i.e. the output is a function of the input and the set of weighted vectors of neuron interconnections). Therefore, they do not help improve the heuristic understanding of operational problems. Consequently, it is difficult for the user to introduce his/her knowledge to the network and to find an intuitive explanation of their results. Some examples of NN controllers and prediction tools to WWTP can be found in (Hunt *et al.*, 1992; Côte *et al.*, 1995; Steyer *et al.*, 1997; Zhao *et al.*, 1999; and Hong *et al.*, 2003). Some applications have been developed in the domain of solids separation problems (Capodaglio *et al.*, 1991 and Belanche *et al.*, 2000). Despite these promising applications, some experts fear that even neural networks will not reach success in wastewater treatment control due to the lack of a physical interpretation of their parameters (Olsson *et al.*,1998).

The application of a single conventional, a single knowledge-based or a single soft-computing technique can also present some limitations because real problems are too complex or risky to be handled by individual systems, which are limited in solving real world problems. In contrast, most real world problems require interdisciplinarity. A new generation of intelligent control systems is emerging based on hybrid architectures that combine varius knowledge-based techniques with soft computing and other conventional tools. The reulting **Knolwedge-based Decision Support Systems**, reduce the time in which decisions are made, and improve the consistency and quality of those decisions to guarantee the quality of the plant's treated effluent (Wen and Vassiliadis, 1998). They simulate the problem-solving behaviour of a group of experts who apply various problem solving methodologies and expertise in relation to a given formulated problem. KBDSS allocate the detailed engineering to numerical computations, while delegating the logical analysis and reasoning to supervisory intelligent systems (Stephanopoulos and Han, 1996).

Despite all these important advances, the use of the individual knowledge-based systems cited above solves only certain aspects of the overall WWTP management process. A deep approach, therefore, is still necessary to obtain an optimal WWTP management, especially when complex problems such as the solids separation problems need to be handled. Some examples of integrated intelligent systems applied to wastewater treatment plants can already be found in the literature (Sànchez-Marrè *et al.*, 1996; Rao *et al.*, 1998; Ohtsuki *et al.*, 1998; Ashraf Islam *et al.*, 1999 and Poch *et al.*, 2004). None of them, however, has been explicitly designed for the specific management of solids separation problems which, concerning the general complexity of wastewater processes, are the most troublesome problems in the process.

CHAPTER 2

OBJECTIVES

The main objective of the present thesis is to develop and validate a Knowledge-Based Decision Support System specifically designed to support plant operators in handling solids separation problems of microbiological origin occurring in activated sludge systems. To achieve this main purpose, the system must comprise the following characteristics:

- Its application must be *viable* and *realistic*. The implementation of the system into full-scale WWTPs must be feasible, the results must be easy to understand, and its routine execution must motivate plant operators.

- It must be *dynamic*. The goals and results of the system have to be continuously adjusted according to the own evolution of the process. The reasoning process followed by the system has to match the dynamics of microbiologically related solids separation problems in activated sludge systems.

- It must be *intelligent*, based on the principles of human reasoning, and including both expert and experiential knowledge.

The fulfilment of the main objective involves the design, development, implementation and validation of the Knowledge-Based Decision Support System, which implies the following sub-objectives:

- Study of the solids separation problems background at a closer (local) scale in order to approach the real magnitude of the problems and to evaluate the feasibility of applying and using the proposed decision support tool. Catalonia is chosen as the local area of study.

- Use of the Domino Model to conceptually design the dynamic KBDSS.

- Development of the Knowledge-Based Decision Support System. This task involves firstly, the selection of the most appropriate techniques from Artificial Intelligence that can provide the system with the required intelligence and, secondly, the development of all the required components for the proper implementation of the KBDSS.

- Implementation of the developed techniques into a structured KBDSS architecture, including the codification of the system into a specific shell, the G2-Gensym environment.

- Validation of the Knowledge-Based Decision Support System in a full-scale WWTP.

CHAPTER 3

APPROACH TO SOLIDS SEPARATION PROBLEMS IN CATALONIA

With the aim of developing a KBDSS that is viable and realistic enough to be implemented in a full-scale WWTP, the domain where it is going to be applied, where the system is expected to accomplish its main goals, must be thoroughly studied. In chapter 1, an exhaustive review of the activated sludge process characteristics and the occurrence and control of solids separation problems has been performed from a theoretical point of view. It has been reported, for example, that the occurrence of solids separation problems can vary geographically from one country to another, mainly depending on the wastewater characteristics or on the type of treatment or process used to treat the water.

In addition, it has been found that several monitoring tools and control methods are nowadays available to manage solids separation problems. The use or application of one or another control technique also differs from one plant to another, mainly depending on two main restrictions: objectives of treatment and cost limitations.

According to this and with the purpose of developing a viable and realistic KBDSS, a closest study of the domain where the system is going to be applied has been carried out. Catalonia has been chosen as the scale where the domain is going to be approached.

3.1 Description of the Study

Catalonia is a region of Spain located at the North-East of the country, bordering on the North with France. During 2003, 13 WWTPs of Catalonia (see location in Figure 3.1) were visited in order to study: (1) the occurrence of solids separation problems, and (2) the management approach to these problems (monitoring and control tools). The study was performed in two main steps: (1) personal visits to the specific facilities to study the main characteristics mentioned above, and (2) meeting session or symposium where all the plant managers of the studied plants were invited to discuss the preliminary results of the study. The 13 WWTPs of Catalonia submitted to study are the following: Figueres, Girona, del Terri (Banyoles), Celrà, Cassà de la Selva-Llagostera, Vic, Taradell, Montornès, La Llagosta, Granollers, Manlleu, Manresa and Lleida.

The compilation of information for each of the visited WWTPs was supported by the design and administration of a personal questionnaire whose main objective was to collect the answers to the following questions.

Regarding the WWTP characteristics:

- design and actual capacity of the plant such as design flow or equivalent inhabitants
- type of activated sludge process and flow schematic
- characterisation of treated wastewater in terms of average wastewater composition: COD, BOD, TSS, forms of N and P, and % of industrial wastewater and types of industries which contribute to this %

- operational conditions such as hydraulic retention time, sludge age, mixed liquor concentration, F/M ratio, average dissolved oxygen set-point, sludge volume index or average summer and winter temperatures

Regarding the occurrence and control of solids separation problems:

- type and occurrence of solids separation problems, including the average length of the episodes and the observation of seasonal changes, common filamentous species and possible causes

- list of parameters that can be used to monitor solids separation problems (availability and frequency of measurement), control methods commonly used and application of information technology techniques

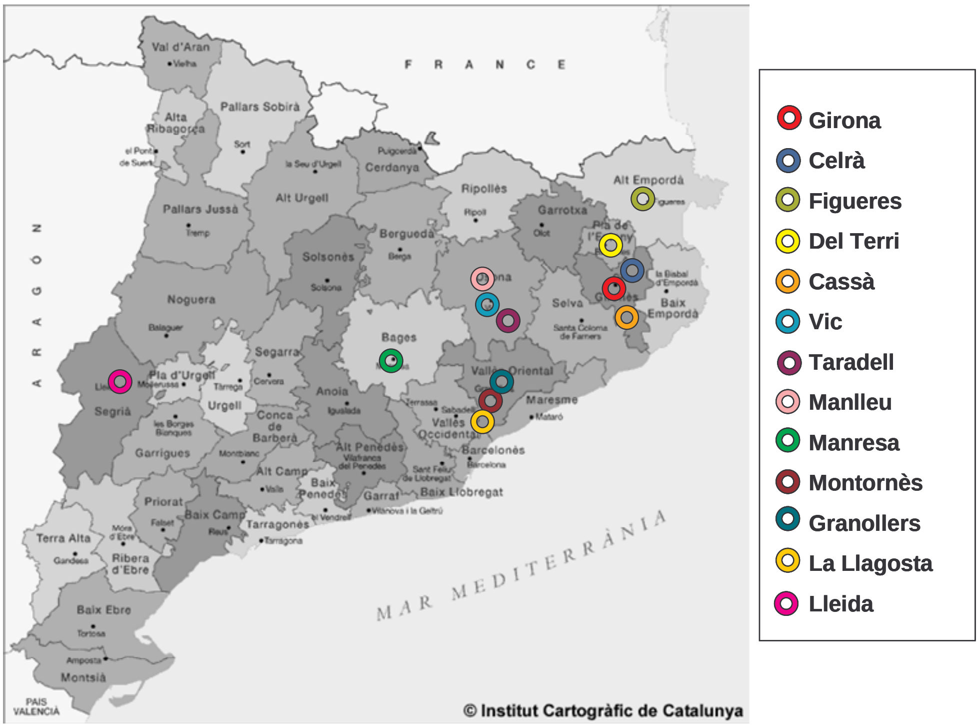


Figure 3.1 Location of the 13 WWTP of Catalonia submitted to study

Based on the answers to the questionnaire, a general list of conclusions was compiled and discussed during the symposium, where representatives of 12 of the 13 studied WWTP and 8 representatives of the Laboratory of Environmental and Chemical Engineering at the University of Girona, included myself, were in attendance. There, the preliminary conclusions were discussed and amplified with some generic and personal comments. Moreover, for some WWTPs we were provided with data from the process; these were used to perform some statistical studies which allowed us to examine correlations between possible causes and occurrence of solids separation problems. A summary of both the conclusions compiled from the interviews and the symposium, and an outline of the results of the statistical analysis of data from the WWTPs of Vic, Taradell and Figueres are presented in the following section.

3.2 Results of the Study

Regarding the type of treatment followed in each one of the 13 facilities under study, 54% were designed to remove COD, 15% to remove COD and Nitrogen, and 31% to perform complete nutrient removal (COD, N and P). Concerning the configuration of the reactor, 38% of the plants used plug-flow configuration, another 38% used oxidation-ditch reactors and just 24 % of them were completely mixed reactors. Finally, 70% of the plants received an important contribution of industrial wastewater (more than 50% of the treated wastewater came from local industries) and the remaining 30% treated only municipal wastewater. Table 3.1 summarizes some of the wastewater characteristics and operational parameters of some of the studied plants, including the average values for inflow, influent COD, SVI, F/M and SRT.

Table 3.1 Average yearly values for some of the wastewater characteristics and operational parameters of 11 studied plants (* F/M in kg COD/kg MLSS,d)

WWTP	Type of treatment	Inflow (m^3/d)	COD-inf. (mg/L)	SRT(d)	F/M*	SVI(mL/g)
Figueres	COD removal	17494	813	3.0	0.55	440
Girona	COD removal	45455	504	7.7	1.60	205
Montornès	COD removal	27230	736	7.0	0.17	175
Cassà	COD removal	2500	637	22	0.01	188
Manresa	COD removal	27122	504	8.9	0.25	144
Granollers	COD and N removal	22124	766	8.3	0.40	120
Banyoles	COD and N removal	10350	1033	18	0.10	154
Manlleu	COD, N and P removal	7400	525	10	0.10	300
Taradell	COD, N and P removal	1878	750	19	0.11	403
Vic	COD, N and P removal	22302	1210	11	0.12	186
Celrà	COD, N and P removal	2000	1200	13	0.03	65

It can be observed the existing variability regarding both the type of treatment and the plant capacity (inflow), varying from less than 2000 up to more than 45000 m^3/d. Likewise, differences between operational parameters such as SRT and F/M are also evident. Finally, the average value for SVI is higher than 150 mg/L in 8 of the 11 plants which may be symptomatic of some solids separation problems such as bulking or foaming. On the other hand, just the Celrà WWTP shows an average value for SVI lower than 75 mg/L, which could be symptomatic of some kind of sludge deflocculation.

A summary of both the conclusions related to solids separation problems, compiled from the interviews and the discussion session, and the most relevant results of the statistical study performed at the WWTPs of Vic, Taradell and Figueres are presented in the following sections.

3.2.1 Occurrence of solids separation problems

Regarding the question concerning the occurrence or frequency of the different solids separation problems, filamentous bulking and foaming are, by far, the most common and bothering separation problems in the process for the studied plants. Filamentous Bulking was considered as a very bothering problem in 85% of the studied plants. Biological Foaming was almost as important as bulking since 75% of the plants suffered from upsetting episodes of uncontrollable foams. Rising sludge seems to be less problematic, since just around 40% of the plants suffered from important loss of solids due to rising sludge in clarifiers, especially during summer. Finally, less than 25% of the plants considered the problems of viscous bulking, dispersed growth or pin-point floc to be a common and problematic source of settleability problems in the facility. On most occasions, the occurrence of severe episodes of solids separation problems had lead to the loss of solids through the effluent which, apart from affecting the efficiency of the activated sludge process due to the loss of biomass, had affected the effluent quality, occasionally exceeding legal standards (TSS: 35 mg/L and COD:125 mg/L).

As commented before, the origin of most solids separation problems is usually an imbalance between the different groups of the microfauna (generally floc-forming and filamentous bacteria). This fact implies that usually, unless the cause of this imbalance is identified and removed, the problem will persist in the process, causing long-lasting episodes of settleability problems. This has been confirmed by studying the average length of the common separation problems in the plants under analysis. The results are the following:

⇒ The studied plants used to undergo 1 to 3 important episodes of filamentous bulking per year with an average duration of 1-3 months.

⇒ Biological foaming episodes also presented an average frequency of 1 to 3 important episodes per year, although their average extent was shorter than bulking and varied from 15 days up to 1-2 months.

⇒ The facilities that experienced viscous bulking problems used to suffer from their nuisances almost permanently in the plant. The most severe episodes though used to last up to 1 month.

⇒ Finally, the problems of rising sludge, dispersed growth and pin-point floc were just sporadically present in the facilities and their average duration was usually less than 1 month. Episodes of deflocculation could even last less than 1 week.

3.2.2 Common filamentous bacteria

The most common filamentous bacteria in the studied plants were, in order of incidence, the following:

1) Nocardioforms (one of the predominant bacteria in 72% of the plants)

2) Type 021N (predominant in 54% of the plants)

3) *M. parvicella* and Type 0041(predominant in 45% of the plants)

4) *N. limicola* (predominant in 36% of the plants)

5) *H. hydrossis*, Type 1863, *Thiothrix* sp., *S. natans*, Type 0675, Type 1851 and Type 1701 (predominant in 27% of the plants)

It can be concluded that the most predominant filamentous bacteria in the 13 plants of Catalonia are: nocardioforms, Type 021N, *M. parvicella*, Type 0041 and *N. limicola*. If we compare these approximated results with the conclusions of several surveys performed in different countries (see Table 1.1), we can notice that the predominat filaments found in the studied plants are more similar to the filaments found in countries such as Australia, USA or Italy than to the filaments found in South-Africa or The Netherlands. On the other hand, according to the results presented in Table 1.1, it can be shown that the 5 most predominant filaments around the world are nocardioforms, Type 021N, Type 0041, *M. parvicella* and Type 0092. Four of these species correspond to the species commonly found in the surveyed plants of Catalonia. Type 0092 is the only filament not found in these studies.

3.2.3 Possible causes of solids separation problems

The analysis of the average values regarding most operational parameters and wastewater characteristics of the studied plants and the discussion with plant operators have let us compile a list of the possible causes of solids separation problems in these plants. For filamentous and viscous bulking or biological foaming, the suspected causes are:

- ⇒ Sudden variations in the organic load, especially at low values of F/M.
- ⇒ Variations in the dissolved oxygen concentration. Usually due to some limitations of the aeration system whenever sudden increases in the organic load do not allow the achievement of the oxygen requirements.
- ⇒ Low temperatures. Especially in winter when the temperature can fall below 10 ºC.
- ⇒ Nutrient deficiency. Specifically in those plants which collect a significant percentage of industrial wastewater.
- ⇒ Readily biodegradable substrate. Usually discharged from industries such as food and dairy industries or slaughterhouses.

The general and common cause for dispersed growth or pin-point floc is the arrival of some toxics to the process (in general heavy metals) which affect the stability of microfauna. Other possible causes are the undesired results of overchlorination or the consequences of an excessive purge which decreases the sludge age excessively.

Finally, the temporary rising episodes usually appear in plants that ordinarily do not remove nitrogen but that during some periods (especially during warm days) undergo some conditions (increase in water temperature and sludge age) which favours nitrification within the reactor, and additional circumstances (high residence time in clarifiers and anoxic conditions), which subsequently favour denitrification in clarifiers.

Figures 3.2 and 3.3 depict the most relevant correlations found between some of the possible causes of bulking and foaming and their incidence on sludge settleability at the Taradell WWTP. In Figure 3.2a, a clear correlation between temperature and sludge settleability can be observed. The periods where the DSVI is very high (>400mL/g) correspond to those months where the temperature of wastewater is very low (<12ºC). On the other hand, the correlation of DSVI and SRT is not as clear as with T, but in general it can be noticed that the episodes with higher SVI are usually observed after periods of time when the SRT has been increased over 10 days. These suppositions can also be related with the fact that the predominant filament during those periods is usually *M. parvicella*, which causes both situations of filamentous bulking and foaming. This filament is usually encouraged by temperatures < 12ºC and conditions of low F/M and high SRT (> 8 days).

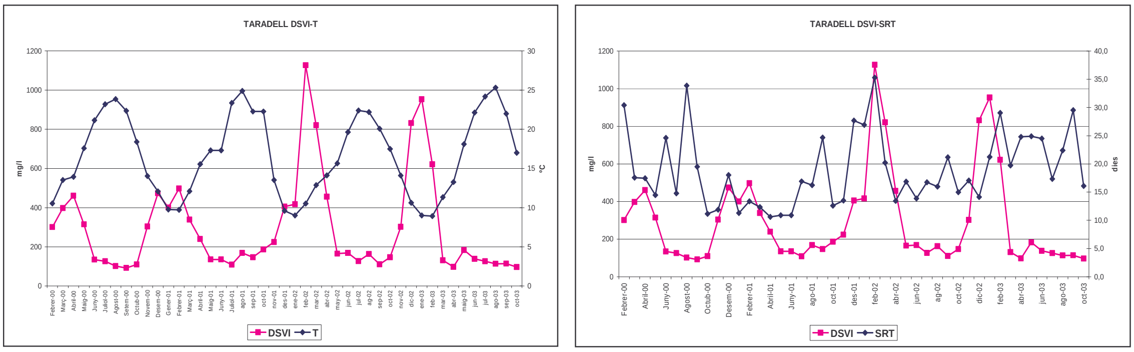


Figure 3.2 Correlations between (a) SVI and Temperature and (b) SVI and SRT at the Taradell WWTP

Figure 3.3 describes some possible correlations between the dissolved oxygen concentration and the F/M ratio with the sludge settleability at the Figueres WWTP.

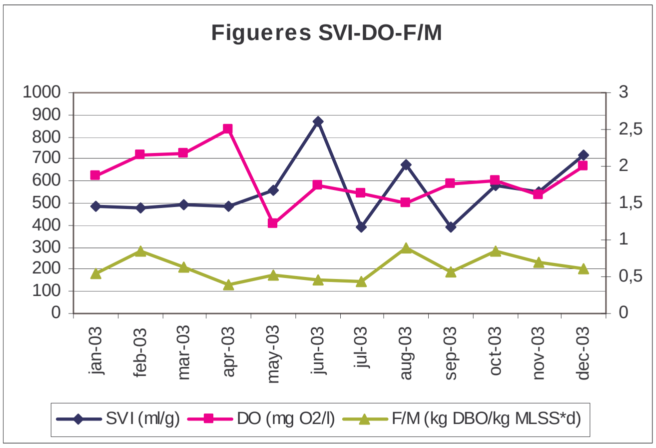


Figure 3.3 Correlation between DO and F/M with SVI

In general, it can be observed that the episodes with higher SVI (>550 mL/g) are correlated with periods of low DO concentration (< 2.0 ppm). Regarding the F/M ratio, a general conclusion can not be stated since very low values are generally not observed. A possible correlation might exist between sudden increases of F/M as the one observed during August and October, and the increase of SVI.

Finally, with the data from the Vic WWTP of 2003, a statistical study using the Principal Components Analysis (PCA) was performed in order to find possible correlations between a set of 20 variables from the process. The selected variables are: Inflow (q_e), TSS-influent (ss_e), TSS-effluent (ss_s), COD-influent (dqo_e), COD-effluent (dqo_s), BOD-influent (dbo_e), BOD-effluent (dbo_s), TKN-influent (tkn_e), TKN-effluent (tkn_s), total P-influent (p_e), total P-effluent (p_s), BOD/COD (c_c), COD/N (c_n), COD/P (c_p), MLSS (mLss), V30 (v30), SVI (ivf), DO (O2), T (t) and F/M (CM). Figure 3.4 represents the factor scores obtained for the selected 20 variables.

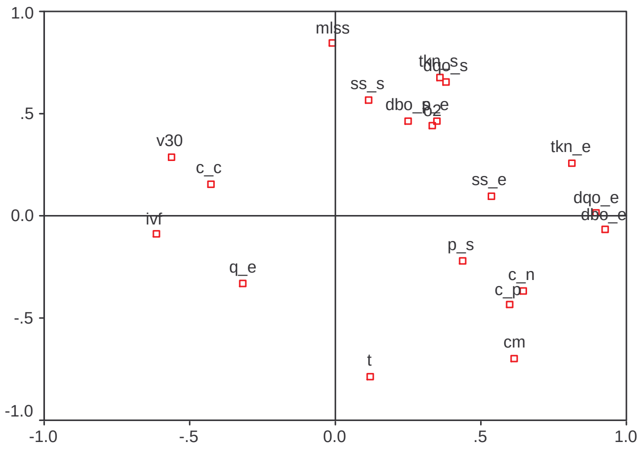


Figure 3.4 Graphical representation of the factor scores obtained by the Principal Components Analysis.

From the graphical interpretation it can be concluded that: (1) there is a direct correlation between SVI and V30 and an indirect correlation between these 2 parameters and the F/M. This means that in general, the lower the F/M, the higher the SVI; (2) there is a direct correlation between SVI or V30 and effluent-TSS, effluent-BOD and effluent-COD, meaning that in general, the higher the SVI, the higher the loss of solids and finally (3) an indirect correlation has been found between SVI and COD/P, which means that the deficiency of P can be related with an increase in the SVI.

Regarding the seasonal changes of bulking and foaming episodes, some correlations have been found between the season and the occurrence of one or another type of filament. Figure 3.5 represents the general incidence of filamentous bulking and foaming at the studied plants related to the season (months) and the specific prevalence of the commonest filamentous bacteria through the year.

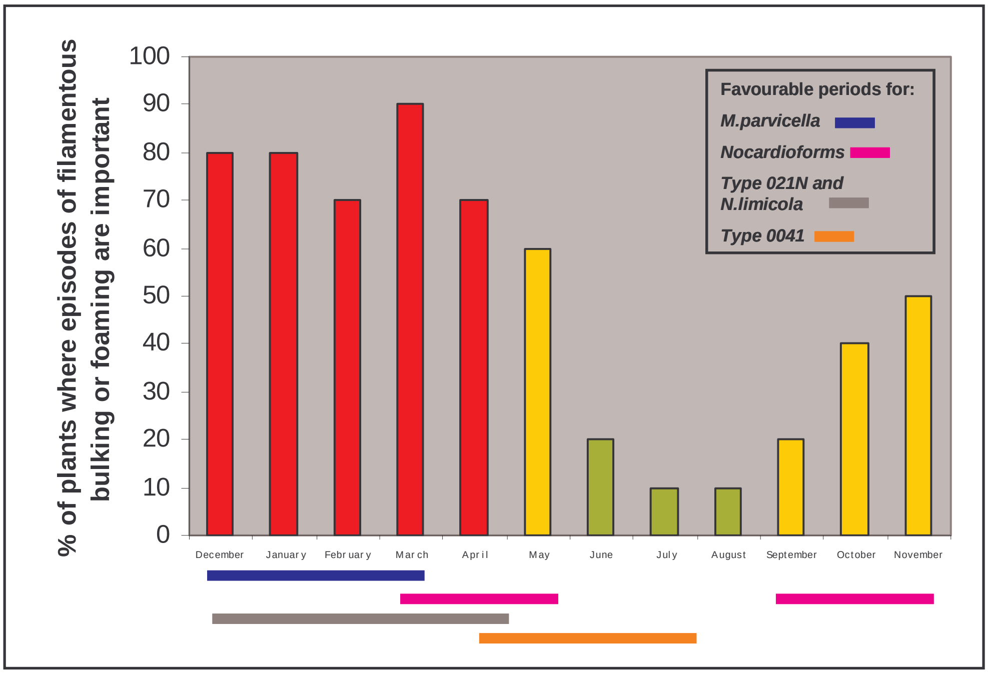


Figure 3.5 Percentage of plants suffering from filamentous bulking and foaming related to the season. The red bars represent the months where the incidence of these problems is higher while the green ones represent the months with less problematic episodes. The coloured lines below represent the periods where *M.parvicella*, nocardioforms, Type 021N, *N.limicola* or Type 0041 usually proliferate.

In general, it can be concluded that the most problematic months regarding filamentous bulking and foaming incidences include the winter and spring seasons (from December to May). In second term, the period of September-November (autumn) is also significant since some filamentous bacteria (especially nocardioforms) proliferate during these months. Finally, it seems that the summer season is the period where less bulking or foaming episodes occur.

3.2.4 Monitoring of solids separation problems

A review of the most used parameters for monitoring activated sludge solids separation problems has been performed for the studied plants.

Regarding the parameters used to identify possible situations of solids separation problems, the following has been found.

- **Sludge settleability**. The analysis of the sludge settleability is usually performed by means of the V30 test (Figure 3.6a) and the subsequent calculation of SVI. 100% of the plants calculate SVI at least once a week and just 50% of them calculate SVI daily (mainly because MLSS are not daily calculated). The use of more accurate measurements such as DSVI is only used by 3 of the studied plants and none of them calculate the SSVI. On the other hand, the sludge blanket depth is not widely used as an indicator of the sludge settleability. It is regularly calculated by just 15% of the facilities. This parameter is generally measured by using a plastic pipe with an anti-return valve (Figure 3.6b). The meter is placed at the clarifier center well and it measures the height of the sludge blanket.

  On the other hand, the measurement of the zone settling velocity is not calculated by any of the plants. Although most of them calculate the upflow velocity (outflow rate/surface area of clarifiers), it does not provide the same kind of information as the ZSV. Likewise, the use of on-line settlometers is not widespread among the studied plants.

Figure 3.6 Sludge settleability measurements: (a) V30 test and (b) sludge blanket measurement

Finally, the use of other kind of information that can be mined from macroscopic observations such as the qualitative interpretation of the V30, or other direct observations on clarifiers or reactors such as the presence of foams by means of different **foaming tests**, is not routinely performed by any of the plants.

- **Microscopic analysis**. 85% of the studied plants perform microscopic analysis with an average frequency of once every 10-15 days (just 15% of these plants measure microscopic characteristics almost daily). 15% of the plants never perform microscopic observations. The low frequency of the microscopic analysis is usually attributed to the lack of time or skills of the plant operators to perform the observation. In fact, it needs to be noticed that in 30% of the plants the observation is done by people external to the facility.

  The type of measured information varies from one plant to another although the most common measured parameters are: general floc characteristics, general overview of the main groups of the microfauna, filamentous bacteria abundance and identification of the predominant filamentous bacteria. Filaments abundance and identification are usually determined by using the subjective scoring methodology proposed by Jenkins *et al.* (2003) and the identification of species by using dichotomous keys and general manuals of the activated sludge microbiology. Most advanced measurements such as the TEFL or the use of more sophisticated techniques such as FISH, are not used at any of the plants.

  The common routine microscopic analysis of most of the plants, however, does not consider the registration of the valuable information obtained from microscopic analysis into the general analytical worksheet where the rest of data from the process is compiled.

- **Analytical data and operational parameters**. Regarding the analytical data that can be used as indicative of solids separation problems, the measurement of effluent TSS and COD is performed by all the studied plants since these are some of the required parameters that the administration uses to monitor plant efficiency. More than 50% of the plants even measure these parameters daily. On the other hand, the analytical determination of effluent turbidity is just measured by 25% of the plants.

  The parameters used for monitoring the possible causes of separation problems, together with their measurement frequencies, are the following:

  -***Sludge age***. This parameter is calculated by 70% of the plants almost every day. The sludge age is calculated by means of the SRT equations, which is usually preferred over the MCRT.

  -***F/M ratio***. It is calculated by 85% of the plants with a frequency of 3-4 days per week. It is usually calculated by using COD instead of BOD in order to get values more often.

  -***COD/N/P***. This ratio is not a parameter commonly measured by most of the plants. Even though they have the values for nutrients and COD, only 30% of the plants calculate this ratio. The frequency of measurement is once a week (usually depending on the frequency of measurement of nutrients at the primary effluent).

-*DO*. The dissolved oxygen concentration within the reactors is daily measured on-line by all the plants.

-*OUR*. The oxygen uptake rate is not regularly measured by any plant.

-*pH*. It is also measured by all the plants but usually at the plant inflow, not at the primary effluent. More than 50% of the plants measure it on-line and the rest measures it analytically.

-*T*. The wastewater temperature is calculated on-line by almost 100% of the plants. They usually use the same location used for the DO sensor.

-*Sulphide*. This parameter is calculated by just 15% of the plants, which usually analyse the influent sulphide once a week.

-*Readily biodegradable substrate*. The % of COD that is readily biodegradable is only calculated at one of the studied plants. It is calculated by the floc/filtration technique with $ZnSO_4$.

3.2.5 Common control methods

The control methods most commonly used whenever the studied plants suffer from solids separation problems are summarized in Table 3.2. The percentage of plants that use each one of the non-specific and specific methods is reported. In general, it can be stated that the use of more advanced or innovative control methods such as the addition of ballasting compounds or the addition of ozone as an oxidant is rare, mainly because of the cost limitations or the existence of constraints in the plant infrastructures.

In Figure 3.7, a common non-specific control method, the physical removal of foams by means of a hose, and an unusual specific control method, the use of an anoxic selector for filamentous bulking control, are represented.

Table 3.2 Specific and non-specific control methods used by the studied plants to solve solids separation problems

Control Method	% of WWTPs that habitually use this method
Non-specific Control Methods	
Chlorination of RAS	38%
Addition of coagulants or flocculants	30%
Physical removal of foams	23%
Addition of antifoam compounds	23%
Recirculation of RAS	15%
Specific Control Methods	
Control of the DO concentration ≥ 2ppm	72%
Control of the SRT or the F/M by means of waste flow	45%
Nutrients addition (in general phosphoric acid or urea)	15%
Use of selectors	8%

Figure 3.7 (a) non-specific control method for foam removal: use of draught water to break the scum on the surface of clarifiers and (b) specific control method: use of anoxic selector for the control of filamentous bacteria at Lleida WWTP

3.2.6 Use of information technology

At present, 92% of the studied plants are equipped with a SCADA (Supervisory Control and Data Acquisition) system, although apparently these are used more often for data acquisition than for control. Regarding process control, the most used control strategies (used by all the studied plants) are:

- ⇒ Automatic control of the biological aeration. Usually controlled by PI controllers, which automatically regulate oxygen supplied as a function of the established set points

- ⇒ Control of the return activated sludge flow rate proportional to the influent flow rate

On the other hand, 3 of the studied plants perform an optimisation of Nitrification/Denitrification by using the on-line DO and RedOx potential measures to maximise nitrogen removal through recirculation flow rate control and establishment of DO set point. Finally, one of the plants also uses a control algorithm based on fuzzy rules for the automatic establishment of the RAS rates as a function of SVI and hydraulic loading in the clarifier.

The use of advanced tools of Information Technology such as model-predictive control or decision support systems is even more limited. Nonetheless, three of the studied plants (Granollers WWTP, Montornès WWTP and La Llagosta WWTP) use a Decision Support System (also developed by the LEQUIA group in previous research projects) for the control of general process problems.

3.3 Discussion

The occurrence of solids separation problems in the Catalan WWTPs is as important and disturbing as in other countries of the world. However, some characteristics concerning problem specificities such as predominant filaments, causes or frequency and length of episodes are slightly different mainly due to wastewater characteristics, process parameters and type of treatment.

The most common parameters generally used to monitor sludge settleability are the V30 test and the calculation of the SVI. These are the most used parameters generally because the results are rapidly obtained. The use of more accurate measurements such as DSVI, SSVI, sludge blanket or the calculation of the ZSV are generally not measured due to these are high time-consuming techniques.

Microscopic analysis is commonly used as a monitoring tool mainly to study floc characteristics and abundance of filaments. The main shortages are, on the one hand, the low frequency in the measurement of these variables, mainly attributed to the lack of skills or time of plant operators, on the other hand, the lack of standardization and regularity in registering the measured parameters. In general, the different parameters are measured, but they are not recorded with the rest of data obtained from wastewater analysis or other parameters of plant operation. In this sense, the development of standard macroscopic and microscopic worksheets to standardize the compilation of qualitative information is suggested. The objective of these worksheets must be to collect (in a standard way) all the valuable information that can be gathered from both macroscopic observations (for example from the settleability test or from the direct observation of reactors and clarifiers) and microscopic analysis. The worksheets must be user-friendly and not time-consuming. A frequency of once a day (at least during working days) is recommended.

Regarding the group of analytical data and operational parameters that can be used to monitor the evolution of separation problems and their possible causes, it can be stated that in general, most of the required variables are commonly measured at least once a week. Sulphide concentration and soluble COD are the exception, given that those parameters are usually not measured in most plants. As a consequence, it could be recommended that, in order to detect or monitor each one of the possible causes, the frequency of measurement of such key variables should be increased (at least during problematic episodes).

It has been confirmed that there are some preferences in using one or another control method to handle the different solids separation problems, usually depending on the cost of their application and on the characteristics or limitations of the facility. In general, plant operators usually apply those plans that in the past were successful to restore the problematic situation. Nevertheless, the consideration of alternative plans is also very important, especially when the ongoing control methods do not produce the expected improvements in the process.

The good results obtained in some of the studied plants from the use of DSS in controlling general problems in activated sludge systems, together with the observed necessities of supervisory control tools to support plant operators in handling solids separation problems, makes that the use of a KBDSS could favour the management of solids separation problems in all of the studied plants of Catalonia.

This study allowed us to develop a most reliable and realistic KBDSS, designed according to the existing demands of activated sludge systems in handling solids separation problems. Nevertheless, despite this approximation, the developed KBDSS must be able to be straightforwardly extended to any activated sludge system all over the world.

CHAPTER 4

DYNAMIC REASONING

Apart from being viable and realistic, it is important to recall the importance of defining a new dynamic approach of the system in order to achieve its main purpose of efficiently supporting plan operators in handling solids separation problems. The approach followed in building the system must match the dynamism of the domain where it is going to be applied. Thus the proposed KBDSS must be dynamic and must use dynamic reasoning.

As commented in section 1.6.3 *Use of Artificial Intelligence Techniques*, there are a number of approaches founded on knowledge-based techniques that have been proposed to improve WWTP operation. The experience gained from these applications has demonstrated that the main bottleneck in developing feasible KBDSS is the representation of all the necessary knowledge from the domain (especially that which is not readily quantifiable) into the Knowledge Base, together with the proper definition of the reasoning process that the system must follow to make a decision. Failure to define properly these key components can directly compromise the feasibility of the system. According to Olsson *et al.*, (1998) the excessive complexity of the system and the inefficiency in capturing all the necessary knowledge into the Knowledge Base are the main reasons why most applications of KBDSS have not really succeeded when implemented into full-scale facilities.

The additional complexity and the corresponding necessities imposed by the dynamic nature of solids separation problems exacerbate the limitations in developing simple KBDSS that follow from an inappropriate approach. In this case, a suitable approach should adjust the reasoning process to the continuously evolving environment and, in case the objective of the system was to recommend a control strategy to solve a specific situation, the proposed plans should also be adjusted to the evolution of the problems.

Taking into account all these necessities, and following the suggestions of Fox and Das (2000) to define an efficient KBDSS, a previous conceptualisation phase was considered in which a conceptual design of the system was set up. The following concepts were tackled during this conceptual design: (a) domain understanding (definition of the knowledge base and its correlation with the reasoning process); (b) characterisation of the reasoning process, and (c) determination of the different objectives. This conceptual phase was followed by the steps of development, implementation and validation of the KBDSS, which are analogous to those followed in conventional software engineering.

The present chapter reports on the use of the Domino Model as a theoretical model for the conceptual design of our KBDSS. A complete description of the model is presented, together with a detailed depiction using a specific case-study from the solids separation problems domain. Finally, the feasibility of the model in facilitating the design of KBDSS in complex domains, such as the dynamic solids separation problems, is discussed.

4.1 Description of the Domino Model

The Domino Model is a framework originally introduced by Fox and Das (2000) to model the interaction among different kinds of agents for problem solving in complex domains. According to the authors, the usability of the model stems from the fact that the development of an efficient Knowledge-Based Decision

Support System operating in complex domains needs to be guided by a high level of understanding of the domain and a well-structured conceptualization of its different components.

In the proposed model, an **agent** can be defined as a system that presents specific objectives and usually operates in environments that are subject to constant and often unexpected changes (dynamic domains). Its functions (such as data acquisition, decision making, plans performance or action scheduling) must work accurately together to ensure that its goals are met. In addition, a rational agent must be able to make only those deductions and take the decisions that are consistent with its prior beliefs and that will maximize the expected value of its actions. Agents are capable of accepting new tasks or aims, and of improving their actions by means of new knowledge about the domain. They must adapt themselves to the changes in the environment and update the relevant knowledge they possess.

The Domino Model (Figure 4.1) provides an intuitive visualization of all the functions or capabilities that the system may require and their interrelationships. The nodes in the model represent **knowledge bases** and the arrows symbolize **computational functionalities**.

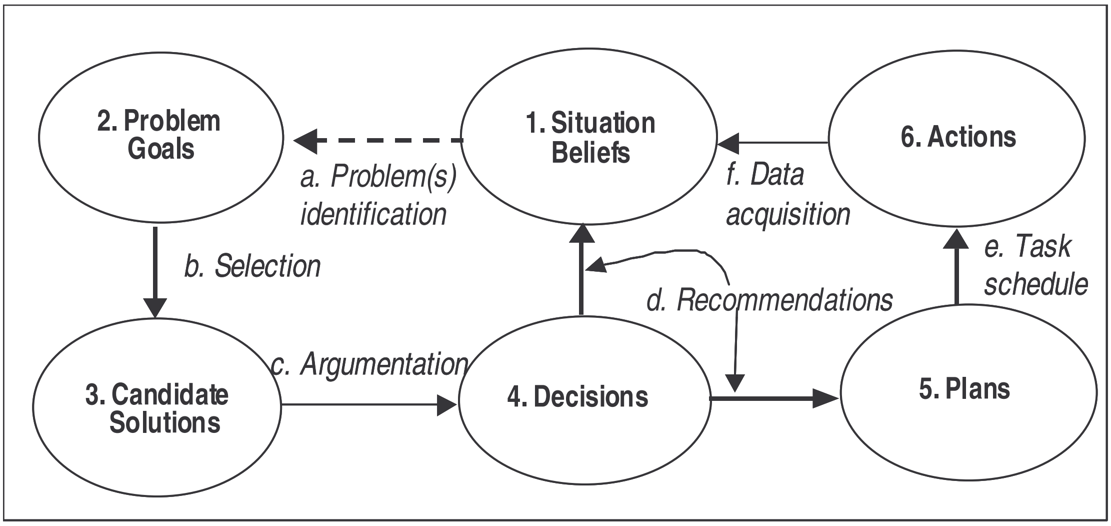


Figure 4.1 The Domino Model (Fox and Das, 2000)

One of the main characteristics of this model is that the decision-making process and the plan performance to solve an identified situation, are represented in the same diagram, where the knowledge bases (nodes) can be a piece of data that the decision procedure can reason about, a set of possible solutions or an action that can be initiated. On the other hand, the computational functionalities (symbolized by arrows) can be the application of any reasoning process (decision trees, arguments, procedures, rules, cases, etc.) or numerical calculations (calculation of trends, data pre-processing, mathematical modelling, etc.) that will be followed until arriving to a conclusion. In the model, steps 1 to 4 would correspond to the classical process followed by any Decision Support System to solve general problems. But, when it is necessary to diagnose and operate on more specific and complex problems (such as solids separation problems in WWTP), the actions to be carried out in order to solve that problem must follow a protocol and therefore a specific task program, and this is exactly what the Domino Model incorporates in steps 5 and 6. Hence, according to the requirements that a KBDSS needs to incorporate to become an efficient tool to give support in the decision-making process, Fox and Das proposed the Domino Model as a conceptual model for designing Knowledge-Based Decision Support Systems.

It seems clear that the Domino Model constitutes one of the most important parts in the logic process followed in the KBDSS engineering, given that on the one hand, it helps to understand the domain and the functions or goals that the system must follow, and on the other hand it characterizes the conceptualization of the different knowledge bases and the computational functionalities, together with their relationships. Therefore, it can be considered of relevant importance during the first stages in the development of these kinds of Intelligent Systems (Conceptual Design).

The functionality of the Domino Model has been validated in a wide range of decision-making domains, especially during the development of KBDSS in the medical domain. Nevertheless, according to Fox and Das, the Domino Model can be applied to any field where decision-making is complex and has important consequences. Wastewater treatment plants management and specifically the control of solids separation problems can also be considered an appropriate field, because of all the characteristics that make its control a rough task. Preliminary studies about the feasibility testing of the Domino Model, as a conceptual model to develop KBDSS for managing solids separation problems, are presented in Martínez (2002), Martínez *et al*. (2003) and Cortés *et al*. (2003), where the Domino Model was validated as a conceptual design of a prototype KBDSS to solve filamentous bulking in a full-scale WWTP of Chicago (Illinois, USA): the North Side Water Reclamation Plant of Chicago. A summary of the results of these preliminary studies are described next.

4.2 Feasibility Testing of the Domino Model in the Solids Separation Problems Domain

Wastewater treatment plants management is, like the medical domain, a good example of a complex system where KBDSS can be considered as a useful tool to improve the control of the process. Actually, both domains possess certain similarities such as: both the doctor and the plant manager need a broad knowledge of the domain in order to diagnose and solve the problems properly, all the processes evolve in real time and some of the large mechanisms involved are difficult to understand due to their complexity. Moreover, the WWTPs domain has some characteristics that increase this complexity, for example, because it is an environmental process it evolves over time and involves many interactions between physical, chemical and biological reactions. As a result, it has to deal with larger amounts of information than the medical domain, including on-line data provided by sensors (e.g. flow, pH or dissolved oxygen), off-line data provided by analytical results (e.g. organic matter, nitrogen or sludge settleability), and qualitative information provided by the process operators (e.g. water color, characteristic odor, foam presence or filamentous proliferation), all of which are difficult to acquire, integrate and understand. In addition, the action plans to solve problems are better defined in the medical domain (where they are very specific) than in the WWTPs management.

Because of all these characteristics, and considering the successful use of the Domino Model in the medical domain (Fox and Das, 2000), it seems clear that the management of WWTP in general, and the management of solids separation problems in particular, are both a clear type of domain where the Domino Model could work as a conceptual model to develop efficient KBDSS to control those dynamic domains.
In a WWTP, the decision-making process to identify and solve a problem relies on a cycle that includes:

1) Identification of problematic situations (diagnosis) or future problematic situations (prediction).
2) Analysis of the possible causes of the problem.
3) Selection of a control method to restore the process.
4) Execution of the selected actions.
5) Verification of the results.

This process could be adapted into the Domino Model (Figure 4.1). As mentioned above, the NODES in the model symbolize knowledge bases that can correspond to a piece of data, a set of possible solutions or an action that can be initiated. In our approach, each node could represent:

1. **Situation beliefs**. Main variables or parameters that characterize the domain. In the case of solids separation problems, these variables would be those directly available from the process (obtained by laboratory analysis, microscopic or macroscopic observations) that can help to identify or predict a problematic situation, i.e. SVI, abundance of filaments, SRT, DO concentration, etc.

2. **Problem goals**. The main objective of the system is to identify and solve any solids separation problem in order to return the process to a “normal” condition. Nevertheless, given that the process is continuously changing and evolving (dynamic domain) as a result, the situation beliefs and goals are continuously changing too. For example, the different goals can vary from: (1) Detect or predict the problem that derives from an anomalous value; (2) Determine the cause of the identified problem or (3) Propose a complete control plan to solve the identified situation.

3. **Candidate solutions**. A set of possible solutions can be defined according to the decision-making process applied and the goal of the system in that moment. For example: if the goal is to identify the problem, the candidate solutions will be all the possible solids separation problems, including: filamentous bulking, non-filamentous bulking, biological foaming, pin-point floc, dispersed growth or rising sludge. On the other hand, if the goal is to identify the proper control method, the candidate solutions will be: addition of toxicants, biological control, use of selectors, etc.

4. **Decisions**. The decision-making process applied will reach a conclusion by choosing from the set of candidates, the most appropriate option: a decision is made. According to the dynamics of the process, the resulting decision will directly affect the characteristics of the process. The Situation Beliefs node have been modified, affecting the next cycle of the system.

5. **Plans**. The result of the last reasoning process of the system results in a definition of the most appropriate method to restore the process from the identified undesired state. Any type of reasoning process and any type of conditions can be applied in order to get the most appropriate control plan. For example, a condition of time-efficiency or cost-efficiency can be applied during the reasoning process in order to get, the most time-effective control plan or the most cost-effective control plan respectively.

6. **Actions**. Considering the complexity of the domain, the suggested control plan must be well structured into a set of scheduled actions, including the order of the different actions or the duration of the plan. The effects of the applied actions on the process are directly revealed in the following Situation Beliefs node, which may lead to further goals and a new cycle will start again.

On the other hand, the ARROWS in the Domino Model represent the computational functionalities or applications of any reasoning process (decision trees, arguments, procedures, rules, cases, etc.) or numerical calculation (calculation of trends, data pre-processing, mathematical modelling, etc.) which will be applied to reach a conclusion. The computational functionalities suggested in the original model could be adapted to the solids separation domain as follows (this is just an example, since any reasoning process or numerical calculation could be applied instead):

a. Goal(s) identification
b. Definition
c. Selection
d. Recommendation(s)
e. Task schedule
f. Data acquisition.

4.2.1 Case study

With the aim of deeply understand the main usefulness of the model in organizing the different knowledge bases and reasoning processes within the dynamic domain of solids separation problems, a case-study based on a hypothetical situation of filamentous bulking is represented in Figure 4.2. The cycle and structure followed by the hypothetical KBDSS is described next.

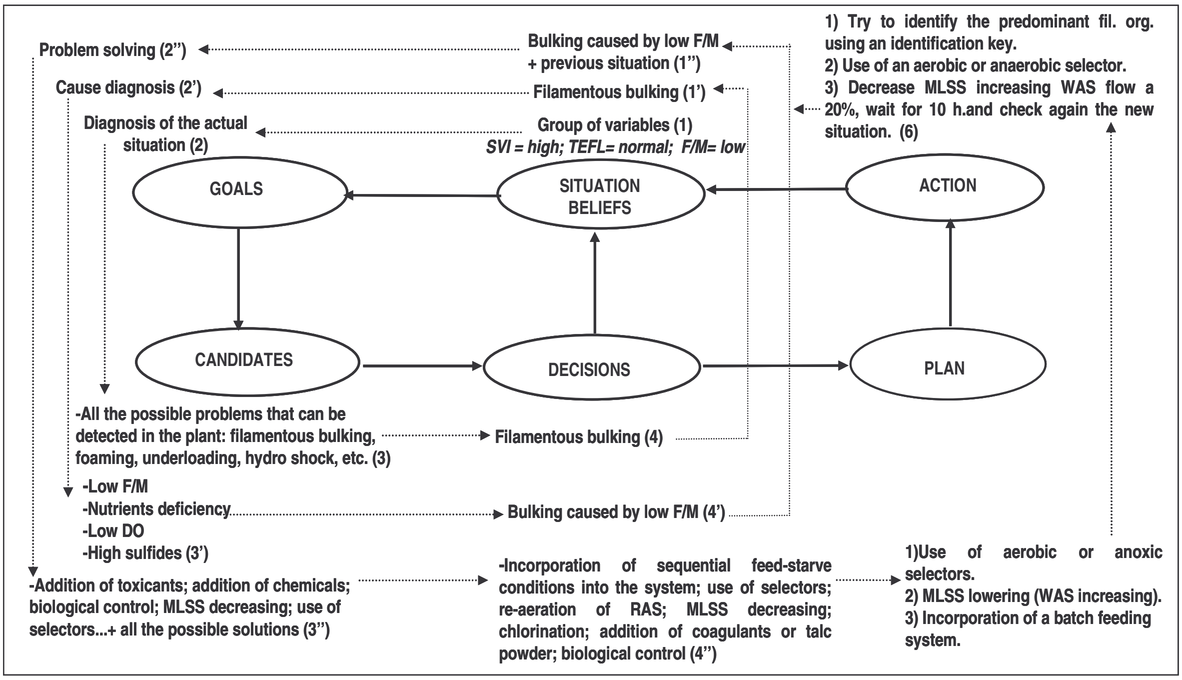


Figure 4.2 Complete process of the Domino Model in the case-study of filamentous bulking

Given that the system believes that the process presents a high SVI, it raises a goal to diagnose the actual situation of the process (2 in Figure 4.2). It then infers from the process knowledge that there are three possible or candidate diagnoses, filamentous bulking, non-filamentous bulking or biological foaming, and activates a decision tree to come to a conclusion. In this case, other data from the process (value of other variables such as the predominant filament or the presence of foams), compiled within the Situation Beliefs node, enables the system to reach the conclusion that there exists a filamentous bulking situation. This conclusion is then added into the system's database of beliefs (Situation Beliefs).

Given this new belief, a second goal is raised: to identify the possible cause of the identified problem (2' in Figure 4.2). As in the previous diagnosis decision, there exist a candidate set of possible causes. Here the F/M value compiled within the Situation Beliefs database helps the decision making process to decide that there exists a low F/M ratio in the biological reactor.

Given this new belief, a third goal is raised: to solve the problem of filamentous bulking caused by low F/M (2" in Figure 4.2). As in a diagnosis decision, there can be any number of decision candidates mainly in relation to the cause that has originated the problem. Here given that there exist a low F/M ratio in the biological reactor, there are five possible treatment options to solve this specific problem: to increase F/M ratio, to use aerobic or anoxic selectors, to incorporate a batch feeding system, to chlorinate or to add coagulants or flocculants, and the system again follows a reasoning process to reach a conclusion (taking into account the expected efficiency, cost, plant characteristics, etc.). On the basis of these considerations, the system recommends a F/M increasing as the most preferred treatment. The adoption of a F/M increasing plan leads to an appropriate schedule of specific actions: (1) Increase a 10 % wasting rate to restore a proper F/M ratio; (2) Try to identify the predominant filamentous microorganism to check if it is one of the low F/M group and study its evolution and (3) Follow up the plan during a period of 3 times the sludge residence time or until the low F/M filamentous bulking situation disappear.

Setting the length of the plan, each time that the cycle is started it detects (checking the process state value) in which phase of the episode the process is, and will not generate a new plan until all the control actions have been executed (unless a new situation is diagnosed, e.g. storm). Meanwhile the system will be monitoring and analyzing the new values of the key variables (SVI or floc characteristics in this case) in order to verify the effects of the plan and to adjust the length and intensity of each action (e.g. the recalculation of the required WAS flow). On the other hand, once the plan has been fully executed, if the process is still in the situation of filamentous bulking, an alternative non-specific plan such as chlorination will be suggested by the KBDSS.

4.3 Discussion

During the present study, the feasibility of the Domino Model, previously confirmed in several approaches within the medical domain, was also proven to be feasible in conceptually designing a KBDSS in the domain of solids separation problems management. The following key issues have been approached during the phase of conceptual design:

⇒ An intuitive visualization and a direct characterization and organization of the different knowledge bases and reasoning processes necessary to develop the KBDSS.

⇒ A dynamic interrelationship between the knowledge bases and the corresponding reasoning processes, due to their cyclic (not linear or static) organization.

- ⇒ Possibility of structuring the suggested control plans into a well-organized and scheduled set of actions.

- ⇒ Possibility of defining any suitable knowledge-based technique (case-based reasoning, rule-based reasoning, fuzzy logic, etc.) to perform the reasoning process that must permit the system to accomplish the pre-defined objectives and reach the appropriate conclusions (problem diagnosis, cause identification and control plan proposal).

The use of the Domino Model also permits to structure the development of the system according to different cycles of operation (each one defined to accomplish the different objectives). The beginning of the first cycle is determined by the acquisition of a new set of data that characterize the new situation of the process (Situation Beliefs). The following cycles are dynamically chained in such a way that the decision of the anterior cycle becomes the beginning (by being part of the new situation beliefs base) of the following cycle. Likewise, different cycles can be running in parallel whenever different objectives need to be achieved. The cyclic process is repeated until accomplishing the pre-defined goals.

CHAPTER 5

DEVELOPMENT OF THE DYNAMIC EXPERT SYSTEM

One of the main goals that the KBDSS development must achieve in order to ensure the feasibility of the system concerns its intelligent component. In order to efficiently fulfil its main tasks, the developed system must be **intelligent** and use intelligent reasoning. According to Stephanopoulos and Han (1996), some of the features and tasks that an intelligent KBDSS developed in the domain of WWTP must be able to carry out include: (a) to distinguish between normal and abnormal operating conditions; (b) to identify the causes of abnormal operating conditions (e.g., external load disturbances, equipment faults, operational degradation, operator-induced mishandling); (c) to evaluate current process trends and anticipate future operational states; and (d) to create action plans and to schedule sequences of operating steps to bring the plant to the desired operating level.

Due to their great potential for emulating human intelligence in decision making processes, expert systems (ES) are among the fastest growing applications of Artificial Intelligence in scientific and engineering fields. An **Expert System** can be defined as an interactive computer program that attempts to emulate the reasoning process of experts in a given domain (the group of processes over which the expert makes decisions). This can be performed through the correct operation of their 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. These rules take the form: **IF** < *a set of conditions (or premises or antecedents) is true* > **THEN** < *certain conclusions (or actions or results or consequences) can (will or should) be drawn (occur)* >. An ES works by applying known facts to the left-hand side of rules; if true, the right-hand side fires and a new fact is discovered. The knowledge sources are usually human experts, statistical data about the domain or technical literature. The **inference engine** is the software that controls the reasoning operation of the ES by chaining the knowledge contained in the knowledge base optimally. The order in which rules are chained depends on the method of inference used (forward or backward). *Forward chaining* is done from conditions, which are already known to be true, towards problem situations that those conditions allow to establish, whereas *backward chaining* is done from a goal state towards the necessary conditions for its establishment.

The following are some of the advantages that the use of Expert Systems presents to the detriment of other techniques: (1) ES facilitate the inclusion and retention of heuristic knowledge from experts and allow qualitative information processing; (2) an ES emulates the reasoning process followed by an expert while avoiding some disturbances like time pressures or stress or other emotional situations that can affect a human but not an ES; (3) it is more easy to transfer or reproduce Expert Systems than to transfer knowledge from one human expert to another (which is laborious, lengthy and expensive); (4) knowledge is represented in an easily understandable form (rules); (5) a well-validated ES offers potentially optimal answers because action plans are systematised for each problematic situation; and finally (6) ES also enable acquisition of a large general knowledge base, with flexible use for any WWTP management.

Given all these potential characteristics and the promising results from their application to WWTP operation such as Maeda (1989); Patry and Chapman (1989); Barnett *et al*. (1993); Ladiges and Kayser (1993); Ozgur and Stenstrom (1994); Zhu and Simpson (1996); Serra *et al.* (1997); Furukawa *et al*. (1998); Ashraf Islam *et al*. (1999); Vouros *et al*. (2000); Puñal *et al*. (2002); Baeza *et al*. (2002) or Rodríguez-Roda *et al*. (2002), an Expert System is the first of the intelligent techniques that we have selected to integrate into our proposed KBDSS.

However, in order to overcome the limitations found in applying KBDSSs to full-scale facilities (see Chapter 4), and specially, in order to cope with the specific characteristics of our domain –the management of solids separation problems–, a further approach is required to ensure that the main goals of the KBDSS are achieved: the ES must be able to reach dynamic diagnoses based on the needs of the process at each moment (Martínez *et al.*, 2005a). The integration of dynamic reasoning must allow the system to achieve the following goals or tasks: (1) the system must be able to understand how the process evolves (whether the current situation is the beginning of a new problem or a continuation of a previously started problem) and to predict in advance the occurrence of future problematic situations; (2) the system must also identify the cause or origin of the present or future problematic situations; and, (3) additionally, it must propose a complete control strategy to restore the process in the most efficient way, based not only on literature suggestions but also continuously adapted to the own evolution of the process. We have called this new approach: **Dynamic Expert System**.

In order to overcome the limitations derived from developing a KBDSS in such a complex domain, and according to the promising results obtained during the feasibility testing of the Domino Model in defining an optimal KBDSS (see Chapter 4), the architecture of the proposed Dynamic Expert System (specifically the organization of the knowledge bases) has been developed following the conceptual design depicted by the Domino Model. The Domino Model can be considered as the key structure that has provided the targeted dynamic approach to the whole system. Therefore, based on the conceptual design suggested by the Domino Model, the development of the Expert System has followed different cycles of reasoning. In our approach (see Figure 4.2) three different cycles have been distinguished, according to the different goals that need to be achieved:

- ⇒ Cycle 1. Diagnosis of the current situation of the process and prediction of future situations.
- ⇒ Cycle 2. Identification of the possible causes of the identified situation.
- ⇒ Cycle 3. Proposal of the most appropriate and complete control plan to restore the process.

The knowledge bases and reasoning processes developed for each of the specific cycles are explained in detail next.

5.1 Diagnosis and Prediction of Solids Separation Problems

The first cycle, depicted in Figure 5.1, consists of 4 knowledge bases and their corresponding computational functionalities. The description of the different knowledge bases, the computational functionalities and the order in which they are executed, are depicted next.

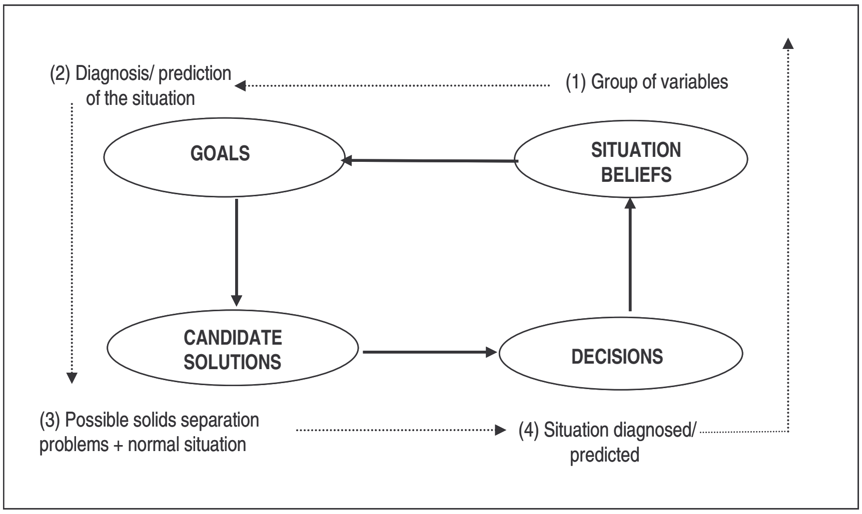


Figure 5.1 First cycle of Expert System architecture

5.1.1 Situation Beliefs

The first knowledge base of the system includes the set of variables or parameters that characterize the domain. In our approach, these are the variables available in a WWTP (obtained by on-line sensors, laboratory analysis, microscopic or macroscopic observations) that can help to identify or predict solids separation problems. Despite the great advances in the development of analytical and microscopic techniques which facilitate the identification and prediction of some solids separation problems (e.g.: FISH techniques, image analysis or on-line settlers), the selection of variables for the Situation Beliefs module has been restricted to the most commonly measured variables in a facility. The approach to the solids separation problems generally followed in Catalonia (chapter 3) has helped to define these variables, which include both quantitative and qualitative variables.

The selected **quantitative variables** are the following:

- **SVI or DSVI**. The Sludge Volume Index is used as indicative of sludge settleability. If DSVI is available, it is preferred over SVI since it has been demonstrated to be more representative of settleability problems (Stobbe, 1964 and Lee *et al.*, 1983).

- **V30.** The value of the V30 will also be used as an estimate of sludge settleability, specially when SVI is not calculated because the MLSS is not available,

- **Effluent-TSS**. The concentration of total suspended solids at the effluent is used as indicative of solids loss due to solids separation problems.

- **% of COD removal**. This parameter indicates the existence of any problem in the process that affects the efficiency of COD removal, such as the presence of pin-point or dispersed growth problems.

These data are commonly calculated by means of laboratory analysis and not more than a value per day is generally obtained, although in some plants both TSS and COD are measured by means of on-line sensors. Nonetheless, we will consider a standard facility where all these variables are calculated daily and obtained off-line. If more than one value were obtained per day, the daily average value would be used.

As previously commented, qualitative information regarding both microscopic and macroscopic observations, is extremely important in the diagnosis of solids separation problems. An assessment of the state-of-the-art of solids separation problems monitoring in general, and in particular in Catalonia (chapter 3), made us aware of the existing limitations in registering this type of information, especially due to a lack of standardization. In this sense, both a macroscopic and a microscopic worksheet have been developed to facilitate the process of compiling **qualitative variables**. The use of these specific worksheets has allowed the transformation of qualitative information into a set of **categorical variables** which will be more easily used by the Expert System.

5.1.1.1 Development of the macroscopic and microscopic worksheets

The aim of the developed worksheets is to facilitate the registration of all the qualitative information that can be useful for the identification of present or future episodes of solids separation problems. The worksheets (Figures 5.2 and 5.3) have been designed to be filled out in about 30 minutes, depending on the skills of the user, in order to promote its daily registration. It is recommended that the worksheet was always filled out by the same person. The information required in the worksheet has also been selected taking into consideration the common microscopic variables measured in a facility (see Chapter 3). Thus, 3 macroscopic variables and 7 microscopic variables have been chosen. Nevertheless, out of these 10 variables, only 2 macroscopic variables and 5 microscopic variables are actually used by the Expert System. The rest of the variables, as well as the possibility of including additional remarks, have been included to allow plant managers to register as much relevant information as possible. Nowadays, the developed worksheets are routinely used in 10 out of the 13 plants studied in Catalonia.

The **macroscopic information** includes those relevant data that can be obtained from observing the sludge settleability characteristics, by means of the V30 test or by directly observing the characteristics of the biological reactor or the clarifiers. The selected variables are:

- **Sludge settleability**. This variable tries to classify the macroscopic characteristics of the V30 in order to study the sludge settleability in clarifiers. Apart from the quantitative value obtained from the measurement of the sludge height –V30 – (also used as a variable of the Situation Beliefs), this variable classifies the way the sludge settles into 7 categories. A schematic drawing (see Figure 5.2) is provided to facilitate this classification. The possible categories are the following:

    1) *The sludge settles well showing a compact sludge layer and a clear supernatant*. This category is indicative of a normal situation where sludge and treated water are correctly separated.

    2) *The sludge does not settle well or it settles very slowly. The supernatant, however, can be very clear*. This category represents a situation of filamentous bulking where sludge settleability is hindered by the presence of high quantities of filamentous bacteria.

3) *The sludge settles well but a scum layer forms on the surface. This scum can be dark-brown, brown or greyish-white*. A problem of biological foaming is characterized by this category. It is known that sometimes the presence of foams can not be noticed during settleability tests; it is for this reason the macroscopic variable regarding presence of foams has been included.

4) *The sludge settles well during the first 30 min., but after 1 or 2 hours, most of the sludge rises to the surface. Gas bubbles are present.* This definition tries to describe a situation of rising sludge which is easily detected by performing the V30 test.

5) Hardly any sludge settling occurs. Most of the flocs remain suspended within the sample. This is a description of a deflocculation problem. Specifically, it tries to define a situation of dispersed growth where deflocculation is extreme. It is known though, that it may be difficult to distinguish between a dispersed growth situation and a pin-point floc situation (category 6); for this reason, other microscopic variables will also be checked before reaching any conclusion.

6) *The sludge settles well but only a thin settled layer is formed. Lots of small suspended flocs are present in the supernatant layer with a "pin-point" form*. This is the category which represents a situation of pin-point floc in clarifiers.

7) *The sludge has a dark-brown colour and sometimes an acid smell. At the beginning it settles well, but after some minutes, large flocs rise to the surface*. The last category includes the possibility of anoxic conditions in the sludge. Although the problems directly related to anoxic sludge have not been directly considered as a type of solids separation problem, mainly because their origin is not related to the microbiological composition of the sludge but with the presence of anoxic conditions, this category has been included in the worksheet just as another option for plant operators to register other cause of loss of solids by the effluent.

- **Presence of foam.** This variable has been included in order to provide important information about the presence of any type of foams on the surface of reactors or clarifiers. It is obvious that the presence of foams is a symptom of a specific type of solids separation problems such as biological foaming or viscous bulking so, when observed it means that the problem was already initiated some time ago. At any rate, by recording this information (it is a parameter not commonly recorded), the knowledge base of Situation Beliefs will receive extra information that will facilitate the diagnosis of specific solids separation problems. The 7 categories considered are:

W) *White frothy foams on aeration basin and/or in pre-treatment processes*. As in category 7 of sludge settleability, this type of scum is not related to any solids separation problem but to the presence of biodegradable or non-biodegradable surfactants, which in future cycles could be used as information related to one of the possible causes of sludge deflocculation.

G) *White or grey foams on aeration basins and/or secondary clarifiers*. This type of foam is usually related to high concentrations of Type 1863.

B) *Stable brown foams on aeration basins and/or secondary clarifiers*. The presence of brown or dark-brown foams is usually originated by the overabundance of some filamentous bacteria such as nocardioforms or *M. parvicella*. Although some experts ara capable of distinguishing between the appearance of foams caused by one or another type of foam-forming bacteria, this small difference has not been considered here to avoid possible misinterpretations.

R) *Scum presence on secondary clarifiers or on reactor (anoxic zone). The sludge is brown and it contains some gas bubbles*. The presence of gas bubbles (of N_2) has been distinguished in this category as indicative of rising sludge.

A) *Sludge blanket on surface of aeration basins and/or secondary clarifiers. The sludge is black and it has a "rotten eggs" smell.* As in category 7 of sludge settleability, despite not being used by the ES, the occurrence of anoxic sludge has been considered also in this category of foam presence.

Z) *Thin layer of sticky, viscous foam on the surface and/or walls of the secondary clarifiers*. This type of foam is characteristic of viscous bulking episodes.

N) Foam is absent or present only in small amounts of fresh, light brown foam. The absence of biological foaming or the presence of very few quantities is indicative of a normal non-foaming situation.

The **microscopic information** includes those relevant data that can be obtained from microscopic observation of the floc and the microfauna. This is the part that possibly requires both more time and more skills to be filled in. The selected variables are:

- **Floc characteristics.** The analysis of floc characteristics helps to identify the existence or beginning of a solids separation problem. It is recommended to determine floc characteristics by observing different fields of a sample of activated sludge at 100x or 400x magnification. The possible categories are:

  C) *The floc components are very small and dispersed. The floc does not show any outline*. These floc characteristics are typical of a dispersed growth situation.

  D) *Absence of filaments. The flocs are small and they are not joined together*. The absence of filaments and the presence of small flocs are indicative of pin-point floc.

  A) *Ideal floc. The floc is quite compact and there exist a balance between flocs and filaments*. This category identifies a normal or ideal situation where the presence of filaments is in balance with the floc-formers microorganisms so their presence is necessary to allow flocs to settle properly.

B) *Filaments present in most flocs but at low concentration < 5 filaments/floc*. In this category, filaments are present at higher quantities although their concentration is not worrying yet.

E) *High abundance of filaments: 5-20 filaments/per floc*. The presence of such a quantity of filaments can originate filamentous bulking or foaming situations.

F) *Very high abundance of filaments: more than 20 filaments/per floc*. The presence of more than 20 filaments per floc indicates serious problems of filamentous bulking or foaming.

- **Presence of *Zoogloea***. This variable has been included to identify the occurrence of *Zoogloea* spp. or the presence of important amounts of extracellular polymeric substances (EPS) which indicates episodes of viscous bulking.

- **Predominant filamentous bacteria.** The identification of the predominant filamentous bacteria during episodes of filamentous bulking or foaming is extremely important on the one hand to identify if the species is going to cause filamentous bulking, foaming or both disturbances, and on the other hand to subsequently identify (in future cycles) the possible cause of its proliferation. This variable is arguably the hardest to determine, since apart from requiring a good microscope, some skills are also necessary to distinguish between the different species. 400x and 1000x magnifications are suggested to identify species. The use of Gram or Neisser stains may be necessary. It is also provided the possibility of registrating the secondary filamentous bacteria observed, although just the predominant species will account during the reasoning process.

- **Predominant group.** The identification of the predominant group can help to detect possible disturbances in the process such as the presence of toxics or low DO concentrations, and some operational conditions such as the SRT or the F/M, while providing a broad indication of treatment efficiency. 100x and 400x magnifications are suggested to determine the predominant groups. Four different categories have been considered:

1) *Flagellates, amoebae (without septa) and small free-swimming ciliates*. This category indicates conditions of low SRT or low oxygenation.

2) *The stalked ciliate species: Opercularia spp. and/or Vorticella microstoma.* The predominance of one or both species of stalked ciliates indicates a low efficiency of treatment due to high and sudden organic loads or low oxygenation.

3) *"Crawling" and stalked ciliates (except from Opercularia spp. and V. microstoma) and/or amoebae with septa*. When these specific groups predominate, it can be stated that efficiency of treatment is good and that no extreme disturbances affect the microfauna.

4) *Stalked ciliates, rotifers and higher invertebrates (especially nematodes)*. The predominance of these groups of microorganisms indicates that the process is being operated at high SRT or low F/M ratios.

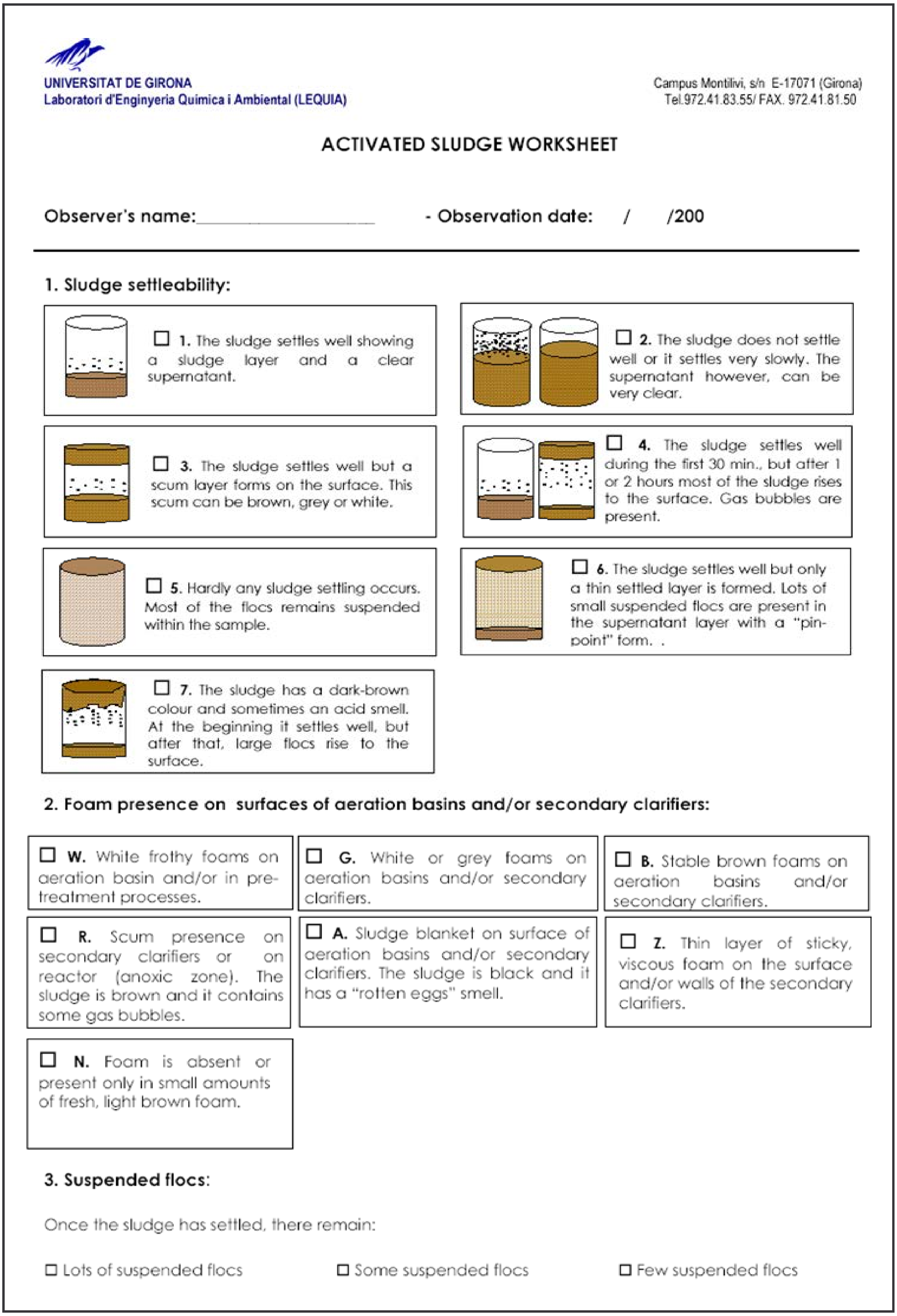
- **Presence of Protozoa**. This last variable is also used to indicate more precisely some specific conditions of the process such as toxicity, very important in the identification of the origin of some solids separation problems such as pin-point floc or dispersed growth. Three categories are identified:

  0) *Very few or none*. Presence of few or no protozoa may indicate 3 different conditions: a high F/M ratio (or very low SRT), an oxygen deficiency, or the presence of a toxic that adversely affects the health of microfauna.

  1) *Present but inactive*. It frequently indicates that a slug of toxic material has recently entered the system.

  2) *Present and active*. Normal and active protozoa indicate the presence of healthy microfauna.

The qualitative abundance of suspended flocs in the macroscopic worksheet and the floc shape or the registration of observed species of microorganisms (apart from filamentous bacteria) in the microscopic worksheet can also be registered as extra information.



UNIVERSITAT DE GIRONA
Laboratori d'Enginyeria Quimica i Ambiental (LEQUIA)

Campus Montilivi, s/n E-17071 (Girona)
Tel.972.41.83.55/ FAX. 972.41.81.50

ACTIVATED SLUDGE WORKSHEET

Observer's name:_____________________ **- Observation date: / /200**

1. Sludge settleability:

☐ **1.** The sludge settles well showing a sludge layer and a clear supernatant.

☐ **2.** The sludge does not settle well or it settles very slowly. The supernatant however, can be very clear.

☐ **3.** The sludge settles well but a scum layer forms on the surface. This scum can be brown, grey or white.

☐ **4.** The sludge settles well during the first 30 min., but after 1 or 2 hours most of the sludge rises to the surface. Gas bubbles are present.

☐ **5.** Hardly any sludge settling occurs. Most of the flocs remains suspended within the sample.

☐ **6.** The sludge settles well but only a thin settled layer is formed. Lots of small suspended flocs are present in the supernatant layer with a "pin-point" form. .

☐ **7.** The sludge has a dark-brown colour and sometimes an acid smell. At the beginning it settles well, but after that, large flocs rise to the surface.

2. Foam presence on surfaces of aeration basins and/or secondary clarifiers:

☐ **W.** White frothy foams on aeration basin and/or in pre-treatment processes.

☐ **G.** White or grey foams on aeration basins and/or secondary clarifiers.

☐ **B.** Stable brown foams on aeration basins and/or secondary clarifiers.

☐ **R.** Scum presence on secondary clarifiers or on reactor (anoxic zone). The sludge is brown and it contains some gas bubbles.

☐ **A.** Sludge blanket on surface of aeration basins and/or secondary clarifiers. The sludge is black and it has a "rotten eggs" smell.

☐ **Z.** Thin layer of sticky, viscous foam on the surface and/or walls of the secondary clarifiers.

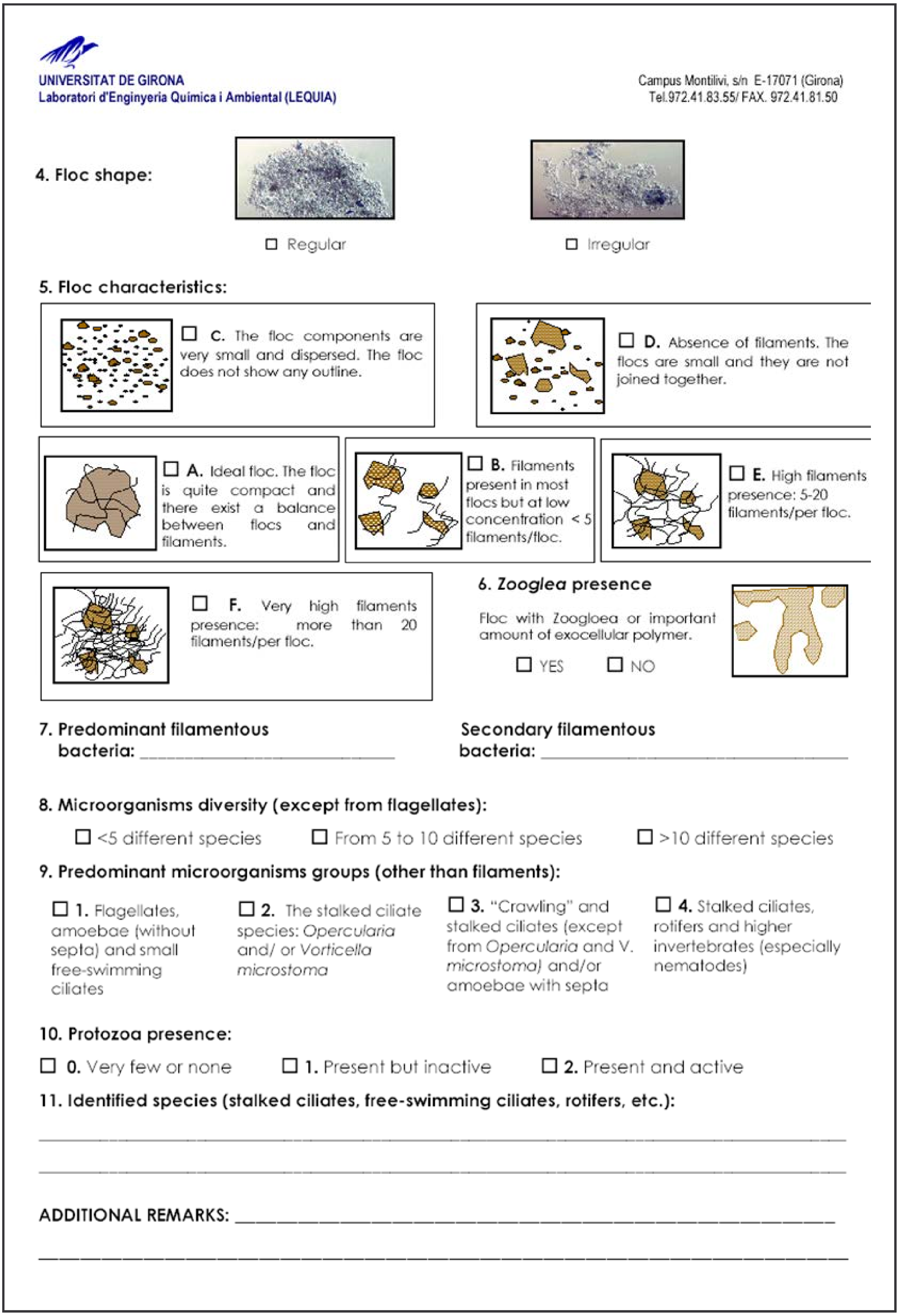
☐ **N.** Foam is absent or present only in small amounts of fresh, light brown foam.

3. Suspended flocs:

Once the sludge has settled, there remain:

☐ Lots of suspended flocs
☐ Some suspended flocs
☐ Few suspended flocs

Figure 5.2 Activated sludge worksheet (macroscopic information)



UNIVERSITAT DE GIRONA
Laboratori d'Enginyeria Química i Ambiental (LEQUIA)

Campus Montilivi, s/n E-17071 (Girona)
Tel.972.41.83.55/ FAX. 972.41.81.50

4. Floc shape:

☐ Regular ☐ Irregular

5. Floc characteristics:

☐ **C.** The floc components are very small and dispersed. The floc does not show any outline.

☐ **D.** Absence of filaments. The flocs are small and they are not joined together.

☐ **A.** Ideal floc. The floc is quite compact and there exist a balance between flocs and filaments.

☐ **B.** Filaments present in most flocs but at low concentration < 5 filaments/floc.

☐ **E.** High filaments presence: 5-20 filaments/per floc.

☐ **F.** Very high filaments presence: more than 20 filaments/per floc.

6. *Zooglea* presence

Floc with Zoogloea or important amount of exocellular polymer.

☐ YES ☐ NO

7. Predominant filamentous bacteria: ______________________________

Secondary filamentous bacteria: ____________________________________

8. Microorganisms diversity (except from flagellates):

☐ <5 different species ☐ From 5 to 10 different species ☐ >10 different species

9. Predominant microorganisms groups (other than filaments):

☐ **1.** Flagellates, amoebae (without septa) and small free-swimming ciliates

☐ **2.** The stalked ciliate species: *Opercularia* and/ or *Vorticella microstoma*

☐ **3.** "Crawling" and stalked ciliates (except from *Opercularia* and V. *microstoma*) and/or amoebae with septa

☐ **4.** Stalked ciliates, rotifers and higher invertebrates (especially nematodes)

10. Protozoa presence:

☐ **0.** Very few or none ☐ **1.** Present but inactive ☐ **2.** Present and active

11. Identified species (stalked ciliates, free-swimming ciliates, rotifers, etc.):

____

____

ADDITIONAL REMARKS: ____

____

Figure 5.3 Activated sludge worksheet cont. (microscopic information)

Apart from the quantitative and qualitative information, indicative of current operational characteristics, other important information will be stored in the Situation Beliefs database in order to define the evolution of the solids separation problems. This information is compiled by means of **numerical counters** which compile the previous situation of the process as well as the number of days that the process has been suffering from a specific problem. In such a way, whenever a diagnosis is made the system can determine if the situation is part of an already initiated problem, if the problem is starting in that moment or if the problem has disappeared from the process.

5.1.1.2 Data gathering process

During Data Gathering, the different data stored in the Situation Beliefs knowledge base are collected and validated using a data reconciliation system. This stage includes data filtering and fault detection together with a discretisation stage (where the numerical values of the different variables are converted into qualitative or categorical values, e.g. high, normal or low) in order to feed the system with the most accurate and useful information to reach a diagnosis. In case on-line data were available, this information would be automatically gathered by means of an automatic data acquisition system. Whenever the necessary information comes from off-line or qualitative data, the user will have to introduce it by means of specific worksheets.

Discretisation

The discretisation module includes a set of rules to conduct a qualitative abstraction of the quantitative variables to qualitative modalities (linguistic labels). This amounts to establishing the limits within which these values are considered: high, normal, low, or other categories according to the standard conditions of the facility. The different categories are usually established first of all by means of a statistical analysis of the historical data of the facility and then, by considering the expertise of plant operators. In general, thus, the determined limits or categories will depend on the specific characteristics of each facility where the KBDSS is being implemented. The modification of these limits (which is straightforward) will be enough to adapt a developed system to another WWTP.

The discretisation must be performed only for the 4 quantitative variables considered in the knowledge base of Situation Beliefs. The definition of the different classes or categories considered during the reasoning-process is performed by means of **fuzzy logic**. The use of fuzzy sets has allowed us take into account the absence of a sharp boundary between the different sets or categories of information considered in our approach. It also allows the system to solve problems with incomplete data or to conclude uncertain solutions.

Figure 5.4 represents an example of a membership function for a specific quantitative variable: the sludge volume index (SVI), where the membership degree for each value of the variable is represented according to the pre-defined fuzzy sets. The rest of membership functions considered for the specific implementation at the Girona WWTP are depicted in Figure 7.6.

There are no strong rational grounds for preferring one membership function to another, and in a real problem there would be hundreds of such functions. In our approach, triangular and trapezoidal functions have been chosen in order to decrease the calculus effort.

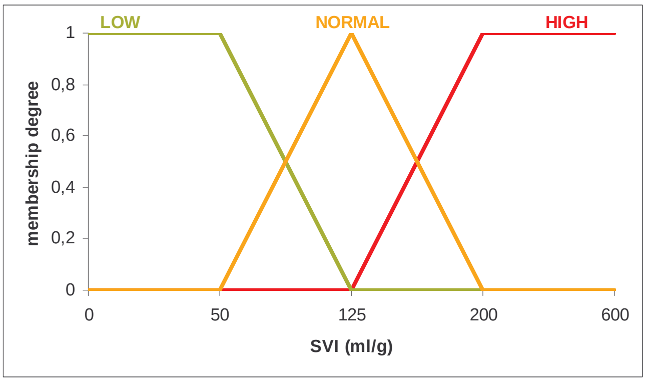


Figure 5.4 Membership function for the Sludge Volume Index variable

The membership degree for each category is used in the problem diagnosis process to determine the percent certainty when no exact diagnoses can be reached. The rest of variables comprising the knowledge base of Situation Beliefs (qualitative data provided by microscopic and macroscopic worksheets) do not need to be discretised by means of fuzzy sets since they are already categorical variables classified into a set of non-ordered categories.

5.1.2 Goals

The main goal of the first cycle of the dynamic ES is first of all, to diagnose the state of the process according to the most recent data, and secondly to study how the identified situation is temporarily located: whether the current situation is the beginning of a new problem or a continuation of a previously started problem and whether the process is evolving to a better situation (improvement), is keeping static or is evolving to a worse situation. This is needed to predict the occurrence of future problematic situations.

According to the current objective, the different candidate solutions are determined.

5.1.3 Candidate solutions

The candidate solutions are the set of solids separation problems that the KBDSS is able to diagnose. The following are the possible diagnosis:

- Filamentous bulking
- Non-filamentous bulking
- Biological foaming
- Dispersed growth
- Pin-point floc
- Rising sludge
- Normal situation. Defined as a situation where no solids separation problem is diagnosed.

Candidate solutions can be concluded to be *true*, *false* or *possible*. The category *possible* is concluded whenever not enough information is available to reach a *true* or *false* diagnosis. By using fuzzy reasoning, the final conclusion will represent a partial truth and all the fractions between truth and falsity will be expressed as percent certainty.

On the other hand, the set of candidate solutions regarding the prediction of future problematic situations includes:

- Risk of sludge bulking or foaming
- Risk of sludge deflocculation
- Risk of solids loss

5.1.3.1 Reasoning process

A **decision** can be defined as a choice between alternative beliefs. Thus, the decision making process implies reasoning on the probability of some unknown present or future situation to occur, according to the available knowledge. In general, when a human expert tries to identify and solve a problem, s/he takes into account all data from the process and uses his/her knowledge on the domain to solve it.

The developed ES must be able to reach a conclusion using the values of the measured variables together with the human expert knowledge compiled by means of **decision trees**, which have been designed to compile all symptoms, facts, procedures and relationships necessary for decision-making. Decision trees have been preferred to other types of computational functionalities such as argumentation mainly because of their easy and understandable representation, which allows an easy and direct codification into heuristic rules by means of the appropriate expert system shell. Decision trees consist of hierarchical, top-down descriptions of the linkages and interactions among any kind of knowledge utilised to describe facts and reasoning strategies for problem solving (objects, events, performance and meta-knowledge). In other words, these logic trees represent the “causal” chain of interactions from symptoms to problems (forward chaining). A knowledge decision tree is also typically depicted as networks of interconnected nodes and arcs. Each node corresponds to a question related to a particular set of information (a single fact, a parameter or a condition), whereas each arc between nodes corresponds to a possible value for that information. In a decision tree, leaf nodes represent a class or diagnosis, and decision nodes specify some test to carry out on a single attribute value, with one branch for each possible outcome of the test. This structured feature of interactions among the nodes of the graph allows the direct interpretation of diagnostic reasoning. From this, the translation of the knowledge contained in a branch of decision trees into a production rule is direct. For example, the arc between nodes A and B in Figure 5.5 identifies that A is the premise of condition 1 while B is its conclusion.

Figure 5.5 A is the premise of condition 1 while B is its conclusion

The whole set of heuristic rules of the process can then be easily generated by traversing each path from the root to every leaf of these decision trees, and representing them as production rules of the form IF-THEN. Therefore, a rule can be defined as a linguistic expression of the knowledge of human expert which relates particular values of the input variables to an output of problem diagnosis. The whole representation of such reasoning processes constitutes the functional basis of the Expert System: **the inference engine**, which defines the way rules are connected to reach a conclusion (inference chain).

All the knowledge acquired in this study was represented as decision trees, as a previous step to the rule codification into the knowledge base. For this purpose, all the general knowledge obtained through literature review, own expert knowledge on the domain, interviews with other experts on the domain and analysis of data from the studied plants of Catalonia, was reviewed, organised and synthesised into a set of decision trees, where all the reasoning processes regarding the diagnosis and prediction of solids separation problems, the identification of causes and the proposal of control strategies are represented for the range of solids separation troubleshooting.

During this first cycle, the values of the different data stored in the knowledge base of Situation Beliefs, used jointly with the expert knowledge represented in the set of heuristic rules, permits the determination of the current state of the process, classifying it between the several options proposed in the Candidate Solutions knowledge base. To accomplish this main objective, the reasoning process has been divided into 3 different modules: (1) meta-diagnosis; (2) diagnosis of solids separation problems, and (3) prediction of future solids separation problems. The reasoning processes followed in all the modules use dynamic reasoning as well as fuzzy logic in order to enhance the efficiency of the Expert System.

Module 1: meta-diagnosis

The diagnosis of solids separation problems, as of any other complex problem, consists in comparing the actual behaviour of the process, as manifested by the values of the operating variables, with the predefined abnormal or threshold values (very low, low, normal, high). This meta-diagnosis process evaluates the impact of problems, selects those that have crossed a predefined threshold and launches the corresponding diagnosis rules to identify the offending problem that resulted in the observed behaviour. This whole process is done by inference through the representation of knowledge.

Therefore, based on the qualitative abstraction, the rules contained in the Meta-Diagnosis module determine which tree and diagnosis paths (i.e. decision rules or procedures) will be explored to infer the situation. The knowledge that decides which rules should be considered and which should be rejected is referred to as meta-knowledge or meta-rules, according to AI nomenclature. As pointed out by (Jackson, 1999), meta-rules are distinguished from ordinary production rules in that their role is to direct the reasoning required to solve the problem, rather than to actually perform that reasoning. In the knowledge representation of the ES (decision trees), the meta-knowledge is symbolised by the intermediate alarms. If an intermediate alarm is fired (any of the signs of the problems is fulfilled), then some other diagnosis trees will be explored. If no diagnosis can be concluded, due for example to a lack of data, the intermediate alarm will be transformed into the conclusion of a **possible** situation. Figures 5.6 and 5.7 represent the decision trees developed to perform the meta-diagnosis.

In case all the necessary data to get a preliminary conclusion were unknown (for example if neither the SVI nor the V30 were available) the system would conclude (by means of an alarm message) that no decisions can be made due to a lack of knowledge.

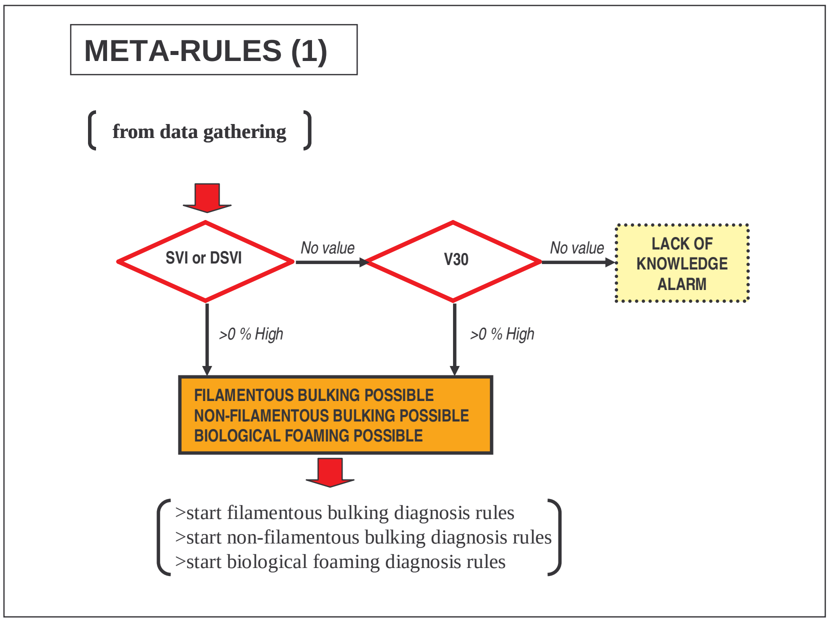


Figure 5.6 Decision tree for the meta-diagnosis of some of the solids separation problems: filamentous bulking, non-filamentous bulking and foaming

From the reasoning process depicted in Figure 5.6, the following rules can be directly generated:

__IF__ SVI > 0% high __THEN__ conclude that Filamentous Bulking=possible and start Filamentous Bulking Diagnosis tree
__IF__ SVI > 0% high __THEN__ conclude that Non-filamentous Bulking=possible and start Non-filamentous Bulking Diagnosis tree
__IF__ SVI > 0% high __THEN__ conclude that Biological-foaming=possible and start Biological Foaming Diagnosis tree

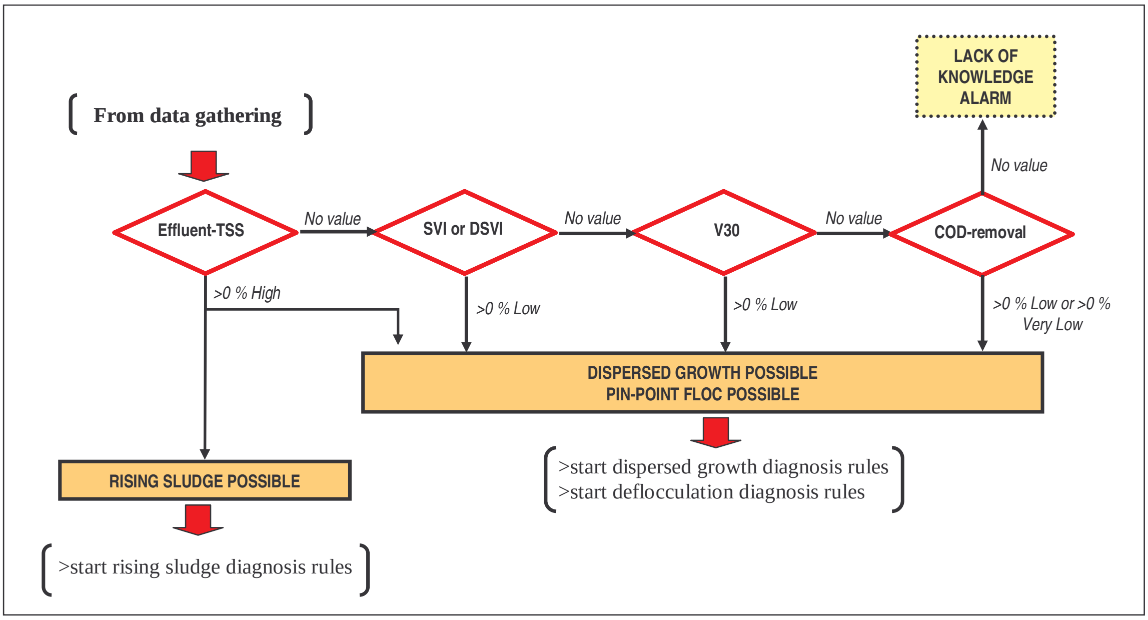


Figure 5.7 Decision tree for the meta-diagnosis of some of the solids separation problems: dispersed growth, pin-point floc and rising sludge

Module 2: diagnosis of solids separation problems

Whenever any of the premises of meta-diagnosis trees are confirmed (a situation is possible) the decision trees of the corresponding solids separation problem are launched in order to confirm the possible situations. Six diagnosis decision trees (see Annex) have been developed to identify each one of the possible solids separation problems: filamentous bulking diagnosis; non-filamentous bulking diagnosis; biological foaming diagnosis; dispersed growth diagnosis; pin-point floc diagnosis and rising sludge diagnosis. All the diagnosis decision trees use both quantitative and qualitative variables to identify the current separation problems. Figure 5.8 shows an example of the decision tree defined to diagnose situations of filamentous bulking.

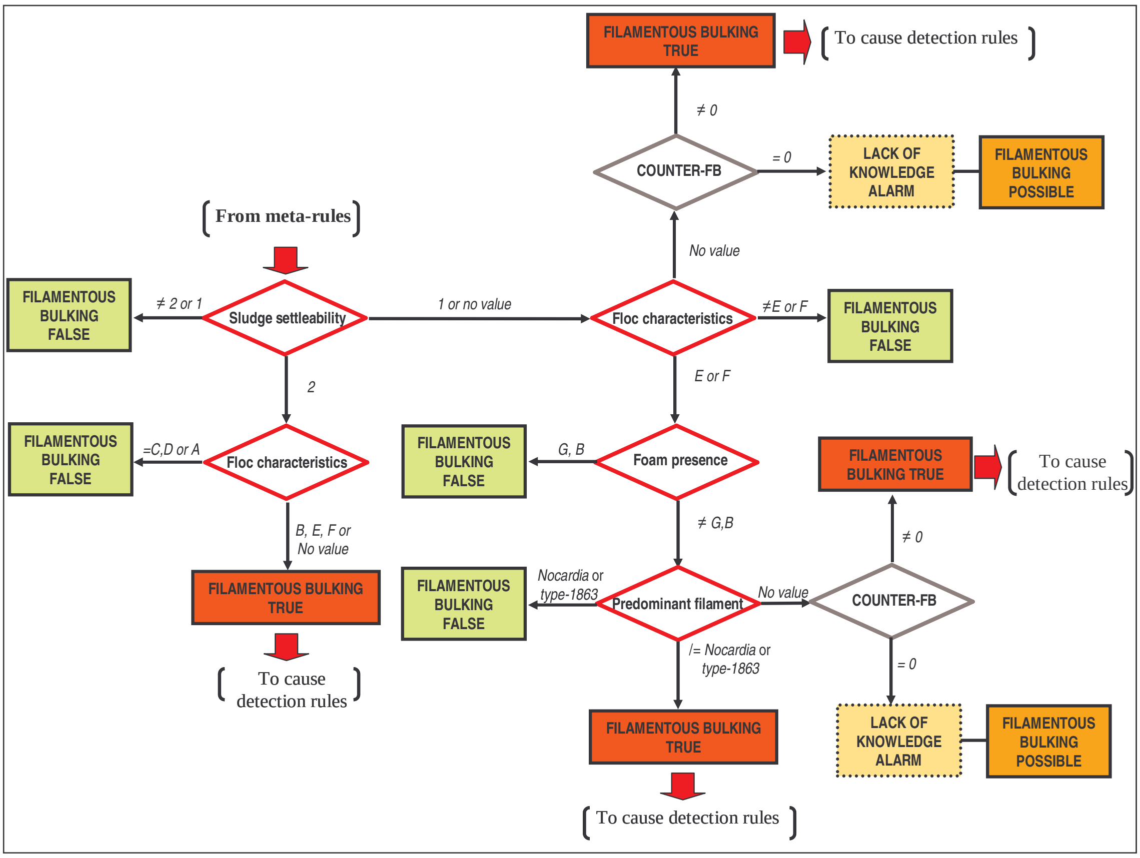


Figure 5.8 Decision tree for the diagnosis of filamentous bulking

Use of dynamic reasoning

In comparison with conventional KBDSSs based on static reasoning, our dynamic approach uses dynamic reasoning in order to identify the state of the process whenever a diagnosis is made. The following are some examples which demonstrate the usefulness of incorporating dynamic reasoning:

- ⇒ Whenever a situation is identified, the reasoning process is capable of identifying if this is the beginning of a new solids separation problem or if it is part of a problem previously initiated and probably already handled.
- ⇒ In some occasions where the lack of data makes it difficult to make a diagnosis, the use of dynamic reasoning can reach an approximate conclusion (see Figure 5.8, where despite the lack

of data concerning sludge settleability and floc characteristics, a conclusion of filamentous bulking diagnosis can be reached).

⇒ Whenever a diagnosis is made, the reasoning process is capable of identifying if the process is evolving to a better state (to a situation of no solids separation problem: improvement), if it is keeping static (no important changes), or if it is evolving to a worse situation (worsening).

In this sense, some specific rules or procedures are usually launched in parallel to gather some extra knowledge about the process, necessary to carry out this dynamic reasoning procedure. Some examples include the use of the status-rules which permits the registration of the number of days that an episode of solids separation problems lasts on the process by means of specific counters. An example is showed next:

IF *Filamentous-bulking=true and the FB-counter ≠ 0* ***THEN*** *conclude that FB-counter = FB-counter + 1*
IF *Filamentous-bulking=false and the FB-counter ≠ 0* ***THEN*** *conclude that FB-counter = 0*

As observed in Figure 5.8 the use of the information registered in such counters is used in different reasoning processes. Apart from the status rules, other type of rules and procedures are usually launched to compile (and sometimes calculate) some extra information from the process. The calculation of SVI-trend or the calculation of the effluent-TSS-trend are some other examples that will be depicted during the description of the prediction module. Apart from being used during the first cycle of the ES, the dynamic reasoning is also part of the reasoning processes followed both in cycle 2 (identification of the possible causes) and in cycle 3 (where a control strategy to solve the diagnosed problematic situation is suggested).

Tackling the uncertainty

The ambiguity found in the linguistic description of a concept during the decision-making process forced us to incorporate a measure of uncertainty into the DSS conclusions to fully exploit the fallible but valuable judgemental knowledge offered by human experts. The uncertainty associated with the description of heuristic knowledge is a feature common to the problems of any domain involving expert knowledge. There are many approaches to cope with this vagueness. In our approach, uncertainty is tackled using the principles of the fuzzy decision theory (Bellmann and Zadeh, 1970). The main idea is to represent the uncertainty of a value by the degree to which it is a member of a certain fuzzy set. With fuzzy truth values the system will be able to conclude, for example, to which extent a filamentous bulking problem occurs (the degree of membership, also known as possibility, in the set or class filamentous bulking) instead of saying unequivocally whether filamentous bulking is occurring or not.

In our approach this uncertainty has been tackled by applying the fuzzy sets of the quantitative variables developed during the discretisation module. Hence, by taking into account the membership degree of the different quantitative variables considered in the diagnosis trees, the final diagnosis would incorporate the percent certainty associated with a conclusion. Figure 5.9 illustrates how certainty is tackled in the decision-making process followed to diagnose a filamentous bulking situation. In case the SVI value was not available, the membership degree of the V30 will be considered instead.

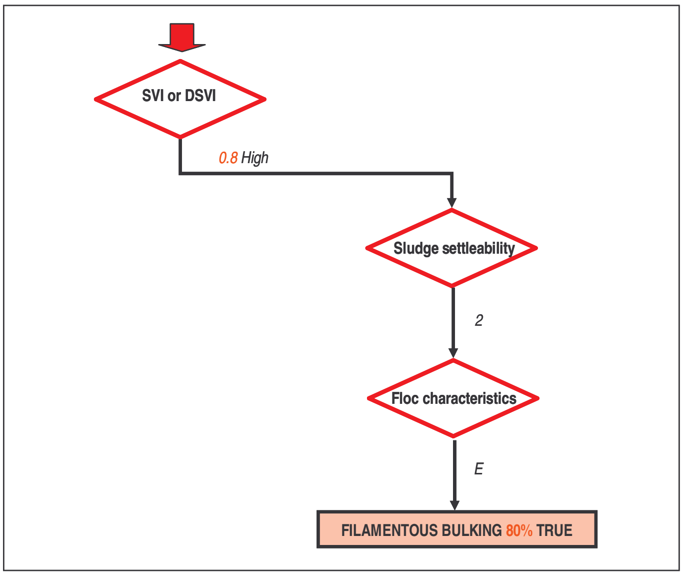


Figure 5.9 Tackling of certainty within the decision-making process for filamentous bulking diagnosis

Module 3: prediction of future solids separation problems

Considering that some relevant data that could of much help in the prediction of solids separation problems, such as the use of image analysis or the measurement of TEFL, among others (see chapter 3), are not commonly measured, the prediction of future solids separation problems is certainly a difficult task even for a KBDSS.

Despite this lack of relevant information, a prediction module has been incorporated into the dynamic ES, which incorporates the calculation of some extra parameters (SVI trend and effluent-TSS trend) that provide the system with some extra information necessary to: (1) predict future solids separation problems, and (2) study the evolution of initiated problems. The latter capability is very useful whenever the effect of a control method is being evaluated.

The evolution of the process is estimated by means of two main quantitative variables directly related to solids separation problems: the SVI and the effluent-TSS. As an example, the estimation of SVI trend is depicted next:

- If the coefficient of variation[2] of the average of the 3 previous values is < 1: SVI trend = (current SVI value - SVI mean of the previous 3 days)
- If the coefficient of variation of the average of the 3 previous values is > 1: SVI trend = (current SVI value - SVI of the previous day)

The coefficient of variation was used in order to consider the dispersion of data whenever the trend is calculated. A coefficient of variation is >1 means that the dispersion around the mean is so large that the value of the trend would be magnified if calculated by using the mean of the previous values; consequently, the trend needs to be calculated by considering only the value of the immediately previous day.

---

[2] Statistical measure of the deviation of a variable from its mean. It is defined as the ratio of the standard deviation to the mean and is calculated as follows: c_v=standard deviation (σ)/mean (μ)

According to the typical values of the studied plants, the following limits have been considered:

⇒ SVI trend values < -30 are considered low (decreasing trend)
⇒ SVI trend values >-30 and ≤ 30 are considered normal (no trend)
⇒ SVI trend values > 30 are considered high (increasing trend)

⇒ Effluent-TSS trend values < -15 are considered low (decreasing trend)
⇒ Effluent-TSS trend values >-15 and ≤ 15 are considered normal (no trend)
⇒ Effluent-TSS trend values > 15 are considered high (increasing trend)

This extra information helps the system to diagnose whether the process is **evolving**, if it is getting better or worse or if it is static in terms of sludge settleability and loss of solids, i.e. in terms of solids separation problems occurrence.

On the other hand, the information provided by these calculated trends can be used together with some other important information from the process in order to **predict future problematic situations**. In the present approach, SVI trend has been combined with the membership degree of the current SVI value in order to predict future situations of solids separation problems resulting from either an increasing SVI or a decreasing SVI. Figure 5.10 depicts the different reasoning process followed to predict future solids separation problems.

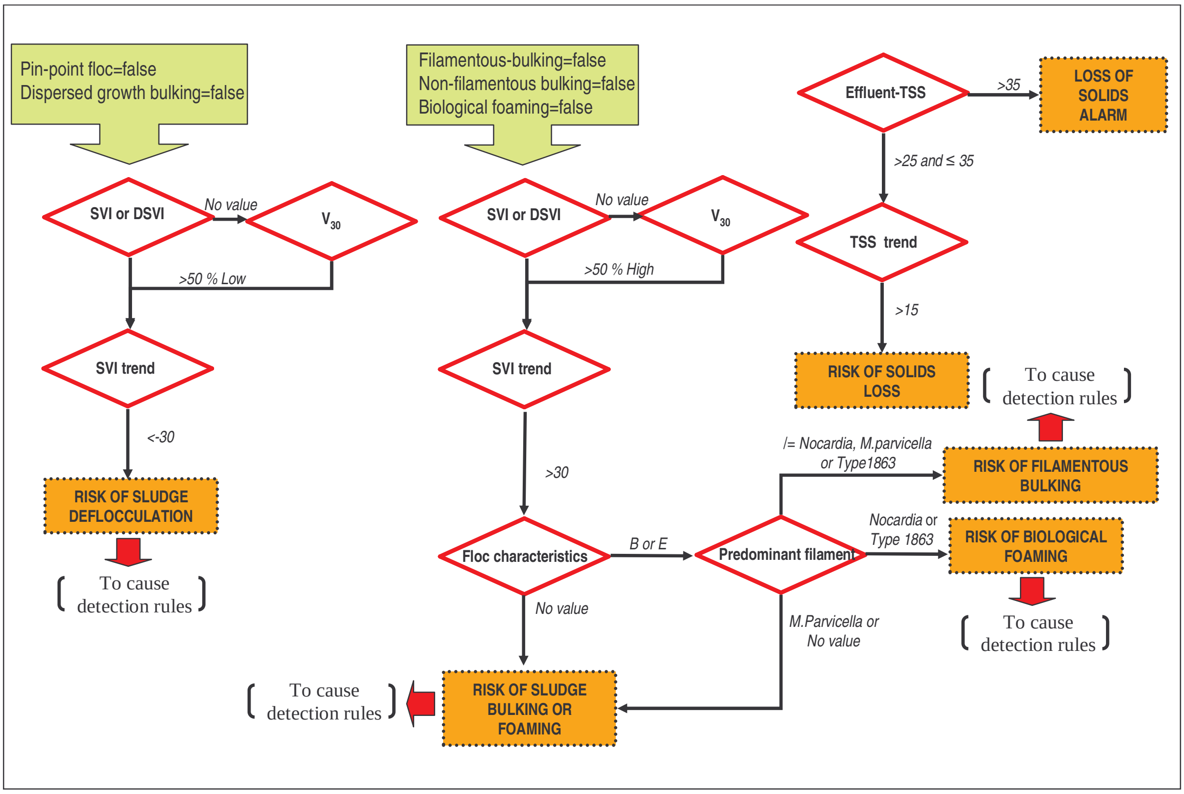


Figure 5.10 Decision trees for the prediction of future solids separation problems

These kinds of predictions have been classified into 3 different categories or alarms:

⇒ Risk of sludge bulking or foaming/ Risk of filamentous bulking/ Risk of biological foaming

- ⇒ Risk of sludge deflocculation
- ⇒ Risk of solids loss

Apart from these 3 types of alarms, an alarm of **loss of solids** (see Figure 5.10) is fired every time the effluent limit of TSS is exceeded, i.e. effluent-TSS > 35 mg/L. Regarding the prediction of rising sludge episodes, it needs to be mentioned that because some relevant variables such as the DO, COD or NO_3^- concentrations at the clarifier influent are not usually measured (see chapter 3), predicting such situations is difficult.

An example of the decision-making process followed to predict a risk of sludge bulking or foaming is depicted next:

1) Whenever none of the possible situations characterized by a high SVI, i.e. filamentous bulking, non-filamentous bulking or biological foaming, have been identified as *true* or *possible* situations, the following evaluation is performed:

   a. IF the membership degree of SVI to the *high* set is > 50% (or the membership degree of V30, in case no value of SVI was available, to the *high* set is >50%) and the SVI-trend=*increasing* THEN a risk of future bulking or foaming problems will be diagnosed.

      i. IF the Floc characteristics = B or E and the Predominant filament ≠ *M. parvicella*, *Nocardia* sp or Type 1863 THEN a risk of filamentous bulking will be predicted.

5.1.4 Decisions

The decisions or conclusions that the ES reaches at the end of the first cycle include:

- ✓ Diagnosis of the current state of the process, according to the most recently available data.
- ✓ Calculation of the percent certainty for the reached diagnosis.
- ✓ Determination of the problem extent, if the current state is new or is part of an already initiated situation.
- ✓ Identification of process trends and prediction of future problematic situations.

5.2 Identification of the Possible Causes

Once the new situation is diagnosed and the first goal is achieved, the second cycle of the ES starts. The second cycle, depicted in Figure 5.11, follows the main structure of the first cycle, but the content of the knowledge bases and the computational functionalities are quite different.

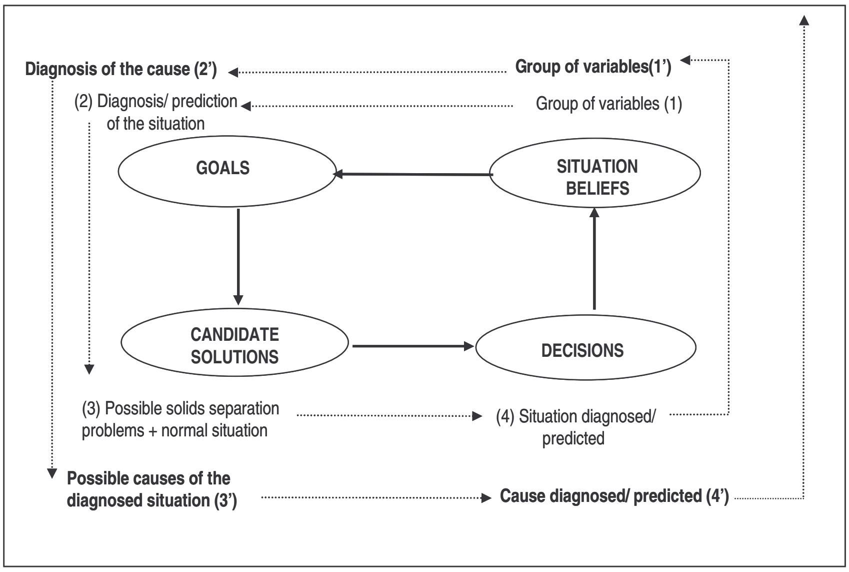


Figure 5.11 Second cycle of Expert System architecture

5.2.1 Situation Beliefs

During the second cycle new data are gathered to become part of the new knowledge base of Situation Beliefs. This is the set of information that the ES requires to identify the possible causes of the previously diagnosed situation. As in the previous cycle, both quantitative and qualitative data are gathered.

Table 5.1 Quantitative values compiled within the knowledge base of Situation Beliefs during the second cycle

Variable	Description
COD/P	Variable calculated by COD and Total P at the primary effluent
COD/N	Variable calculated by COD and Total N at the primary effluent
SRT or MCRT	Operational variable, if MCRT is available it is preferred to SRT
F/M	Operational variable
% of soluble COD	Variable calculated by total COD and soluble COD at primary effluent and at effluent
DO	On-line variable. The daily average value is used. It is preferred to the DO set-point to measure the real oxygen concentration within the reactor
Aeration	Variable calculated by air-flow and DO set-point
% COD removal	Variable calculated by influent COD and effluent COD
pH	On-line or analytical variable. If on-line, the daily average value is used
Sulphide	Analytical variable measured at the influent or primary effluent (preferred)
Temperature*	On-line variable. The daily average value is compiled.
	(*) Although temperature is gathered with the rest of possible causes, this variable will not be identified as the unique specific cause of any solids separation problem (but as one of the possible causative factors) since no specific control action is usually carried out to increase or decrease the temperature of wastewater.

The quantitative data consist of a set of analytical and operational data. All these variables are supposed to be measured at least once a day; if more than one value per day was gathered, the average value would be used. Table 5.1 shows the set of quantitative variables chosen for this application. In case any of the calculated or operational variables (COD/N, COD/P, MCRT or F/M) was not measured, despite having at its disposal all the required data, the specific variable will be automatically calculated by the system during the data gathering process.

Besides, four qualitative variables, gathered by means of the macroscopic and microscopic worksheet, are also part of the current situation beliefs knowledge base used to identify some of the possible causes. These qualitative variables are:

- Presence of foam
- Presence of protozoa
- Predominant filament
- Predominant group

5.2.1.1 Data gathering

As in the previous cycle, the data gathering process is preformed by means of a data reconciliation system including the discretisation of the quantitative variables. For each of these variables, except for COD removal (already defined for problem diagnosis) and temperature, a membership function was developed. The defined fuzzy sets are used later on, during the diagnosis of the possible causes. As commented with regards to the problem diagnosis cycle, a change in the limits for each of the defined fuzzy sets will be enough to update the current approach. Figure 5.12 shows as an example of the membership function developed for the SRT variable.

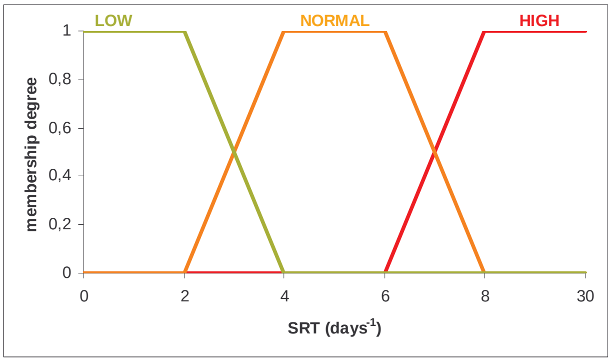


Figure 5.12 Membership function for the quantitative variable of SRT

The membership degree for each category is used in the cause diagnosis process to determine the percent certainty when a possible cause is identified, especially in the prediction of future events.

Frequency of data

Given that the objective of the dynamic ES is to offer a daily support to handle solids separation problems of microbiological origin, characterized by their slow dynamics, the daily variations, such as the diurnal dynamics, of most of the variables identified as possible causes (F/M, DO, SRT, etc.) are considered as not significantly affect the occurrence of solids separation problems. Given the slow dynamics of this type of situations, any sudden change within the process, except for the presence of toxics within the influent, will not immediately involve changes on the sludge settleability. For example, a specific decrease on the DO concentration or on the F/M ratio will not directly favour the occurrence of bulking. Given that the origin of most solids separation problems is the imbalance between floc-forming and filamentous bacteria, a specific condition is not enough to encourage for example the overgrowth of filamentous bacteria and thus to provoke an episode of filamentous bulking. The suitable conditions must occur for a longer period of time to have a direct effect on the biomass. According to this, we consider that the frequency of measurement (usually one measurement per day) used in our approach is suitable enough to reflect those conditions that significantly persist for long periods of time and thus are clear requisites for the occurrence of solids separation problems.

5.2.2 Goals

The goals of the second cycle of the dynamic ES are two-fold: (1) to identify the possible cause(s) of the previously identified situation, according to the most recent data and taking into consideration the problematic conditions already present in the process or the trend to future problematic conditions, and (2) when future problematic situations are predicted, to identify the origin of such situations. Candidate solutions are determined according to these goals.

5.2.3 Candidate solutions

The candidate solutions are the set of conditions that can be inferred as the possible cause of the different solids separation problems. Table 5.2 shows the set of possible causes that the ES is able to identify for each of the problems.

Table 5.2 Set of possible causes (candidate solutions) for each of the identified solids separation problems

PROBLEM	POSSIBLE CAUSES (CANDIDATE SOLUTIONS)	
Filamentous bulking	▪ Nitrogen deficiency ▪ Phosphorous deficiency ▪ Low pH ▪ High sulphide ▪ High readily biodegradable substrate	▪ Low DO ▪ Low F/M ▪ High SRT ▪ Cause not identified
Non-filamentous bulking	▪ Nitrogen deficiency ▪ Phosphorous deficiency ▪ High F/M	▪ Low SRT ▪ Cause not identified
Biological foaming	▪ Low F/M ▪ High SRT ▪ High F/M ▪ High readily biodegradable substrate	▪ Low SRT ▪ Low DO ▪ Cause not identified
Dispersed growth	▪ Overaeration ▪ Poorly biodegradable surfactants ▪ High F/M	▪ Low SRT ▪ Toxic shock ▪ Cause not identified
Pin-point floc	▪ Overaeration ▪ Poorly biodegradable surfactants ▪ High F/M ▪ Low SRT ▪ Toxic shock	▪ Low DO ▪ Low pH ▪ Low F/M ▪ Overoxidized sludge ▪ Cause not identified
Rising sludge	Set of necessary conditions* for the occurrence of denitrification in secondary settlers: ▪ High SRT ▪ High nitrate concentration in clarifiers ▪ High BOD concentration in clarifiers ▪ Low DO in clarifiers (*)Since all these conditions are required for the occurrence of denitrification, they have been considered as a whole cause. Thus, whenever rising occurs, it is supposed that the main cause is the set of these conditions (which usually can not be checked because not any analysis is performed at the reactor effluent).	

5.2.3.1 Reasoning process

During this second cycle, the values for the different variables stored in the knowledge base of Situation Beliefs are used jointly with the human expert knowledge represented in the decision trees specifically developed for cause identification. To accomplish the main objective of troubleshooting the possible causes of settleability, the reasoning process has been divided into 3 different modules: (1) identification of the possible cause(s) of the diagnosed solids separation problems; (2) identification of future problematic conditions, and (3) identification of the possible cause(s) of the predicted solids separation problems.

Module 1: identification of the possible cause(s) (diagnosed problems)

During the second cycle, a meta-diagnosis module to direct the reasoning to identify the possible causes is not necessary since the corresponding decision trees are already launched by the specific decision trees of diagnosis, once a situation is identified. Therefore, during the second cycle, only the decision trees to identify the cause(s) of each diagnosed solids separation problem are necessary to reach a conclusion. Eight decision trees (see Annex) have been developed to identify the possible causes. Regarding rising sludge, a direct conclusion about the possible cause(s) of undesired denitrification in clarifiers is reached because the whole set of conditions (including high SRT, high DO, high COD and high NO_3^- concentrations) must be met for its occurrence. On the other hand, for the identification of the possible causes of foaming, two different decision trees have been developed. The first one (see Figure 5.13) incorporates a kind of meta-diagnosis or discriminating procedure by which the predominant filament is initially checked in order to delimit the set of possible causes (candidate solutions). If information about the predominant filament is not available, the generic decision tree for the diagnosis of foaming causes will be launched (see Annex).

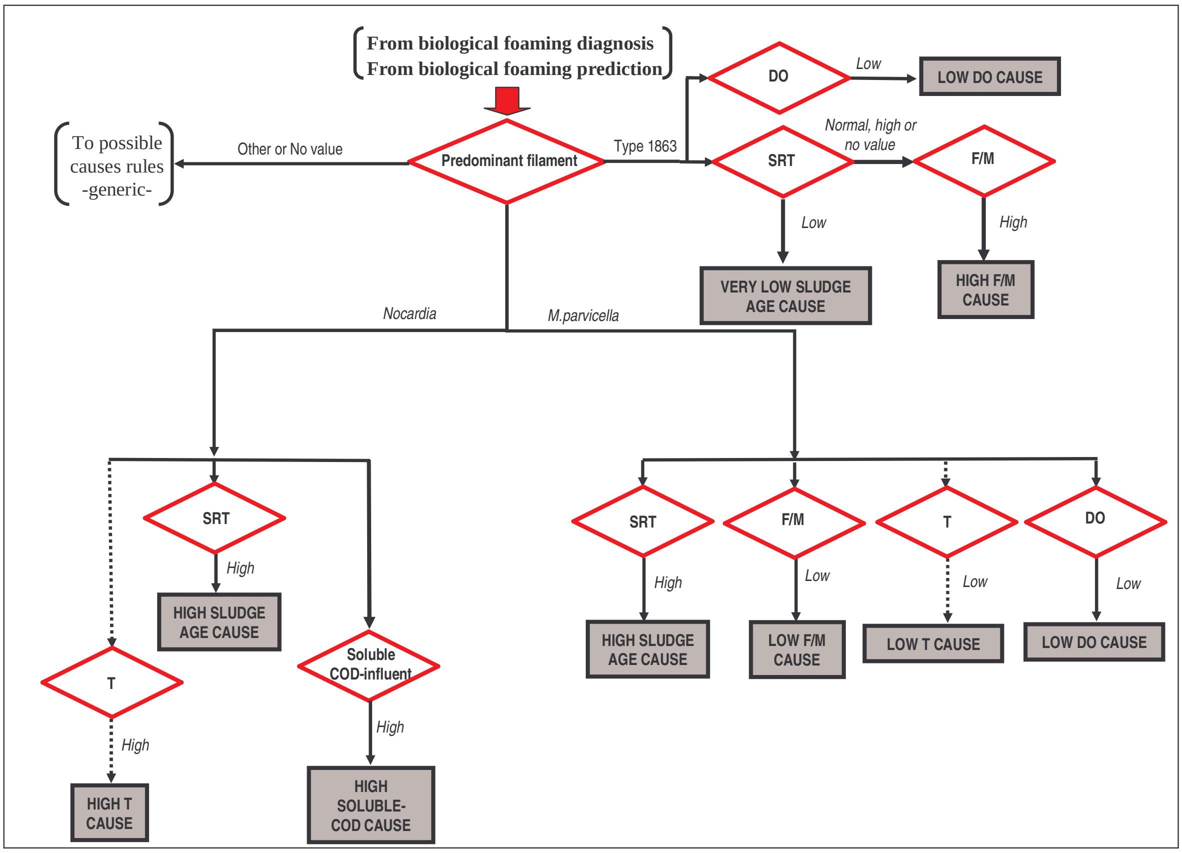


Figure 5.13 Decision tree for the identification of the possible causes of biological foaming

If none of the candidate solutions is identified as the possible cause, the concluded diagnosis will be: cause not identified. This can happen both when all causes are false or when some causes are false and some of them have not been identified due to a lack of available data. In such circumstances, the user will be advised to measure the required variable(s).

Module 2: identification of future problematic conditions

During the second cycle, the percent of certainty obtained through the application of the membership functions is used in combination with the calculation of the evolution of all the variables related to the possible causes or problematic conditions. That is, whenever a candidate solution is not identified as a possible cause because the membership degree of its value is not directly classified within the "problematic" category (membership degree ≠1) the following evaluation is performed:

1) Whenever the membership degree of the value to the problematic category is >0.5 and <1, the trend of this variable is calculated.

a. If the trend shows that the variable is evolving towards the problematic category (increasing or decreasing, depending on the variable) then the system will predict a risky condition that might result in a specific solids separation problem.

The trend is estimated by means of the following procedure:

- If the coefficient of variation of the average of the values for the last 7 days was < 1:

Variable trend = (current value – mean value of the previous 7 days)

- If the coefficient of variation of the average of the values for the last 7 days was > 1:

Variable trend = (current value – last value of the variable)

Trends for the variables associated with the possible causes are not calculated with just the 3 previous values but considering the values for the previous 7 days. The purpose is to overcome those situations where the variable is not measured daily. By determining a range of 7 days we assume that at least more than 1 value will be available.

Module 3: identification of the possible cause(s) (predicted problems)

As soon as a problematic situation is predicted (e.g., risk of filamentous bulking, risk of biological foaming or risk of deflocculation), the specific decision trees to identify the possible cause are launched. For example, if a biological foaming situation is not diagnosed but predicted as a future problem the biological foaming cause detection decision tree (see Figure 5.13) will be launched. As in the case of a diagnosed situation, the user will receive the conclusions from this reasoning process, but in this case, the following cycle concerning control strategy proposal will not be launched.

5.2.4 Decisions

The decisions or conclusions that the ES reaches at the end of the second cycle include:

- ✓ Diagnosis of the possible causes of the identified or predicted solids separation problems, according to the most recent available data
- ✓ Prediction of future problematic conditions (possible causes) that can lead to solids separation problems

5.3 Proposal of a Control Strategy

Once the first and second cycles of the Expert System reach their conclusions about the solids separation problem present on the process and about its cause, the system starts the last cycle. The third cycle, depicted in Figure 5.14, is the longest one, since this time the Expert System must propose a complete control plan to restore the process. Moreover, given that planning is involved, the whole cycle proposed by the Domino Model must be completed. The 6 knowledge databases and the corresponding computational functionalities involved in the cycle are described next.

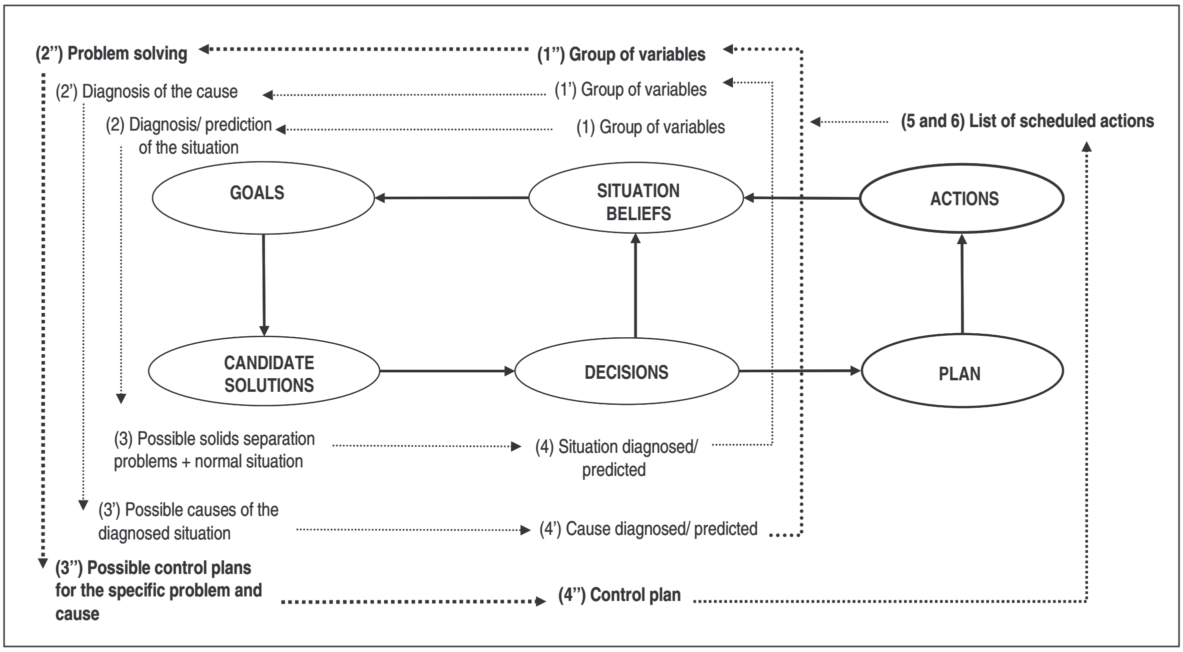


Figure 5.14 Third cycle of the Expert System architecture

5.3.1 Situation beliefs

The set of data gathered during the last cycle constitutes the set of information that the ES needs to make the final decision concerning the selection of the most appropriate control strategy to restore the process. This set of data can be divided into two main groups:

⇒ Information about the current problem and cause that the system is undergoing (identified during the previous cycles) in order to define, as far as possible, a specific control method. The required information stored in the knowledge base of Situation Beliefs includes both the diagnosed problem and the identified cause.

⇒ Information about the evolution of the problem and the effect of the applied control actions (in case any control plan has already been initiated). This information is compiled by means of calculated trends and specific counters specifically defined to monitor the evolution of the process and the number of days that a specific action has been applied.

5.3.1.1 Data gathering

The data gathering process is performed according to the evolving necessities of the process. Thus, as soon as specific information is required during the reasoning process (i.e. SVI-trend, counter-PD, SRT, etc.), the specific variables are calculated and compiled within the knowledge base of Situation Beliefs. This is another characteristic that the use of dynamic reasoning incorporates in our system.

5.3.2 Goals

Due to the special characteristics of the solids separation problems and the disturbing consequences of their occurrence in activated sludge systems, the goal of the third cycle of the dynamic ES is to choose the most appropriate control method to restore the problematic situation as fast as possible in order to minimize the drawbacks caused by the problems and if possible, to definitively eliminate them from the process. The main criterion followed to define the control plan is, first of all, to propose the specific control method that can most efficiently solve the situation. Only when those specific actions are not effective, or when the cause has not been identified, a non-specific control method is suggested. According to the current objective, the different candidate solutions are defined.

5.3.3 Candidate Solutions

The candidate solutions are the set of possible control methods that can be applied to solve each of the solids separation problems.

The existing control methods that can be used to solve solids separation problems are summarized in Tables 1.7 and 1.8. Nevertheless, not all of these have been considered in our application. According to the study presented in chapter 3, only those control actions that can be directly o potentially applied in a full-scale facility (i.e. those control plans that have a high proven efficiency in the day-to-day operation) have been considered in the present approach. Table 5.3 summarizes the set of possible actions (both specific and non-specific) that can be considered to solve each of the possible problems.

Table 5.3 Set of possible control plans (specific and non-specific) representing the candidate solutions of the 3rd cycle

PROBLEM	POSSIBLE CONTROL PLANS (CANDIDATE SOLUTIONS)	
	SPECIFIC	NON-SPECIFIC
Filamentous bulking	▪ Nitrogen addition ▪ Phosphorous addition ▪ Increase aeration ▪ Decrease SRT or increase F/M ▪ Increase pH ▪ Reduce septicity ▪ Use of selectors (aerobic, anoxic or anaerobic)	▪ Chlorination ▪ Coagulants/flocculants addition
Non-filamentous bulking	▪ Nitrogen addition ▪ Phosphorous addition ▪ Increase SRT or decrease F/M	▪ Coagulants/flocculants addition
Biological foaming	▪ Decrease SRT or increase F/M ▪ Increase SRT or decrease F/M ▪ Increase aeration ▪ Use of selectors (aerobic, anoxic or anaerobic)	▪ Chlorination ▪ Chlorination of foams ▪ Coagulants/flocculants addition ▪ Foam removal
Dispersed growth	▪ Increase SRT or decrease F/M ▪ Inhibition of toxics ▪ Surfactants removal ▪ Decrease aeration or avoid excessive shearing	▪ Coagulants/flocculants addition
Pin-point floc	▪ Increase SRT or decrease F/M ▪ Inhibition of toxics ▪ Surfactants removal ▪ Decrease aeration or avoid excessive shearing ▪ Decrease SRT ▪ Increase pH	▪ Coagulants/flocculants addition
Rising sludge	▪ Favour denitrification in reactors ▪ Avoid nitrification in reactors ▪ Avoid denitrification in clarifiers	

5.3.3.1 Reasoning process

The first reasoning process followed to reach a primary conclusion regarding the type of control strategy to apply is launched in order to decide between applying specific or non-specific control methods. In general, when the problem and the cause have been correctly identified the reasoning process will decide to apply a specific control method. Only when the cause is not identified, or when the specific control methods are not effective are non-specific control methods suggested.

5.3.4 Decisions

The primary decisions made during the first reasoning process reach a primary decision regarding the kind of control method to apply:

- ✓ A **specific control method**, dependent on the specific problem and identified cause and specially defined to eliminate the specific cause that induces the specific solids separation problem in order to definitely remove it.

- ✓ A set of possible **non-specific control methods**, dependent of the diagnosed problematic situation, and defined to temporarily remove the symptoms of the solids separation problem.

The suggested control method must be presented in the form of an organized and scheduled control strategy, understood as a set of detailed control actions where the execution order of the actions, the expected deadline of effectiveness and a set of alternative actions are provided. In order to specify and plan for this complete control strategy, the system must follow further reasoning processes. These are discussed next.

5.3.5 Plan and actions

Once a decision about using specific or non-specific control methods has been made in the previous reasoning process, further steps in the reasoning process lead to a decision about the type of control strategy proposed to restore the process in the most efficient way. As represented in Figure 5.15, the possible strategies that can be applied are the following:

- ⇒ To **define** a new control plan.
- ⇒ To **continue** with the previously applied control plan. Due to the complexity of the solids separation problems, the application of a control plan can last several days.
- ⇒ To **modify** the applied control plan since it has exceeded the pre-defined term of effectiveness.
- ⇒ To **provide** a preventive control plan to be applied while the information that is not available is being checked.

5.3.5.1 Reasoning process

During the following reasoning process, a feed-back of information between the knowledge base of Situation Beliefs and the reasoning process is carried out. The values of different variables (e.g. counter-PD, 3MCRT, etc.) stored in the knowledge base of Situation Beliefs are used jointly with the human expert knowledge represented in the specific decision tree developed to decide the type of control strategy to apply. Figure 5.15 depicts schematically this reasoning process. Blue arrows symbolise the reasoning process followed to suggest specific control strategies while the orange arrows depict the reasoning process followed to propose non-specific strategies.

Following the schematic reasoning process represented in Figure 5.15, 39 decision trees (see Annex) have been developed to define the complete control strategy that, according to the expert criteria of the system, is the most appropriate to restore the process.

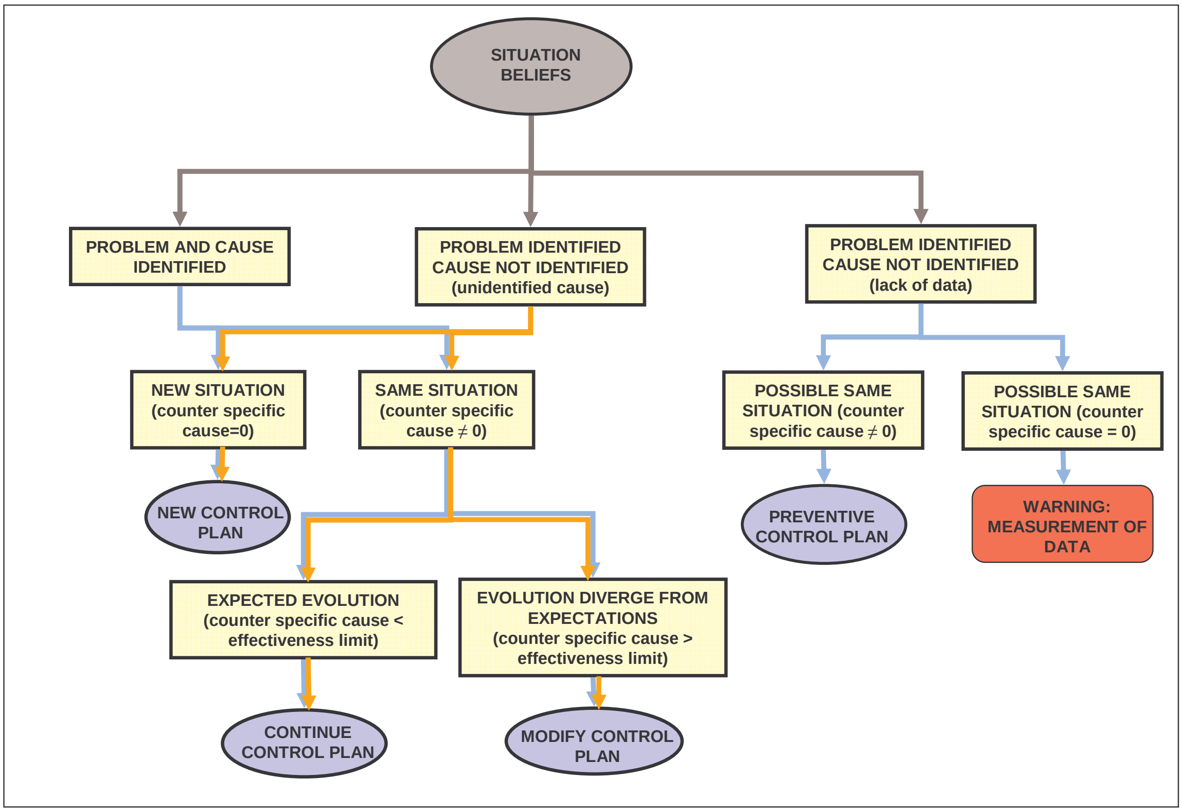


Figure 5.15 Reasoning process followed by the dynamic ES to choose between the possible control strategies (blue arrows reach specific control strategies while orange arrows lead to non-specific control plans)

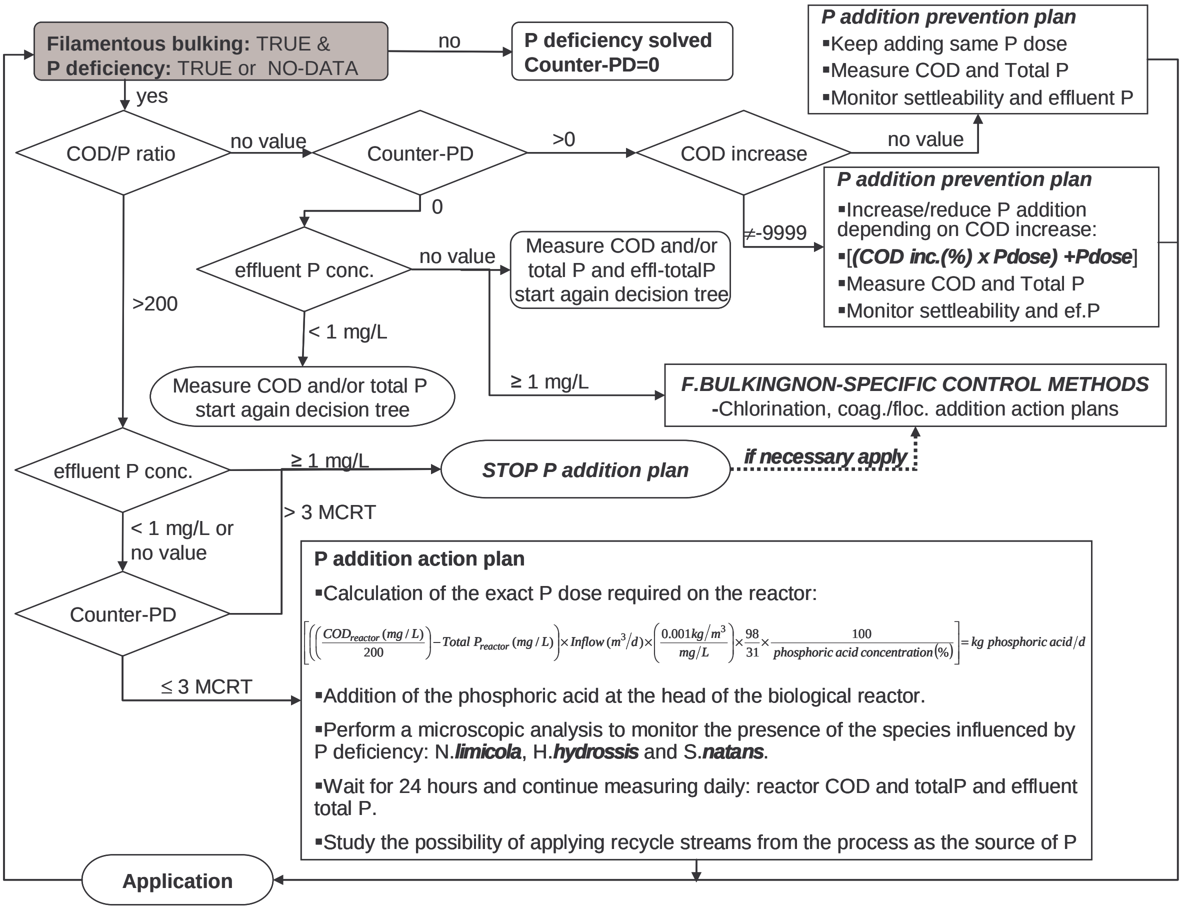
Figure 5.16 represents the reasoning process followed to decide on the most appropriate control method to solve an identified situation of filamentous bulking caused by phosphorous deficiency. Notice that the decision tree starts with the information enclosed in the Situation Beliefs knowledge base regarding the previously diagnosed situation and its cause (grey box in the tree). Next, extra data from the Situation Beliefs (symbolized as parallelograms in the decision tree), including data from the process or specific counters, are checked in order to suggest the most appropriate control strategy. The final conclusion of the reasoning process can involve:

⇒ A set of planned actions (complete plan) that must be immediately applied in order to restore the process as soon as possible.

⇒ A set of intermediate actions whose objective is to compile the necessary information before reaching a conclusion (e.g. “Measure COD and/or total N and effluent-total N”). As soon as this information is available, the reasoning process can be started again.

⇒ Set of actions suggested as a preventive measure while the necessary information is gathered or calculated.

⇒ Invocation of non-specific control methods. They are invoked whenever data from the process demonstrate that the specific control method would not be suitable for the process (e.g., it would not be suitable to add phosphoric acid to the process if the system shows some problems in removing the present phosphorous) or whenever the specific control plan exceeds the expected term of efficiency.

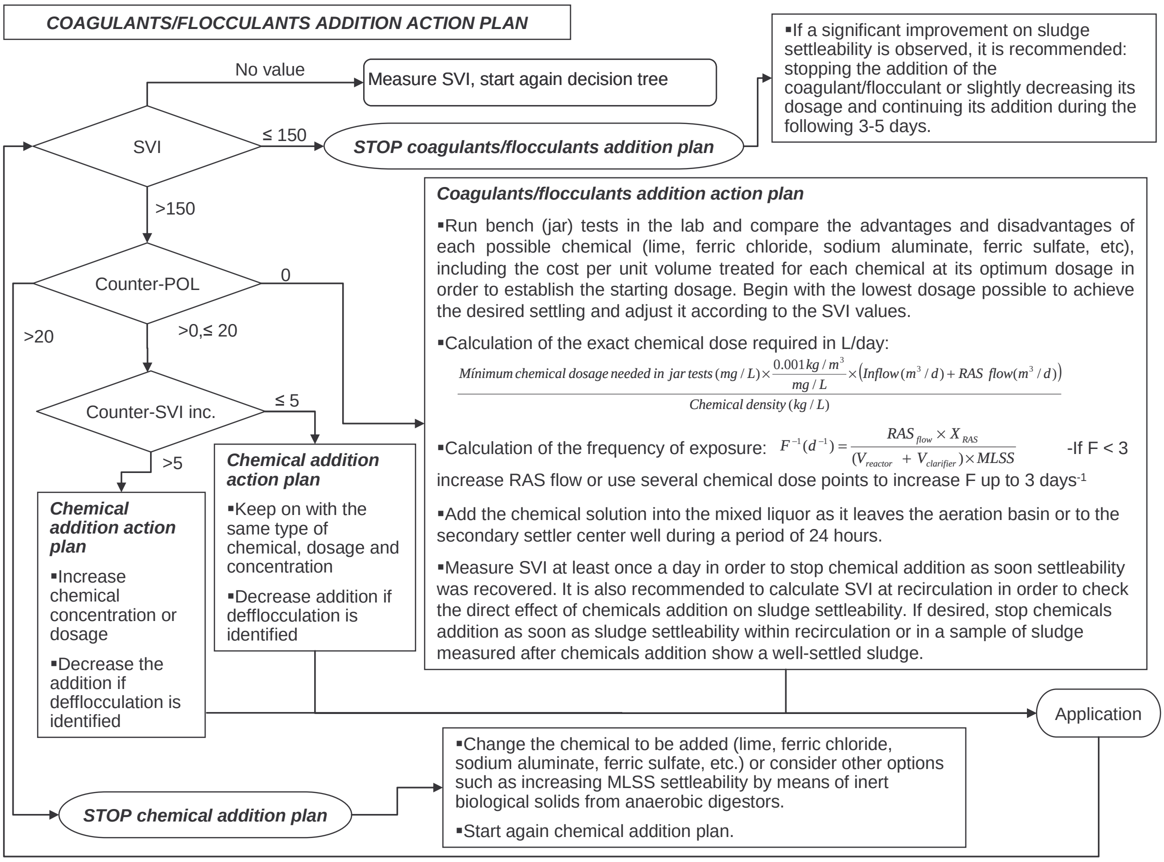


$$\left[\left(\left(\frac{COD_{reactor}\,(mg/L)}{200}\right)-Total\ P_{reactor}\,(mg/L)\right)\times Inflow\,(m^3/d)\times\left(\frac{0.001 kg/m^3}{mg/L}\right)\times\frac{98}{31}\times\frac{100}{phosphoric\ acid\ concentration\,(\%)}\right]=kg\ phosphoric\ acid/d$$

Figure 5.16 Decision tree for the selection of the most appropriate control method to solve a situation of filamentous bulking caused by phosphorous deficiency

As commented above, the application of non-specific control methods is only suggested under one of the following premises: (1) the cause of the solids separation problem has not been properly identified; (2) the specific control method has not produced any improvement in the process (during the expected term); or (3) long specific control methods are applied and fast solutions are necessary (e.g. while a selector is being built).

Figure 5.17 depicts a decision tree developed to define a non-specific control method for the solution of a pin-point episode. Notice that throughout the reasoning process followed by the system, the monitoring of several data stored into the knowledge base of Situation Beliefs (i.e. SVI, counter-POL or counter-SVI inc.) is also required.



▪If a significant improvement on sludge settleability is observed, it is recommended: stopping the addition of the coagulant/flocculant or slightly decreasing its dosage and continuing its addition during the following 3-5 days.

Coagulants/flocculants addition action plan

▪Run bench (jar) tests in the lab and compare the advantages and disadvantages of each possible chemical (lime, ferric chloride, sodium aluminate, ferric sulfate, etc), including the cost per unit volume treated for each chemical at its optimum dosage in order to establish the starting dosage. Begin with the lowest dosage possible to achieve the desired settling and adjust it according to the SVI values.

▪Calculation of the exact chemical dose required in L/day:

$$\frac{Mínimum\ chemical\ dosage\ needed\ in\ jar\ tests\ (mg/L)\times\frac{0.001\,kg/m^3}{mg/L}\times\left(Inflow\,(m^3/d)+RAS\ flow(m^3/d)\right)}{Chemical\ density\,(kg/L)}$$

▪Calculation of the frequency of exposure: $F^{-1}(d^{-1})=\frac{RAS_{flow}\times X_{RAS}}{(V_{reactor}+V_{clarifier})\times MLSS}$ -If F < 3 increase RAS flow or use several chemical dose points to increase F up to 3 $days^{-1}$

▪Add the chemical solution into the mixed liquor as it leaves the aeration basin or to the secondary settler center well during a period of 24 hours.

▪Measure SVI at least once a day in order to stop chemical addition as soon settleability was recovered. It is also recommended to calculate SVI at recirculation in order to check the direct effect of chemicals addition on sludge settleability. If desired, stop chemicals addition as soon as sludge settleability within recirculation or in a sample of sludge measured after chemicals addition show a well-settled sludge.

Chemical addition action plan

▪Keep on with the same type of chemical, dosage and concentration

▪Decrease addition if defflocculation is identified

Chemical addition action plan

▪Increase chemical concentration or dosage

▪Decrease the addition if defflocculation is identified

▪Change the chemical to be added (lime, ferric chloride, sodium aluminate, ferric sulfate, etc.) or consider other options such as increasing MLSS settleability by means of inert biological solids from anaerobic digestors.

▪Start again chemical addition plan.

Figure 5.17 Decision tree for the selection of the most appropriate control method to solve a situation of pin-point floc

Once the three cycles are completed, the conclusions they reached are communicated to the plant operator in a user-friendly way. With this information at hand, the user should be able to make a decision about how to manage a solids separation problem in the facility.

5.4 Discussion

The structure of the knowledge base, considered as the core of the Expert System, and the necessary reasoning processes, have been efficiently organised by means of the three cycles of operation pre-defined during the preliminary phase of conceptual design (Domino Model). The three cycles have been established basically according to the pre-defined objectives that the ES must accomplish, which are: (1) diagnosis or prediction of solids separation problems; (2) identification of the possible causes, and (3) proposal of a control strategy. This modular structure will facilitate the integration of the ES with other control tools or Artificial Intelligence techniques as well as its implementation in other facilities or domains.

By incorporating dynamic reasoning, the Expert System is capable of adapting its reasoning to the evolving necessities of the domain. Hence, during the dynamic reasoning process, and before making a decision, the ES integrates different types of information from the process (analytical data, microscopic information, trends, counters, etc.) regarding the evolution of the process. Due to this dynamic reasoning the system is aware, for example, of the number of days that a particular problem or cause has been diagnosed (now the system may not just conclude that the process is suffering from a filamentous bulking situation but also inform the user that the process is in the 12th day of an episode of filamentous bulking and that no recovery of the process has been observed). On the other hand, the system is also aware of the number of days that a specific strategy has been applied, as well as the response of the process to a specific control action (if the sludge settleability is improving or getting worse, if the expected deadline of effectiveness has been exceeded, etc.).

Whenever a diagnosis is reached, the dynamic reasoning process is capable of identifying if the process is evolving to a better state (to a situation of no solids separation problem: improvement), if it is keeping static (no significant changes are observed), or if it is evolving to a worse situation (worsening). By considering this information the ES, apart from being aware of the evolution of the process, can also predict the occurrence of future problematic situations. The use of dynamic reasoning also allows the system to reach a diagnosis even when not all the necessary information is available (i.e. by checking the previous diagnosis, and based on partial current information, a hypothetical diagnosis about the current state of the process can be reached).

The application of the Domino Model allows the system to develop complete control strategies (instead of simple control actions) including a set of scheduled actions (specific or non-specific), where the order of execution and the expected deadline of effectiveness are provided, together with a set of alternative control actions to be applied in case the facility presented some limitations in directly applying such suggested actions. With this new approach, the suggested control strategy evolves with time. The use of dynamic reasoning avoids the repetition of the same control strategy whenever the same situation persists in the process. Now the dynamic reasoning can choose to continue with the same control strategy, to modify it (even suggesting alternative control strategies) or to suggest a preventive control plan whenever the lack of information prevents the system from confirming the diagnosis and/or the possible cause. Furthermore, the expected deadline of effectiveness, usually used by the system to change from a control strategy to another (e.g. 3MCRT) is configurable, in such a way that can be adjusted by the user.

Eight decision trees have been developed to diagnose any type of solids separation problems, three decision trees have been developed to predict future solids separation problems, eight more to identify the possible causes and finally thirty-nine decision trees have been developed to suggest the control strategy to be applied.

The quantity of required information (compiled within the knowledge base of Situation Beliefs) has been limited to the set of necessary variables that the reasoning process requests to reach a conclusion. In the present approach, only 22 variables are required throughout the reasoning process represented in the decision trees. Hence, just by checking the value of one quantitative variable (SVI or V30) and a qualitative variable (sludge settleability or floc characteristics) the system is capable of reaching a diagnosis about the type of solids separation problem affecting the process. Regarding the identification of the possible cause, from 5 to 8 different variables (depending on the type of problem) must be checked to reach a conclusion. Finally, whenever a control strategy is proposed, just by checking 2 or 3 different variables (including trends, counters or other analytical variables) a conclusion can be reached.

In order to facilitate the acquisition of qualitative knowledge (considered as a very valuable type of information due to the microbiological origin of the solids separation problems), which includes microscopic and macroscopic observations of the activated sludge, both a microscopic and a macroscopic worksheet have been developed. They have been designed to be filled out in about 30 minutes, so that a daily registration is promoted.

The use of qualitative knowledge during the reasoning process gives the final conclusion a certain degree of uncertainty as a consequence of the subjectivity related to the measurement of such variables.

The use of fuzzy logic during the reasoning process permits the conclusion of uncertain diagnosis (very usual in human reasoning). The development of fuzzy sets for each of the quantitative variables used during the reasoning process allows the system to include the percentage of certainty associated with a diagnosis. This fact provides the user with extra information about how the process is closer or farther to a specific situation.

Given that the objective of the KBDSS is to handle solids separation problems of slow dynamics, the diurnal dynamics of those variables involved in the occurrence of solids separation problems have not been considered in the present approach. We are aware of the effects that the diurnal dynamics of some specific variables have on the sludge properties. A clear example is the diurnal dynamics of the inflow. In those plants where flow equalisers are not used, the diurnal variations of the influent can have direct consequences on the sludge settleability properties, and therefore effluent quality, especially if solids separation problems are present. For example, whenever the sedimentation zone of the secondary settling tank is full of poorly compacted sludge, as a consequence of solids separation problems, a sudden increase in the inflow can directly result in an increase in effluent TSS. In such situations, a real-time control, specifically designed to manipulate recirculation flow according to these sudden flow variations (an considering sludge settleability), would be the most suitable control tool to handle them. In this sense, futher work of our research group is addressed to the integration of real-time control actions (Fiter, 2005b and Fiter, 2006), based on these diurnal dynamics, within the decision support tool.

Because of the complexity of solids separation problems means that there are still a lot of open questions left, related to their possible causes and effects. Thus, the efficiency of the ES in diagnosing and predicting problematic situations and their potential causes is directly affected by this fact.

On the other hand, some limitations derived from the use of Expert Systems must be identified:

- ⇒ Part of the knowledge introduced in the knowledge base, which is acquired from literature revision, is empirical, very generic and sometimes contradictory. The little specific knowledge included in the knowledge base that was extracted from scheduled interviews with plant managers and operators, generally results in some discrepancies or imprecision.

- ⇒ Complex problems require many (hundreds of) rules, causing long development time and creating potential problems in both using and maintaining the system.

- ⇒ The knowledge base is static. Once developed, it is not an easy task, at least for the expert or final user, neither to modify some rules nor to adapt the knowledge base to new specifications. Thus, the system is unable to learn from new experiences. This last fact could provoke the systematic repetition of errors in diagnosis and proposed actions. The cooperation between a knowledge engineer and an expert is therefore usually necessary to uptake the knowledge base.

CHAPTER 6

DEVELOPMENT OF THE EPISODE-BASED REASONING SYSTEM

In order to increase the accuracy and efficiency of the proposed KBDSS, designed to deal with solids separation problems, a second knowledge-based tool, known as Case-Based Reasoning Systems, has been chosen to complement the functionality of the dynamic Expert System. The main objective of the Case-Based Reasoning System (CBRS) is to improve the feasibility of the KBDSS by overcoming the main limitations encountered when using only the Expert System as a decision support tool. As commented during the discussion section of chapter 5, one of the main limitations of Expert Systems is the static nature of the knowledge base which, once developed, cannot learn from new experiences. The use of a CBRS should improve the reasoning of the KBDSS by continuously learning from new experiences from the domain. This fact will allow the KBDSS to dynamically increase its reasoning capacity.

Case-Based Reasoning Systems can be defined as knowledge-based techniques that permit the use of past experiences to solve new problems that arise in a process. The basic idea behind its functionality is 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 (Kolodner, 1993). For example, although there may be a quite well-established plan of chlorination whenever the facility is facing severe bulking episodes, every time that this solution is applied there are some process particularities that the operator takes into account to adjust the dose, the dosing point, the frequency, etc. This experience would be very useful in the future whenever the process was affected by a similar upset, especially on those occasions where the Expert System cannot make reliable decisions due to the existence of incomplete or missing data.

The core of any Case-Based Reasoning System is the case structure. A **case** can be described as a conceptualised piece of knowledge representing an experience that teaches a fundamental lesson to achieve the goals of the reasoner (Kolodner, 1993). In our domain, a case would correspond to an experience regarding a solids separation problem occurring in the process.

The traditional case-based reasoning cycle (Figure 6.1) consists in a four-step process: retrieve, reuse, revise and retain (Aamodt and Plaza, 1994). A new case is solved by retrieving one or more previously experienced cases (the most similar to the input problem), reusing those experiences to produce a solution, revising that solution through simulation or test execution, and retaining the new experience by incorporating it into the existing case library for future use. This process reflects the way how problems are usually solved manually by the human beings. By incorporating a CBRS, we try to automate the expert steps in a straightforward way.

In our approach, once a new problem is identified, the aim of the CBRS will be to complement the solution provided by the Expert System by suggesting a complementary control strategy, this time based on the experience accrued at the facility in facing similar situations (experiential knowledge). One key advantage of using Case-Based Reasoning is that inexperienced operators can draw on the knowledge of more experienced colleagues, including those outside the organisation, to face problematic situations.

Case-Based Reasoning Systems offer a number of **advantages** that allow improvements in the management and control of complex processes like the occurrence of solids separation problems in activated sludge processes. Their most important features are the following: (1) most of the knowledge contained in the CBRS is totally specific to the domain (i.e., experiential knowledge based on historical information of the WWTP itself) and, therefore, the proposed actions will have a greater chance of success when confronting the operational problems or cases that usually occur at the facility; (2) CBRS are easier to develop (historical cases normally already exist as corporate documentation) and less expensive to maintain than Expert Systems; and finally the most important feature is that (3) CBRS can be considered by themselves, **dynamic systems** with respect to the knowledge contained and enhance their behaviour with time, since they learn from new experiences, both from successful and failed events.

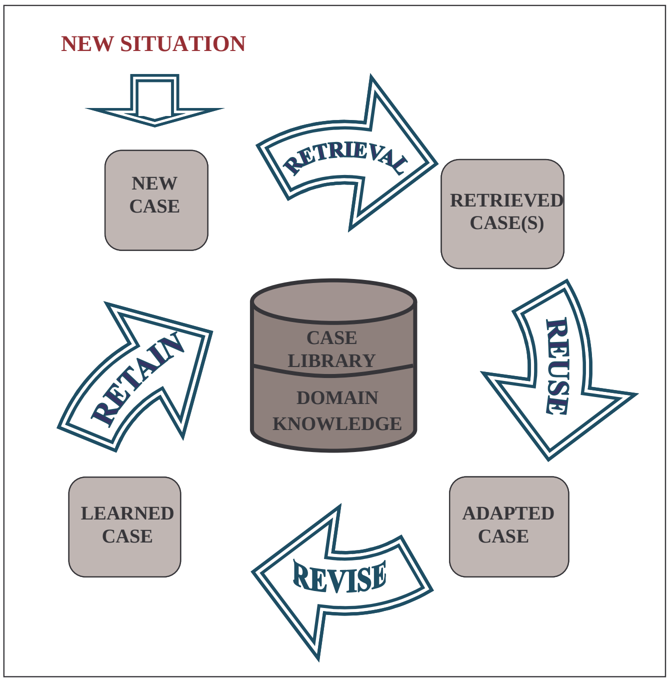


Figure 6.1 The Case-Based Reasoning cycle (adapted from Aamodt and Plaza, 1994)

Case-Based Reasoning Systems can therefore be properly defined as dynamic systems. Actually, according to Schank (1982) the cognitive model used by CBRS is based on Dynamic Memory. This model states that human memory is dynamic because it is continuously changing according to the new experiences one is exposed to. These individual experiences, or cases, in the CBR terminology, encompass lessons learned in a specific context that can be used to face new situations appearing in the domain.

Due to their optimal results and relatively easy development, CBRS have recently been used in several environmental domains such as: information retrieval from historical meteorological databases, planning forest fire fighting, prediction for rangeland pest management advisories, etc. CBRS have also been proposed as support tools in the domain of wastewater treatment (Krovvidy and Wee, 1993; Kraslawski *et al.*, 1995; R-Roda *et al.*, 1999; Sànchez-Marrè *et al.*, 1999 and Wiese *et al.*, 2003). However, most of these proposals were presented as simplified daily case studies, and exposed limitations especially when facing complex problems with slow dynamics where the retrieval of single cases or experiences is not very useful. To guarantee the success of a CBRS it is necessary to consider a flexible approach, especially when dealing with continuous and dynamic environments. The solutions suggested by the CBRS must also be dynamic and continuous.

6.1 Temporal Case-Based Reasoning Approach: Episode-Based Reasoning System

Continuous or dynamic or temporal domains such as WWTPs control and, in particular, solids separation problems, commonly involve a set of characteristics which make them really difficult to work with, such as: (1) a large amount of new valuable experiences are continuously generated, (2) the current state or situation of the domain depends on previous temporal states or situations of the domain, and (3) states have multiple diagnoses. This means that the classical individual case retrieval process would not be very accurate, given that the dynamic domain is structured as a temporally related stream of cases rather than as single cases. Hence, the CBRS solutions should also be dynamic and continuous, and temporal dependencies among cases should be taken into account.

In CBRS, this temporal reasoning in continuous or dynamical domains was not studied until recently. Ma and Knight (2003) proposed a theoretical framework to support historical CBR, based on a relative temporal knowledge model. Similarity evaluation was based on two components: non-temporal similarity, based on elemental cases, and temporal similarity, based on graphical representations of temporal references. Most related publications, such as those of (Jaczynski, 1997) and (Nakhaeizadeh, 1994) use temporal models with absolute references. Jaere *et al*. (2002) use a qualitative model derived from the temporal interval logic from Allen (1984). Martin and Plaza (2004) propose a model which considers unsegmented sequences of observational data stemming from multiple coincidental sources. However, according to Sànchez-Marrè *et al*. (2005), none of these models offers an answer for temporal episodes because they focus more on predicting numerical values, which can be described as time series, than on using the correlation among cases forming an episode. Also, the feature dependency among the individual cases forming an episode is not addressed by the principal known approaches, which rather provide temporal logic reasoning mechanisms, which cannot solve all related problems.

In order to take proper advantage of using CBR in the domain of solids separation problems, a new approach is necessary. In this sense, we propose a temporal approach to CBR: an **Episode-Based Reasoning System (EBRS)**. This new approach is based on the abstraction of temporal sequences of cases, or **episodes**, rather than in the abstraction of single cases. Long problematic experiences should be coded as single episodes, and the EBR algorithm should be able to compare current data with those real experiences by determining the relevant phase of the problem. The goal of this dynamic approach is also to obtain and reuse a fundamental lesson from each new experience, once the results of the control strategy are fully evaluated. The results of the applied control strategy, the daily comments and the final lesson are automatically stored in a historical library to be reused in the future when facing similar problems.

The main bottleneck in classical CBRS is the representation of cases, i.e., deciding what to store in a case, finding an appropriate structure for describing case contents, and deciding how the case memory should be organized and indexed for effective retrieval and reuse (Aamodt and Plaza, 1994). In order to understand the role of our proposed temporal approach, the key components of the EBRS are described next: the episode and case structures, the episode library and the whole dynamic EBR cycle.

6.2 Episode structure

The core of any Case-Based Reasoning System is the case structure. A **case** can be described as a conceptualised piece of knowledge representing an experience that teaches a fundamental lesson to achieve the goals of the reasoner (Kolodner, 1993). In our approach, the main piece of knowledge that the reasoning process addresses is the complete experience of a solids separation problems event in the process, defined as an **episode**. From here on, we will differentiate between cases and episodes mainly by considering that an episode is composed by an ordered set of cases.

6.2.1 Episode definition

An **episode** can be defined as a temporal sequence of cases of length *l*, representing the full sequence of events (cases) associated with a single solids separation problem within the activated sludge process. In order to characterise an episode, a set of attributes have been defined:

(	:episode-identifier	EI
	:initial-time	t
	:episode-length	l
	:episode-description	ED
	:episode-diagnosis	d
	:episode-lesson	EL
	:initial-case	C_t
	:final-case	C_{t+l-1}
)		

- **Episode identifier***:* specific label which identifies each episode. In our approach it is an acronym of the episode diagnosis, i.e. if the diagnosis is filamentous bulking the episode identifier will be FB followed by the number of the episode, e.g. FB-6.
- **Initial time:** indicates the specific instant of time when the episode begins. Due to the low-dynamics of solids separation problems, the initial time does not need to be very precise. In general, the definition of the initial **day** is accurate enough.
- **Episode length:** number of days since the beginning of the episode or passed from the initial case (first day of the episode) to the final case (final day of the episode).
- **Episode description:** the episode is described by the set of cases that compose it.
- **Episode diagnosis:** diagnosis of the specific solids separation problem: filamentous bulking, non-filamentous bulking, foaming, dispersed growth, pin-point floc, rising sludge or normal situation.
- **Episode lesson:** this is the most important piece of knowledge stored in an episode. It summarise the control strategy carried out during the whole episode, including the detailed set of control actions as well as the results obtained from their application (evaluation). It can be defined as the fundamental lesson learnt from the application of a particular control strategy.

- **Initial case:** identifier of the first case of the episode.
- **Final case:** identifier of the last case of the episode. As soon as the KBDSS identifies the occurrence of a new episode, the previous case will be identified as the final case of the previous episode.

Formally, an episode with diagnostic *d*, length *l*, which starts at a given instant in time *t* is:

$$E^{d}_{t,\,l} = \langle EI,\ t,\ l,\ ED,\ d,\ EL,\ C_{t},\ C_{t+l-1}\rangle$$

From a temporal point of view, an episode with diagnostic *d*, length *l*, which starts at initial time *t* can be described as the sequence of *l* temporal consecutive cases:

$$E^{d}_{t,\,l} = (C_{t},\ C_{t+1},\ C_{t+2},\ \ldots,\ C_{t+l-1})$$

Thus, an episode of solids separation problem will be described by the sequence of consecutive cases that belong or have been identified as being part of the same episode. Simplifying the episode definition, an episode comprises those successive situations (defined as cases or isolated cases) which share the same diagnosis. Therefore, an episode is defined or started as soon as a new problem is diagnosed in the process and lasts until the identified situation disappears from the process (usually when the diagnosis changes to normal situation or to another type of problem). Hence, the main characteristic of this new approach is the **undefined length** of episodes, which implies that neither static nor pre-defined lengths are set for the episodes.

An example of episode definition is shown in Figure 6.2, where each of the components which define a real episode stored in the Girona WWTP episode library is presented. From here on, the different components which define an episode or a case will be named **attributes** or **features**.

:episode-identifier	((	FB-8	))
:initial-time	((	2/21/2005	))
:episode-length	((	21	))
:episode description	((	C_1, C_2, C_3, C_4, C_5,…,C_{21}	))
:episode-diagnosis	((	Filamentous Bulking	))
:episode-lesson	((	During this episode, the predominant filament was *S. natans* that has proliferated as a consequence of 2 main conditions. On the one hand because of the oxygen deficiency occurred during several days and on the other hand due to the high concentration of readily biodegradable substrate. Regarding the specific control methods applied to solve the oxygen deficiency, no specific actions could be carried out due to the air system limitations that are going to be repaired in the near future. On the other hand, it seems that the application of the flocculant ZETAG-7689 for 15 days at a dose of 600L/h and a concentration of 0.6% has improved sludge settleability. It is suspected though that since the specific cause has not been solved, future episodes of *S. natans* or Type 021N will appear.	))
:initial-case	((	2/21/2005	))
:final-case	((	3/13/2005	))

Figure 6.2 Definition of a real episode of filamentous bulking in the Girona WWTP

6.2.2 Case definition

A **case** can be defined as the minimal structure capable of describing several features of a temporal domain at a given moment *t*. This time stamp could be measured in any unit of time, depending on the temporal domain at issue. Thus, it could be a month, a day, an hour, a minute, a second or any other unit. In our approach, the minimal structure has been defined to match the frequency of time at which the system is able to define a situation. According to our preliminary studies (see chapter 3), a minimum of 24 hours will be required to obtain most of the necessary variables that the KBDSS needs to make a diagnosis. Thus, 24 hours, 1 day, is the **fixed length** defined for each one of the cases which form an episode. This time, a fixed length is therefore necessary in order to ensure that the minimum information to make a diagnosis is available.

Consequently, each case reflects a properly characterised 24 h-long situation within the activated sludge process. As in the definition of episode, each case is defined by a set of attributes:

(	:episode-identifier	EI
	:case-identifier	CI
	:temporal-identifier	*t*
	:case-situation-description	CD
	:case-problem-diagnosis	CPD
	:case-cause-diagnosis	CCD
	:case-parallel-problems	CPP
	:case-control-strategy	CS
	:case-comments	CC

)

- **Episode identifier**: identifier of the episode that the specific case belongs to. All the cases belonging to the same episode share the same episode identifier.
- **Case identifier**: every case is identified with a specific label. In our approach, this label corresponds to the 24-hours period when the specific situation occurred. This day is identified by the date as month/day/year.
- **Temporal identifier:** identifier of the specific position of the case within the episode (case 1, 2 ,3, etc.).
- **Case situation description**: this is arguably the most important attribute which defines the specific situation characterized in a case. The selection of the set of variables necessary to describe a case is a tough task even for WWTP experts, who find it quite complex to determine the most relevant variables to characterise the activated sludge process. In our approach, out of all the available measures (either quantitative or qualitative values) that are frequently compiled in a WWTP, we kept only those that are relevant to define first of all a situation of solids separation problems and secondly the possible causes of a given situation. The number of variables though is an option that can be customized according to the characteristics of the domain. In our approach, **22 variables** have been chosen as the key to describe a case:

  - SVI, V30 and the qualitative category of the sludge settleability, as indicators of the capacity of sludge to settle
  - Floc characteristics, predominant filament, presence of *Zoogloea* spp. and predominant group, to describe the microscopic characteristics of the floc
  - Presence of foams and effluent TSS to identify any kind of solids separation problem
  - Nine variables have been used to define the potential cause: SRT, F/M, sulphide and pH in the influent, and COD/N, COD/P, DO, temperature, and percentage of soluble COD in the reactor
  - % of COD removal has been used to indicate the general process efficiency
  - Influent flow identifies one of the wastewater characteristics which directly affects the consequences of solids separation problems
  - Waste flow and air flow represent some of the control actions carried out in the process

The information about each situation is codified in a storable and easily retrievable form for future use. As in the Expert System, when more than one measurement is performed for a specific variable, the daily average value will be used. On the other hand, the EBRS has been developed to handle the lack of information so that even if some of the variables lacked a value, the reasoning process would keep on with the reasoning process.

- **Case problem diagnosis**: diagnosis of the specific solids separation problem: filamentous bulking, non-filamentous bulking, foaming, dispersed growth, pin-point floc or rising sludge.
- **Case cause diagnosis**: diagnosis of the potential cause: nitrogen deficiency, phosphorous deficiency, low pH value, high influent sulphide concentration, low dissolved oxygen concentration, low or high food to micro-organism ratio (F/M), high readily biodegradable substrate (measured as soluble COD), low or high SRT value, low or high temperature, high nitrate concentration, toxic shock load, or overaeration.
- **Case parallel problems:** whenever more than one solids separation problem coincide or overlap in the process, these are identified as different but parallel situations (different cases) which evolve separately in time. The function of this specific attribute is to compile this important information.
- **Case control strategy**: full report of the set of actions carried out during that specific unit of time to solve the problem.
- **Case comments:** any comment that the operator registers as additional information to characterise that specific day, once the results of the applied control strategy have been evaluated.

An example of case definition is represented in Figure 6.3. It represents a case of non-filamentous bulking caused by phosphorous deficiency at the Girona WWTP, the second day of an episode of non-filamentous bulking.

:episode-identifier	((	NFB-3		))
:case-identifier	((	10/22/2004		))
:temporal-identifier	((	2		))
:case-situation-description	((	SVI (mL/g)	253	)
	(	V30 (mL)	762	)
	(	effluent-TSS (mg TSS/L)	26	)
	(	influent-sulphide (mg S^{2-}/L)	-	)
	(	COD/total Nitrogen (mg O2 mg total N/L)	32	)
	(	COD/total Phosphorous (mg O2 mg total P/L)	227	)
	(	SRT (days)	3.19	)
	(	F/M (kg DBO kg MLSS,d)	0.23	)
	(	influent-pH	6.57	)
	(	dissolved oxygen (mg O2/L)	4.72	)
	(	soluble COD (m COD/L)	202	)
	(	temperature (ºC)	10.3	)
	(	sludge settleability (qualitative category)	2	)
	(	predominant filament (qualitative category)	*Nostocoida limicola*	)
	(	presence of Zoogloea (YES/NO)	YES	)
	(	Presence of foams (qualitative category)	N	)
	(	floc characteristics (qualitative category)	E	)
	(	predominant group (qualitative category)	3	)
	(	COD removal (%)	81.9	)
	(	waste flow (m^3/d)	1560	)
	(	air flow (m^3/d)	2855	)
	(	influent flow (m^3/d)	26415	))
:case-problem-diagnosis	((	Non-filamentous bulking		))
:case-cause-diagnosis	((	Phosphorous deficiency		))
:case-parallel-problems	((	None		))
:case-control-strategy	((	1. Addition of phosphoric acid during 10 hours at 2.15 ppm		)
	(	2. 2 microscopic analysis a day to monitor the presence of Zoogloea		)
	(	3. Monitoring of sludge settleability in order to detect sudden improvements		))
:case-comments	((	It is too early to identify improvements in sludge settleability. Up to now it seems that *Zoogloea* still predominates in the biomass.		))

Figure 6.3 Definition of an hypothetical case of non-filamentous bulking caused by phosphorous deficiency at the Girona WWTP

6.3 Episode Library

Apart from defining the episode and case structures, the basic configuration of the EBRS also requires a historical database to store the set of relevant episodes covering the wide range of problems that arise in the process. This database is called the episode library, and in our approach, it stores the episodes of solids separation problems occurring in the activated sludge process. The organisation of the episode library is crucial because it has great impact on the system efficiency both in terms of response time and in the successful retrieval of the most suitable episodes to solve the current situation. The structure of a general case library for a typical CBRS can be based on two main approaches: flat memories and hierarchical memories. In **flat memories** every case in the memory is compared with the current case, implying more time consumption. On the other hand, in **hierarchical memories** the matching process and retrieval time are more efficient due to a prior discriminating search in the hierarchical structure, which results in the consideration of only a few cases for similarity assessment purposes.

In our proposal, the following outstanding features have been considered before defining the episode library: (1) the same case can belong to different episodes; (2) the description or state depicted by an episode can correspond to several situations or problems (multiple diagnostics) at the same time, and not only one, as it is assumed by most CBRS; (3) episodes can overlap, and this fact should not imply a redundancy in the episode base representation of the common cases overlapped by the episodes; and (4) episode retrieval should be as efficient as possible.

Taking into account the different possibilities in defining the historical memory of episodes and considering the specific characteristics of our domain, we have decided to develop an episode library which integrates hierarchical formalisms to represent the episodes, and flat representations for the different cases belonging to those episodes. Thus, both episode and case retrieval will be fast enough. This representation model will set an abstraction process that allows splitting the temporal episode concept and the real case of the domain. Discrimination trees for the episodes and a flat structure for cases are proposed. The **discrimination tree** (see Figure 6.4) enables the system to search which episodes should be retrieved according to the feature values of the current episode description, and specifically according to the attribute of episode-diagnosis. On the other hand, episodes contain the necessary information to retrieve all cases belonging to them.

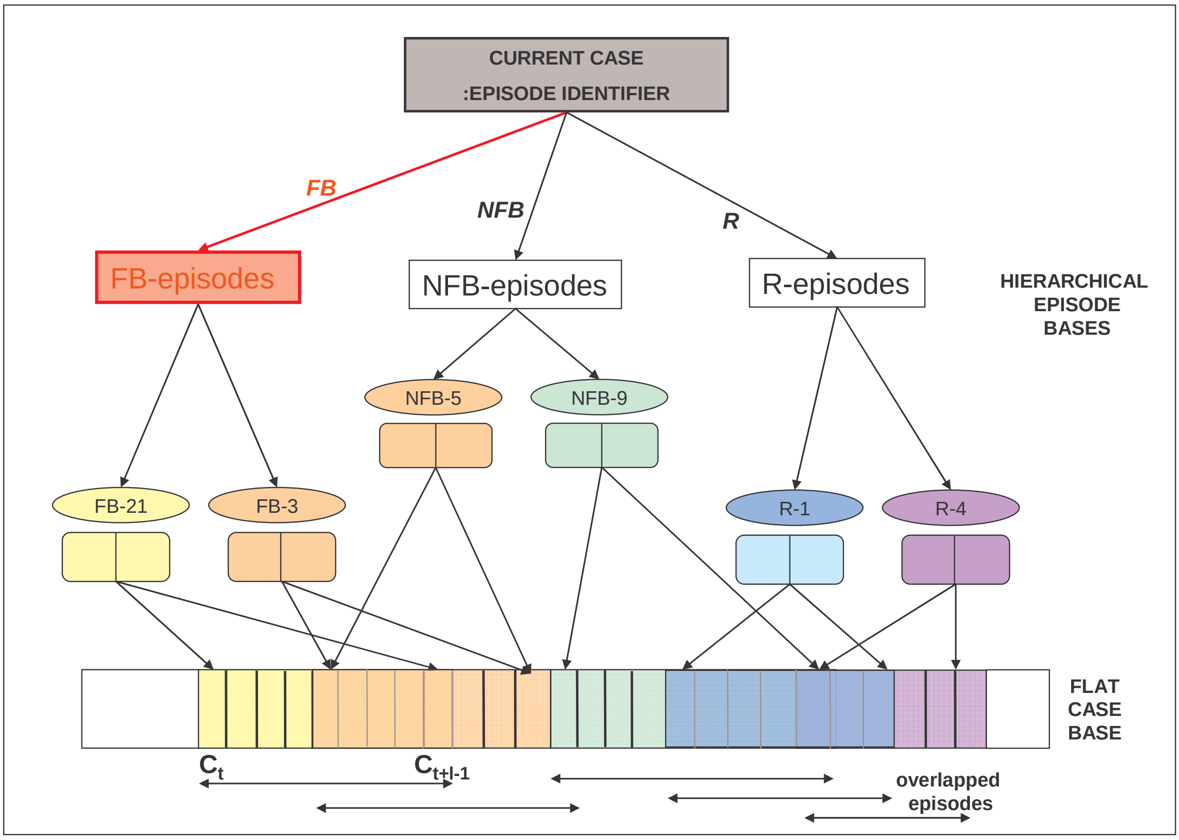


Figure 6.4 Proposal of database organization including both the hierarchical episode base and the flat case base

The episode library is therefore hierarchically organised according to the episode-identifier, which classifies the episodes according to their diagnoses. In this way, considering the current situation of the plant, only those episodes that match the problem diagnosis, e.g. filamentous bulking, will be explored. Once this first discriminating process is performed, all the cases belonging to the discriminated episodes that match the discriminating issue and thus share the same diagnosis will be compared with the current case following the flat memory. This key step optimises the retrieval phase.

In case the current situation could not be classified (i.e. the problem is not properly diagnosed), the first discriminating process will be omitted and the whole case library will be explored.

6.3.1 Initial seed

Once the architecture of the episode library is organized, the initial library can be built. A set of initial episodes is necessary for this purpose; this is called the initial seed. The range of episodes to be covered by the initial seed must be defined according to the typical situations occurring in the domain, including normal operation and, in our approach, all the possible solids separation problems: filamentous bulking, non-filamentous bulking, biological foaming, dispersed growth, pin-point floc and rising sludge. These specific episodes can be obtained from different sources such as the historical database of the facility, existing reports or as a result of personal interviews with plant operators. Since historical data are often not complete enough to generate robust episodes (i.e., rarely all relevant observations are listed), interviews with process experts allows to capitalize on their experience to add other essential information (e.g., specific action proposed in front of a certain situation). After this step, it is recommended to upgrade the library with other common situations (or general episodes) obtained from technical books, troubleshooting guides or provided by general process experts. This group of collected episodes represents the initial seed of the episode-library.

Once the initial seed is defined, the EBRS is ready to propose adequate solutions to detected problems similar to those archived in the initial "seed". Without this seed, the EBRS would have to learn a significant number of actual process episodes before it became useful, lengthening the set up time before the EBRS could be used successfully. Therefore, the long-term success of an episode library demands a continuous updating by adding new episodes, refining the existing ones and forgetting the useless ones. The library keeps evolving with new episodes as knowledge about the process grows; so, the EBRS evolves into a better reasoning system and, as a consequence, system accuracy improves.

6.4 The Dynamic Working Cycle

Once the episode and case structures are defined, and the episode library is built and organised, the dynamic working cycle of the EBRS is designed in order to define the reasoning process followed by the system to reach a conclusion. Following the conceptualisation proposed by the Domino Model (see Figure 6.5), in the EBRS, the first *Situation Beliefs* knowledge base is composed by the set of 22 variables which define the current situation of the process. Continuing within the cyclic structure, now the *Goal* that the EBRS must accomplish is just one: to retrieve the most similar case and its related episode from the episode library to obtain a possible control method to restore the current situation. *Candidate Solutions* are thus all the candidate episodes and their associated cases from the episode library that the discriminating tree has concluded are possible solutions. Once the candidate solutions are set, the *reasoning process* is launched to achieve the main goal of the system. Given the complexity of the type of problems handled by the system, the proposal of a single control action (supposing that just individual cases were used) wouldn't be feasible enough in our domain and thus, as when running the Expert System, the proposal of a complete control strategy is necessary. Following the dynamic cycle suggested by the Domino Model, the retrieval of the most similar case would provide the user with a general *Plan*, and the subsequent retrieval of the complete episode would finally provide the user with he whole and scheduled set of *Actions* to restore the process. The results obtained once the selected control strategy is applied to the process will be gathered in the definition of the next situation (Situation Beliefs) and the cycle will start again.

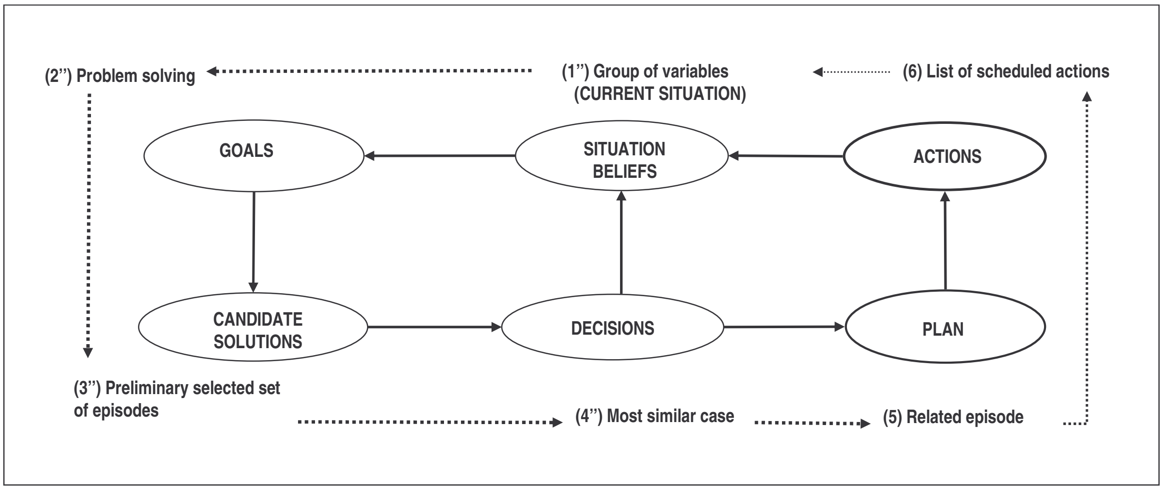


Figure 6.5 Organization of the EBRS architecture according to the Domino Model

The **reasoning process** followed by the EBRS and represented in Figure 6.6 is in turn cyclic and dynamic and involves the completion of 12 main steps which can be grouped within the following reasoning processes: (1) Definition of the current situation -steps 1 and 2-; (2) Searching and retrieval of the most similar case and its related episode -steps 3 and 4-; (3) Episode reuse, adaptation and application -steps 5, 6 and 7- and (4) Evaluation -steps 8, 9, 10 and 11 and (5) Learning -step 12-. Each one of these main processes is depicted next.

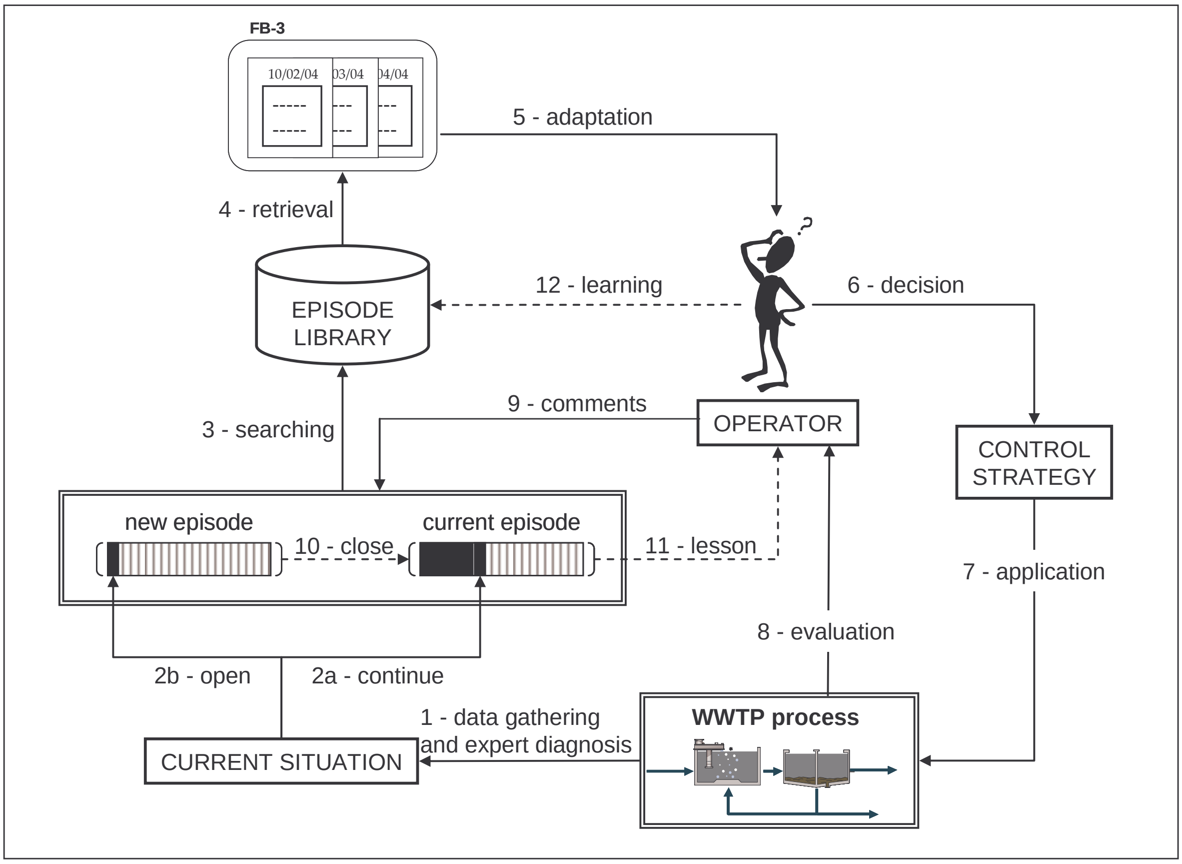


Figure 6.6 Reasoning process followed by the EBRS

6.4.1 Definition of the current situation

The cycle begins as soon as a new situation of solids separation problems arises in the process: the current situation. The first step is getting a diagnosis of the current situation. This can be obtained in different ways. For example, a set of inference rules can be used that are capable of diagnosing the state or the situation of the domain from their relevant features. Another way is by using meta-cases (Sànchez-Marrè *et al.*, 2000). In this method, the similarity between the current case and the meta-cases is evaluated, and the current case is then be labelled with the diagnostic labels of the most similar meta-cases. Meta-cases can be obtained in the same way as rules: from experts or from an inductive clustering process. In our proposal, heuristic rules are used. In order to simplify KBDSS operation, the current situation and its possible cause (in case the situation was a solids separation problem) are identified by means of the data gathering and heuristic rules modules already developed for the Expert System.

Once the current situation has been identified, it is identified either as a new day in the current episode (current episode, 2a), or as the first day of a new episode (a new episode is then opened, 2b). Note that the duration of the episodes is not pre-defined in our EBR approach (in Figure 6.6, the black areas in the episode boxes represent the number of days that the process has been experiencing that situation, one for a new episode, and several for the current episode).

It needs to be mentioned that whenever more than one solids separation problem occurs in the process, parallel cycles are launched and different experiences are obtained for each new situation.

6.4.2 Searching and retrieval

The phases of search and retrieval within the EBR cycle can be considered the key steps for achieving optimal performance of the system. If we manage to retrieve the historical experience that is most similar to the current problematic situation, there will be a greater chance that the control plan stored in the historical experience will be useful enough to be considered before deciding the control strategy to apply for restoring the current situation.

Usually, the retrieval step of case-based reasoning is performed in two steps. In the first step, the case library is scanned in search of the most useful cases to solve the current situation. In the second step, the similarity degree between the current situation and each of the candidate historical cases is evaluated in order to find the most similar case (score computation).

In our approach, for each of the identified current situations, the episode library is examined (searching step in Figure 6.6) in order to obtain (retrieval step in Figure 6.6) the most similar case and its corresponding episode. The retrieval task proceeds with the hierarchical episode base and the subsequent flat case base depicted in section 6.3, where a set of discriminating rules allow the system to pre-select a set of candidate episodes. After that, the whole group of cases belonging to these candidate episodes are examined and compared with the current situation. The comparison task to find the most similar case consists in matching the attributes of the current situation with those of each historical case, and then computing a composite similarity metric for each case. Thus, the comparison task is performed at a same scale, i.e. **case to case**, matching the daily attributes of the current situation with the daily attributes of each previously selected historical case.

When working with flat memory, the system retrieves the most similar cases while in hierarchical memory the number of cases retrieved depends on which branches are explored. Since this primary step is executed according to the specific diagnostic label, in situations where more than one diagnosis is identified (e.g., when a current situation of both filamentous bulking and non-filamentous bulking is identified), parallel reasoning processes will be launched for each one of the identified situations and different solutions, and the corresponding solutions obtained.

6.4.2.1 Similarity calculation

Once the candidate episodes and cases are set, the similarity calculation among cases is performed. Most CBR applications use a case representation based on attribute-value vectors. Similarity measures are intrinsically related to this case representation formalism. Within the attribute vector approach, the similarity measures are normally computed as an aggregation of attribute differences between two cases. If similarity measures fail to capture the actual differences between cases, the retrieval step, and the whole CBR performance, will be inefficient (Núñez *et al.*, 2004). Thus, the selection of an appropriate similarity measure is a key point in CBR systems, and also in our EBR approach.

Most Case-based reasoners use a generalised weighted dissimilarity measure such as,

$$diss(C_i, C_j) = \frac{\sum_{k=1}^{n} w_k * atr\_diss(C_{ik}, C_{jk})}{\sum_{k=1}^{n} w_k}$$

where,

C_i and C_j are two compared cases; w_k is the weight or importance assigned to attribute k; and $atr\_diss(C_{ik}, C_{jk})$ is the dissimilarity degree between the value of attribute k in cases i and j.

The similarity between two cases, and as a consequence the overall efficiency of the case retrieval and the subsequent retrieval of the whole episode, depends on the following factors:

1) The similarity measure used during case retrieval
2) The type of attributes and their management
3) The weight (w_k) assigned to each attribute (feature weighting)

1. The similarity measure

There exist several similarity algorithms used in CBR systems. Table 6.1 summarizes some of the most used ones. See reference Núñez *et al.*, (2004) for more details about the possible similarity measures.

Table 6.1 Summary of some of the most used similarity measures in case-based reasoning systems

Type of similarity measure	**Similarity measures**	
Measures derived from **_Minkowski 's metric:_** $d(C_i,C_j)=\left(\sum_{k=1}^{n}\left|C_{ik}-C_{jk}\right|^{r}\right)^{1/r}$ $r\geq 1$	***Manhattan*** **or** ***City-Block distance*** $d(C_i,C_j)=\frac{\sum_{k=1}^{n} w_k^r * d(C_{ik},C_{jk})}{\sum_{k=1}^{n} w_k^r}$	***Euclidean distance*** $d(C_i,C_j)=\left(\frac{\sum_{k=1}^{n} w_k^2 * \left|d(C_{ik},C_{jk})\right|^2}{\sum_{k=1}^{n} w_k^2}\right)^{1/2}$
Unweighted similarity measures	***Clark distance*** $d(C_i,C_j)=\sum_{k=1}^{n}\frac{\left|C_{ik}-C_{jk}\right|^2}{\left|C_{ik}+C_{jk}\right|^2}$	***Canberra distance*** $d(C_i,C_j)=\sum_{k=1}^{n}\frac{\left|C_{ik}-C_{jk}\right|}{\left|C_{ik}+C_{jk}\right|}$
Heterogeneous similarity measures	***Heterogeneous value difference metric*** **(HVDM)** $HVDM(C_i,C_j)=\sqrt{\sum_{k=1}^{n} d_k^{\,2}(C_{ik},C_{jk})}$	***Interpolated value difference metric*** **(IVDM)** $IVDM(C_i,C_j)=\sum_{k=1}^{n} ivdm_k\left(C_{ik},C_{jk}\right)^2$

There exist too some comparative studies among similarity measures (Wilson and Martínez (1997); Liao and Zhang (1998) or Núñez *et al.* (2004)). Based on the results of these comparative studies and on the experience gained from previous applications of CBRS in WWTP domains, such as Comas (2000), an adaptation of the *Eixample* distance (Sànchez-Marrè *et al.*, 1996) has been chosen as the distance function for the retrieval process. The Eixample distance is a normalised exponential weight-sensitive distance function derived from a general weighted-distance function (typical Hamming distance), and has been selected mainly for two reasons. Firstly, it is an exponential weight-sensitive algorithm that gives additional value to the weight of attributes. This exponential transformation allows amplifying the differences among attributes when the number of attributes, *n*, is very high, as it happens in our domain. Secondly, it considers both quantitative and qualitative variables, the type of variables used in our case definition.

The Eixample distance is defined as:

$$d(C_i,C_j)=\frac{\sum_{k=1}^{n} e^{w_k}\times d(C_{ki},C_{kj})}{\sum_{k=1}^{n} e^{w_k}}$$

where,

$$d(A_{ki},A_{kj})=\begin{cases}\dfrac{\left|\,quantval\,(A_{ki})-quantval\,(A_{kj})\,\right|}{upperval\,(A_k)-lowerval\,(A_k)} & \text{if } A_k \text{ is a quantitative attribute}\\[2ex] 1-\delta_{qualval(A_{ki}),qualval(A_{kj})} & \text{if } A_k \text{ is a qualitative non - ordered attribute}\\[2ex] \dfrac{1-\delta_{qualval(A_{ki}),qualval(A_{kj})}}{\#\,\mathrm{mod}(A_k)-1} & \text{if } A_k \text{ is a qualitative ordered attribute}\end{cases}$$

and,

C_i is case i; C_j is case j; W_k is the weight of variable k; A_{ki} is the value of the variable k in case i; A_{kj} is the value of the variable k in case j; *quantval* (A_{ki}) is the quantitative value of A_{ki}; *quantval* (A_{kj}) is the quantitative value of A_{kj}; A_k is variable k; *upperval* (A_k) is the upper quantitative value of A_k; *lowerval* (A_k) is the lower quantitative value of A_k; *qualval* (A_{ki}) is the qualitative value of A_{ki}; *qualval* (A_{kj}) is the qualitative value of A_{kj}; #mod(A_k) is the number of modalities (categories) of A_k; δ *qualval* (A_{ki}), *qualval* (A_{kj}) is the Kronecker delta.

Whenever two qualitative values are compared, such as the attributes related to macroscopic and microscopic variables, the distance is calculated by using the Kronecker δ. In such cases, when the value for a qualitative attribute in case i, *qualval* (A_{ki}), is different from the value of the same qualitative attribute in case j, *qualval* (A_{kj}), the Kronecker $\delta = 0$ and the distance between the attributes will be the maximum (1). When both cases have the same value for the same qualitative attribute, the Kronecker $\delta = 1$ and the partial distance will be the minimum (0). An exception is the attribute for floc characteristics, where the categories A, B, E, and F are ordered. In this case, the distance is calculated by dividing the partial distance between the qualitative values by the number of modalities (#mod(A_k)) – 1.

Once the similarity algorithm was set, the assignment of weights (w_k) was performed. This process is called **attribute** or **feature weighting**. Since the selected similarity algorithm is exponentially weight-sensitive (i.e., the values assigned to each weight have a high influence on the results of the retrieval step), the process of feature weighting has been performed very accurately in order to maximize the efficiency of case and episode retrieval.

2. Case attributes and their management

As depicted in the case definition, 22 different variables (including quantitative and qualitative variables) have been selected to describe a case situation (:case-situation-description), that is to describe the solids separation problem and its possible cause. From here on, this set of variables will be defined as the set of attributes (or features) used during the similarity calculation. During similarity calculation, the set of attributes and their corresponding values will be represented as attribute-value vectors.

On the other hand, the efficiency of episode retrieval also depends on the **quality of data**. If the episode library contains a high percentage of unknown values (missing), the retrieval stage will definitely be less efficient than when using episode libraries containing values for all their attributes. Due to the constraints already mentioned regarding the compilation of real data from the activated sludge process, it is often the case that some of the variables required for defining a case are unknown. Some of the strategies that can be followed to deal with these situations include: (a) deleting those cases with unknown values; (b) not considering those attributes with missing values during the similarity calculation; (c) replacing the missing values with calculated values by means of heuristics, and (d) assigning the maximum value (1) to the partial distance when one of the compared values is unknown, and the minimum value (0), when both compared values are unknown.

In our approach, strategies (a) and (d) are used. Thus, on the one hand, cases with more than 10 unknown values do not get stored in the case library. On the other hand, the partial distance of a pair of attributes containing an unknown value is calculated by assigning a partial distance = 1 when one of the values is unknown and 0 when both values are unknown.

3. Feature weighting

When making a CBR search query, the weight of certain variables or attributes can be increased in order to give them a heightened importance during the similarity calculation. In fact, this is something that people do instinctually when they search their minds for applicable past experiences: they focus on the most important aspects of the situation. For instance, in a situation where an operator is facing a filamentous bulking or foaming episode, s/he might give greater importance to the predominant filamentous microorganism data than if the problem is dispersed growth or pin-point floc. Likewise, when the cause identified is nitrogen deficiency in the influent, the attribute (COD/total N) that defines this cause would be more relevant than for example the pH in the reactor. In CBR, the equivalent of this mental weighting is the process of weighting attributes, i.e. feature weighting.

Feature weighting is a very important but complex task that allows us to improve the feasibility of our retrieval algorithm and thus of our EBRS. The complexity of the process is basically due to the fact that deciding on the importance that an attribute has compared to the others requires in-depth knowledge of the domain (knowledge-based approach) or a good machine learning mechanism to determine the attribute's importance automatically from the stored cases (observation-based approach). In observation-based modelling probability plays a central role. Due to probability estimations, these approaches determine whether a relation exists between the elements of the model and, if sufficient historical or behavioural data are available, provide a quantitative measure of this likelihood (Mérida-Campos and Rollón, 2003). Observation-based approaches, however, need a huge number of experiences to effectively extract the relationships. Therefore, they cannot be successfully applied when dealing with domains characterised by a huge number of attributes and few available cases, as is the case in our domain. The alternative under these circumstances is to use knowledge-based approaches in which the knowledge to establish the relevance relationships comes from the experts in the domain.

The method selected to find and represent all the relevance relationships existing among the different attributes that define our domain was a **Relevance Matrix** (see Figure 6.7). When developing the Relevance Matrix, the first step was to define and quantify the Relevance of each attribute. **Relevance** is not a well understood concept in spite of the huge amount of research into this topic, mainly due to inconsistently used terminology. According to Mizzaro (1998), there are many kinds of valid Relevance definitions, not just one, which can be defined in relation to four different dimensions (information resources, representation of the user's problem, time, and component). Hence, Relevance can be defined in different ways and each definition refers to the performance of different tasks (e.g. Relevance of A compared to B for performing task X, Relevance of A compared to B for performing task Y). In this study, we have defined **Relevance *R*** of an attribute *f* in a certain instant of time, related to the problem characterization task and given that the other attributes are assigned certain values (*Context C*), as the importance that attribute *f* has compared to the other attributes in the domain in the time it takes to perform the particular task.

Relevance quantification is also a very prolific topic in the literature. Although the existence of degrees or regions of relevance is accepted, there is no agreement on what is the best method for defining these degrees (Schamber, 1994). However, it is commonly accepted that the relevance values should be expressive enough to capture the particularities of the domain knowledge. The range selected for determining the relevance in our framework uses the results of the study in Tang *et al*. (1999), who empirically observed that a scale of around seven values is the optimum choice for expressing relevance.

In order to simplify the representation, we propose a scale of six categories where the relevance values considered are: 0=No Relevance, 1=Neutral Relevance, 2=Slight Relevance, 3=Medium Relevance, 4=High Relevance and 5= Extreme Relevance. This ranking classifies the relevance of attributes from 0 (the attribute does not help to define the current situation) up to 5 (the attribute is extremely important to define the current situation). Once the term Relevance and the range of relevancies were defined, the whole Relevance Matrix was developed according to our own domain expertise about the types of relevancies that an attribute can be affected by (Martínez *et al.*, 2005c).

The last step of the feature weighting process is the final translation or mapping of the relevance values assigned by the relevance network into the set of feature weights used during case retrieval. In order to simplify this procedure, the mapping method assumes that there is a linear relationship between degrees of relevance and feature weights so that the correlation between relevance and weights is direct. Figure 6.7 represents the Relevance Matrix developed to compile the different weights (relevant weights) assigned to the selected 22 attributes according to the specific solids separation problem and their corresponding cause.

The different weights represented in the Relevance Matrix can be directly introduced in the searching stage by means of attribute-weights vectors. Depending on the problem and cause diagnosis, the assignment of the specific weights is performed by means of simple deductive rules, e.g. "IF the current diagnosis= filamentous bulking and the current cause = P deficiency conclude that the SVI-weight=5 and conclude that the V30-weight=4".

In order to test the feasibility of using relevant weights to improve the accuracy and usefulness of case retrieval, the suggested searching step has been performed using the two possible retrieval procedures: (1) with relevant weights, varying according to the context (type of solids separation problem and cause), and (2) with constant weights for the different attributes, which are kept constant whatever the context is. The set of weights for the second testing procedure was determined according to the typical weights assigned in previous applications of CBR in the WWTP domain (Rodríguez-Roda *et al.*, 2002).

The testing procedure was performed by using the leave-one-out cross-validation (LOOCV) technique that consists in randomly selecting a case from the episode library and calculating its distance to the rest of the cases from the episode library. Once the first test is done, the previously selected case becomes part of the episode library again and a new case is randomly selected. We have considered 10% of the total number of cases stored in the case-library as the number of testing procedures to analyse case retrieval efficiency. The episode library used for the preliminary study consists of 221 real cases that occurred in the Girona WWTP from January 2004 to February 2005. During this 14-month period 15 different episodes of solids separation problems occurred in the process including: 7 episodes of filamentous bulking, 3 episodes of non-filamentous bulking, 2 episodes of biological foaming and 3 episodes of rising sludge. Dispersed growth and pin-point floc, the two other possible problems, could not be considered in this test because they did not occur during this period of study.

PROBLEMS	SVI	V$_{30}$	eTSS	%DQO	i-S^2	COD/N	COD/P	SRT	F/M	i-pH	DO	COD$_S$	T	s.set.	p.fil.	f.pres.	f.char	z.pres	p.group	w-flow	air-flow	in-flow
FILAMENTOUS BULKING (*FB*)	5	4	3	2	1	1	1	1	1	1	1	1	2	4	4	2	4	1	2	2	2	3
-Nitrogen deficiency (-*ND*)	5	4	3	2	1	5	1	1	2	1	1	1	2	4	4	2	4	1	2	2	2	3
-Phosphorous deficiency (-*PD*)	5	4	3	2	1	1	5	1	2	1	1	1	2	4	4	2	4	1	2	2	2	3
-Low pH (-*LPH*)	5	4	3	2	1	1	1	1	1	5	1	1	2	4	4	2	4	1	2	2	2	3
-High sulphide (-*HS*)	5	4	3	2	5	1	1	1	1	1	1	1	2	4	4	2	4	1	2	2	2	3
-Low DO (-*LDO*)	5	4	3	2	1	1	1	1	3	1	5	1	2	4	4	2	4	1	2	2	5	3
-Low F/M (-*LFM*)	5	4	3	2	1	1	1	4	5	1	2	1	2	4	4	2	4	1	2	3	3	3
-High SRT (-*HSRT*)	5	4	3	2	1	1	1	5	4	1	1	1	2	4	4	2	4	1	2	3	2	3
-High CODs (-*HSCOD*)	5	4	3	2	1	1	1	1	3	1	1	5	2	4	4	2	4	1	2	2	2	3
NON FIL. BULKING (*NFB*)	5	4	3	2	0	1	1	1	1	0	1	1	1	4	1	3	4	4	2	2	2	3
-Nitrogen deficiency (-*ND*)	5	4	3	2	0	5	1	1	2	0	1	1	1	4	1	3	4	4	2	2	2	3
-Phosphorous deficiency (-*PD*)	5	4	3	2	0	1	5	1	2	0	1	1	1	4	1	3	4	4	2	2	2	3
-Low SRT (-*LSRT*)	5	4	3	2	0	1	1	5	4	0	1	1	1	4	1	3	4	4	2	3	2	3
-High SRT (-*HSRT*)	5	4	3	2	0	1	1	5	4	0	1	1	1	4	1	3	4	4	2	3	2	3
-Low F/M (-*HFM*)	5	4	3	2	0	1	1	4	5	0	2	1	1	4	1	3	4	4	2	3	3	3
-High F/M (-*HFM*)	5	4	3	2	0	1	1	4	5	0	2	1	1	4	1	3	4	4	2	3	3	3
-High CODs (-*HSCOD*)	5	4	3	2	0	1	1	1	3	0	1	5	1	4	1	3	4	4	2	2	2	3
FOAMING (*F*)	5	4	3	2	0	0	0	1	1	0	1	1	2	4	4	4	4	1	2	2	2	3
-Low SRT (-*LSRT*)	5	4	3	2	0	0	0	5	4	0	1	1	2	4	4	4	4	1	2	3	2	3
-High F/M (-*HFM*)	5	4	3	2	0	0	0	4	5	0	2	1	2	4	4	4	4	1	2	3	2	3
-Low DO (-*LDO*)	5	4	3	2	0	0	0	1	3	0	5	1	2	4	4	4	4	1	2	2	5	3
-High SRT (-*HSRT*)	5	4	3	2	0	0	0	5	4	0	1	1	2	4	4	4	4	1	2	3	2	3
-High CODs (-*HSCOD*)	5	4	3	2	0	0	0	1	3	0	1	5	2	4	4	4	4	1	2	2	2	3
-Low F/M (-*LFM*)	5	4	3	2	0	0	0	4	5	0	1	1	2	4	4	4	4	1	2	3	2	3
-Low T (-*LT*)	5	4	3	2	0	0	0	1	1	0	1	1	4	4	4	4	4	1	2	2	2	3
-High T (-*HT*)	5	4	3	2	0	0	0	1	1	0	1	1	4	4	4	4	4	1	2	2	2	3
DISPERSED GROWTH (*DG*)	4	4	5	4	0	1	1	1	1	0	1	0	2	4	1	1	4	0	2	2	2	3
-Toxic shock (-*TS*)	4	4	5	4	0	1	1	1	1	0	3	0	2	4	1	1	4	0	5	3	3	3
-Surfactants presence (-*SP*)	4	4	5	4	0	1	1	1	1	0	3	0	2	4	1	4	4	0	4	2	2	3
-Overaeration (-*OA*)	4	4	5	4	0	1	1	1	3	0	4	0	2	4	1	1	4	0	2	2	5	3
-Low SRT (-*LSRT*)	4	4	5	4	0	1	1	5	4	0	1	0	2	4	1	1	4	0	2	3	2	3
-Low F/M (-*LFM*)	4	4	5	4	0	1	1	4	5	0	2	0	2	4	1	1	4	0	2	3	3	3
PIN-POINT FLOC (*PP*)	4	4	5	4	0	0	0	1	1	1	1	0	2	4	1	1	4	0	2	2	2	3
-Overaeration (-*OA*)	4	4	5	4	0	0	0	1	3	1	4	0	2	4	1	1	4	0	2	2	5	3
-Low DO (-*LDO*)	4	4	5	4	0	0	0	1	3	1	5	0	2	4	1	1	4	0	2	2	5	3
-High SRT (-*HSRT*)	4	4	5	4	0	0	0	5	4	1	1	0	2	4	1	1	4	0	2	3	2	3
-Toxic shock (-*TS*)	4	4	5	4	0	0	0	1	1	1	3	0	2	4	1	1	4	0	5	3	3	3
-Surfactants presence (-*SP*)	4	4	5	4	0	0	0	1	1	1	3	0	2	4	1	4	4	0	4	2	2	3
-Low pH (-*LPH*)	4	4	5	4	0	0	0	1	1	5	1	0	2	4	1	1	4	0	2	2	2	3
RISING SLUDGE	1	1	1	2	1	1	0	5	4	1	1	0	1	4	0	4	0	0	2	3	3	2
NORMAL SITUATION	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Figure 6.7 Relevance matrix compiling the specific weights (relevant weights) of the 22 attributes to be used within the distance algorithm during case retrieval

Taking into account the size of the episode library, 22 case studies (10% of all case studies) were selected for the test, i.e. they were compared to the rest of the cases using the LOOCV procedure. These cases were randomly selected in proportion to the total percentage of each of the five different situations within the case-library. Therefore, given that almost 50% of the cases corresponded to filamentous bulking, from the targeted 22 case-studies, 10 cases corresponded to filamentous bulking episodes. Likewise, 4 cases of non-filamentous bulking, 4 cases of biological foaming and 4 cases of rising sludge were also randomly selected. Once the testing procedure was completed, the efficiency of the case retrieval process was measured by determining its precision. We computed the precision of case retrieval in both procedures as the percentage of the retrieved cases that were relevant to experts based on the knowledge they could provide when retrieved. Therefore, for each retrieved case a precision value of 0 (a useless case), 0.5 (a case quite useful to be consulted) or 1 (a case highly useful so that the same control strategy can be applied in the current situation) was determined. To determine this precision value, the Girona WWTP plant manager was consulted. As an example, Table 6.2 represents the results obtained from the first test using case 12/01/2004, a case of filamentous bulking caused by low dissolved oxygen.

Table 5.5 Retrieved cases obtained from the first test. The second column shows the case retrieved using the relevant weights and the third column shows the case retrieved using constant weights. The weights used in both contexts are shown.

	CASE-STUDY	RETRIEVED CASE (relevant weights)		RETRIEVED CASE (ct. weights)	
episode-id.	FB-6	FB-4	WEIGHTS	FB-2	WEIGHTS
case-id.	12/01/2004	09/02/2004		02/05/2004	
temporal-id.	1	2		5	
SVI	196	209	5	213	3
V30	396	288	4	275	1
e-TSS	15	25	3	91	1
i-sulphide	-	-	0	-	1
COD/N	6.3	5.6	1	12.1	1
COD/P	65.2	59.0	1	56.11	1
SRT	4.5	5.6	1	15.2	1
F/M	1.3	1.8	3	0.51	1
i-pH	7.5	7.5	1	7.6	2
DO	1.0	1.0	5	2.0	1
COD soluble	-	-	0	-	1
T	9.7	8.9	2	9.5	1
Sludge set.	1	1	4	1	1
Pred.fil.	*S.Natans*	*S.Natans*	4	Type 021N	1
Foam.pres.	N	N	4	N	1
Floc char.	E	E	2	E	1
Pred. Group	1	1	2	2	1
Zoog.pres.	NO	NO	1	NO	1
%COD rem.	89.4	81.3	2	72.9	2
W-Flow	1440	960	3	1120	1
Air-Flow	131490	130540	2	146950	1
In-flow	46279	45150	5	52160	3
c-diagnosis	Filamentous bulking caused by low DO	Filamentous bulking caused by low DO		Filamentous bulking caused by high SRT	
c-control-strategy	-	DO set-point increasing up to 2.0 ppm during 7 days		Increase purge flow up to 1120 m^3/d in order to reduce SRT	
c-comments	-	No improvements observed		The number of filaments is decreasing	
%similarity	-	96.1%		76.4%	
precision	-	1		0	

The results obtained from the first test using case 12/01/2004 show that, by using the relevance network, case 09/02/2004 is retrieved. This is a case of filamentous bulking caused by low DO, which is the same problem and cause as the case study. Actually, if we compare the most relevant variables (those with weight 5 or 4) we can see that they are quite similar so it is not surprising that the degree of similarity (calculated by using the *Eixample* distance) was 96%. Due to its high similarity to the case study, it is reasonable that the control method applied on the previous occasion was worth consulting; hence it obtained a precision value = 1. On the other hand, the retrieved case in procedure 2 was also a case of filamentous bulking, but on this occasion the cause of the problem was different from the case study and consequently the retrieved reported control method could not be considered a possible solution for the current situation; hence it obtained a precision = 0.

Once the LOOCV testing was performed, the total precision of each procedure was computed obtaining the following results:

- ⇒ 90.9 % of the cases retrieved by using the relevant weights were considered highly relevant by experts according to the knowledge these cases provided. This represents a mean precision of 0.91 per retrieved case.

- ⇒ 63.3 % of the cases retrieved using constant weights independent of the context were considered highly relevant by experts according to the knowledge these cases provided. This represents a mean precision of 0.63 per retrieved case.

Therefore, the use of relevant weights before the similarity calculation results in a remarkable improvement in the efficiency of case retrieval and thus in the overall effectiveness of case-based reasoning as a support tool for plant operators when dealing with complex problems in the process such as the common solids separation problems.

6.4.2.2 Score computation

Once the deductive rules are inferred to assign the corresponding weights, the similarity calculation is the final step in the searching phase. Using the *Eixample* algorithm the differences between each of the attributes (more or less relevant according to their weights) in the current case and each historical case are found. As soon as the searching step is completed, the Retrieval stage is inferred and **score computation** is performed in order to quantify the distances between each of the selected cases and the current situation. Each retrieved case is added to a sorted list of cases by decreasing degree of similarity. The first case of the list, the historical case with the smallest distance to the current case, is the one considered to be the most similar to the current situation.

6.4.2.3 Episode retrieval

The retrieval of episodes is the final step in the stage of searching and retrieval. Once the most similar case is retrieved, the retrieval of the complete episode to which the most similar case belongs is direct. Using the episode membership information contained in the identifier within the retrieved case (:episode-identifier), all historical cases with the same episode-identifier (meaning that they are temporal members of the same episode) will also be retrieved in order to provide the user with the maximum information of the complete episode, especially information regarding the control strategy carried out and the final lesson obtained from its application.

6.4.3 Episode reuse, adaptation and application

Once the plant operator receives all the information provided by the retrieved episode, s/he has the option to make the final decision about which control strategy to apply to restore the process. We must keep in mind that despite being an “intelligent” tool, the EBRS (like the ES) is a Decision Support tool; hence the final decision is considered as being task of the plant operator.

In general, if the most similar case and its corresponding episode retrieved by the system are significantly different from the current problem, an adaptation of the solution of the retrieved case may be required in order to reuse it. We must consider, though, that in reality the control parameters that are usually modified in a facility are only a few, mainly recycle and wasting sludge flow rates, dissolved oxygen set-points, and chemical additions in case the treatment involves chemical treatment. The actions to be taken involve "maintaining", "rejecting", "increasing" or "decreasing", and the adjustment involves a slight modification of the parameters. If the similarity distance is small enough (meaning a high similarity between the cases), no adaptation of the solution may be required. For example, for two similar situations of “low oxygen level in the bioreactor”, the proposed action plan is identical: “raise the aeration of the bioreactor” and its application modifies the value of a parameter that, in this case, is the dissolved oxygen set-point or the airflow rate.

Once the proper modifications of the suggested control strategies are performed, a final decision about the control method to be applied is made by the plant operator (step 6 in Figure 6.6). Imediately after, a set of actions (step 7) will be applied on the activated sludge process in order to restore the problematic situation in the most efficient way.

6.4.4 Evaluation

Once the control actions are applied, a final evaluation or revision of the results by a human expert is usually required. There are different ways to evaluate these results: by direct expert or operator assessment, by simulating the proposed actions and evaluating its consequences, by directly getting feedback from the results of the proposed solution, or by a combination of some of the previous. In general, an evaluation of the partially applied actions is usually completed by the operator (step 8 in Figure 6.6). On the other hand, the main variables compiled in the case also help to characterize both the applied actions and the evolution or response of the activated sludge process.

In the current approach, the evaluation process will be performed by using two different procedures:

- ⇒ The daily control actions (case-control-strategy) and the corresponding evaluation of its effect on the process (case-comments) automatically are registered daily (step 9 in Figure 6.6) together with the rest of required attributes which compose a case: episode-identifier, case-identifier, temporal-identifier, case-situation-description, case-diagnosis and case-control-strategy.

⇒ Once a new episode is identified, as a consequence of a change in case-diagnosis, the episode that has been opened on previous days must be closed (step 10 in Figure 6.6). The operator has to summarise and conclude the fundamental piece of knowledge (episode-lesson) that has been learnt (step 11). Once the episode-lesson is noted down, the rest of attributes which define an episode (episode-identifier, initial-time, episode-length- episode-description, episode-diagnosis, initial-case and final-case) will be automatically registered.

6.4.5 Learning

This final phase concerns the dynamic process of knowledge learning and retaining (step 12). In our approach, whenever an episode is finished and all the required information is being compiled, it will be automatically retained in the episode library. The knowledge contained in the episode library is continuously updated with both successful and failed new experiences that can be useful for future problem solving. An episode that is solved successfully should be incorporated to the case library so that the EBRS can benefit from this information in future situations. This is known as learning from success. Likewise, if the proposed solution fails, it should also be recorded to prevent similar mistakes in the future. This is known as learning from failure.

The learning capability is one of the outstanding features of an EBRS. Thanks to the learning step, the episode library is continuously updated with new episodes as soon as new experiences of solids separation problems emerge. Thus, the EBRS evolves over time, improving its reasoning capability and system accuracy from new experiences. In general, the more experiences the episode library includes, the wider the range of possible scenarios that could be tackled. Nevertheless, in order to optimize episode retrieval, if the number of stored cases and episodes becomes too large to hinder the retrieval process, the number of stored episodes can be reduced. If necessary, it is recommended to delete those episodes which contain a high percentage of missing data (incomplete knowledge) (see section 6.4.2.1).

6.5 Discussion

The new approach illustrated at the present chapter, has allowed the classical Case-Based Reasoning System to go one step further and to consider the basic piece of knowledge, the cases, at a larger and undefined length. Thus, instead of dealing with single daily cases, the new temporal approach, named Episode-Based Reasoning System deals with episodes, defined as a temporal sequence of cases. This temporal approach has allowed the system to handle efficiently solids separation problems, characterized by their variable length.

An exponential-weight sensitive algorithm (*Eixample* distance) has been selected as the similarity algorithm due to the importance that the weight assigned to each attribute has during similarity calculation. Consequently, a relevance matrix has been built to assign different weights to each one of the 22 attributes, according to the context (the type of solids separation problem and its cause). In this way, the retrieval of similar cases (sharing the same context) is favoured. Moreover, the use of relevant weights before the similarity calculation results in a remarkable improvement in the accuracy of case retrieval and thus in the overall efficiency and usefulness of the EBRS.

The set of 22 attributes used during the similarity calculation, as well as the different relevant weights have been chosen according to its relevance and its usual availability in most plants. Regarding the transferability of the system to different facilities, the permanent lack of an attribute does not affect the performance of the system. Likewise, the set of attributes and weights could be easily adapted to each facility, if necessary, by modifying the relevance matrix.

In order to streamline the searching process, a discriminating tree was developed. The resulting heuristic rules use the diagnosis concluded by the Expert System to discriminate and finally to select the set of episodes that the system must explore during the searching step.

In order to simplify the retrieval process, the similarity calculation between the current situation and each one of the historical experiences has been performed at a same scale approach (case to case). Once a set of the possible similar episodes has been selected, each one of the historical cases belonging to those episodes, defined as attribute-value vectors, are compared with the current situation, also represented as an attribute-value vector defining a 24-hours situation. Further work on this topic will be based on the consideration of the temporal feature within the similarity calculation. The temporal identifier, the total length of the episode and the position of the case within the episode will be considered during the searching process. A new approach is currently being studied (Sànchez-Marrè *et al.*, 2005), where the total length of the episode is considered during the comparison between episodes. However, the preliminary results are not presented in this thesis.

The historical library of experiences was hierarchically organized at two different levels: the first (and superior) level includes all the stored episodes related to solids separation problems. The second (and inferior) level incorporates all the sequence of cases belonging to each episode. The innovative structure of the historical library, named as **episode library**, presents the following chacteristics: (1) the same case can belong to different episodes; (2) the description or state depicted by an episode can correspond to several situations or problems (multiple diagnostics) at the same time, and not only one, as it is assumed by most CBRS; (3) episodes can overlap, and this fact does not imply a redundancy in the episode base representation of the common cases overlapped by the episodes; and (4) the length of the episodes is variable.

In order to increase the usefulness of the retrieved information contained in the historical experiences with regards to the applied control strategy, the retrieval of complete episodes (instead of single cases) has been provided. The retrieval of the complete episode allows the user not only to consult the applied control strategy, but also to study the whole effect of its application on the process, as well as the lesson learned by the expert in applying such control strategy.

The direct sharing of historical episodes between plants (without any modification or adaptation) is just recommended for very similar plants, especially to build up the initial seed. If possible, the historical experiences belonging to other facilities should be replaced with the own historical experiences, as soon as the episode library is representative enough.

In order to increase the transferability of historical experiences between facilities, future work on episode-based reasoning should be based on the extraction of generic knowledge from specific experiences, i.e. the extraction of generic knowledge that could be reused by any facility to solve similar problems. Likewise, the development of a generic case library covering all the possible solids separation problems and their appropriate solutions would facilitate most facilities being able to benefit from all these experiences, having an initial library with at least one historical episode for each problematic situation.

In spite of these evident advantages with respect to classical CBRS, the developed EBRS still presents a couple of limitations that need to be further studied:

⇒ The steps of episode reuse, adaptation and evaluation are not automatic. Given the complexity of the domain and the limited information available, an automatic adaptation, reuse and evaluation of the control strategy is difficult to perform. The performance of such steps is currently a task of the final user, which also implies a kind of risk concerning human errors (lack of registration, registration of subjective evaluations, errors in adapting the suggested control strategy, etc.).

⇒ EBRS can not cope with situations that have never happened before in the process since the episode library will not include reference episodes to compare to. The EBRS needs to have experienced many past examples of problematic situations in order to increase its accuracy. Nevertheless, this fact could be overcome by the complementation offered by the ES. Whenever the EBRS is not capable of retrieving a similar historical experience to handle the current problematic situation, the suggestion of the ES will provide the user with a control strategy based on expert knowledge. Once the user decides whether to consider the suggested control strategy or to apply a different one, the new experience will immediately become part of the episode library and will be available for future use.

CHAPTER 7

IMPLEMENTATION OF THE DYNAMIC KNOWLEDE-BASED DECISION SUPPORT SYSTEM

Once the conceptual design and the development of the dynamic KBDSS have been completed, the next step is to implement the developed system into a full-scale WWTP, in order to test the feasibility and functionality of the tool in a day-to-day operation while supporting plant operators in handling solids separation problems. The following chapter describes the implementation step, including both a description of the WWTP where the system has been implemented and a depiction of the final KBDSS architecture together with some examples of its functionality.

7.1 Experimental System

The wastewater treatment plant selected to implement the dynamic KBDSS prototype is located in Girona, in the Ter river basin (Catalonia, NE of Spain). Built in 1986, this plant initially included preliminary and physical-chemical treatment for organic matter and suspended solids removal. Nowadays, the facility is being enlarged to remove nitrogen, but at the time of the present study it only provided preliminary, primary and secondary treatment to remove the organic matter and suspended solids contained in the raw water of about 160,000 inhabitant-equivalents. The raw influent comes from a sewer that collects urban wastewater and a fraction of industrial wastewater generated at the following municipalities: Girona, Salt, Sarrià, Sant Julià de Ramis, Vilablareix and part of Fornells and Bescanó. A current aerial picture of the Girona WWTP and a detailed picture of the biological reactor and secondary settler are shown in Figure 7.1.

Figure 7.1 (a) Aerial picture of the Girona WWTP (Photo: J.S.Carrera); (b) detailed picture of the biological reactor and (c) detailed picture of the secondary settler

7.1.1 Current plant configuration

The overall treatment process of wastewater in the Girona WWTP, as in any other plant, can be divided into two main treatment lines: water and sludge. Three phases can be distinguished in the water treatment line: preliminary treatment, primary treatment, and secondary (biological) treatment. On the other hand, the treatment line for sludge encompasses the following steps: thickening, anaerobic digestion, dewatering and reuse in agriculture or composting. A flow sheet of the water and sludge line of the Girona WWTP is presented in Figure 7.2.

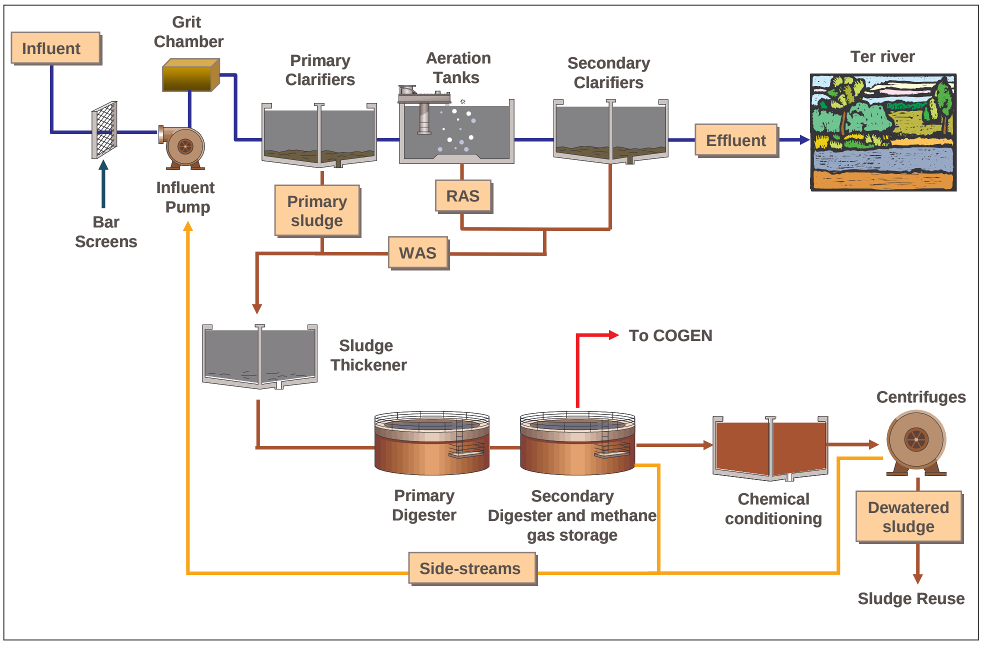


Figure 7.2 Flow diagram of the Girona WWTP including water and sludge lines

The **preliminary treatment** includes screening for coarse particles removal (manual bar racks), 2 grit chambers where sand, foam and oil are removed and a Parshall flume where the influent flow is measured. Wastewater is then impelled to the primary and secondary treatment units by means of 7 pumps. During heavy storms, part of the influent will likely be by-passed to avoid sludge losses from the biological reactor.

The **primary treatment** consists of three circular primary sedimentation tanks (total volume = 4180 m^3) and includes a physical-chemical treatment where ferric sulphate, at 35 mg/L, and sometimes polyelectrolyte are added in order to decrease the COD before entering the secondary treatment. COD is usually decreased by 30%. Primary excess sludge is intermittently wasted at an average rate of 25 m^3/h.

From the primary treatment, the wastewater is derived to the **secondary (biological) treatment**. Wastewater flow rate by-pass is fixed to a set point around 1800 m^3/h. Flow rates exceeding this set point (i.e., during rainy or storm episodes) by-pass the biological treatment and are only subjected to primary treatment. The outflow from primary treatment units is distributed into three parallel lines of biological treatment. The biological treatment consists of the conventional suspended growth system (activated sludge) involving a biological reactor and a secondary settler. The three biological reactors of the Girona WWTP (total volume=7980 m^3) are designed as plug-flow reactors. A dissolved oxygen control enables to minimise the required air supplied. Each line has a dissolved oxygen sensor located at the end of each reactor, and its control loop is based on a PI controller. Downstream from the aeration tank, a set of three circular secondary settlers (total volume: 5885 m^3) are responsible from the separation of the biomass and the treated water, which is discharged into the Ter river.

The excess sludge generated at the primary and secondary settlers are mixed and pumped to the **sludge line**. The excess sludge wasting system includes two pumps which impel a maximum of 53 m^3/h of wasted sludge. The primary sludge is firstly treated by a set of rotary grids and then mixed with the secondary sludge. The mixture of primary and secondary sludge is thickened by means of two circular thickeners (total volume: 1200 m^3). The thickening sludge is then sent to the anaerobic digestion unit. The Girona WWTP is equipped with two anaerobic digesters (total volume: 7000 m^3), where, at the same time that organic matter is degraded, a portion of biogas is produced. This biogas is used as fuel to maintain the desired temperature in digesters and to generate part of the energy used by the facility. Finally, after stabilising the sludge with ferric chloride and lime, the excess sludge is dewatered and dried. This process is conducted by means of 3 centrifuges. Side-streams from the centrifuges (very highly loaded) are mixed with the influent raw wastewater and reinitiate the treatment process. Finally, the dried sludge is reused in agriculture or composting.

7.1.2 Wastewater characteristics

Average values for the main parameters describing the Girona WWTP are listed in Tables 7.1 and 7.2. These values correspond to the historical data recorded during the year 2004.

Table 7.1 Typical wastewater concentrations at influent and primary effluent (or biological reactor inflow)

Parameter	Mean	Q1	Q3
Design inflow rate (m^3/ d)	40000		
Inflow rate (m^3/ d)	45455	41919	49200
COD-influent (mg/L)	504	415	594
BOD-influent (mg/L)	193	150	232
COD-primary effluent (mg/L)	341	281	391
COD removal (%)	83.9	80.9	90.7
NTK-influent (mg/L)	55.9	47.7	63.3
Total P-influent (mg/L)	8.4	7.5	9.4
TSS-influent (mg/L)	203	162	232
pH-primary effluent	7.4	7.4	7.5

Table 7.2 Typical operational and control parameters at the Girona WWTP

Parameter	Mean	Q1	Q3
HRT (hours)	5.7	4.8	6.1
SRT (days)	7.7	4.2	12.1
V30 (mL)	436	330	540
MLSS (mg/L)	1609	1338	1835
SVI (mL/g)	205	133	246
F/M (kg COD/kg MLSS, d)	1.6	0.5	1.2
DO set-point (mg O_2/L)	2.0	1.5	2.0

Table 7.3 Typical wastewater concentrations at effluent

Parameter	Mean	Q1	Q3
TSS-effluent (mg/L)	33.5	12.4	32.2
COD-effluent (mg/L)	81.1	47.2	88.1
NH_4 -effluent (mg/L)	20.7	16.3	27.1
NO_3^- -effluent (mg/L)	2.6	0.3	3.3
Total P-effluent (mg/L)	3.1	2.3	3.8

As observed, the Girona WWTP usually operates at high F/M ratio, this is mainly due to the high organic load entering the process (COD>500 mg/L) and the low concentration of microorganisms present within the activated sludge system (MLSS: 1300-1800 mg/L). In general, a MLSS concentration superior to 2000 mg/L is avoided in order to prevent a solids loss with the effluent whenever problems with sludge settleability occur. Actually, a high SVI (Sludge Volume Index) is quite common in the process, mainly due to filamentous bulking problems (SVI mean > 200 mL/g). On the other hand, apart from the high organic load, the concentration of nitrogen and phosphorous at the influent is relatively high, which avoids a nutrient deficiency for microorganisms. Finally, it can be observed that the SRT is usually high, considering that the facility just operates for organic matter and suspended solids removal.

7.1.3 Available data

In order to characterize and monitor the operation of the WWTP, plant and laboratory operators routinely carry out an exhaustive characterisation of the wastewater at different points of the process: sewer system, influent, primary effluent, aeration tank and effluent. Both water and sludge quality and process state variables, which include both quantitative and qualitative information, are considered. Table 7.3 summarises the quantitative data available in the Girona WWTP, including the source, sampling location, and frequency of each measurement.

Table 7.3 Quantitative data measured in the Girona WWTP (I: influent; PE: primary effluent, E: effluent, RF: return flow, PSF: primary sludge flow, WF: waste flow, AT: aeration tanks, SS: sewer system)

SOURCE	VARIABLE	SAMPLING LOCATION	FREQUENCY OF ANALYSIS
Sensors (on-line)	Water flow rates	I, PE, E	Continuous
	Sludge flow rates	RF, PSF, WF	
	Aeration flow	AT	
	DO concentration	AT (final compartment)	
	pH	SS, I, PE,E	
	Temperature	AT (final compartment)	
	Conductivity	SS, I, PE, E	
Analytical (off-line)	SS	SS, I, PE, E	6 measurements/week
	VSS	SS, I, PE, E	6 measurements/week
	COD	SS, I, PE, E	Daily
	BOD_5	SS, I, PE, E	6 measurements/week
	Soluble COD	PE, E	2 measurements/week
	TKN	SS, I, PE, E	2 measurements/week
	NO_3^-	SS, I, PE, E	2 measurements/week
	MLSS	AT, RF	Daily
	MLVSS	AT, RF	Daily
	V30	AT, RF	Daily
Global (calculated)	HRT	-	Daily
	SRT		6 measurements/week
	F/M		Daily
	SVI		Daily
	Upflow velocity		Daily

7.1.3.1 Quantitative data

Quantitative data measured in the Girona WWTP can be divided into the following types:

- On-line data provided by sensors. These data are continuously gathered by means of system of PLCs which compiles data for a number of parameters, including flow rates of water, sludge and air, at different points along the process, as well as physical parameters such as pH, temperature, conductivity or dissolved oxygen concentration in the aeration tank. The doses of flocculants or polyelectrolyte are also registered together with some extra information regarding the state of different pumps, valves or other devices of the system.

  Since the WWTP does not have at its disposal a SCADA system, only the daily average values of on-line parameters are registered.

- Off-line data are obtained by analyzing integrated samples (from the last 24 hours) at 4 different locations: the sewer system, the WWTP influent, the primary effluent (or aeration tank influent) and the final effluent. Current regulations establish that SS, COD and BOD_5 measured at the influent, primary effluent and final effluent comprise the minimum set of parameters that must be monitored. This implies the routine measurement of COD and the 6 measurements per week of SS, VSS and BOD_5. Although the facility does not yet remove N and P (at the moment is being enlarged to remove N), nutrients are analysed to register the regular values; TKN, NO_3^- and total phosphorus are measured 3 times per week at the 4 sampling points (see Table 7.3). MLSS and MLVSS are measured daily in order to control the biomass concentration in aeration tanks. V30 is measured 3 times per day to monitor sludge settleability. Finally, the soluble COD is measured periodically (usually 2 times per week) in order to study the biodegradability of the substrate.

- Global parameters such as Food to Microorganisms Ratio, Sludge Residence Time, Sludge Volume Index and Hydraulic Residence Time are calculated daily using different quantitative data. They also help to understand how the process is working.

7.1.3.2 Qualitative data

At the Girona WWTP, microscopic observations are usually performed 5 times a week (everyday except from weekends). These observations routinely include a general observation of flocs at 100x and 200x magnifications and a more specific observation of samples at 400x and 1000x magnification destined to identify main species or groups of protozoa, metazoa and if possible, the predominant filamentous bacteria and its general abundance.

Despite the high importance and quality of this kind of information, not any special protocol was routinely carried out to register this daily information. It was periodically taken down in a piece of paper but usually being lost since it was not registered into the computer with the rest of quantitative information.

To avoid this loss of information, the macroscopic and microscopic worksheets specially developed to compile relevant data for the dynamic KBDSS (see Figures 5.2 and 5.3) were provided to the Girona WWTP, where they have been efficiently used from February of 2003.

7.1.4 Operation and control

Despite the exhaustive analytical characterization of wastewater and the treatment process (see Table 7.3), the options to act on the process are not so numerous. There are two automatic control systems: (1) a control loop based on a set-point and established for the dissolved oxygen within the aeration tanks, and (2) an automatic control to monitor the temperature within the anaerobic digestors. Regarding the recirculation flow, the hydraulic design of the cofferdam determines a recirculation flow of approximately 100% of the inflow.

The rest of actions are initiated by the plant manager, who, depending on the evolution of the process (as inferred from the study of the quantitative and qualitative data) usually responds by adjusting the WAS flow or the amount of chemicals (ferric chloride, polyelectrolyte, flocculants, chlorine, etc.) added.

Based on expertise, experience and heuristics, the plant operators have devised specific control strategies to properly operate whenever the facility does not suffer from any unforeseen event. Some examples are:

- ⇒ the dissolved oxygen concentration is kept around 2.0 mg/L within the aeration tanks (1.5 during summer to avoid nitrification)
- ⇒ the recirculation flow is maintained around 100% of the inflow (limited by the by-pass to 39,000 m^3/d)
- ⇒ the WAS flow is calculated according to the target SRT (between 4 and 7 days) and according to the desired MLSS concentration, which is usually maintained below 2,000 mg/L in order not to exceed the capacity of secondary settlers
- ⇒ other set-points are used to fix an approximate dose of 50 mg/L of ferric salts to be added during primary settling or to fix the quantity of polyelectrolyte to be added during sludge handling.

7.1.5 The facility at a glance

The following points summarize the most relevant plant characteristics that warrant the application of a dynamic KBDSS to assist plant operators in the management of solids separation problems:

- The facility undergoes frequent solids separation problems, especially filamentous bulking episodes caused by Type 021N, *H. hydrossis* or *S. natans*. Episodes of non-filamentous bulking or foaming caused by Type 1863 are quite common in the process too. During warmer periods the facility can also suffer from rising sludge problems. On the other hand, solids separation problems caused by deflocculation events are not very common in the process.
- Whenever an episode of filamentous bulking (or less often non-filamentous bulking or foaming) occurs, it usually lasts for several days (even some months). The episodes usually imply high SVI and, on most occasions, loss of solids with the effluent due to the limited capacity of clarifiers, which is even more reduced when a solids separation problem occurs. The most severe episodes of filamentous bulking have been controlled either by manipulating recirculation flow and WAS

flow in order to reduce the sludge concentration in clarifiers, or by shock chlorination at the recycle sludge stream.

- The causes that usually lead to these undesired situations are not well understood. According to plant operators, the important contribution of industrial wastewater makes the influent to be quite variable. An example are the fluctuations of the organic load which makes the F/M to vary from values below 0.7 to values above 1.5, with direct repercussions on the available dissolved oxygen concentration. On the other hand, due to a limitation of the air supplying system, whenever the organic load is too high, the OD set-point of 2.0 mg/L is difficult to maintain, particularly during peak loads (usually registered at night). On such occasions, the supplied oxygen is insufficient for all the biomass, and this affects the overall efficiency of the process both in terms of filaments proliferation and of substrate removal.

  These are, according to plant operators, the possible causes of several operational problems. An argument in favour of this hypothesis is that in summer, when most industries stop their activity, the biologic reactor works better and is more stable than during the rest of the year.

- A significant amount of quantitative and qualitative data (see Table 7.3) are routinely measured in the facility implying a high availability of daily information that could be used by the dynamic KBDSS to monitor the performance of the process.

- The Girona WWTP has a low level of automation due to a lack of a SCADA system capable of collecting on-line data and controlling most of the plant operations.

- There exists a significant amount of historical records corresponding to the characterisation of the plant operation for over more than one year. These records include both quantitative information (e.g., analytic determinations of sludge and water quality at different locations in the plant, and daily average values from different sensors) and qualitative data (e.g., microscopic and macroscopic observations of mixed liquor performed 5 times a week). The examination of these data enables the acquisition of specific and objective knowledge, which is necessary to ensure the sound development of our dynamic KBDSS.

- High level of specialisation of plant experts who have been working in the plant from the beginning of its operation. They know perfectly all type of details that constitute the heuristic knowledge of the plant.

- The proposed dynamic KBDSS can result in a useful consulting tool for plant operators when handling complex solids separation problems. Moreover, this tool could also be used to identify the unknown cause of the severe episodes of solids separation problems and to register the historical memory of the process, which would be very useful both for current plant operators and for new ones, in case the operating company or the plant operators were changed.

7.2 Specifications of the Implementation

The shell used to implement our System was G2®, a real-time rule engine platform (Gensym, 2000). G2 is an object-oriented shell that already incorporates the inference engine. Among other abilities, G2 can: (1) reason about and control events in a continuously changing environment; (2) apply both procedural knowledge and rule-based heuristics; (3) acquire information from any number of data sources, both local and remote; (4) express and make use of relationships between objects, and (5) relieve developers from low-level programming, allowing them to move efficiently into the development of the engineering application.

G2 sustains and optimally controls the reasoning algorithms that run the operation of both the ES (Inference Engine), by chaining the knowledge contained in the knowledge base, and the EBRS (Episode-Based Reasoning System). It also performs the tasks of on-line and off-line data acquisition, database management, temporal reasoning, expert system reasoning, and case-based system reasoning. Moreover, G2 is able to scan the application, focusing on the relevant areas in the same way as a human expert would do, and provides easy connectivity with different data sources and other applications such as SCADA and PLC networks (to acquire data and/or send an order to on-line actuators), or other external applications such as mechanistic or neural network models by means of the G2 Standard Interface (GSI) capabilities (Gensym, 2000).

Knowledge representation in G2 is maintained and extended through classes (in the G2 class hierarchy), rules of different types (if, when, whenever, unconditionally, for, initially), methods and procedures (portions of code containing the details of how to perform any operation). Every class within the G2 class hierarchy is either a system-defined class or a user-defined class. G2 includes a large set of system-defined classes, many of which can be used as the foundation of customised, user-defined classes. Classes have attributes, which define the inherited and locally defined properties of the class. Defining a hierarchy of objects and classes of objects saves time and space, as subclasses can inherit attributes from the classes above in the hierarchy. Attributes common to different objects only need to be defined once in the definition of the superior class, and this efficient organization saves time. The structured organization in classes and objects is one of the characteristics of G2 that contributes to the ease with which a developed KBDSS can be implemented in other facilities.

7.3 The Dynamic KBDSS Architecture

In the following section, the implementation of the dynamic KBDSS approach is described in detail. The process of implementation can be defined as the procedure by which the initial conception of a Knowledge-Based Decision Support System is converted into a useful and effective tool (Klein and Methlie, 1995). In our approach, the key characteristic that we target to provide feasibility to the system is a high efficiency in helping plant operators when a solids separation problem occurs in the process.

Once the different tools or modules which comprise the global KBDSS have been developed, the next step is to integrate and implement it by means of a feasible architecture. In our approach, the proposed KBDSS was structured into a multi-layer architecture (see Figure 7.3), which brings the WWTP process close to the final user (i.e. the plant operator). Fixed abnormal situations, such as the solids separation problems tackled by our system, can be solved more efficiently with a predetermined plan or strategy with the proposed multi-layer architecture than with other types of architectures applied in Artificial Intelligence such as blackboard systems or message passing.

The use of a multi-layer architecture enables the system to: (1) have the appropriate modularity and independence to guarantee the re-design and transferability of the system to other facilities; (2) deal with any kind of data gathered from the process (quantitative and qualitative); (3) include an array of different kinds of knowledge (numerical, heuristic, experiential, and predictive), and (4) integrate the advantages of different intelligent technologies. The use and interaction of both ES and EBRS techniques increases the machine-learning capabilities, improves the reasoning capabilities (deeper knowledge) and upgrades the decision making reliability (Comas, 2000). Finally, the integration of these AI technologies provide the dynamic KBDSS with an **intelligent** behaviour since it should be able to reason both based on the expert knowledge from literature and experts (ES) and based on experiential knowledge accumulated through years of operation in the facility (EBRS). This is really essential for diagnosing and solving abnormal or problematic situations. The combined use of these tools also enables the system to make predictions about possible future events. This intelligent feature of the system allows overcoming the limitations of classical control in this kind of processes.

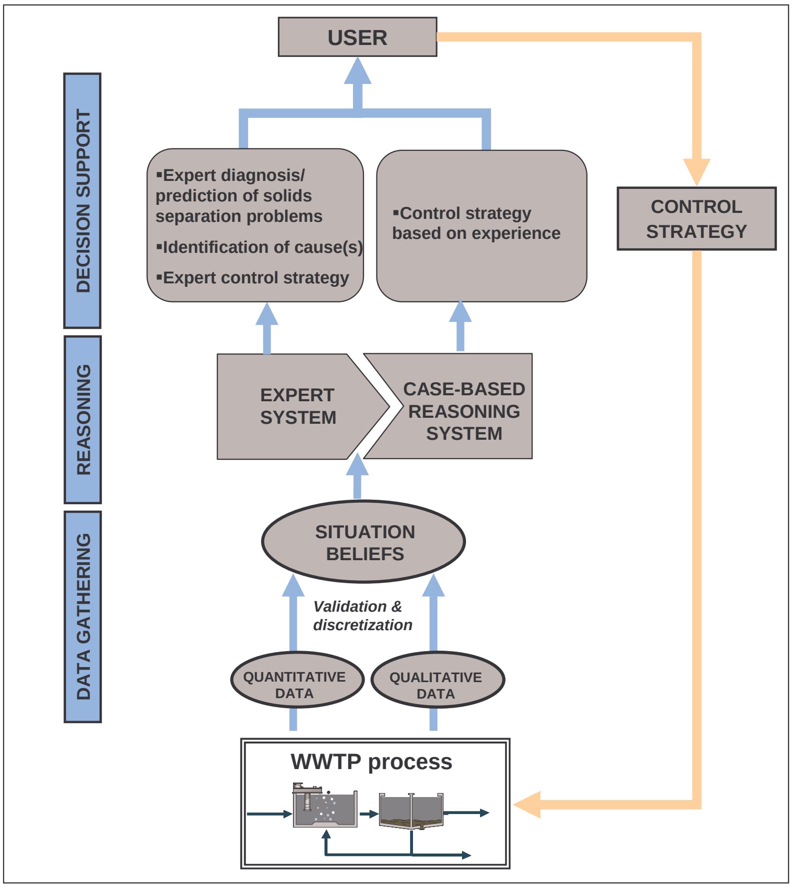


Figure 7.3 Multi-layer architecture of the dynamic KBDSS. Blue arrows represent the actions carried out by the system. Orange arrows represent the actions carried out by the user.

The different tasks carried out by the dynamic KBDSS can be summarised within a three-level or three-step cycle, represented in Figure 7.3, which include the following layers or levels:

- Level 1: Data gathering
- Level 2: Reasoning
- Level 3: Decision Support

The cycle is routinely executed once a day, although it could also be started manually at any time, i.e. whenever an alarm symptom is met. Once the actions of the KBDSS are completed (blue arrows in Figure 7.3) the user receives the results and finally decides, based on this supporting information, which is the most suitable control strategy to be applied (orange arrows in Figure 7.3).

In the present section, the implementation of the KBDSS is explained in detail. The way in which levels 1, 2 and 3 are organised, integrated and implemented is depicted, together with some examples of the resulting interfaces of communication with the user.

7.3.1 Data gathering

The first and lower layer of the dynamic KBDSS encompasses the tasks involved in data gathering and registration into the real-time database. The data acquisition module gathers all kind of data collected at the Girona WWTP that are necessary for the reasoning module to make a decision. It includes both quantitative (on-line data providing from sensors and off-line data providing from biological, chemical and physical water and sludge analyses) and other qualitative data from the process (transformed into categorical variables by means of the macroscopic and microscopic worksheets).

According to the domain where a KBDSS is implemented, there exist a minimum set of variables (the basic information or situation beliefs) that must be updated in order to make a reliable diagnosis of the current state of the process. In the current approach, 27 quantitative data items and 7 qualitative data items are gathered for future use by the reasoning processes of the ES or the EBRS. The data gathering process is performed by means of an EXCEL® spreadsheet. The specific spreadsheet is connected with the databases and daily updated by the plant operators; hence, whenever a new value is introduced, the specific spreadsheet will be automatically updated. On the other hand, in case a SCADA system was available in a facility, it could also be connected to the KBDSS in order to automatically gather all the on-line data. At the Girona WWTP, the on-line data is not gathered by using an SCADA system but by means of a daily registration of daily-average values into a specific database. Figure 7.4 shows the 3 different databases used at the Girona WWTP and the spreadsheet specially developed to collect those data necessary for the upper level of the system architecture (Dades Girona.xls).

Once all the required data are compiled into the specific spreadsheet, they are transferred to the G2 shell by means of the ActiveXlink bridge, a tool also developed by the G2 Gensym Corporation which permits the communication between the G2 shell and other programs. This communication can be achieved even between programs codified in different languages, such as EXCEL® or WORD®. The bridge has been codified in Visual Basic language by a set of subroutines included in a macro within the spreadsheet. Thus, whenever all the required information is introduced in the corresponding spreadsheets and the ActiveXlink bridge is activated (it can be either automatically activated whenever new data is introduced or it can be directly activated by the user), all the registered data are transferred to the G2 by means of data vectors. Figure 7.5 represents an example of the Excel spreadsheet and the ActiveXlink bridge with its corresponding Visual Basic subroutine. Once the necessary data are transferred to the G2, a group of procedures validate, discretise and organize the data into different objects.

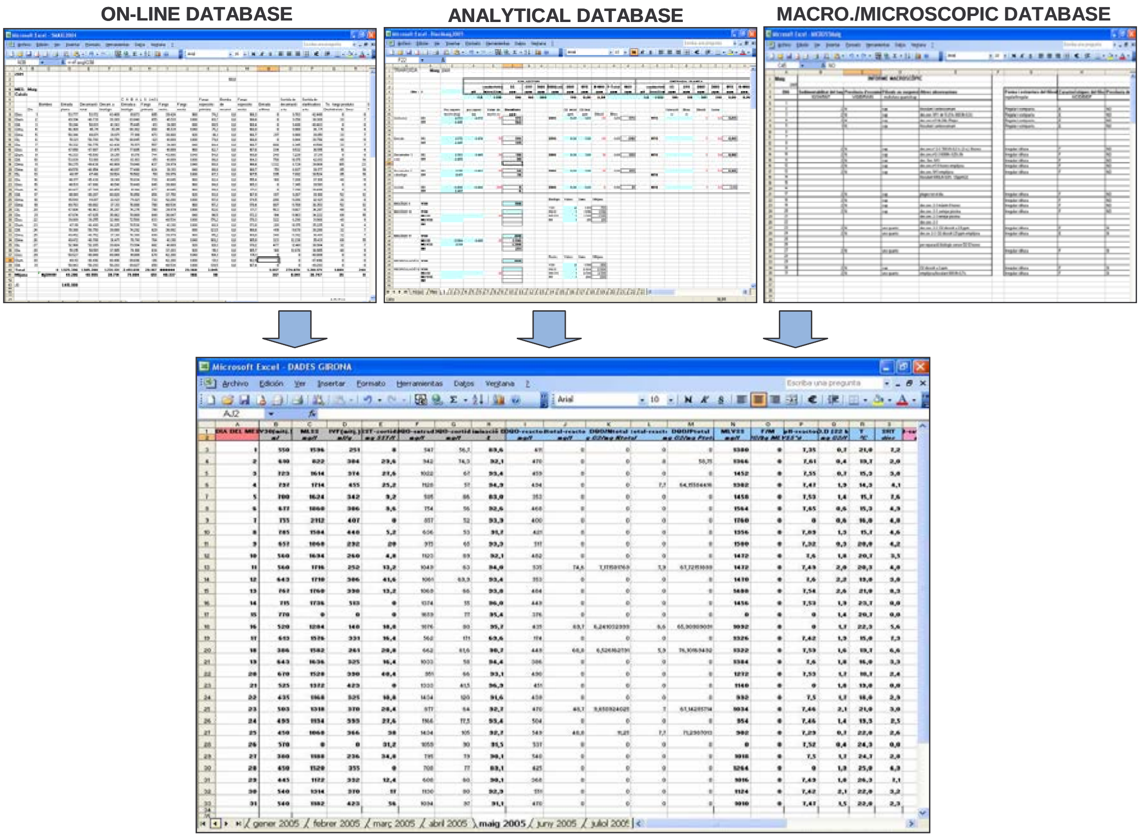
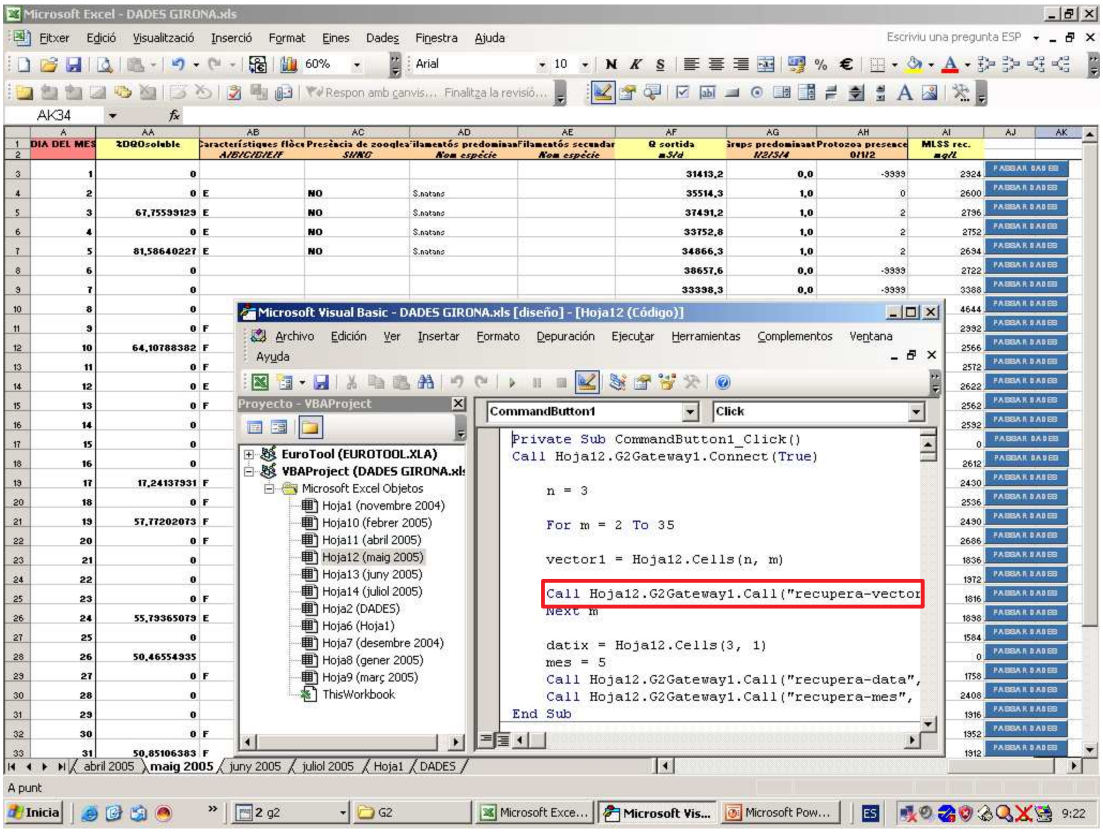


Figure 7.4 Different databases used at the Girona WWTP: on-line database; analytical database and microscopic and macroscopic databases, and the spreadsheet developed for the Girona application



```
Private Sub CommandButton1_Click()
Call Hoja12.G2Gateway1.Connect(True)

    n = 3

    For m = 2 To 35

    vector1 = Hoja12.Cells(n, m)

    Call Hoja12.G2Gateway1.Call("recupera-vector
    Next m

    datix = Hoja12.Cells(3, 1)
    mes = 5
    Call Hoja12.G2Gateway1.Call("recupera-data",
    Call Hoja12.G2Gateway1.Call("recupera-mes",
End Sub
```

Figure 7.5 The Girona spreadsheet with the Visual Basic subroutine to establish a connection with the G2 shell through the ActiveXlink bridge

7.3.1.1 Validation and discretisation of data

Original raw data are often defective, requiring the application of a number of pre-processing procedures before there are committed to the corresponding G2 objects in an understandable and interpretable way. The first step is **data validation**, which includes filtering out incorrect values (outliers), noise, and missing values. Filtered values are then converted into a homogeneous unit and time-scale. Since our KBDSS performs dynamic reasoning, the optimal monitoring of the process state must also include information about the evolution of the main variables. Thus, during data gathering, the system also calculates and analyses changes in some specific variables to detect sudden deviations, trends and periodicities.

The **discretisation** step, as explained in chapter 5, consists of a qualitative abstraction of the quantitative variables into qualitative modalities. The use of qualitative modalities facilitates the reasoning process and the transferability of the architecture to other facilities. The membership functions developed for such application are represented in Figures 7.6 and 7.7, where the membership degree for each value of the variable is represented according to the pre-defined fuzzy sets. Figure 7.6 shows the membership functions developed for each one of the quantitative variables considered during problem diagnosis, while Figure 7.7 represents the membership functions developed for each one of the quantitative variables considered during cause identification.

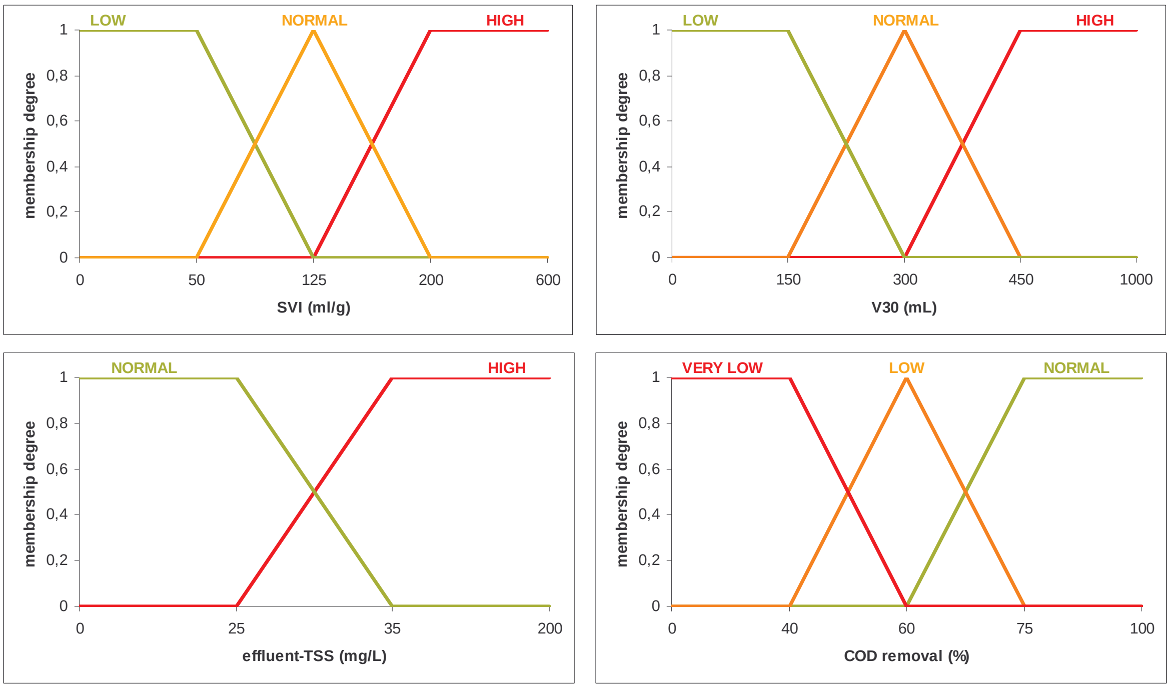


Figure 7.6 Membership functions for the quantitative variables considered during problem diagnosis

As explained in chapter 5, the development of such memebership functions was done by both performing an statistical study of the historical values of the quantitative variables and by the supervision of the plant manager.

The application of the developed membership functions was carried out by the direct inference of the specific equations derived from each one of the different fuzzy sets. For example, according to the membership function of SVI represented in figure 7.6, the equation inferred from SVI values between 125 and 200 mL/g (y=1.33x-167) can be incorporated into the knowledge base by means of the following heuristic rule:

IF *SVI* ≥ 125 and SVI < 200 **THEN** conclude that the *membership-degree of SVI -high*= (SVI * 1.33) - 167

Hence, during data discretisation both the qualitative values and the membership degree (expressed as the % of certainty that the value of a specific variable belongs to one or another category) are obtained for those quantitative variables that will be used later, during the reasoning process of the Expert System.

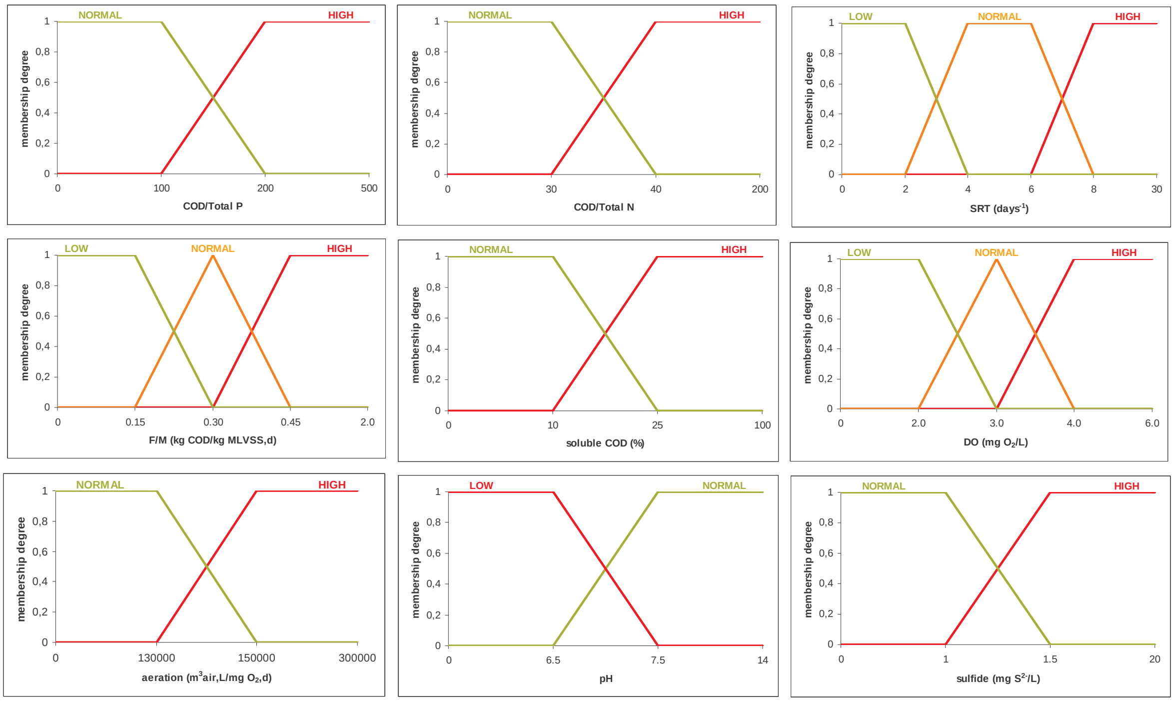


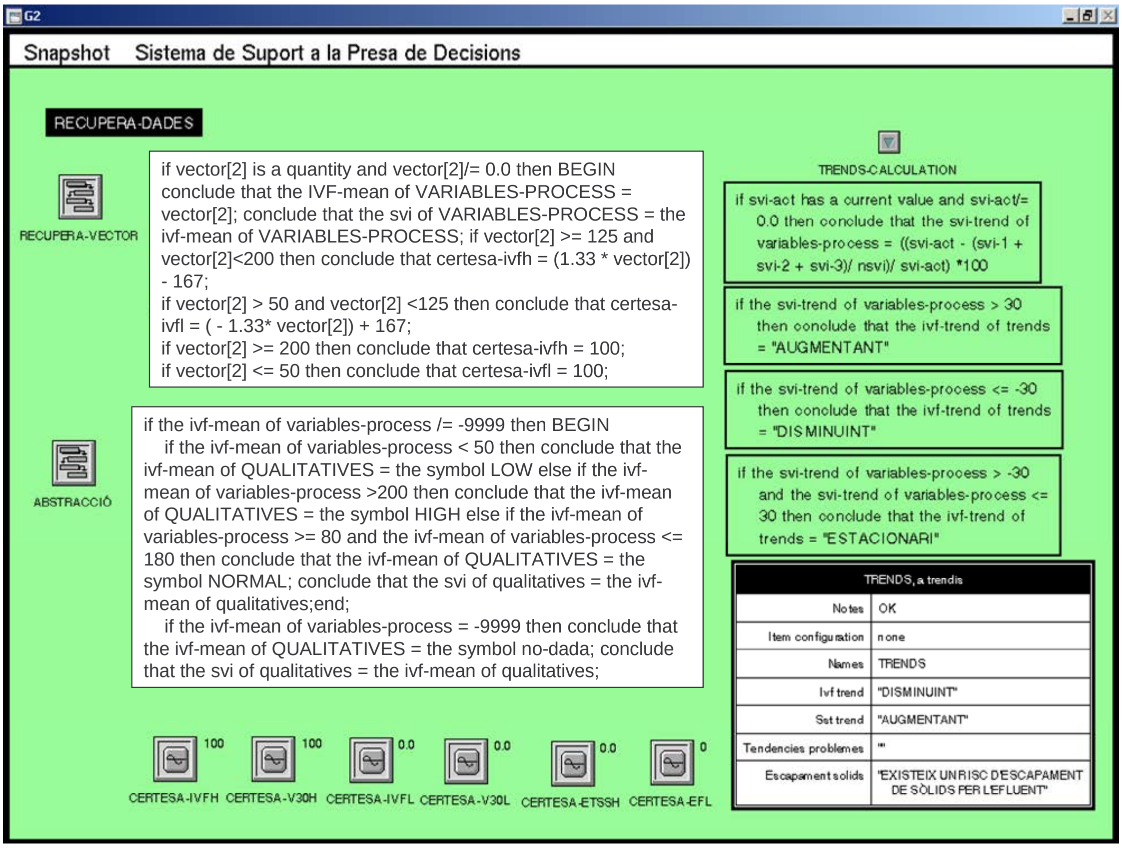
Figure 7.7 Membership functions for the quantitative variables considered during cause identification

Regarding the EBRS, it needs to be mentioned that it does not need to use qualitative categories given that the reasoning processes of the EBRS directly uses the quantitative values of the selected variables.

The results of the validation and discretisation processes are stored in specific objects of G2.

Figure 7.8 represents some of the procedures (“Recupera-vector” and “Abstracció”) developed to perform the validation and abstraction of data together with some calculation rules developed to estimate the trends of the most relevant variables. Figure 7.9 shows some examples of the different objects of G2 which compile all the necessary data used by the Expert System. Finally, Figure 7.10 shows the specific object defined to compile the necessary data which characterizes the current situation in the Episode-Based Reasoning System.

Once all the necessary data are validated, discretised, and compiled into their corresponding objects, both the ES and the EBRS have at its disposal all the necessary information to perform the reasoning process.



if vector[2] is a quantity and vector[2]/= 0.0 then BEGIN conclude that the IVF-mean of VARIABLES-PROCESS = vector[2]; conclude that the svi of VARIABLES-PROCESS = the ivf-mean of VARIABLES-PROCESS; if vector[2] >= 125 and vector[2]<200 then conclude that certesa-ivfh = (1.33 * vector[2]) - 167;
if vector[2] > 50 and vector[2] <125 then conclude that certesa-ivfl = (- 1.33* vector[2]) + 167;
if vector[2] >= 200 then conclude that certesa-ivfh = 100;
if vector[2] <= 50 then conclude that certesa-ivfl = 100;

if the ivf-mean of variables-process /= -9999 then BEGIN
if the ivf-mean of variables-process < 50 then conclude that the ivf-mean of QUALITATIVES = the symbol LOW else if the ivf-mean of variables-process >200 then conclude that the ivf-mean of QUALITATIVES = the symbol HIGH else if the ivf-mean of variables-process >= 80 and the ivf-mean of variables-process <= 180 then conclude that the ivf-mean of QUALITATIVES = the symbol NORMAL; conclude that the svi of qualitatives = the ivf-mean of qualitatives;end;
if the ivf-mean of variables-process = -9999 then conclude that the ivf-mean of QUALITATIVES = the symbol no-dada; conclude that the svi of qualitatives = the ivf-mean of qualitatives;

if svi-act has a current value and svi-act/= 0.0 then conclude that the svi-trend of variables-process = ((svi-act - (svi-1 + svi-2 + svi-3)/ nsvi)/ svi-act) *100

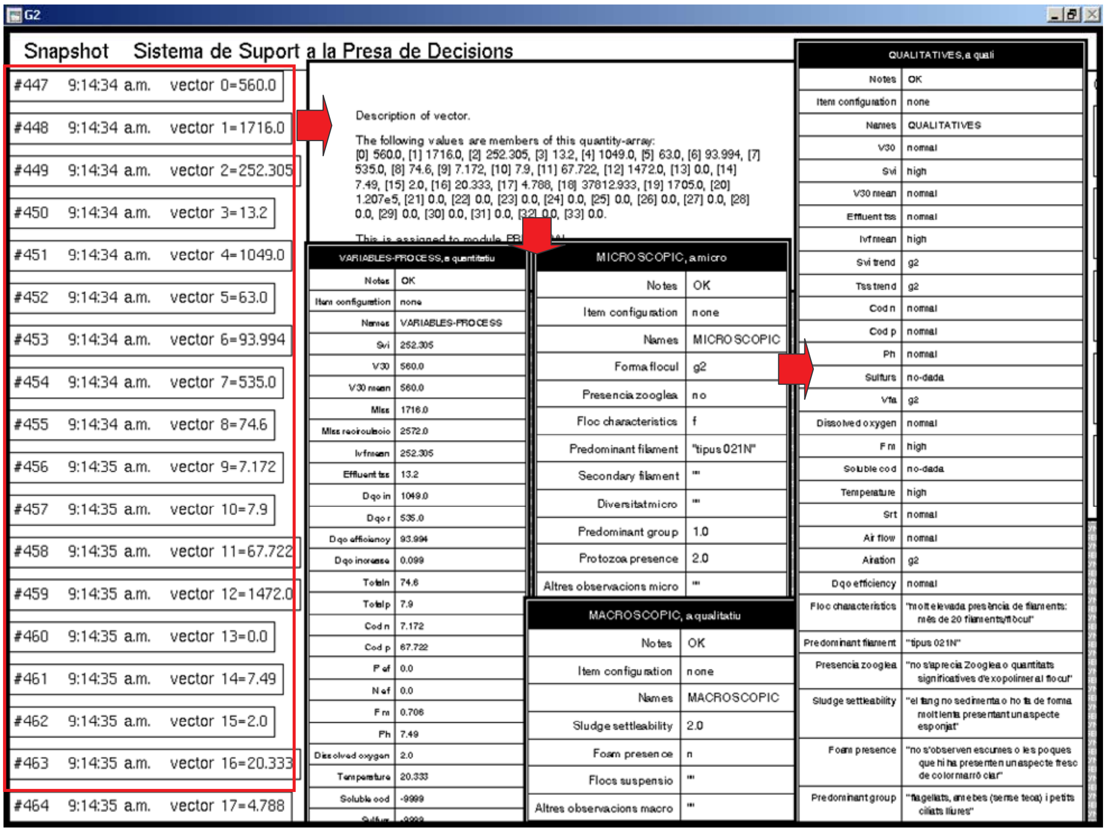
if the svi-trend of variables-process > 30 then conclude that the ivf-trend of trends = "AUGMENTANT"

if the svi-trend of variables-process <= -30 then conclude that the ivf-trend of trends = "DISMINUINT"

if the svi-trend of variables-process > -30 and the svi-trend of variables-process <= 30 then conclude that the ivf-trend of trends = "ESTACIONARI"

TRENDS, a trendis	
Notes	OK
Item configuration	none
Names	TRENDS
Ivf trend	"DISMINUINT"
Sst trend	"AUGMENTANT"
Tendencies problemes	""
Escapament solids	"EXISTEIX UN RISC DESCAPAMENT DE SÒLIDS PER L'EFLUENT"

Figure 7.8 Part of the procedures developed to perform the validation and discretisation of data and some of the rules developed to calculate the trends of the most relevant variables



```
#447 9:14:34 a.m. vector 0=560.0
#448 9:14:34 a.m. vector 1=1716.0
#449 9:14:34 a.m. vector 2=252.305
#450 9:14:34 a.m. vector 3=13.2
#451 9:14:34 a.m. vector 4=1049.0
#452 9:14:34 a.m. vector 5=63.0
#453 9:14:34 a.m. vector 6=93.994
#454 9:14:34 a.m. vector 7=535.0
#455 9:14:34 a.m. vector 8=74.6
#456 9:14:35 a.m. vector 9=7.172
#457 9:14:35 a.m. vector 10=7.9
#458 9:14:35 a.m. vector 11=67.722
#459 9:14:35 a.m. vector 12=1472.0
#460 9:14:35 a.m. vector 13=0.0
#461 9:14:35 a.m. vector 14=7.49
#462 9:14:35 a.m. vector 15=2.0
#463 9:14:35 a.m. vector 16=20.333
#464 9:14:35 a.m. vector 17=4.788
```

Description of vector.

The following values are members of this quantity-array:
[0] 560.0, [1] 1716.0, [2] 252.305, [3] 13.2, [4] 1049.0, [5] 63.0, [6] 93.994, [7] 535.0, [8] 74.6, [9] 7.172, [10] 7.9, [11] 67.722, [12] 1472.0, [13] 0.0, [14] 7.49, [15] 2.0, [16] 20.333, [17] 4.788, [18] 37812.933, [19] 1705.0, [20] 1.207e5, [21] 0.0, [22] 0.0, [23] 0.0, [24] 0.0, [25] 0.0, [26] 0.0, [27] 0.0, [28] 0.0, [29] 0.0, [30] 0.0, [31] 0.0, [32] 0.0, [33] 0.0.

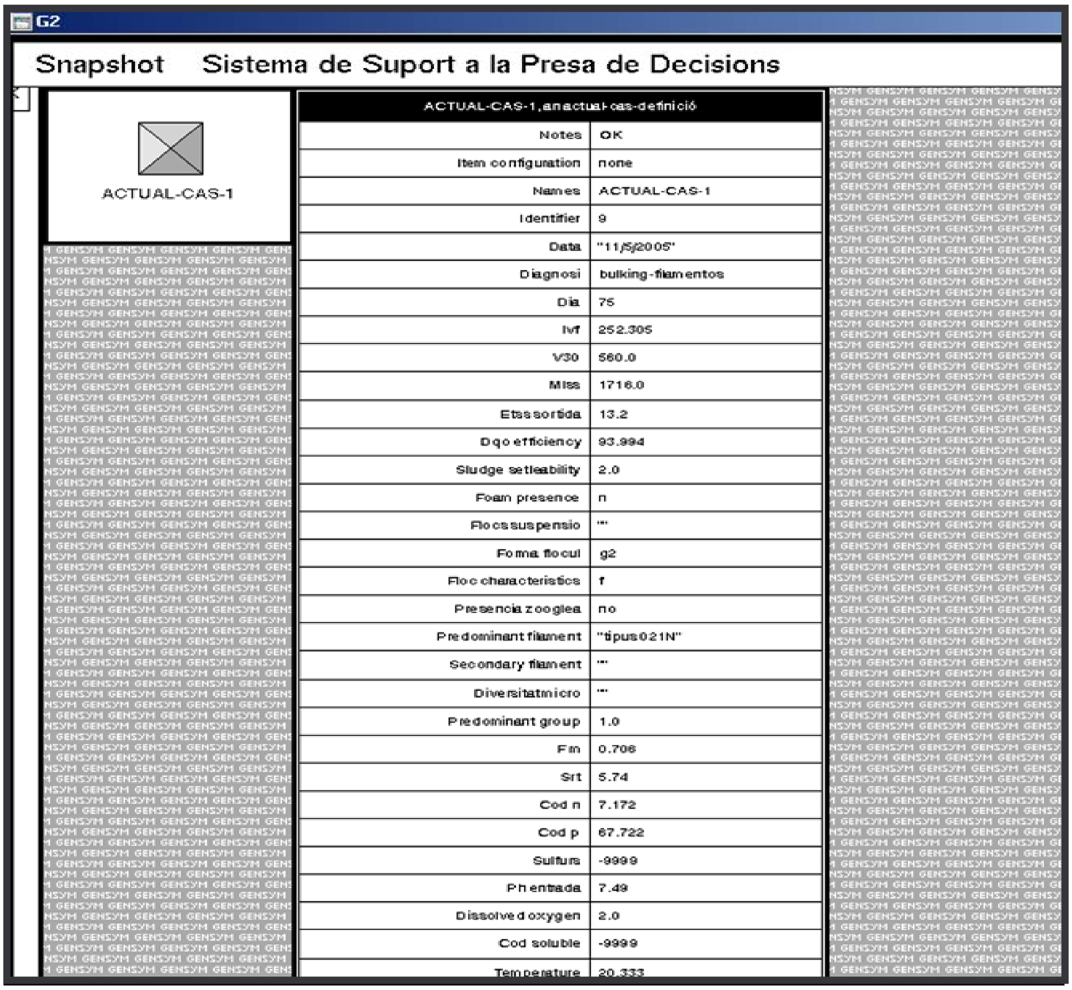
VARIABLES-PROCESS, a quantitatiu	
Notes	OK
Item configuration	none
Names	VARIABLES-PROCESS
Svi	252.305
V30	560.0
V30 mean	560.0
Mlss	1716.0
Mlss recirculacio	2572.0
Ivfmean	252.305
Effluent tss	13.2
Dqo in	1049.0
Dqo r	535.0
Dqo efficiency	93.994
Dqo increase	0.099
Totaln	74.6
Totalp	7.9
Cod n	7.172
Cod p	67.722
P ef	0.0
N ef	0.0
F m	0.706
Ph	7.49
Dissolved oxygen	2.0
Temperature	20.333
Soluble cod	-9999

MICROSCOPIC, a micro	
Notes	OK
Item configuration	none
Names	MICROSCOPIC
Forma flocul	g2
Presencia zooglea	no
Floc characteristics	f
Predominant filament	"tipus 021N"
Secondary filament	""
Diversitat micro	""
Predominant group	1.0
Protozoa presence	2.0
Altres observacions micro	""

MACROSCOPIC, a qualitatiu	
Notes	OK
Item configuration	none
Names	MACROSCOPIC
Sludge settleability	2.0
Foam presence	n
Flocs suspensio	""
Altres observacions macro	""

QUALITATIVES, a quali	
Notes	OK
Item configuration	none
Names	QUALITATIVES
V30	normal
Svi	high
V30 mean	normal
Effluent tss	normal
Ivf mean	high
Svi trend	g2
Tss trend	g2
Cod n	normal
Cod p	normal
Ph	normal
Sulfurs	no-dada
Vfa	g2
Dissolved oxygen	normal
F m	high
Soluble cod	no-dada
Temperature	high
Srt	normal
Air flow	normal
Airation	g2
Dqo efficiency	normal
Floc characteristics	"molt elevada presència de filaments: més de 20 filaments/flòcul"
Predominant filament	"tipus 021N"
Presencia zooglea	"no s'aprecia Zooglea o quantitats significatives d'exopolimer al flocul"
Sludge settleability	"el fang no sedimenta o ho fa de forma molt lenta presentant un aspecte esponjat"
Foam presence	"no s'observen escumes o les poques que hi ha presenten un aspecte fresc de color marró clar"
Predominant group	"flagellats, amebes (sense teca) i petits ciliats lliures"

Figure 7.9 Some of the objects in G2 which compile the necessary data needed by the Expert System



Sistema de Suport a la Presa de Decisions

ACTUAL-CAS-1,anactual-cas-definició	
Notes	OK
Item configuration	none
Names	ACTUAL-CAS-1
Identifier	9
Data	"11/5/2005"
Diagnosi	bulking-filamentos
Dia	75
Ivf	252.305
V30	560.0
Mlss	1716.0
Etss sortida	13.2
Dqo efficiency	93.994
Sludge setleability	2.0
Foam presence	n
Flocs suspensio	""
Forma flocul	g2
Floc characteristics	f
Presencia zooglea	no
Predominant filament	"tipus021N"
Secondary filament	""
Diversitatmicro	""
Predominant group	1.0
Fm	0.706
Srt	5.74
Cod n	7.172
Cod p	67.722
Sulfurs	-9999
Ph entrada	7.49
Dissolved oxygen	2.0
Cod soluble	-9999
Temperature	20.333

Figure 7.10 Specific object defined in G2 to compile the necessary information about a case in the EBRS. Just some of the attributes are showed.

7.3.2 Reasoning

The data gathering module provides all the necessary information to the second level of the architecture, the Reasoning level, which is in charge of inferring diagnosis regarding the state of the process in order to reach a reasonable proposal of action to support the whole plant supervision. As described in previous sections, two knowledge-based tools were used. The dynamic Expert System, based on expert knowledge, is structured in knowledge bases and heuristic rules. The Episode-Based Reasoning System, based on the experiential knowledge stored within the historical experiences, consists of a main episode-library which includes a set of past experiences of the process grouped into different episodes. Both systems are codified in the G2 shell. A fair approximation of the necessary volume of code is 120,000 lines corresponding to 578 rules, 1954 procedures, 116 text vectors, 66 quantity vectors, 1022 variables or parameters, and 71 objects.

7.3.2.1 Dynamic Expert System implementation

The conceptual design and development of the dynamic ES has been widely explained in chapter 5. In the present section, the knowledge acquisition and the implementation in G2 are specifically depicted.

The **knowledge acquisition** process was performed by means of a deep literature review and a series of interviews and symposia with selected experts in the domain from which both heuristic and tacit knowledge was acquired. With all this knowledge on the specific domain of solids separation problems, 57 decision trees were developed. Every path or branch of the final decision trees was codified as a production rule (see Figure 5.6), often heuristic, and implemented into the G2 shell.

The collection of extracted rules constitutes the whole Knowledge Base (KB) of the Expert System. As depicted in chapter 5, the structure of the Expert System KB is modular. Each module has a specific task and consists of different sets of rules, methods and/or procedures. Figure 7.11 shows the modular structure of the KB, which includes the following modules: meta-diagnosis, diagnosis of solids separation problems, tackling of uncertainty and risk of future situations, identification of the possible cause(s) and proposal of the control strategy.

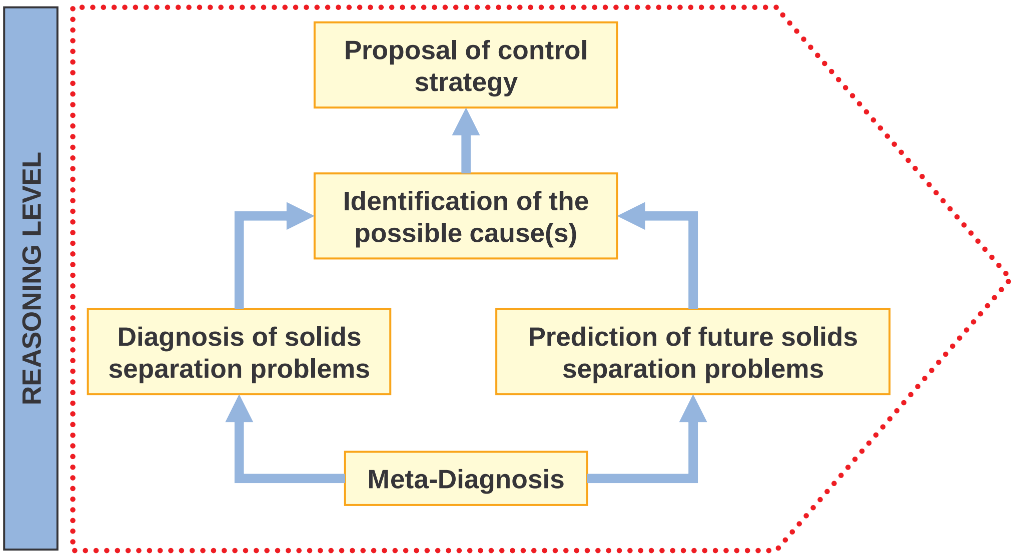


Figure 7.11 Modular structure of the dynamic Expert System

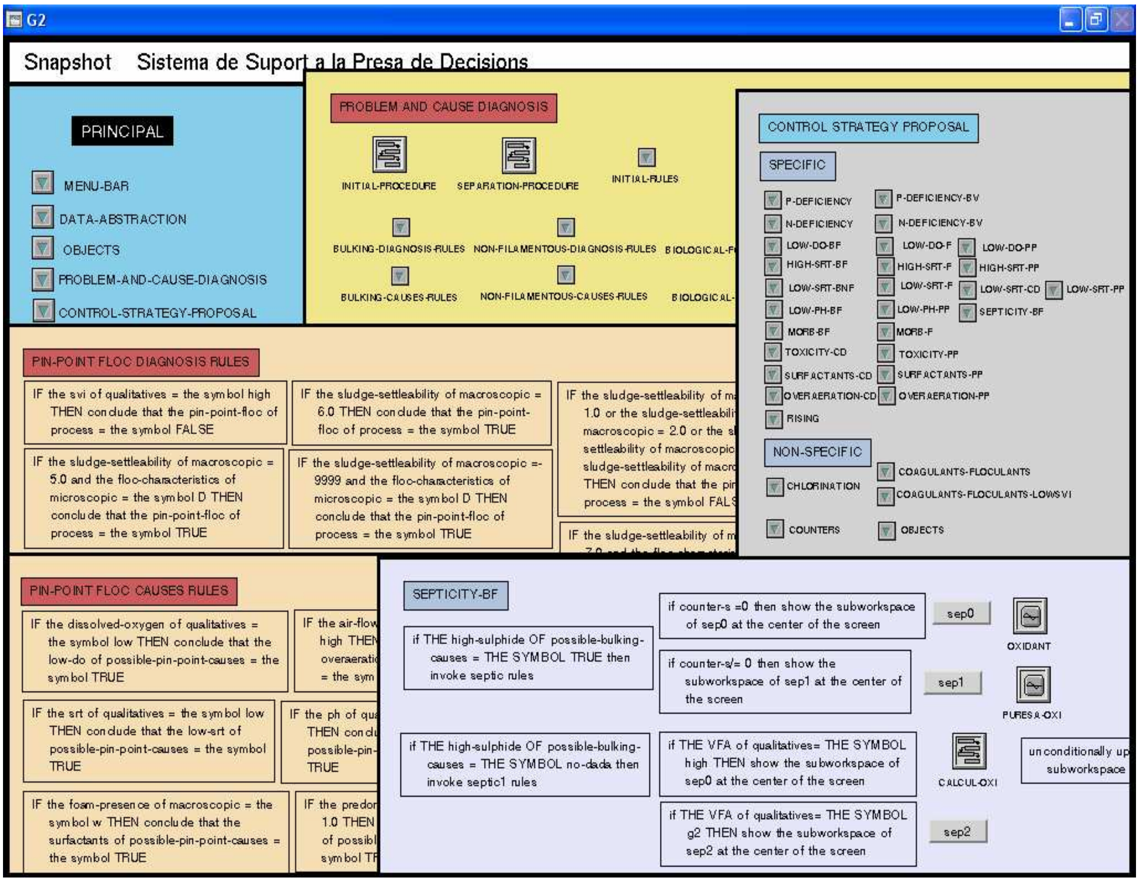
The set of rules related to each module were implemented by defining different categories of rules or procedures. First of all, two general procedures named “Initial-procedure” and “Separation-procedure” were developed to organise the inference of the different rule modules. This can be considered as the meta-diagnosis module and is, in terms of ES architecture, an important part of the inference engine. In Figure 7.12, a fragment of the “Separation-procedure” is presented, where different rules to diagnose a filamentous bulking situation and its possible causes are inferred.

Thus, different set of rules, as represented in Figure 7.13, were defined to perform the diagnosis of the different solids separation problems: bulking-diagnosis-rules, non-filamentous-diagnosis-rules, etc. and their possible causes: bulking-causes-rules, non-filamentous-causes-rules and so on. Likewise, a set of rules was developed to determine the most suitable control strategy. These rules were classified into different categories regarding whether the cause was identified (specific) or not identified (non-specific). Thus, for each specific cause (i.e. nutrient deficiency, phosphorous deficiency, low DO, etc.) a set of rules was developed. Moreover, a set of rules was also developed to propose non-specific control methods, such as chlorination or the addition of flocculants/coagulants, in case the ES does not reach a conclusion about the specific cause.

Separation-procedure()

```
…"IF the ivf-mean of variables-process=-9999 and the v30-mean of variables-process=-9999 then BEGIN
  conclude that the filamentous-bulking of process = the symbol diagnosi-no-possible;
  conclude that the non-filamentous-bulking of process = the symbol diagnosi-no-possible;
  conclude that the biological-foaming of process = the symbol diagnosi-no-possible;END;

  IF certesa-ivfh/= 0.0 or certesa-v30h/= 0.0 then BEGIN
  conclude that the filamentous-bulking of process = the symbol POSSIBLE;

  IF the sludge-settleability of macroscopic = -9999 and the floc-characteristics of microscopic=the symbol
  g2 or the sludge-settleability of macroscopic = -9999 and the foam-presence of macroscopic = the symbol
  g2 and the predominant-filament of microscopic= "" then begin conclude that falta-bf=1; show the
  subworkspace of lack-of-data;end;

  IF falta-bf=0 then invoke bulking-diagnosis rules;
  wait for 0.001 seconds;
  if contador-bf/= 0 and the filamentous-bulking of process=the symbol false then begin
  conclude that epiacabat-bf=1;
  conclude that bf-acabat=episodi-bf;
  conclude that episodi-bf = episodi-bf + 1; end;

  IF the filamentous-bulking of process = the symbol true then BEGIN
  IF contador-bf/= 0 then begin
        if data-2 /= data-ante then
        conclude that contador-bf = contador-bf + 1;
        if data-2 = data-ante then conclude that contador-bf= contador-bf;end;
  if contador-bf = 0 then conclude that contador-bf= contador-bf + 1;

  IF the filamentous-bulking of process = the symbol true or the filamentous-bulking of process = the symbol
  possible then invoke bulking-causes rules;"…
```

Figure 7.12 Fragment of the Separation-procedure used during the meta-diagnosis module



IF the svi of qualitatives = the symbol high THEN conclude that the pin-point-floc of process = the symbol FALSE

IF the sludge-settleability of macroscopic = 6.0 THEN conclude that the pin-point-floc of process = the symbol TRUE

IF the sludge-settleability of ma
1.0 or the sludge-settleabili
macroscopic = 2.0 or the sl
settleability of macroscopic
sludge-settleability of macro
THEN conclude that the pi
process = the symbol FALS

IF the sludge-settleability of macroscopic = 5.0 and the floc-characteristics of microscopic = the symbol D THEN conclude that the pin-point-floc of process = the symbol TRUE

IF the sludge-settleability of macroscopic =-9999 and the floc-characteristics of microscopic = the symbol D THEN conclude that the pin-point-floc of process = the symbol TRUE

IF the sludge-settleability of m

IF the dissolved-oxygen of qualitatives = the symbol low THEN conclude that the low-do of possible-pin-point-causes = the symbol TRUE

IF the air-flow
high THEN
overaeratio
= the sym

IF the srt of qualitatives = the symbol low THEN conclude that the low-srt of possible-pin-point-causes = the symbol TRUE

IF the ph of qua
THEN conclu
possible-pin-
TRUE

IF the foam-presence of macroscopic = the symbol w THEN conclude that the surfactants of possible-pin-point-causes = the symbol TRUE

IF the predom
1.0 THEN
of possibl
symbol TF

if THE high-sulphide OF possible-bulking-causes = THE SYMBOL TRUE then invoke septic rules

if THE high-sulphide OF possible-bulking-causes = THE SYMBOL no-dada then invoke septic1 rules

if counter-s =0 then show the subworkspace of sep0 at the center of the screen

if counter-s/= 0 then show the subworkspace of sep1 at the center of the screen

if THE VFA of qualitatives= THE SYMBOL high THEN show the subworkspace of sep0 at the center of the screen

if THE VFA of qualitatives= THE SYMBOL g2 THEN show the subworkspace of sep2 at the center of the screen

unconditionally up
subworkspace

Figure 7.13 Different modules of the ES regarding problem and cause diagnosis and control strategy proposal

These modules of rules also include several procedures that calculate some of the control parameters suggested at the final proposal of control strategy. The following is an example of a procedure used to calculate the required purge flow, once the "Decreasing SRT strategy" is selected as the most suitable control strategy to be applied.

```
Calculation-purge-flow()
BEGIN
conclude that the purge-flow of operator = (((the reactor-volume of facility * the MLSS of variables-process)/ SRT-des) - (the out-flow of variables-process * the effluent-tss of variables-process))/ the MLSS-recirculation of variables-process;
invoke update-purge-flow rules;
END
```

7.3.2.2 Integration of ES and CBRS

Once all the necessary information has been gathered and processed, the Expert System infers all the required heuristic rules to accomplish its main goals: (1) identification of new episodes of solids separation problems; (2) analysis of the evolution of the problem; (3) identification of the end of a problem; (4) prediction of possible problematic situations; (4) identification of the possible causes of present or future problems, and (5) proposal of complete control methods to restore the process efficiently. This is then followed by the application of a second knowledge-based tool, the EBRS, which provides the user with some complementary information, this time gathered from experience, which will help him/her to make the final decision.

In order to speed up the reasoning process of the EBRS, integration with the ES tool is performed. As explained in chapter 6, in order to perform the first action of the EBRS cycle (i.e., the retrieval of the most similar case and episode), a first definition of the current situation (current case) is necessary. To accomplish this first objective, some of the results obtained by the dynamic Expert System are used.

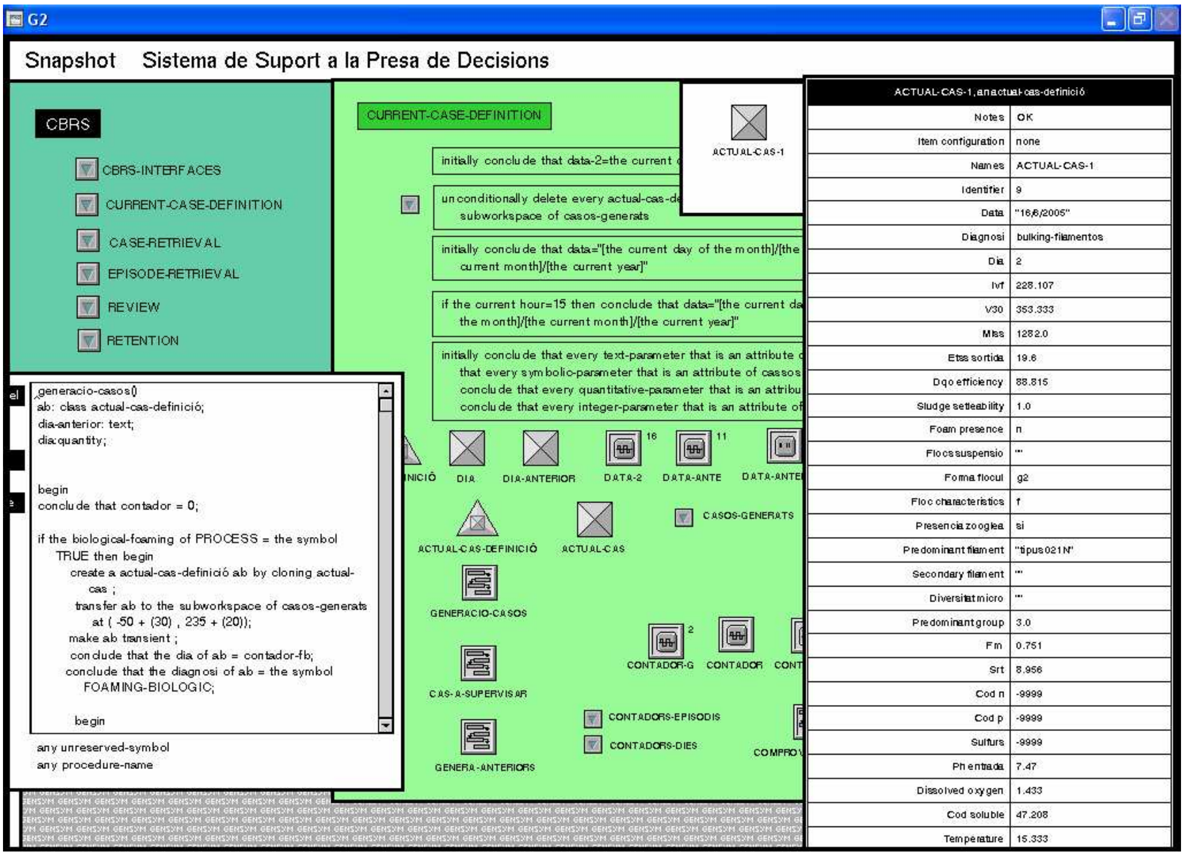
The definition of the current case incorporates the following information provided by the results of the Dynamic Expert System:

⇒ Diagnosis of the current state (:case-problem-diagnosis)
⇒ Temporal situation of the case: new episode or current episode (:temporal-identifier)

Apart from this important information, the collection of the 22 key variables performed during the data-gathering step, provides the case description with the last necessary attribute (:case-situation-description) to initiate the reasoning process.

7.3.2.3 Episode-Based Reasoning System implementation

The developed EBRS was implemented as well into the G2 shell. The proposed working cycle (depicted in 6.4 *The dynamic working cycle*), requires the accomplishment of several steps including: definition of the current situation; searching and retrieval of the most similar case and its related episode; episode reuse, adaptation and application; evaluation and learning. All the knowledge required to perform this working cycle was implemented by defining different modules. Figure 7.14 shows the different modules used during the implementation of the dynamic CBRS together with some of the objects, procedures and rules defined.



```
initially conclude that data-2=the current
```

```
unconditionally delete every actual-cas-de
    subworkspace of casos-generats
```

```
initially conclude that data="[the current day of the month]/[the
    current month]/[the current year]"
```

```
if the current hour=15 then conclude that data="[the current da
    the month]/[the current month]/[the current year]"
```

```
initially conclude that every text-parameter that is an attribute c
    that every symbolic-parsmeter that is an attribute of cassos
    conclude that every quantitative-parameter that is an attribu
    conclude that every integer-parameter that is an attribute of
```

```
generacio-casos()
ab: class actual-cas-definició;
dia-anterior: text;
dia:quantity;


begin
conclude that contador = 0;

if the biological-foaming of PROCESS = the symbol
    TRUE then begin
       create a actual-cas-definició ab by cloning actual-
          cas ;
        transfer ab to the subworkspace of casos-generats
           at ( -50 + (30) , 235 + (20));
      make ab transient ;
       conclude that the dia of ab = contador-fb;
     conclude that the diagnosi of ab = the symbol
        FOAMING-BIOLOGIC;

        begin
```

any unreserved-symbol
any procedure-name

ACTUAL-CAS-1, an actual-cas-definició	
Notes	OK
Item configuration	none
Names	ACTUAL-CAS-1
Identifier	9
Data	"16/6/2005"
Diagnosi	bulking-filamentos
Dia	2
Ivf	228.107
V30	353.333
Mlss	1282.0
Etss sortida	19.6
Dqo efficiency	88.815
Sludge setteability	1.0
Foam presence	n
Flocs suspensio	""
Forma flocul	g2
Floc characteristics	f
Presencia zooglea	si
Predominant filament	"tipus 021N"
Secondary filament	""
Diversitat micro	""
Predominant group	3.0
Fm	0.751
Srt	8.956
Cod n	-9999
Cod p	-9999
Sulfurs	-9999
Ph entrada	7.47
Dissolved oxygen	1.433
Cod soluble	47.208
Temperature	15.333

Figure 7.14 Modules of the CBRS implemented in G2 including objects, tables of attributes, rules and procedures

As depicted in Figure 7.14, six main modules were developed:

⇒ CBRS-interfaces, containing all the specific interfaces of the CBRS tool.

⇒ Current-case-definition, in charge of compiling all the attributes required to define the current situation.

⇒ Case-retrieval, the key step on the CBRS cycle, where the episode library is scanned and the most similar case is retrieved.

⇒ Episode-retrieval, a module which is inferred immediately after case-retrieval is concluded. It provides the user with the global experience about the episode that the retrieved case belongs to.

⇒ Review, the penultimate step within the cycle, where the user evaluates the results of applying the selected control strategy.

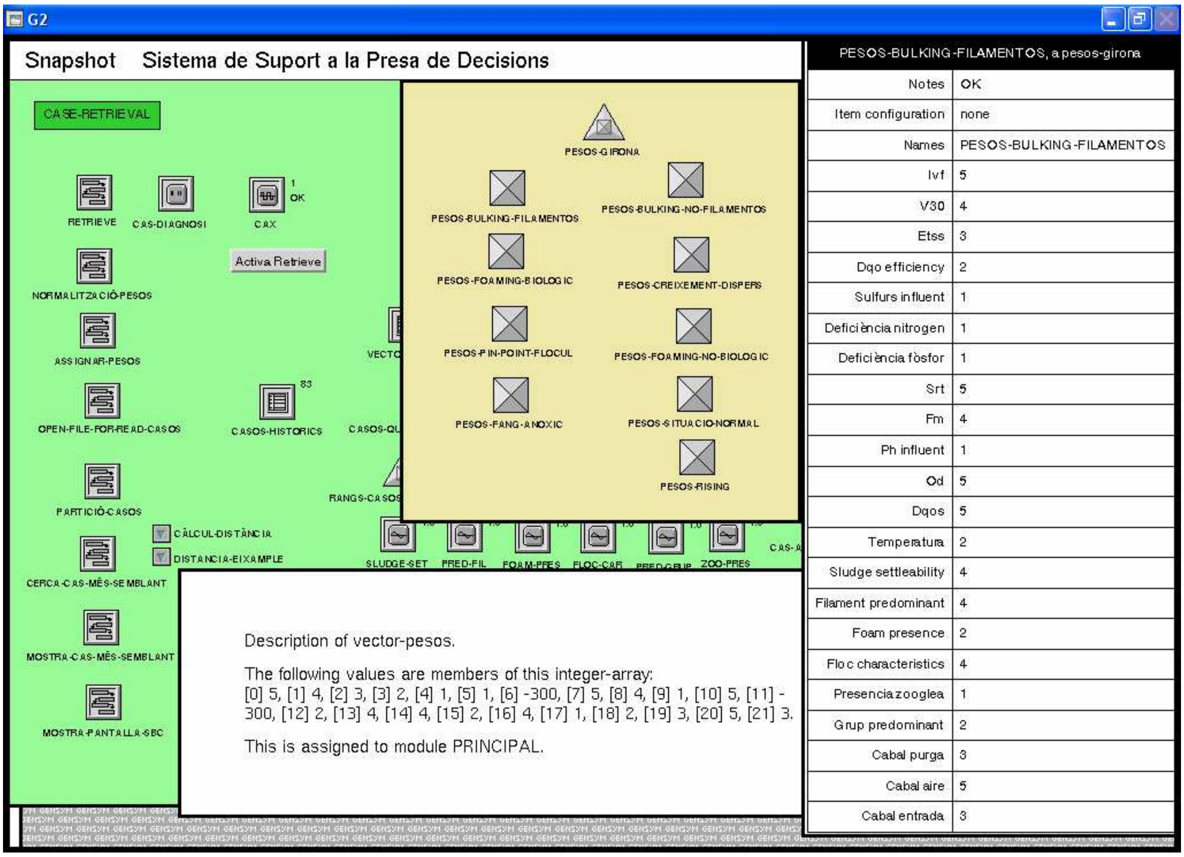
⇒ Retention, a module in charge of retaining the new experience within the episode library. All the required attributes are compiled during the execution of this module.

It needs to be mentioned that the step of reusing and adapting the retrieved experiences was left to be performed by the final user in order to completely adjust the results of the system to the user requirements.

As previously commented, once the data gathering and the reasoning processes of the ES are completed, the EBRS has at its disposal all the required knowledge to start the reasoning cycle. Specifically, the cycle is started by the Separation-procedure() (see Figure 7.12) whenever the ES reaches a conclusion about the current problem and cause. The **current-case-definition** module defines the current situation (current case), incorporating all the required attributes that characterize a case, and at the end, the case-retrieval module is invoked.

The **case-retrieval** module carries out different subtasks until reaching a conclusion about the most similar case of the case-library. The different tasks are invoked, in order, by means of a general procedure, the Retrieve() procedure. The different subtasks are:

1. Assignment of the specific weights (depending on the specific problem and cause). This task is performed by the assignar-pesos() procedure. The specific weights used during the implementation process are represented in the relevance matrix of Figure 6.7. Figure 7.15 shows the different objects of G2 that compile the initial weights of the 22 key variables (pesos-bulking-filamentos, pesos-foaming-biologic, etc.), according to the identified solids separation problem. During the assignar-pesos() procedure, these initial weights are modified also according to the identified cause. The definitive weights are finally stored in a specific vector: vector-pesos.



Description of vector-pesos.

The following values are members of this integer-array:
[0] 5, [1] 4, [2] 3, [3] 2, [4] 1, [5] 1, [6] -300, [7] 5, [8] 4, [9] 1, [10] 5, [11] -300, [12] 2, [13] 4, [14] 4, [15] 2, [16] 4, [17] 1, [18] 2, [19] 3, [20] 5, [21] 3.

This is assigned to module PRINCIPAL.

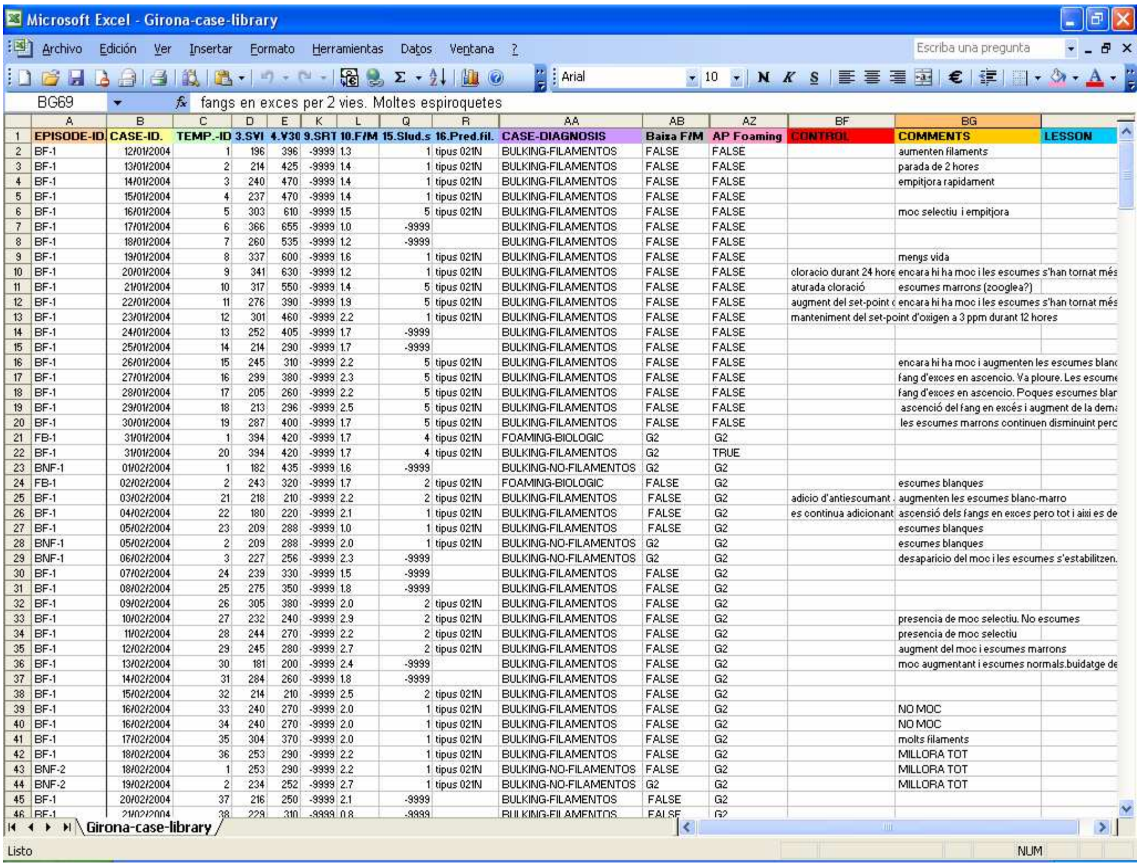
PESOS-BULKING-FILAMENTOS, a pesos-girona

Notes	OK
Item configuration	none
Names	PESOS-BULKING-FILAMENTOS
Ivf	5
V30	4
Etss	3
Dqo efficiency	2
Sulfurs influent	1
Deficiència nitrogen	1
Deficiència fòsfor	1
Srt	5
Fm	4
Ph influent	1
Od	5
Dqos	5
Temperatura	2
Sludge settleability	4
Filament predominant	4
Foam presence	2
Floc characteristics	4
Presencia zooglea	1
Grup predominant	2
Cabal purga	3
Cabal aire	5
Cabal entrada	3

Figure 7.15 Case-retrieval module with an example of the initial weights for a situation of filamentous bulking and the vector-pesos that shows the final specific weights obtained once the assignar-pesos procedure is applied.

2. Exploration of the case-library. The important searching step of the EBRS cycle involves the exploration of the episde library.

    The implementation of an EBRS implies the previous definition of an episode library including a set of past experiences at the facility where the system is to be implemented. The initial episode library (or initial seed) of Girona WWTP contained experiences from January 2004 to September 2004. It included a representative set of 11 real episodes (4 episodes of filamentous bulking, 4 episodes of non-filamentous bulking, 2 episodes of biological foaming and 1 episode of rising sludge) involving a total of 259 individual cases. This initial seed was developed by performing an exhaustive review of different databases from the facility and by carrying out some interviews with the plant operators. Figure 7.16 shows part of the initial case-library of Girona WWTP. As observed, it was implemented on an EXCEL® spreadsheet. Other types of database environments such as ACCESS® could be used instead.



BG69 fangs en exces per 2 vies. Moltes espiroquetes

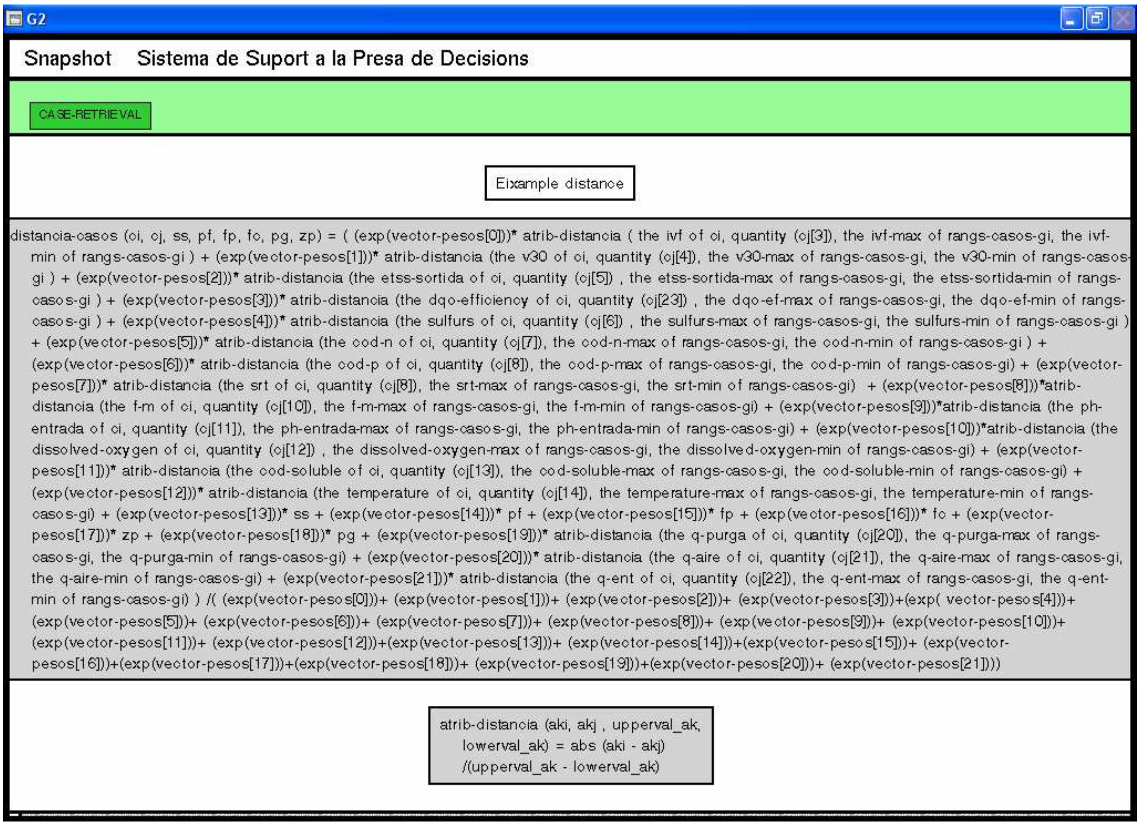
	A	B	C	D	E	K	L	Q	R	AA	AB	AZ	BF	BG	
1	EPISODE-ID.	CASE-ID.	TEMP.-ID	3.SVI	4.V30	9.SRT	10.F/M	15.Slud.s	16.Pred.fil.	CASE-DIAGNOSIS	Baixa F/M	AP Foaming	CONTROL	COMMENTS	LESSON
2	BF-1	12/01/2004	1	196	396	-9999	1.3	1	tipus 021N	BULKING-FILAMENTOS	FALSE	FALSE		aumenten filaments	
3	BF-1	13/01/2004	2	214	425	-9999	1.4	1	tipus 021N	BULKING-FILAMENTOS	FALSE	FALSE		parada de 2 hores	
4	BF-1	14/01/2004	3	240	470	-9999	1.4	1	tipus 021N	BULKING-FILAMENTOS	FALSE	FALSE		empitjora rapidament	
5	BF-1	15/01/2004	4	237	470	-9999	1.4	1	tipus 021N	BULKING-FILAMENTOS	FALSE	FALSE			
6	BF-1	16/01/2004	5	303	610	-9999	1.5	5	tipus 021N	BULKING-FILAMENTOS	FALSE	FALSE		moc selectiu i empitjora	
7	BF-1	17/01/2004	6	366	655	-9999	1.0	-9999		BULKING-FILAMENTOS	FALSE	FALSE			
8	BF-1	18/01/2004	7	260	535	-9999	1.2	-9999		BULKING-FILAMENTOS	FALSE	FALSE			
9	BF-1	19/01/2004	8	337	600	-9999	1.6	1	tipus 021N	BULKING-FILAMENTOS	FALSE	FALSE		menys vida	
10	BF-1	20/01/2004	9	341	630	-9999	1.2	1	tipus 021N	BULKING-FILAMENTOS	FALSE	FALSE	cloracio durant 24 hore	encara hi ha moc i les escumes s'han tornat més	
11	BF-1	21/01/2004	10	317	550	-9999	1.4	5	tipus 021N	BULKING-FILAMENTOS	FALSE	FALSE	aturada cloració	escumes marrons (zooglea?)	
12	BF-1	22/01/2004	11	276	390	-9999	1.9	5	tipus 021N	BULKING-FILAMENTOS	FALSE	FALSE	augment del set-point (	encara hi ha moc i les escumes s'han tornat més	
13	BF-1	23/01/2004	12	301	460	-9999	2.2	1	tipus 021N	BULKING-FILAMENTOS	FALSE	FALSE	manteniment del set-point d'oxigen a 3 ppm durant 12 hores		
14	BF-1	24/01/2004	13	252	405	-9999	1.7	-9999		BULKING-FILAMENTOS	FALSE	FALSE			
15	BF-1	25/01/2004	14	214	290	-9999	1.7	-9999		BULKING-FILAMENTOS	FALSE	FALSE			
16	BF-1	26/01/2004	15	245	310	-9999	2.2	5	tipus 021N	BULKING-FILAMENTOS	FALSE	FALSE		encara hi ha moc i augmenten les escumes blanc	
17	BF-1	27/01/2004	16	299	380	-9999	2.3	5	tipus 021N	BULKING-FILAMENTOS	FALSE	FALSE		fang d'exces en ascencio. Va ploure. Les escume	
18	BF-1	28/01/2004	17	205	260	-9999	2.2	5	tipus 021N	BULKING-FILAMENTOS	FALSE	FALSE		fang d'exces en ascencio. Poques escumes blar	
19	BF-1	29/01/2004	18	213	296	-9999	2.5	5	tipus 021N	BULKING-FILAMENTOS	FALSE	FALSE		ascenció del fang en excés i augment de la dema	
20	BF-1	30/01/2004	19	287	400	-9999	1.7	5	tipus 021N	BULKING-FILAMENTOS	FALSE	FALSE		les escumes marrons continuen disminuint perc	
21	FB-1	31/01/2004	1	394	420	-9999	1.7	4	tipus 021N	FOAMING-BIOLOGIC	G2	G2			
22	BF-1	31/01/2004	20	394	420	-9999	1.7	4	tipus 021N	BULKING-FILAMENTOS	G2	TRUE			
23	BNF-1	01/02/2004	1	182	435	-9999	1.6	-9999		BULKING-NO-FILAMENTOS	G2	G2			
24	FB-1	02/02/2004	2	243	320	-9999	1.7	2	tipus 021N	FOAMING-BIOLOGIC	FALSE	G2		escumes blanques	
25	BF-1	03/02/2004	21	218	210	-9999	2.2	2	tipus 021N	BULKING-FILAMENTOS	FALSE	G2	adicio d'antiescumant .	augmenten les escumes blanc-marro	
26	BF-1	04/02/2004	22	180	220	-9999	2.1	1	tipus 021N	BULKING-FILAMENTOS	FALSE	G2	es continua adicionant	ascensió dels fangs en exces pero tot i aixi es de	
27	BF-1	05/02/2004	23	209	288	-9999	1.0	1	tipus 021N	BULKING-FILAMENTOS	FALSE	G2		escumes blanques	
28	BNF-1	05/02/2004	2	209	288	-9999	2.0	1	tipus 021N	BULKING-NO-FILAMENTOS	G2	G2		escumes blanques	
29	BNF-1	06/02/2004	3	227	256	-9999	2.3	-9999		BULKING-NO-FILAMENTOS	G2	G2		desaparicio del moc i les escumes s'estabilitzen.	
30	BF-1	07/02/2004	24	239	330	-9999	1.5	-9999		BULKING-FILAMENTOS	FALSE	G2			
31	BF-1	08/02/2004	25	275	350	-9999	1.8	-9999		BULKING-FILAMENTOS	FALSE	G2			
32	BF-1	09/02/2004	26	305	380	-9999	2.0	2	tipus 021N	BULKING-FILAMENTOS	FALSE	G2			
33	BF-1	10/02/2004	27	232	240	-9999	2.9	2	tipus 021N	BULKING-FILAMENTOS	FALSE	G2		presencia de moc selectiu. No escumes	
34	BF-1	11/02/2004	28	244	270	-9999	2.2	2	tipus 021N	BULKING-FILAMENTOS	FALSE	G2		presencia de moc selectiu	
35	BF-1	12/02/2004	29	245	280	-9999	2.7	2	tipus 021N	BULKING-FILAMENTOS	FALSE	G2		augment del moc i escumes marrons	
36	BF-1	13/02/2004	30	181	200	-9999	2.4	-9999		BULKING-FILAMENTOS	FALSE	G2		moc augmentant i escumes normals.buidatge de	
37	BF-1	14/02/2004	31	284	260	-9999	1.8	-9999		BULKING-FILAMENTOS	FALSE	G2			
38	BF-1	15/02/2004	32	214	210	-9999	2.5	2	tipus 021N	BULKING-FILAMENTOS	FALSE	G2			
39	BF-1	16/02/2004	33	240	270	-9999	2.0	1	tipus 021N	BULKING-FILAMENTOS	FALSE	G2		NO MOC	
40	BF-1	16/02/2004	34	240	270	-9999	2.0	1	tipus 021N	BULKING-FILAMENTOS	FALSE	G2		NO MOC	
41	BF-1	17/02/2004	35	304	370	-9999	2.0	1	tipus 021N	BULKING-FILAMENTOS	FALSE	G2		molts filaments	
42	BF-1	18/02/2004	36	253	290	-9999	2.2	1	tipus 021N	BULKING-FILAMENTOS	FALSE	G2		MILLORA TOT	
43	BNF-2	18/02/2004	1	253	290	-9999	2.2	1	tipus 021N	BULKING-NO-FILAMENTOS	FALSE	G2		MILLORA TOT	
44	BNF-2	19/02/2004	2	234	252	-9999	2.7	1	tipus 021N	BULKING-NO-FILAMENTOS	G2	G2		MILLORA TOT	
45	BF-1	20/02/2004	37	216	250	-9999	2.1	-9999		BULKING-FILAMENTOS	FALSE	G2			
46	BF-1	21/02/2004	38	229	310	-9999	0.8	-9999		BULKING-FILAMENTOS	FALSE	G2			

Figure 7.16 Part of the initial case-library of the Girona WWTP including some of the attributes which define a case: episode-id.; case-id.; temporal-id.; case-situation-description (just 6 of the 22 variables are showed); case-problem-diagnosis; case-cause-diagnosis (just one of the possible causes is shown); case-parallel-problems (just one of the possible problems is shown); case-control-strategy; case-comments, and episode-lesson.

The long-term success of the episode library demands that it continuously evolves by adding new episodes, refining the existing ones, and forgetting the useless ones. As the library is updated with new episodes and knowledge about the process grows, the EBRS evolves into a better reasoner system, and system accuracy improves. From the beginning of the EBRS implementation and during the seven-month field validation phase, the episode library has increased with new episodes. As a result of the system implementation, the registration of the different attributes which define the new cases and episodes was more regular and hence, the new episodes incorporated more accurate information. Thus, as new, informationally richer experiences were being incorporated into the episode library, those cases from the initial episode library with more than 10 missing attributes (especially those which did not incorporate the key attributes of the control strategy or the lesson) were automatically deleted.

The exploration of the episode library is performed by means of the inference of 2 specific procedures, open-file-for-read-casos() and particio-casos(), which transform each of the stored cases into vectors, the type of structure necessary to calculate the distance between the current situation and the historical experiences of the episode library. Once all the historical cases are converted into vectors and stored in their corresponding object (i.e. casos-historics), the following step is to delimit the number of cases that need to be explored in order to speed up the searching process. Thus, the last subtask of the particio-casos() procedure consists of a selection of those episodes and cases that are going to be explored. This selection is done according to the case-diagnosis of the current situation. Thus, only those historical cases with the same case-diagnosis as the current situation are selected to be explored (i.e. if the case-diagnosis of the current case is filamentous bulking, only those historical cases with the case-diagnosis attribute= filamentous bulking are selected). The candidate cases are stored in another object (casos-acceptats).

3. <u>Distance calculation</u>. The *Eixample* distance, used to calculate the distances between the current situation and each of the candidate historical cases, was implemented by means of a function-definition. The calculated distances are stored in a specific vector (distances-array). Figure 7.17 shows its implementation within the G2 shell.



```
distancia-casos (ci, cj, ss, pf, fp, fc, pg, zp) = ( (exp(vector-pesos[0]))* atrib-distancia ( the ivf of ci, quantity (cj[3]), the ivf-max of rangs-casos-gi, the ivf-
   min of rangs-casos-gi ) + (exp(vector-pesos[1]))* atrib-distancia (the v30 of ci, quantity (cj[4]), the v30-max of rangs-casos-gi, the v30-min of rangs-casos-
   gi ) + (exp(vector-pesos[2]))* atrib-distancia (the etss-sortida of ci, quantity (cj[5]) , the etss-sortida-max of rangs-casos-gi, the etss-sortida-min of rangs-
   casos-gi ) + (exp(vector-pesos[3]))* atrib-distancia (the dqo-efficiency of ci, quantity (cj[23]) , the dqo-ef-max of rangs-casos-gi, the dqo-ef-min of rangs-
   casos-gi ) + (exp(vector-pesos[4]))* atrib-distancia (the sulfurs of ci, quantity (cj[6]) , the sulfurs-max of rangs-casos-gi, the sulfurs-min of rangs-casos-gi )
   + (exp(vector-pesos[5]))* atrib-distancia (the cod-n of ci, quantity (cj[7]), the cod-n-max of rangs-casos-gi, the cod-n-min of rangs-casos-gi ) +
   (exp(vector-pesos[6]))* atrib-distancia (the cod-p of ci, quantity (cj[8]), the cod-p-max of rangs-casos-gi, the cod-p-min of rangs-casos-gi) + (exp(vector-
   pesos[7]))* atrib-distancia (the srt of ci, quantity (cj[8]), the srt-max of rangs-casos-gi, the srt-min of rangs-casos-gi)  + (exp(vector-pesos[8]))*atrib-
   distancia (the f-m of ci, quantity (cj[10]), the f-m-max of rangs-casos-gi, the f-m-min of rangs-casos-gi) + (exp(vector-pesos[9]))*atrib-distancia (the ph-
   entrada of ci, quantity (cj[11]), the ph-entrada-max of rangs-casos-gi, the ph-entrada-min of rangs-casos-gi) + (exp(vector-pesos[10]))*atrib-distancia (the
   dissolved-oxygen of ci, quantity (cj[12]) , the dissolved-oxygen-max of rangs-casos-gi, the dissolved-oxygen-min of rangs-casos-gi) + (exp(vector-
   pesos[11]))* atrib-distancia (the cod-soluble of ci, quantity (cj[13]), the cod-soluble-max of rangs-casos-gi, the cod-soluble-min of rangs-casos-gi) +
   (exp(vector-pesos[12]))* atrib-distancia (the temperature of ci, quantity (cj[14]), the temperature-max of rangs-casos-gi, the temperature-min of rangs-
   casos-gi) + (exp(vector-pesos[13]))* ss + (exp(vector-pesos[14]))* pf + (exp(vector-pesos[15]))* fp + (exp(vector-pesos[16]))* fc + (exp(vector-
   pesos[17]))* zp + (exp(vector-pesos[18]))* pg + (exp(vector-pesos[19]))* atrib-distancia (the q-purga of ci, quantity (cj[20]), the q-purga-max of rangs-
   casos-gi, the q-purga-min of rangs-casos-gi) + (exp(vector-pesos[20]))* atrib-distancia (the q-aire of ci, quantity (cj[21]), the q-aire-max of rangs-casos-gi,
   the q-aire-min of rangs-casos-gi) + (exp(vector-pesos[21]))* atrib-distancia (the q-ent of ci, quantity (cj[22]), the q-ent-max of rangs-casos-gi, the q-ent-
   min of rangs-casos-gi) ) /( (exp(vector-pesos[0]))+ (exp(vector-pesos[1]))+ (exp(vector-pesos[2]))+ (exp(vector-pesos[3]))+(exp( vector-pesos[4]))+
   (exp(vector-pesos[5]))+ (exp(vector-pesos[6]))+ (exp(vector-pesos[7]))+ (exp(vector-pesos[8]))+ (exp(vector-pesos[9]))+ (exp(vector-pesos[10]))+
   (exp(vector-pesos[11]))+ (exp(vector-pesos[12]))+(exp(vector-pesos[13]))+ (exp(vector-pesos[14]))+(exp(vector-pesos[15]))+ (exp(vector-
   pesos[16]))+(exp(vector-pesos[17]))+(exp(vector-pesos[18]))+ (exp(vector-pesos[19]))+(exp(vector-pesos[20]))+ (exp(vector-pesos[21])))
```

```
atrib-distancia (aki, akj , upperval_ak,
   lowerval_ak) = abs (aki - akj)
   /(upperval_ak - lowerval_ak)
```

Figure 7.17 Implementation of the Eixample distance within the G2 shell

4. Display of the most similar case. The final step of the case-retrieval module was implemented in the mostra-cas-mes-semblant() procedure, which selects the lowest distance value from the distances-array vector and retrieves the corresponding case from the candidate historical cases vector. This final conclusion is then transferred both to the decision support level of the KBDSS architecture in order to show the final user the most similar case, and to the episode-retrieval module.

Consequently, during the execution of the **episode-retrieval module**, the complete episode, consisting of those historical cases belonging to the same episode as the retrieved most similar case are also transferred to the upper level of decision support and communicated to the final user.

Finally, the modules of review and retention of the new cases have been also implemented within the G2 shell. They are inferred once the results of both the dynamic ES and CBRS are communicated to the final user.

As soon as the **review module** is started, the plant operator is presented with a specific interface where s/he is asked to introduce the last attributes necessary to complete the current-case definition: the control strategy applied (:case-control-strategy), and the corresponding comments (:case-comments). Likewise, whenever the dynamic reasoning process identifies that an episode is concluded, a specific interface is provided in order to register the final lesson of the episode. Once all the required attributes are registered, the new case (or episode) is ready to be retained into the episode library.

The **retention module** consists of a set of procedures which transform the set of attributes into a vector (current-case) and automatically include the current case in the episode library.

7.3.3 Decision support

The uppermost level of the system architecture is where the dynamic KBDSS exerts its decision support function to the user. At this level, the user receives the conclusions from the reasoning level of both the dynamic ES and the EBRS. The final diagnosis (regarding the problem, causes and trends) together with the suggested control strategy (based on expert or experiential knowledge), are sent to the user through a user-friendly computer interface, who will finally decide on the control strategy to be taken (Figure 7.3).

A menu-based interface has been developed in the G2 environment to provide a simple and feasible way for users to run the Decision Support System and to know the state and evolution of the process over time at any time. In this sense, the conclusions of the system can be displayed at any moment until a new supervisory cycle is started. The dynamic system can also be asked for explanations for the conclusions reached and deductive processes followed, or requested to retrieve certain quantitative or qualitative values. Moreover, several trend charts, graphics or readouts of the most important values facilitates the understanding of the large amounts of information generated by the process. Finally, a dynamic interaction between the system and the user is also established, the user, at any time, can modify some key parameters such as the considered fuzzy sets, the limits of effectiveness, etc.

In view of the current or predictive results, the final user evaluates the conclusions and suggestions, checks their validity, and finally decides which kind of control should be carried out.

Regarding the system interface, some considerations have been taken into account in order to facilitate an efficient communication between the system and the final user. According to Walker *et al.* (2001), an effective human-machine interface should reduce information stress and quickly provide the information required to locate and resolve the source of the abnormal situation.

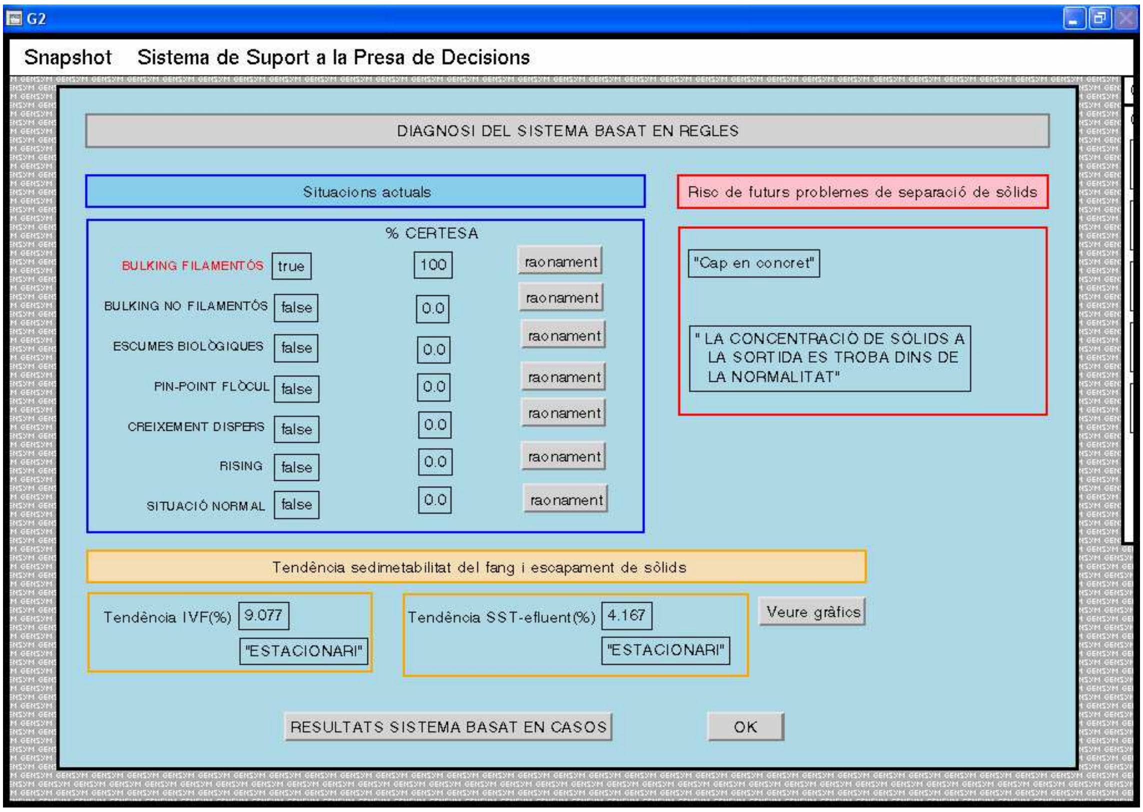
In the next section, a case-study of a real situation of filamentous bulking in the Girona WWTP is presented in order to show the developed user interface and to depict the communication between the lower levels of the architecture and the final user. All the interfaces that are showed are in Catalan language so a translation is provided whenever the content is not sufficiently understood.

7.3.3.1 Communication with the user

The KBDSS cycle begins as soon as the user introduces the daily information in the corresponding spreadsheets and decides to run the Knowledge-Based Decision Support System to obtain expert and experiential information that may help him/her to identify or predict possible solids separation problems and to decide how to act on the process. The system is initiated by clicking on the second button of the menu bar of the KBDSS, labelled "Sistema de Suport a la Presa de Decisions".

<u>Role of the dynamic Expert System</u>

The first interface showed to the user, Figure 7.18, summarizes the results of the Expert System.



DIAGNOSI DEL SISTEMA BASAT EN REGLES

Situacions actuals

		% CERTESA
BULKING FILAMENTÓS	true	100
BULKING NO FILAMENTÓS	false	0.0
ESCUMES BIOLÒGIQUES	false	0.0
PIN-POINT FLÒCUL	false	0.0
CREIXEMENT DISPERS	false	0.0
RISING	false	0.0
SITUACIÓ NORMAL	false	0.0

Risc de futurs problemes de separació de sòlids

"Cap en concret"

" LA CONCENTRACIÓ DE SÒLIDS A LA SORTIDA ES TROBA DINS DE LA NORMALITAT"

Tendència sedimetabilitat del fang i escapament de sòlids

Tendència IVF(%) 9.077 "ESTACIONARI"

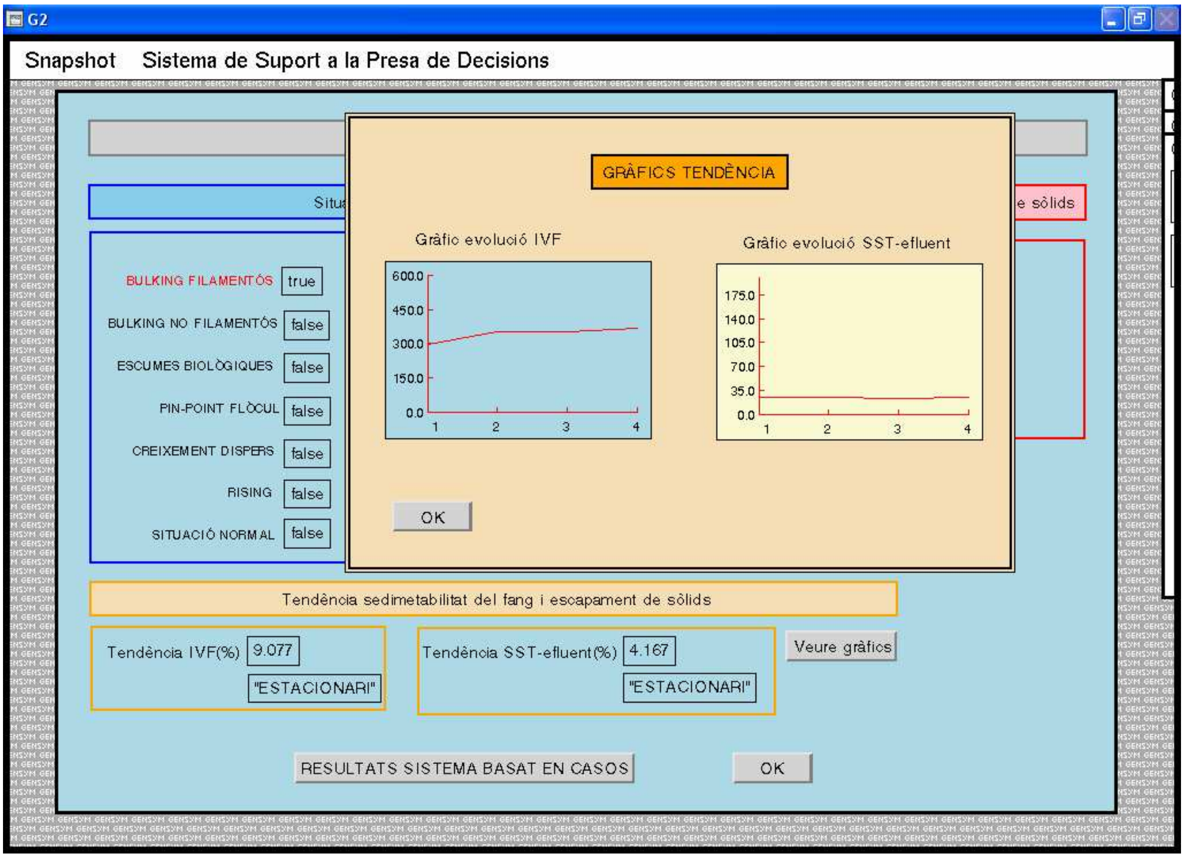
Tendència SST-efluent(%) 4.167 "ESTACIONARI"

Figure 7.18 G2 interface showing the results of the dynamic Expert System

The list on the left shows the results of the ES regarding the diagnosis of the current solids separation problem that the process is suffering from. It can be observed that, in this case, the system has concluded that the process is affected by a filamentous bulking problem and that the % of certainty of this conclusion is 100%.

The red box on the right summarizes the results of the dynamic ES regarding the trends or risk of future solids separation problems. In the present case-study, no risk of other solids separation problems (non-filamentous bulking, foaming, dispersed growth, pin-point floc or rising sludge) has been identified. The second message in this frame informs the plant operator that at the moment the process is not suffering from a loss of solids at the effluent.

Finally, the orange boxes at the bottom of the screen provide the user the results of the analysis of the trends in both SVI and effluent-TSS. The results are presented by means of quantitative and qualitative values. In the present example, both the SVI and the TSS have a qualitative value equal to “estacionari” (stationary). In order to study in depth these specific trends, the user is given the option to graphically represent the values of SVI and TSS used in these calculations. The graphs are obtained by clicking the button on the left, labelled “veure gràfics” (represent graphs). Figure 7.19 shows an example of these graphs.

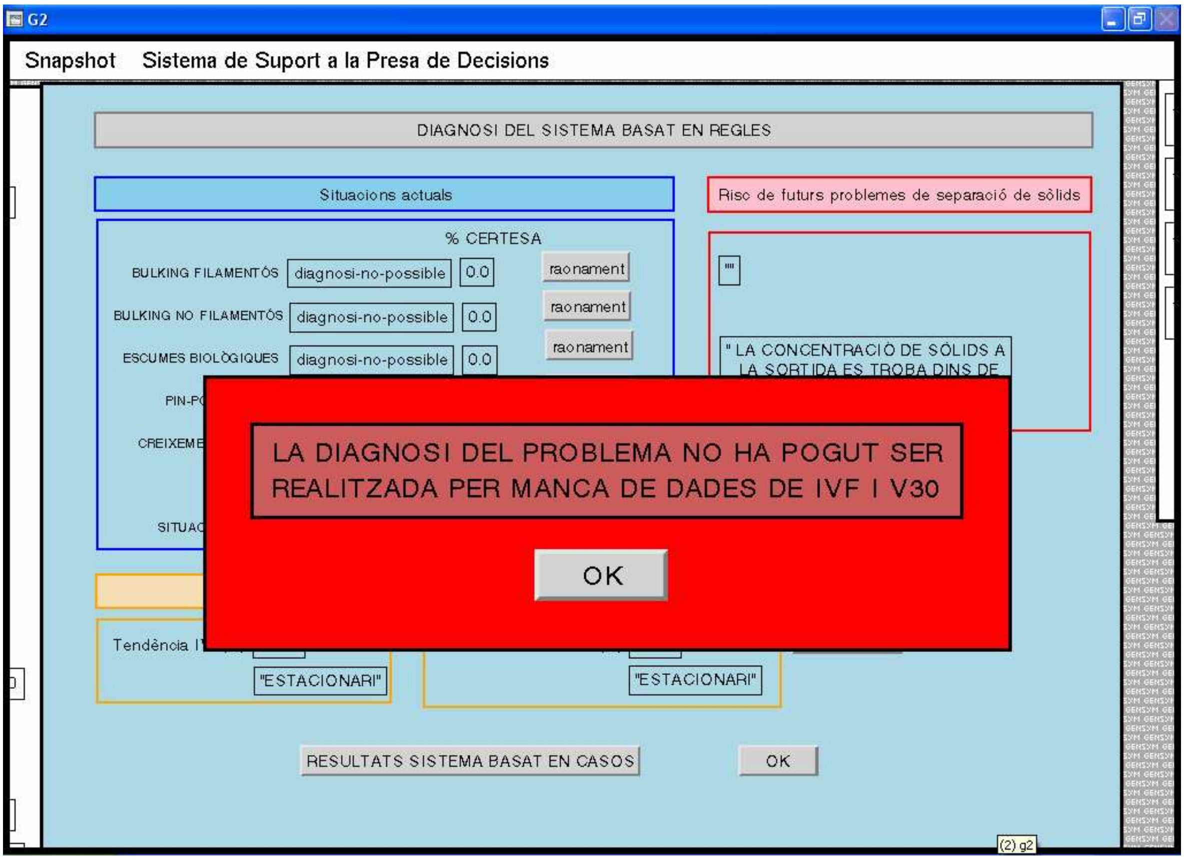


BULKING FILAMENTÓS true
BULKING NO FILAMENTÓS false
ESCUMES BIOLÒGIQUES false
PIN-POINT FLÒCUL false
CREIXEMENT DISPERS false
RISING false
SITUACIÓ NORMAL false

Tendència IVF(%) 9.077 "ESTACIONARI"
Tendència SST-efluent(%) 4.167 "ESTACIONARI"

Figure 7.19 G2 interface showing the graphs of SVI trend (left) and effluent TSS (right)

In case the system did not have at its disposal all the necessary information to reach a conclusion, e.g. SVI and V30 or none of the qualitative variables, the system would show the user an alarm message requesting the necessary data. Figure 7.20 shows an example of these alarm messages inferred by a lack of data. In the hypothetical example, filamentous bulking, non-filamentous bulking and foaming could not be identified due to a lack of SVI and V30 values.



BULKING FILAMENTÓS diagnosi-no-possible 0.0

BULKING NO FILAMENTÓS diagnosi-no-possible 0.0

ESCUMES BIOLÒGIQUES diagnosi-no-possible 0.0

PIN-P

CREIXEME

SITUAC

""

" LA CONCENTRACIÓ DE SÒLIDS A LA SORTIDA ES TROBA DINS DE

Tendència I

"ESTACIONARI"

"ESTACIONARI"

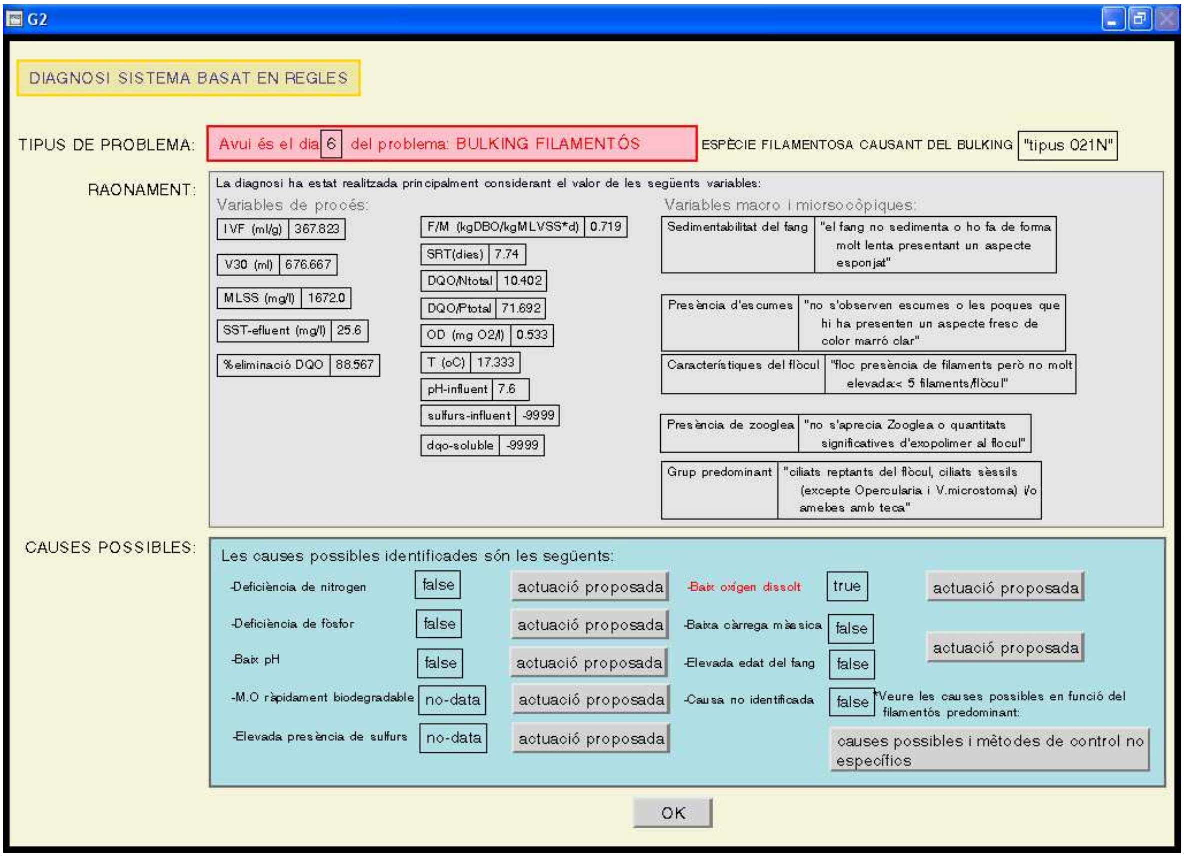
LA DIAGNOSI DEL PROBLEMA NO HA POGUT SER REALITZADA PER MANCA DE DADES DE IVF I V30

Figure 7.20 G2 interface showing an alarm message requesting the missing data necessary to reach a conclusion.

Once this first information is obtained, the user has the option of knowing the reasoning process followed by the ES to reach this first conclusion, i.e. to know the values of the most important variables that have been considered as information inputs for the decision trees. This action is performed by clicking the corresponding button located on the right of each diagnosed solids separation problem. For each type of problem, this button, labelled “raonament” (reasoning), is only active if a concluded diagnosis is true or possible.

In addition to the reasoning process, the user will receive the results of the dynamic ES regarding the possible cause or causes of the current situation (Figure 7.21). The red box on the top informs the operator about the number of days that the plant has been suffering from the same situation. In the present case-study, the red box on the top informs that “Today is the 6th day that the plant is suffering from a filamentous bulking problem”. Next, it also informs that the type of filamentous bacteria that is causing the present situation is type 021N.

The grey frame located in the middle of the screen and labelled “RAONAMENT” (reasoning), summarizes the set of variables that has been used by the ES to reach the diagnosis and detect the causes of the solids separation problem. The set of variables includes: SVI, V30, MLSS, effluent-TSS, %COD removal, F/M, SRT, COD/N, COD/P, DO, T, pH, sulphide and %COD soluble. A set of the most important qualitative variables is also showed, including: sludge settleability, foam presence, floc characteristics, presence of *Zoogloea* and predominant group. These categorical variables have been transformed into their qualitative values in order to provide the user with more understandable information.



TIPUS DE PROBLEMA: Avui és el dia 6 del problema: BULKING FILAMENTÓS

ESPÈCIE FILAMENTOSA CAUSANT DEL BULKING "tipus 021N"

RAONAMENT: La diagnosi ha estat realitzada principalment considerant el valor de les següents variables:

Variables de procés:

IVF (ml/g) 367.823
V30 (ml) 676.667
MLSS (mg/l) 1672.0
SST-efluent (mg/l) 25.6
%eliminació DQO 88.567
F/M (kgDBO/kgMLVSS*d) 0.719
SRT(dies) 7.74
DQO/Ntotal 10.402
DQO/Ptotal 71.692
OD (mg O2/l) 0.533
T (oC) 17.333
pH-influent 7.6
sulfurs-influent -9999
dqo-soluble -9999

Variables macro i micrsocòpiques:

Sedimentabilitat del fang	"el fang no sedimenta o ho fa de forma molt lenta presentant un aspecte esponjat"
Presència d'escumes	"no s'observen escumes o les poques que hi ha presenten un aspecte fresc de color marró clar"
Característiques del flòcul	"floc presència de filaments però no molt elevada:< 5 filaments/flòcul"
Presència de zooglea	"no s'aprecia Zooglea o quantitats significatives d'exopolimer al flocul"
Grup predominant	"ciliats reptants del flòcul, ciliats sèssils (excepte Opercularia i V.microstoma) i/o amebes amb teca"

CAUSES POSSIBLES: Les causes possibles identificades són les següents:

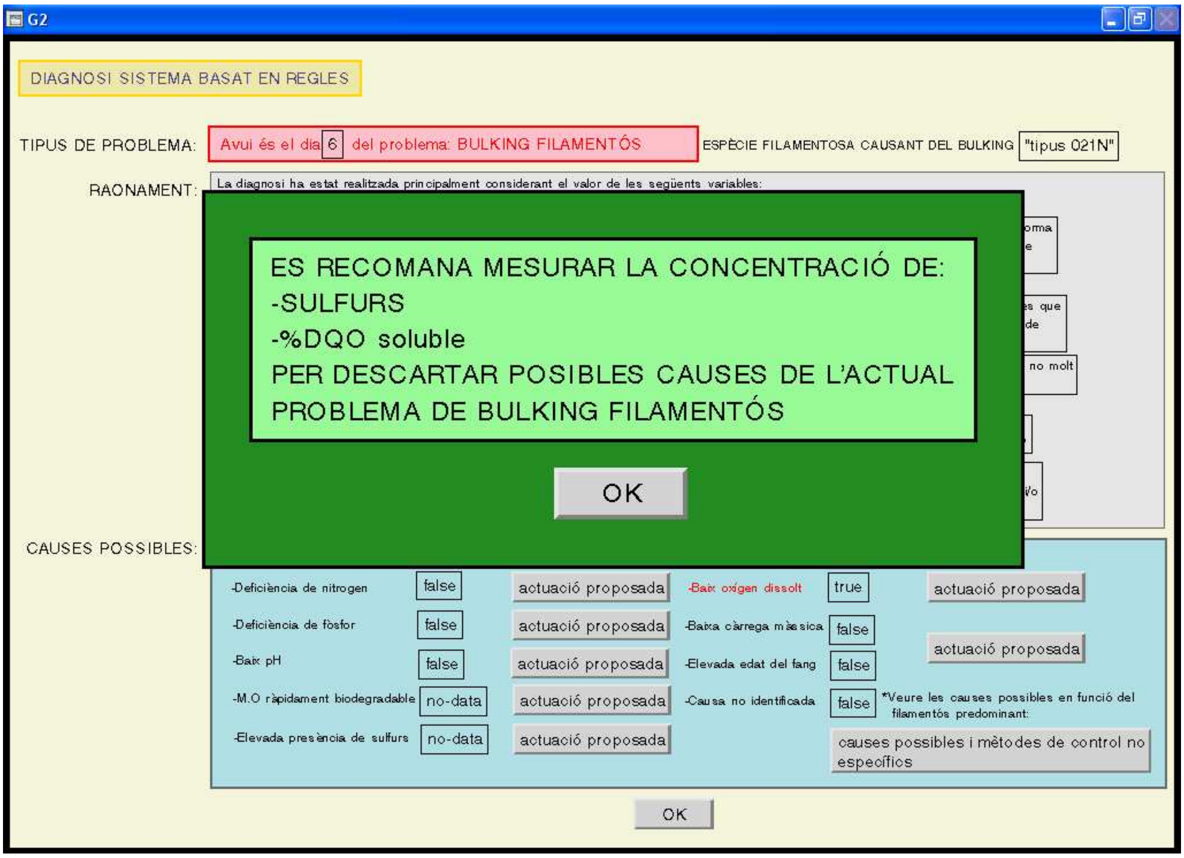
-Deficiència de nitrogen false
-Deficiència de fòsfor false
-Baix pH false
-M.O ràpidament biodegradable no-data
-Elevada presència de sulfurs no-data
-Baix oxigen dissolt true
-Baixa càrrega màssica false
-Elevada edat del fang false
-Causa no identificada false

*Veure les causes possibles en funció del filamentós predominant:

Figure 7.21 G2 interface showing the results of the reasoning process concerning the diagnosis of the possible causes and a compilation of the most relevant variables used during the reasoning process

Finally, the blue frame at the bottom of the screen shows the results of the ES regarding the possible causes of the identified situation. In our example, since the identified problem is filamentous bulking, the possible causes considered by the system are: nitrogen deficiency, phosphorous deficiency, low pH, readily biodegradable substrate, low DO, low F/M, or high SRT. In the present case-study, low DO has been identified as one of the possible causes of the filamentous bulking problem. On the other hand, readily biodegradable substrate and high sulphide have been identified only as potential causes because the data available was insufficient to reach a conclusion. These potential causes are marked with the symbol “no-data”. The rest of causes have also been checked and identified as not possible causes, “false”. This means that their values are within their “normal” ranges, where they usually do not induce filamentous bulking problems. For each of the detected causes, the user can request a control strategy to solve the current situation that specifically addresses the identified cause. This is done by clicking on the button next to each identified cause and labelled “actuació proposada” (suggested action).

Figures 7.22 and 7.23 show the results obtained after requesting a control strategy. First of all, a warning message (Figure 7.22) is showed in order to advise the operator that, in order to perform an exhaustive diagnosis about the possible causes of the present situation, it is suggested to measure the concentration of sulphide and the % of soluble COD.



DIAGNOSI SISTEMA BASAT EN REGLES

TIPUS DE PROBLEMA: Avui és el dia 6 del problema: BULKING FILAMENTÓS

ESPÈCIE FILAMENTOSA CAUSANT DEL BULKING "tipus 021N"

RAONAMENT: La diagnosi ha estat realitzada principalment considerant el valor de les següents variables:

ES RECOMANA MESURAR LA CONCENTRACIÓ DE:
-SULFURS
-%DQO soluble
PER DESCARTAR POSIBLES CAUSES DE L'ACTUAL PROBLEMA DE BULKING FILAMENTÓS

CAUSES POSSIBLES:

-Deficiència de nitrogen false
-Deficiència de fòsfor false
-Baix pH false
-M.O ràpidament biodegradable no-data
-Elevada presència de sulfurs no-data
-Baix oxigen dissolt true
-Baixa càrrega màssica false
-Elevada edat del fang false
-Causa no identificada false

*Veure les causes possibles en funció del filamentós predominant:

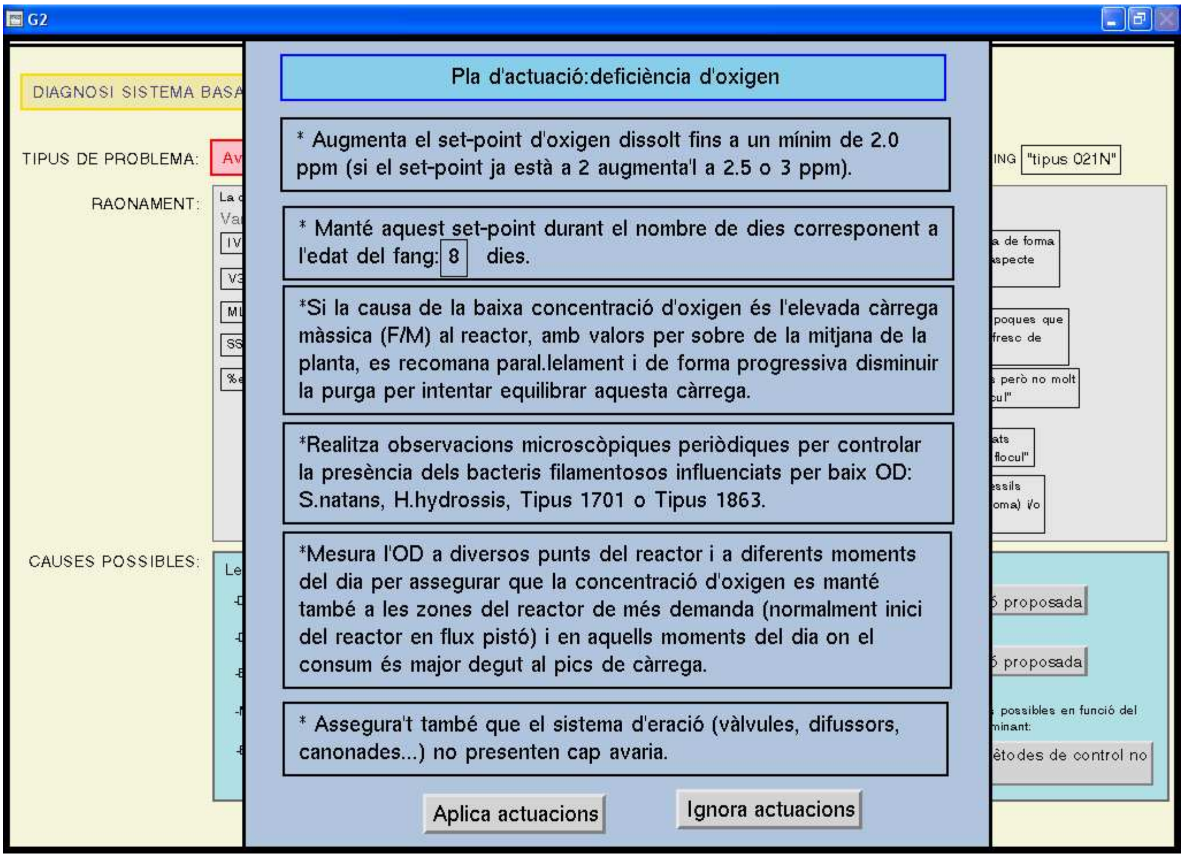
causes possibles i mètodes de control no específics

Figure 7.22 G2 interface showing a warning message informing the operator that the sulphide concentration and the % of soluble COD are required to perform a complete diagnosis about the possible causes of the present situation

After that, the complete control strategy related to the identified cause, in our example low DO, is obtained. In the present case-study the cause of low DO had already been identified several days ago, but the user had decided not to consider the action suggested by the ES and the DO set-point was not increased. Hence, the counter regarding this specific control strategy is=0 and the system suggests again a control strategy based on increasing the DO set-point. Figure 7.23 shows the control plan provided to the user.

The suggested control strategy includes the following set of actions:

- ⇒ Increase the dissolved oxygen set-point up to a minimum of 2.0 ppm.
- ⇒ Maintain this set-point during the same number of days as the sludge age = 8 days.
- ⇒ If the origin of this low DO concentration is the high F/M within the reactor (values over the standard values of the process), it is recommended to progressively decrease waste flow.
- ⇒ Perform a microscopic analysis to monitor the presence of the species influenced by low DO concentration: *S. natans*, *H. hydrossis*, Type 1701 and Type 1863.
- ⇒ Measure the DO concentration at various locations throughout the reactor and at different moments of the day in order to ensure that the desired DO set-point is maintained at all the zones of the reactor and throughout those moments of the day when peak loads increases oxygen consumption.
- ⇒ Make sure that the aeration system (valves, diffusers, etc.) is working properly.



Pla d'actuació:deficiència d'oxigen

* Augmenta el set-point d'oxigen dissolt fins a un mínim de 2.0 ppm (si el set-point ja està a 2 augmenta'l a 2.5 o 3 ppm).

* Manté aquest set-point durant el nombre de dies corresponent a l'edat del fang: 8 dies.

*Si la causa de la baixa concentració d'oxigen és l'elevada càrrega màssica (F/M) al reactor, amb valors per sobre de la mitjana de la planta, es recomana paral.lelament i de forma progressiva disminuir la purga per intentar equilibrar aquesta càrrega.

*Realitza observacions microscòpiques periòdiques per controlar la presència dels bacteris filamentosos influenciats per baix OD: S.natans, H.hydrossis, Tipus 1701 o Tipus 1863.

*Mesura l'OD a diversos punts del reactor i a diferents moments del dia per assegurar que la concentració d'oxigen es manté també a les zones del reactor de més demanda (normalment inici del reactor en flux pistó) i en aquells moments del dia on el consum és major degut al pics de càrrega.

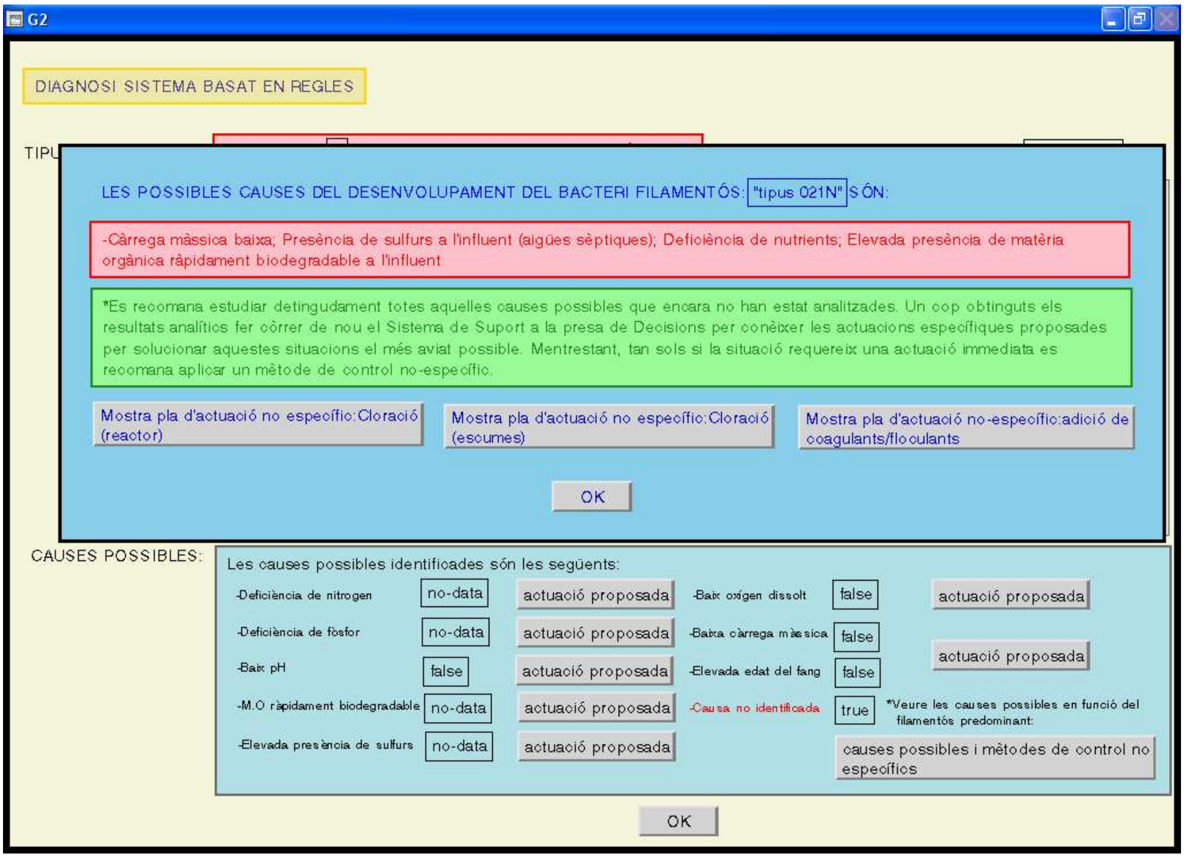
* Assegura't també que el sistema d'eració (vàlvules, difussors, canonades...) no presenten cap avaria.

Figure 7.23 G2 interface showing the control strategy related to a identified situation of filamentous bulking caused by low DO

Whenever the user decides to apply the suggested control strategy and the button of “Aplica actuacions” (apply actions) is clicked, the dynamic ES adds up 1 day to the specific counter of the low DO strategy. Conversely, if the user decides not to carry out the suggested strategy by clicking on the “Ignora actuacions” (ignore actions) button, the specific counter will remain equal to 0. On the following days, if the same cause is identified the same control strategy will be suggested up to day 8 (the established limit of efficiency). After day 8, a modified control strategy will be suggested (see control strategy decision trees in Annex).

When no specific cause is properly identified due to a lack of data or because possible causes have proven false, an alternative workspace (Figure 7.24) is showed to the user.

The alternative workspace first of all suggests a set of possible causes related to the identified filamentous bacteria. For example, if the identified filamentous bacteria were Type 021N, the system would suggest low F/M, high sulphide, nutrient deficiency or readily biodegradable substrate as possible causes (red frame). Accordingly, the green frame below suggests studying all those possible causes that had not been measured (once this is done, the system will restart the reasoning process). Meanwhile though, if the identified solids separation problem is quite severe, the system will recommend applying a non-specific control method. Three action buttons related to possible non-specific methods are showed: chlorination (mixed liquor), chlorination (foams) or addition of coagulants/flocculants. By clicking on one or another of these non-specific options, the user will receive a complete control plan to quickly restore the process.



LES POSSIBLES CAUSES DEL DESENVOLUPAMENT DEL BACTERI FILAMENTÓS: "tipus 021N" SÓN:

-Càrrega màssica baixa; Presència de sulfurs a l'influent (aigües sèptiques); Deficiència de nutrients; Elevada presència de matèria orgànica ràpidament biodegradable a l'influent

*Es recomana estudiar detingudament totes aquelles causes possibles que encara no han estat analitzades. Un cop obtinguts els resultats analítics fer còrrer de nou el Sistema de Suport a la presa de Decisions per conèixer les actuacions específiques proposades per solucionar aquestes situacions el més aviat possible. Mentrestant, tan sols si la situació requereix una actuació immediata es recomana aplicar un mètode de control no-específic.

Les causes possibles identificades són les següents:

-Deficiència de nitrogen no-data
-Deficiència de fòsfor no-data
-Baix pH false
-M.O ràpidament biodegradable no-data
-Elevada presència de sulfurs no-data
-Baix oxígen dissolt false
-Baixa càrrega màssica false
-Elevada edat del fang false
-Causa no identificada true

*Veure les causes possibles en funció del filamentós predominant:

Figure 7.24 G2 interface showing alternative control strategies (non-specific) to the user

Figure 7.25 shows the results of requesting the chlorination control plan. The complete control plant includes the following set of actions:

⇒ It is recommended to perform a respirometric test in order to check the effects of the selected chlorine dose on the biomass. Test different doses between 1 to 10 gCl_2/kg MLSS,d. If respirometric tests can not be performed, an initial dose of $2gCl_2$ /kg MLSS,d is suggested.

⇒ Calculation of the quantity of chlorine necessary per day (the required chlorine quantity is calculated when the user introduces the desired initial dose, e.g. 2.0, and then clicks on the button “càlcul” (calculation). The default units (kg/d) can be changed by the user to L/d or L/h.

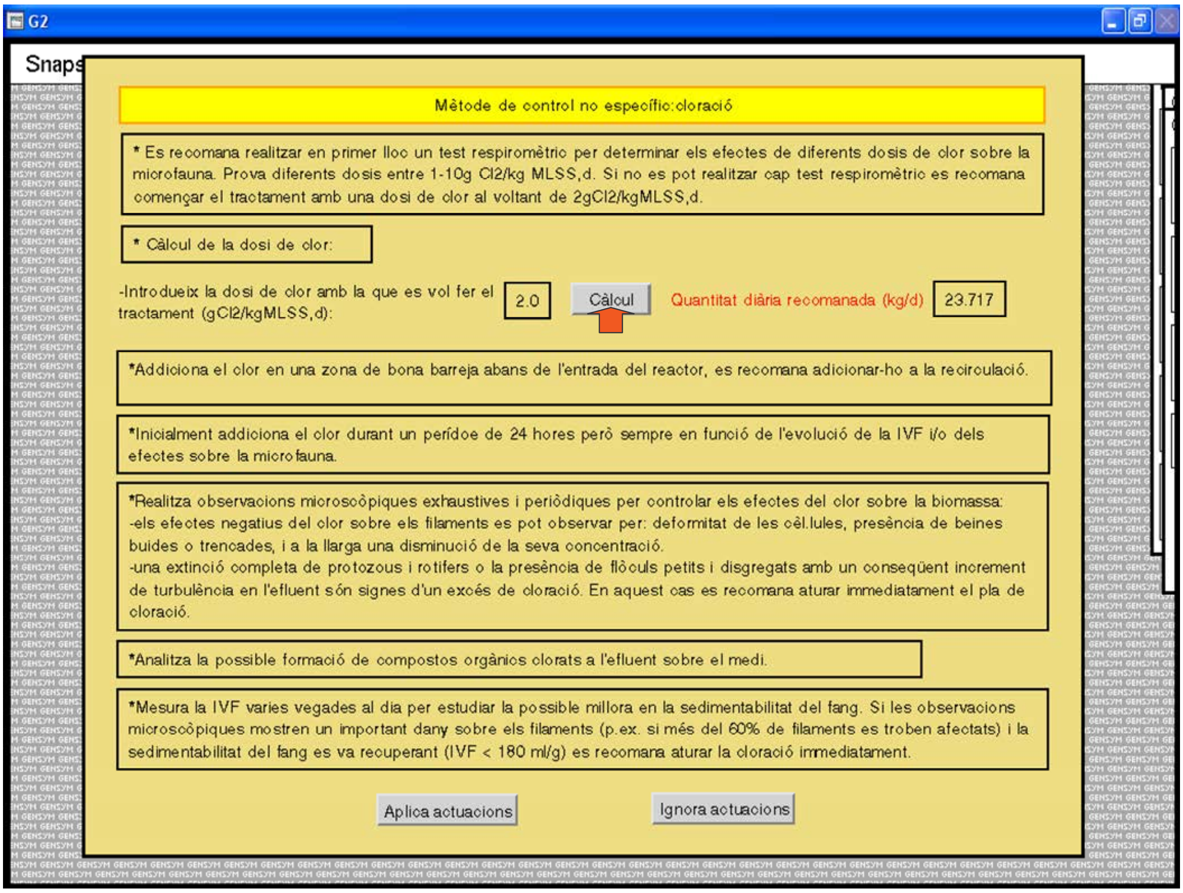
⇒ Dose chlorine solution on a well-mixed zone such as the RAS line.

⇒ Initially add chlorine during a period of 24 hours (depending always on the results of SVI and microscopic analysis).

⇒ Perform an exhaustive control on the responses of the biomass to chlorine addition: do monitor the effects on the filaments structure (cell’s deformity, empty and broken sheaths, etc.) and also on the rest of microorganisms. A complete extinction of protozoa and rotifers or a presence of small, broken-up flocs with a consequence on the effluent turbulence are signs of chlorine overdosing and, if observed, chlorination must be stopped immediately.

⇒ Monitor the formation of chlorine organic compounds such as chloramines in the effluent.

⇒ Measure SVI several times/day to check for any settleability improvement. If microscopic observations show filaments damage (if approx. 60-70% of filaments show observable damage) and settleability is recovered (SVI <180 mL/g) stop chlorination immediately.



Mètode de control no específic:cloració

* Es recomana realitzar en primer lloc un test respiromètric per determinar els efectes de diferents dosis de clor sobre la microfauna. Prova diferents dosis entre 1-10g Cl2/kg MLSS,d. Si no es pot realitzar cap test respiromètric es recomana començar el tractament amb una dosi de clor al voltant de 2gCl2/kgMLSS,d.

* Càlcul de la dosi de clor:

-Introdueix la dosi de clor amb la que es vol fer el tractament (gCl2/kgMLSS,d): 2.0

Quantitat diària recomanada (kg/d) 23.717

*Addiciona el clor en una zona de bona barreja abans de l'entrada del reactor, es recomana adicionar-ho a la recirculació.

*Inicialment addiciona el clor durant un perídoe de 24 hores però sempre en funció de l'evolució de la IVF i/o dels efectes sobre la microfauna.

*Realitza observacions microscòpiques exhaustives i periòdiques per controlar els efectes del clor sobre la biomassa:
-els efectes negatius del clor sobre els filaments es pot observar per: deformitat de les cèl.lules, presència de beines buides o trencades, i a la llarga una disminució de la seva concentració.
-una extinció completa de protozous i rotifers o la presència de flòculs petits i disgregats amb un conseqüent increment de turbulència en l'efluent són signes d'un excés de cloració. En aquest cas es recomana aturar immediatament el pla de cloració.

*Analitza la possible formació de compostos orgànics clorats a l'efluent sobre el medi.

*Mesura la IVF varies vegades al dia per estudiar la possible millora en la sedimentabilitat del fang. Si les observacions microscòpiques mostren un important dany sobre els filaments (p.ex. si més del 60% de filaments es troben afectats) i la sedimentabilitat del fang es va recuperant (IVF < 180 ml/g) es recomana aturar la cloració immediatament.

Figure 7.25 G2 interface showing a non-specific control strategy: chlorination

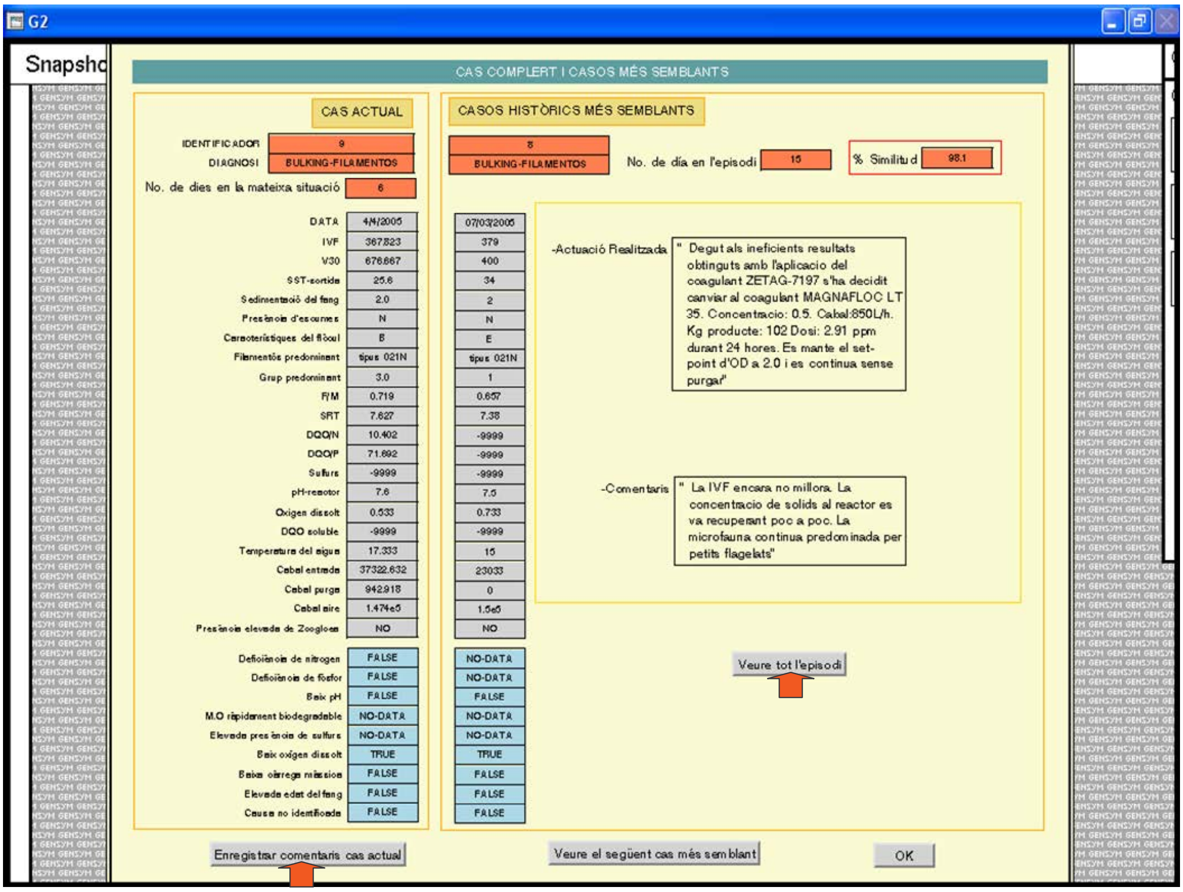
Role of the Episode-Based Reasoning System

Once the user has checked the results of the dynamic ES he/she has also the opportunity to complement these expert results with the knowledge derived from past experiences.

The EBR cycle is initiated by clicking on the action button labelled “Resultats Sistema Basat en Casos” (Results of the EBRS) that is showed once the first conclusions of the ES, regarding the problem and cause diagnosis, are reached (see Figure 7.18).

Figure 7.26 shows the first results obtained by the EBRS. The present workspace summarizes, by means of 2 display tables, a set of the most important attributes for both the current case “cas actual”, on the left, and the most similar case retrieved from the case-library, “casos històrics més semblants”, on the right. Regarding the retrieved case, all the attributes which characterize a historical case are showed. The red table on the top displays the episode-identifier, “identificador”, the case-problem-diagnosis, “diagnosi”, and on the left the temporal-identifier, “nº de dia en l’episodi”. The grey table below displays both the case-identifier as the date of the case, “data”, and the set of variables which describe the situation, case-situation-description. The blue table on the bottom lists the case-problem-causes and finally, two text displays show the user the case-control-strategy, “actuació realitzada”, and the case-comments, “comentaris”.

The red box on the left also displays the distance between the current case and the retrieved case, calculated by means of the Eixample distance algorithm, as the % of similarity, "% similitud".



	CAS ACTUAL	CASOS HISTÒRICS MÉS SEMBLANTS
IDENTIFICADOR	9	8
DIAGNOSI	BULKING-FILAMENTOS	BULKING-FILAMENTOS
No. de dies en la mateixa situació	6	
No. de día en l'episodi		15
% Similitud		98.1
DATA	4/4/2005	07/03/2005
IVF	367.823	379
V30	676.667	400
SST-sortida	25.6	34
Sedimentació del fang	2.0	2
Presència d'escumes	N	N
Característiques del flòcul	B	E
Filamentós predominant	tipus 021N	tipus 021N
Grup predominant	3.0	1
F/M	0.719	0.657
SRT	7.627	7.38
DQO/N	10.402	-9999
DQO/P	71.692	-9999
Sulfurs	-9999	-9999
pH-reactor	7.6	7.5
Oxigen dissolt	0.533	0.733
DQO soluble	-9999	-9999
Temperatura del aigua	17.333	15
Cabal entrada	37322.632	23033
Cabal purga	942.918	0
Cabal aire	1.474e5	1.5e5
Presència elevada de Zoogloea	NO	NO
Deficiència de nitrogen	FALSE	NO-DATA
Deficiència de fosfor	FALSE	NO-DATA
Baix pH	FALSE	FALSE
M.O ràpidament biodegradable	NO-DATA	NO-DATA
Elevada presència de sulfurs	NO-DATA	NO-DATA
Baix oxigen dissolt	TRUE	TRUE
Baixa càrrega màssica	FALSE	FALSE
Elevada edat del fang	FALSE	FALSE
Causa no identificada	FALSE	FALSE

-Actuació Realitzada: " Degut als ineficients resultats obtinguts amb l'aplicacio del coagulant ZETAG-7197 s'ha decidit canviar al coagulant MAGNAFLOC LT 35. Concentracio: 0.5. Cabal:850L/h. Kg producte: 102 Dosi: 2.91 ppm durant 24 hores. Es mante el set-point d'OD a 2.0 i es continua sense purgar"

-Comentaris: " La IVF encara no millora. La concentracio de solids al reactor es va recuperant poc a poc. La microfauna continua predominada per petits flagelats"

Figure 7.26 G2 interface showing the current case and the most similar case from the case library

In our case-study of filamentous bulking caused by low DO, the historical case 07/03/05 was retrieved as the most similar case of the current case-library. This is a case of filamentous bulking, specifically the day 15 of an episode of filamentous bulking caused by low DO. It can be seen that the variables describing the case are quite similar, in particular the most relevant variables: SVI, sludge settleability, predominant filament, and dissolved oxygen. Consequently, the % of similarity is quite high (98.1%). Under the assumption that similar cases require similar solutions, the knowledge stored for that historical case (control strategy applied and comments), should be useful in helping the plant operator to decide the control strategy to apply. Even if the applied control strategy of the historical case were not effective, the user could learn from failed solutions in order not to make the same mistakes.

In the current example, the retrieved control strategy includes the following information:

⇒ "Due to the inefficient results obtained with the application of the coagulant ZETAG-7197 it has been decided to change to the coagulant Magnafloc-LT35. The dosage applied is 850L/h at 0.5% and at a concentration of 2.91 mg/L. It has been applied for 24 hours. Due to a limitation with the aeration system and despite the low DO, the set-point is maintained at 2.0 ppm. The purge flow was =0 today."

The information compiled in the comments box is:

⇒ “The SVI does not show any sign of recovery. The MLSS is increasing due to the low WAS. The biomass though is still dominated by flagellates.”

In case the user considers that the retrieved case does not provide relevant information, he/she has the option of retrieving from the case-library the next most similar case (this procedure can be performed a maximum of 3 times). This process can be initiated by clicking on the second button at the bottom: “Veure el següent cas més semblant” (get the next most similar case).

Once these important pieces of knowledge have been presented to the user, he/she has also the opportunity of retrieving all the information stored in the episode that the retrieved case belongs to. Thus the user will have at his/her disposal the maximum quantity of useful information before making the final decision. To access all these information the user just has to click on the action button located below the case-control-strategy and comments labelled “Veure tot l’episodi” (Get the complete episode). Automatically, the user will receive a new interface (see Figure 7.27 for a summary of the most relevant information regarding the complete episode).

LA LLIÇÓ ENREGISTRADA DE L'EPISODI ÉS LA SEGÜENT:

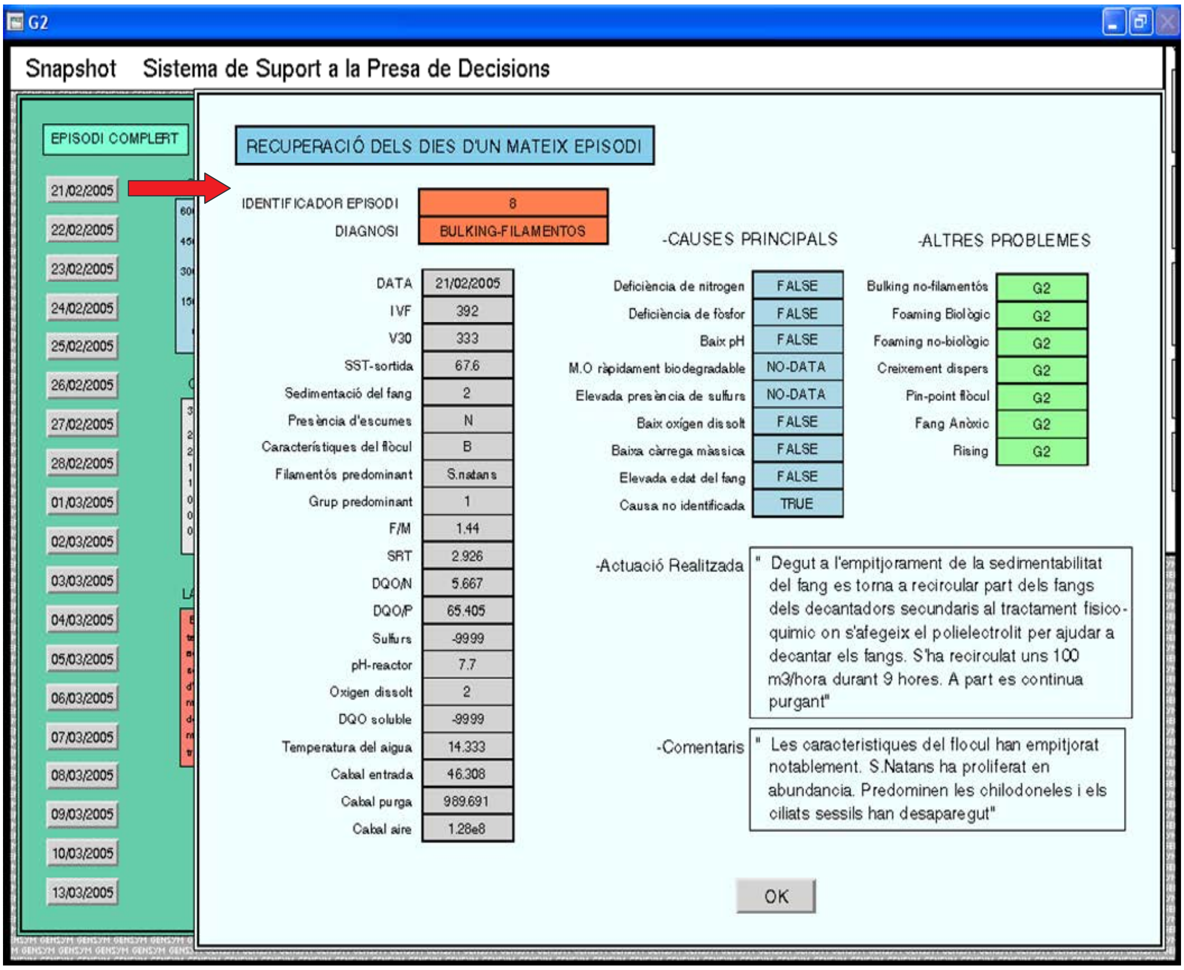
El problema del bulking filamentos no ha estat solucionat sino que ha passat a ser tambe un problema de foaming per un canvi en el filamentos predominant. Durant aquest episodi de bulking el filamentos predominant ha estat S.natans que ha proliferat segurament com a consequencia de 2 situacions principals. Per un banda el deficit d'oxigen que s'ha donat durant varis dies i per altra banda l'elevada concentracio de materia organica facilment biodegradable. Pel que fa als metodes de control sobre el deficit d'oxigen no hi ha molt mes a fer ja que el sistema d'aeracio ja esta treballant al maxim i pujar el set-point no seria viable. Per altra banda sembla que encara no s'ha trobat la dosi i/o tipus de floculant idoni per millorar la sedimentabilitat del fang

Figure 7.27 G2 interface showing a summary of the retrieved complete episode

The set of buttons on the left represent each one of the cases that belong to the retrieved episode. By clicking on any of these buttons, the user obtains all the relevant information corresponding to that particular case of the episode (Figure 7.28). All this information allows the user to study the evolution of the problem together with the specific set of daily actions carried out and the corresponding comments regarding the results of applying such control methods. Next to these buttons, a set of graphs representing the evolution of the most relevant variables (i.e., SVI, TSS, DO, F/M and MCRT) are shown. Finally, the red box at the bottom presents what is arguably the most important piece of knowledge about the episode: the lesson learnt.

In the present case study, regarding a filamentous bulking episode that lasted from 02/21/05 to 03/13/05, the lesson learnt was the following:

⇒ "The current filamentous bulking episode has not been properly solved but turned into a problem of biological foaming due to a change of the predominant filament to Type 1863. The low DO concentration present on most days, mainly caused by high F/M ratios, and the readily biodegradable substrate have made type 021N and *S. natans* the dominant filaments during the episode. Regarding the applied control methods, the DO concentration will never reach the required set-point until the aeration system is repaired, especially during high F/M events. For this reason the set-point has not been increased to 2.5 ppm. On the other hand it seems that none of the applied coagulants (ZETAG-7197 and Magnafloc-LT35) caused significant improvements in sludge settleability. Maybe the applied dosages were not adequate."



Snapshot Sistema de Suport a la Presa de Decisions

RECUPERACIÓ DELS DIES D'UN MATEIX EPISODI

IDENTIFICADOR EPISODI	8
DIAGNOSI	BULKING-FILAMENTOS
DATA	21/02/2005
IVF	392
V30	333
SST-sortida	67.6
Sedimentació del fang	2
Presència d'escumes	N
Característiques del flòcul	B
Filamentós predominant	S.natans
Grup predominant	1
F/M	1.44
SRT	2.926
DQO/N	5.667
DQO/P	65.405
Sulfurs	-9999
pH-reactor	7.7
Oxigen dissolt	2
DQO soluble	-9999
Temperatura del aigua	14.333
Cabal entrada	46.308
Cabal purga	989.691
Cabal aire	1.28e8

-CAUSES PRINCIPALS

Deficiència de nitrogen	FALSE
Deficiència de fòsfor	FALSE
Baix pH	FALSE
M.O ràpidament biodegradable	NO-DATA
Elevada presència de sulfurs	NO-DATA
Baix oxígen dissolt	FALSE
Baixa càrrega màssica	FALSE
Elevada edat del fang	FALSE
Causa no identificada	TRUE

-ALTRES PROBLEMES

Bulking no-filamentós	G2
Foaming Biològic	G2
Foaming no-biològic	G2
Creixement dispers	G2
Pin-point flòcul	G2
Fang Anòxic	G2
Rising	G2

-Actuació Realitzada: " Degut a l'empitjorament de la sedimentabilitat del fang es torna a recircular part dels fangs dels decantadors secundaris al tractament fisico-quimic on s'afegeix el polielectrolit per ajudar a decantar els fangs. S'ha recirculat uns 100 m3/hora durant 9 hores. A part es continua purgant"

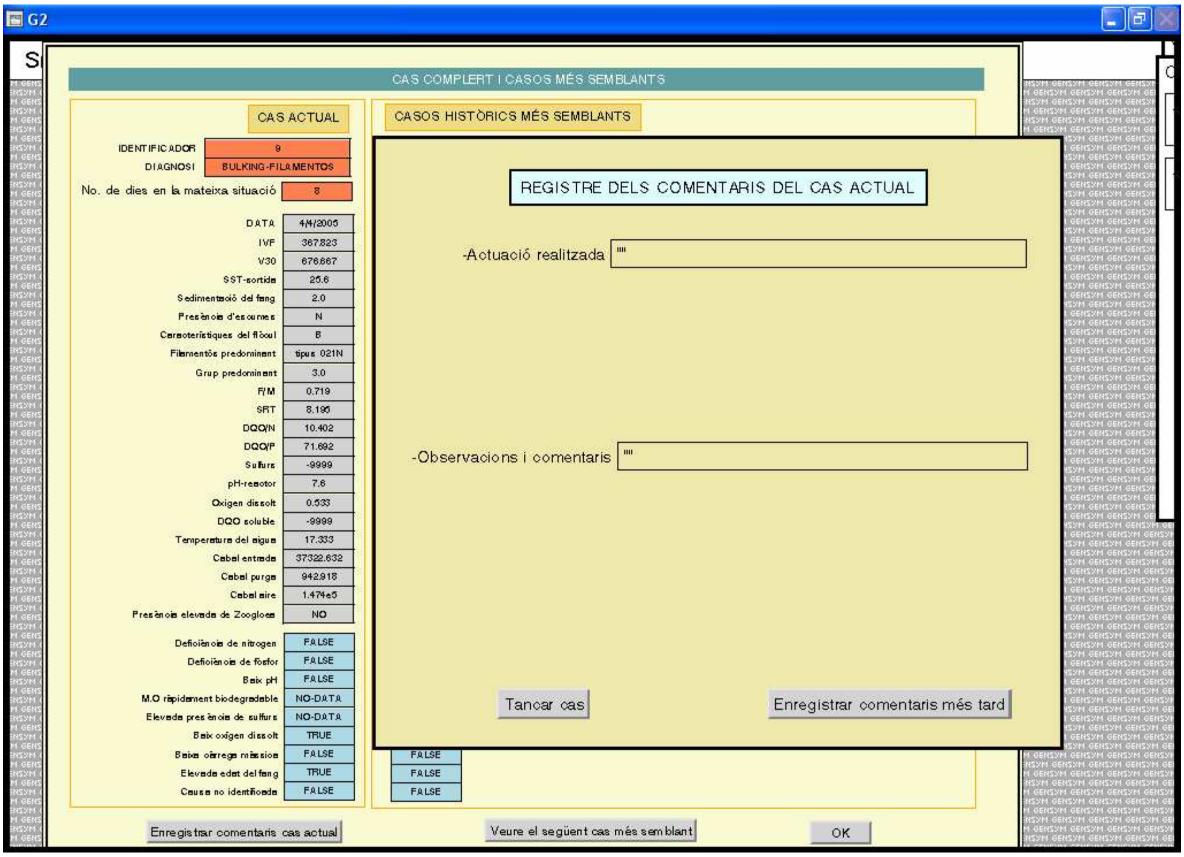
-Comentaris: " Les caracteristiques del floculs han empitjorat notablement. S.Natans ha proliferat en abundancia. Predominen les chilodoneles i els ciliats sessils han desaparegut"

Figure 7.28 G2 interface showing the most important attributes of each one of the cases that compose the retrieved episode

Registration of new cases and episodes

Once the user has received the complete results of both systems, it is then his/her turn to use all this provided knowledge to make a decision about which is the best control strategy to restore the process in the most efficient way. When this decision is made and a control method is applied then it is also time to partially validate the first results of the applied control actions. The registration of this valuable information into the currently initiated new case is a very important step to close the cycle of the dynamic CBRS and of the whole KBDSS.

A specific interface is provided to the user with the purpose of registering such information. By clicking on the first action button on the bottom-left of the first interface of the EBRS (Figure 7.26), "enregistrar comentaris cas actual" (register the comments of the current case), the user will receive an interface such as the one displayed in Figure 7.29.



IDENTIFICADOR: 9
DIAGNOSI: BULKING-FILAMENTOS
No. de dies en la mateixa situació: 8

DATA	4/4/2005
IVF	367.823
V30	676.667
SST-sortida	25.6
Sedimentació del fang	2.0
Presència d'escumes	N
Característiques del flòcul	B
Filamentós predominant	tipus 021N
Grup predominant	3.0
F/M	0.719
SRT	8.195
DQO/N	10.402
DQO/P	71.692
Sulfurs	-9999
pH-reactor	7.6
Oxigen dissolt	0.533
DQO soluble	-9999
Temperatura del aigua	17.333
Cabal entrada	37322.632
Cabal purga	942.918
Cabal aire	1.474e5
Presència elevada de Zoogloea	NO
Deficiència de nitrogen	FALSE
Deficiència de fosfor	FALSE
Baix pH	FALSE
M.O ràpidament biodegradable	NO-DATA
Elevada presència de sulfurs	NO-DATA
Baix oxigen dissolt	TRUE
Baixa càrrega màssica	FALSE
Elevada edat del fang	TRUE
Causa no identificada	FALSE

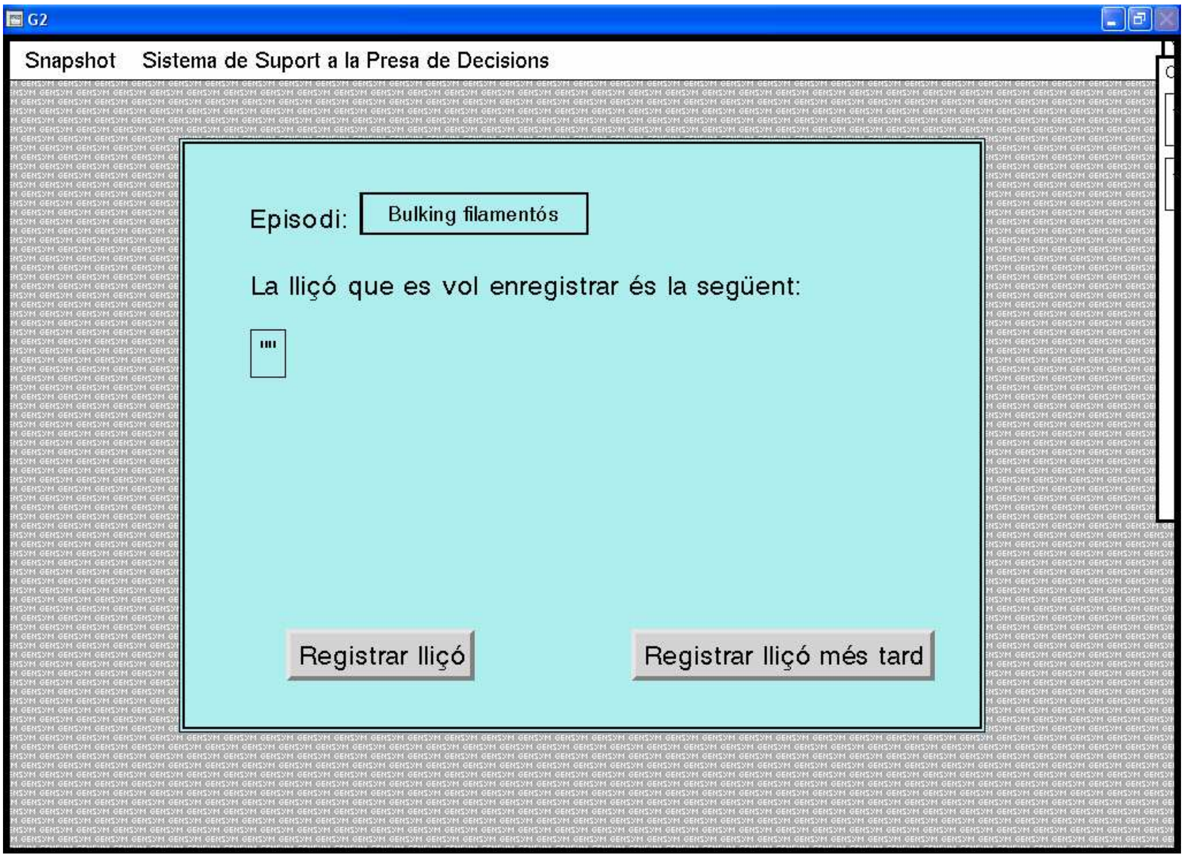
-Actuació realitzada: ""

-Observacions i comentaris: ""

Figure 7.29 G2 interface showing the specific workspace to register the applied control strategy and the related comments

Once the applied control strategy and the corresponding comments are registered, clicking on the action button labelled "Tancar cas" (close case), will cause the new case to be incorporated into the episode library together with all its related attributes.

When the dynamic ES identifies that an episode has been finished, the user is presented with a specific workspace to register the lesson of the episode (Figure 7.30). This lesson can be registered whenever the user decides to; however, as long as the lesson is not registered each time that the KBDSS is executed, the system will remind the user to register the lesson.



Episodi: Bulking filamentós

La lliçó que es vol enregistrar és la següent:

""

Figure 7.30 G2 interface showing the specific workspace to register the lesson of the episode

Clicking on the action button labelled "registrar llicó" (register lesson) the episode is closed and saved as a new experience in the episode library.

7.4 Discussion

The Girona WWTP was an appropriate candidate facility to implement the developed KBDSS. The occurrence of frequent solids separation problems, mainly filamentous bulking, non-filamentous bulking and biological foaming, the complex identification of the possible causes, and the lack of registration of historical experiences (especially the applied control strategies and the results of its application) are some of the characteristics that support the use of the KBDSS as a decision support tool to handle solids separation problems of microbiological origin.

The developed KBDSS was implemented by means of the G2-Gensym environment, an object-oriented shell, which presents the following set of capabilities that facilitated the implementation of the system: (1) reasoning and control of events in a continuously changing environment; (2) application of both procedural knowledge and rule-based heuristics; (3) acquisition of information from any number of data sources, both local and remote; (4) establishment of relationships between classes and objects, which saves time and space and contributes to its easily transferability to other facilities and (5) user-friendly use.

The implementation process has been carried out through three different modules of operation organised by means of a feasible three-layer architecture composed by data gathering, reasoning and decision support modules. The use of a multi-layer architecture enabled the system to: (1) have the appropriate modularity and independence to guarantee the re-design and transferability of the system to other facilities; (2) to deal with any kind of data gathered from the process (quantitative and qualitative); (3) to include an array of different kinds of knowledge (numerical, heuristic, experiential, and predictive), and (4) to integrate the advantages of different intelligent technologies.

The minimum set of data that a WWTP should gather daily to ensure that the goals of the KBDSS are met is the following: SVI or V30; effluent TSS; percentage of COD removal; the qualitative analysis of sludge settleability and the presence of foams; the microscopic analysis of floc characteristics, the predominant filament, the presence of *Zoogloea* spp. and the predominant group; measurement of SRT, F/M, sulphide and pH in the influent, COD/N, COD/P, DO, temperature, and percentage of soluble COD in the reactor; influent waste flow and air flow. The presented approach is based on daily measurements, considering that, according to the results of the study presented in Chapter 3, for most variables at least a daily value (usually measured from an integrated sample) was available. Whenever more than one measurement per day exists, the daily average value is calculated. On the other hand, for those facilities where some key variables such as TSS were available on-line, the reasoning process could be adapted and, consequently, the system could make on-line diagnoses.

The system has been implemented to be routinely executed once a day, although it could also be launched manually at any time, whenever an alarm symptom is detected. According to the slow dynamics of the domain, the daily execution of the system is adequate enough to fulfil the main objectives of the system.

The use and interaction of both ES and EBRS techniques improves the reasoning capabilities (deeper knowledge) and upgrades the decision making reliability of the KBDSS. The integration of these AI technologies also provides the dynamic KBDSS with an intelligent behaviour. Hence, it is now able to reason both based on expert knowledge from literature and experts (ES) and based on experiential knowledge accumulated through years of operation in the facility (EBRS), which is crucial for diagnosing, predicting and solving abnormal or problematic situations. This intelligent feature of the system allows it to overcome the limitations of classical control in this kind of processes.

The integration of both ES and EBRS techniques has allowed the system to provide the user with different alternatives whenever a decision on the control strategy needs to be made: one based on knowledge of the domain and another one based on the plant's own experiences. Given this information, the user will decide which the best option is. Daily operation has demonstrated that the optimal solution is generally a combination of different control actions. Currently, the developed KBDSS do not integrate more than one control strategy; like specific and non-specific actions, they are provided separately. Nevertheless, the registration of the applied control strategy within the episode library will permit the registration of this optimal combination for future use. Future work in this topic will be based on the automatic combination of the suggestions of the ES and EBRS, as well as on the combination of different specific control strategies whenever more than one possible cause is identified.

Apart from showing the results of both the ES and the EBRS, the implemented KBDSS can be requested for explanations of the conclusions reached and the deductive processes followed, or requested to retrieve certain quantitative or qualitative values. Moreover, several trend charts, graphics or readouts of the most important values facilitates the understanding of the large amounts of information generated by the process. Furthermore, a dynamic interaction between the system and the user is also facilitated. The user, at any time, can calibrate or customize the system by modifying some key parameters such as the considered fuzzy sets, the limits of effectiveness, etc. In order to promote and assist the interaction of the user, several user-friendly interfaces were developed.

The maintenance of the KBDSS is a tough task. For example, whenever the activated sludge process undergoes some modifications (i.e. enlargement of the facility, integration of new sensors, etc.) most of the modules should be modified. Regarding the maintenance of the episode library, it directly depends on the user, who finally registers each new experience.

CHAPTER 8

VALIDATION OF THE DYNAMIC KNOWLEDGE-BASED DECISION SUPPORT SYSTEM

The present chapter refers to the final step of the KBDSS implementation: the validation of the system. The validation process can be defined as establishing, by objective evidence, that all the system requirements have been implemented correctly and completely, and are traceable to system requirements (Tanzio, 2001). Moreover, it needs to be remarked that validation covers the entire life cycle of the system implementation, and not just its end use. The present section is therefore focused on the single validation of each technology involved in the KBDSS architecture and on the field validation of the overall system as a real time Supervisory System for the management of solids separation problems.

The validation process must ensure that the system solves problems correctly, with both adequacy and accuracy. The **adequacy** measures to which extent the system covers the domain knowledge, while the **accuracy** can be defined as the proportion of “acceptable” answers that a system generates, and can be a quantitative statistical measure. A validated KBDSS ensures a correct representation of the acquired knowledge and thus, correct solutions to the target problems. Therefore, this process also deals with questions of correctness: e.g., are the conclusions drawn by the system correct?, does the system fit well into the decision making process?. The answers to these questions depend on the internal reasoning of the system. For example, if an Expert System cannot give the right responses, it means that the expert reasoning process was misinterpreted during knowledge acquisition. Validation should be conducted both during the development phase and throughout the implementation of the final prototype. One expects to get some level of acceptable performance in the early stages of the prototype development, and this is expected to increase as the system development progresses. One of the most important requisites to perform an optimal validation is that the expert(s) in the domain (in our approach, the plant operators) contribute to the validation process.

Validation thus refers to the whole set of tasks that involve checking, detecting anomalies, assuring adequacy, etc., and are conducted in parallel to the development phase. The main objective of the validation process is to ensure that the proposed system provides the correct answer in the correct form. To do this, we should discover and eliminate from the system any sources of error or inadequacies that might have been introduced during the different phases of the system development (system design, knowledge acquisition, and implementation). The major causes of errors are as follows:

⇒ Lack of system specifications and poor understanding of the problem. Also semantic and syntactic errors introduced during the implementation.

⇒ Erroneous solutions or inability to find any solution to the problem (in the ES, due to incorrect representation of the domain knowledge; in the EBRS, because of inefficient case and episode retrieval). It includes *errors of commission* and *errors of omission*. Errors of commission occur when the KBDSS reaches an incorrect conclusion for a given set of input values. They affect the accuracy of the system, and while they are easy to detect, they are often difficult to locate and correct. Errors of omission occur when a set of input values fails to cause the system to reach a conclusion because the knowledge necessary to solve a particular problem is not found in the KB. This error is more difficult to detect and affects the adequacy of the system.

⇒ Unsatisfactory relationship with the user.

In order to perform a complete validation process, each one of these error types must be addressed. The whole process of validation implies the accomplishment of different validation activities including: (1) verification and (2) ease of use of the system (usability) and readiness for use (usefulness).

The **verification** process examines the first and second causes of error: checks compliance to the specifications in order to ensure consistency and completeness and avoid redundancy. In other words, verification refers to building the system “right”, substantiating that the system correctly implements its specifications (González and Dankel, 1993). During the development phase and once the prototype of the system was up and running, design specifications were tested, revised, and new specifications were added to accomplish requirements that were not initially known. The main premise is that a consistent system should behave in a non-contradictory way. Normally, the verification process is accomplished by the developer of the system (the knowledge engineer) and the expert on the domain. In the present approach, we have acted both as knowledge engineers and as expert on the domain. The collaboration of various plant operators of the Girona WWTP also helped us during this process.

Other validation activities look at the **usefulness** of the KBDSS, which implies the validation of the system capability to perform its main tasks with efficiency and effectiveness, i.e. the results of the system must be reliable and easily available within an acceptable response time. Finally, the validation of **usability** looks for the functional aspects of the system, regarding for example the communication between the user and the system’s interface the accessibility of the results or the difficulties in running the system. In general, the validation of usefulness and usability are related to the third cause of error: unsatisfactory relationship with the user. Hence, they cannot be evaluated until the system is tested by the final user.

Regarding the planning and execution of any validation process, three distinct parts can generally be distinguished:

- Preparation of the validation. Definition of all the items that are part of a validation: (1) the objective (why is the validation performed); (2) the criteria based on the objectives (what aspects or characteristics should be checked; e.g., check the knowledge base for accuracy and adequacy); (3) the tests or techniques to be performed (how the validation is going to be performed); and (4) the reference standard (what standard is selected for comparison), which can consist in either decisions developed by human experts or previously known results.

- Validation performance. The test workload is performed and the criteria are judged and assessed one by one, detecting errors, checking system responses, comparing the system responses with standards for agreement (comparing with expert responses or known results), and interpreting the feedback obtained. The result of the validation could be either a quantitative value or a qualitative value (good/bad, valid, satisfactory, accurate, etc.). Several formal validation methods or techniques are generally used in the validation tests. Examples include comparison testing, criticality analysis, consistency checks, algorithm analysis, control flow analysis, etc. (Tanzio, 2001)

- Decision making on the basis of the result of the validation. The results of the validation process must result in several decisions concerning the modification or the continuity of the system.

8.1 Procedure for KBDSS Validation

The first step before performing a system validation is to define all the items that will be part of the validation process (preparation of the validation). Hence, first of all the objective of the validation must be defined.

The **objective** of the validation process is to ensure that the developed system has been implemented correctly and completely, and accomplishes all the defined requirements. In our approach, these requirements are related to the management of solids separation problems, including: (1) diagnosis of solids separation problems episodes; (2) prediction of future problematic situations; (3) identification of the possible causes that have led to the problematic situations; and (4) proposal of complete control strategies to restore the process in the most efficient way and based on both expertise and experience on the domain.

Secondly, the **criteria** or list of aspects that, based on the objectives, must be validated is also defined. In our approach, the following aspects have been identified as objects of validation:

- Reliability of communication and data gathering
- Validation of the Expert System, including:
    - Consistency and completeness of decision trees
    - Accuracy of Expert System diagnosis
    - Usefulness of the suggested control plans proposed by the Expert System
- Validation of the Episode-Based Reasoning system, including:
    - Adequacy of the episode library
    - Accuracy of case and episode retrieval
    - Usefulness of the retrieved episodes by the Episode-Based Reasoning System
- Global validation of the overall KBDSS

The different **techniques** applied to validate the KBDSS can be classified into two main categories: (1) **experimental validation**, where the consistency, accuracy, adequacy and usefulness of the system is validated by means of different experimental techniques including questionnaires and interviews, laboratory (prepared) test cases, historical cases with known solutions, laboratory experiments in which the system output and solution were directly compared with an expert prediction and solution (user assessment), etc.; and (2) **field validation**, where the usefulness and usability of the KBDSS results were validated within its real environment (the activated sludge process) and with real problem-types (real episodes of solids separation problems in the Girona WWTP). Field experiments are the most effective of all validity tests if the situations permit them (can sometimes be quite expensive) (Borenstein, 1998).

Finally, the last step in the preparation of the validation is to **select reference standards** or norms with which we can compare the actual behaviour of the system in order to define the agreement of the system's results. Different possibilities have been used during the present approach:

- Comparison against known results (historical or prepared cases): these test cases should be based on real situations and should cover a range of levels of difficulty and test as many aspects of the system as possible. If prepared, they should be generated by unbiased experts.

- Comparison against human expert performance: norms prescribed by the experts for problem solving heuristics. They must be free of bias or prejudice.

Other reference standards (not used in the present study) include: comparison against theoretical norms (if the KBS is modelling a physical process) or sensitivity analysis to input variations and input errors.

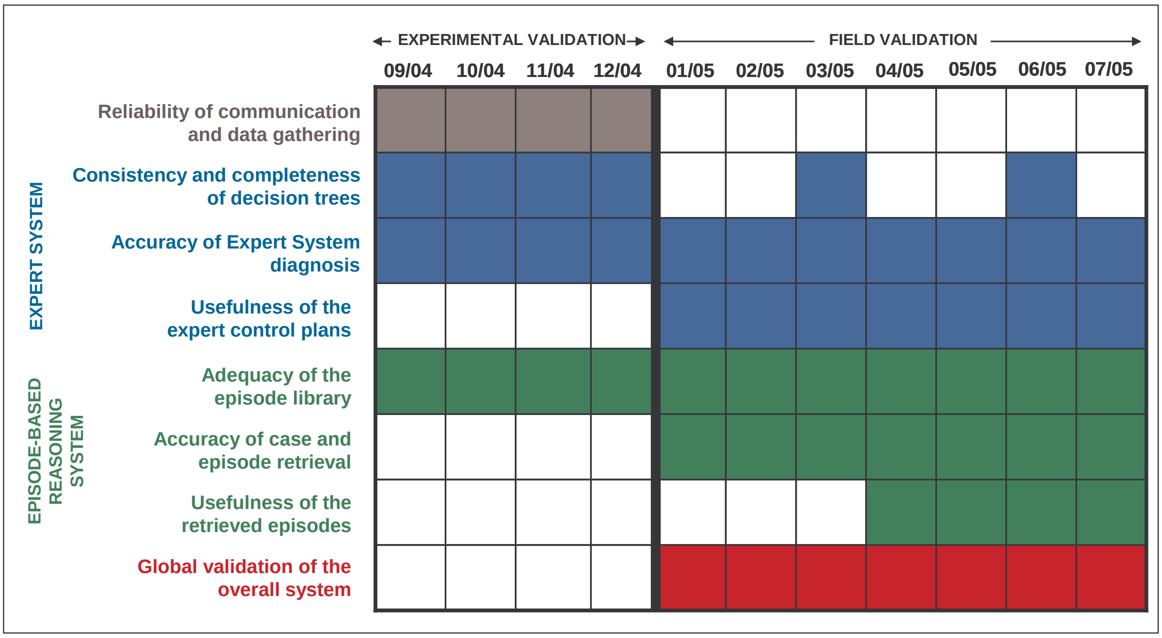
Once the preparation of the validation phase is completed, the **validation performance** is carried out according to the objectives, criteria, techniques and reference standards specified above.

Finally, after all the required validation procedures are performed, a **final decision** needs to be made as to the extent to which the system can be considered consistent enough to be ready for its final use. In our approach, the validation procedure was finished once the KBDSS proved to effectively accomplish all the objectives which it was developed to meet.

In the present study, the proposed validation procedure was carried out iteratively throughout the system development and implementation phases. The results from any stage (or sub-stage) have implied changes (reformulations, redesign, and refinements) in the prototype. Whenever the prototype was modified or expanded, the system was re-evaluated. So, the system validation was not understood as a linear process but it was best conceived as a cyclic process were part of the validation procedure was naturally implemented as part of the system development. Hence, the changes derived from the whole validation procedure were already included in the final prototype presented in chapters 5 and 6. During the final validation stage, the experimental and field validation stages of the system were initiated as soon as a first prototype of the dynamic KBDSS was available. The problem in using a poorly evolved prototype for the field validation is that the system may lose credibility before the users get a chance to see the final product. To avoid this, we had to make it clear to the users that a prototype might not be as complete or as intelligent as they might have expected. The final validation stage started in September 2004, when the first prototype was implemented into a full-scale plant, the Girona WWTP, where it was validated for a total period of 11 months. Figure 8.1 depicts the temporal distribution of each validation phase, including both experimental and field validation.

The following sections describe each of the phases carried out to validate the KBDSS for supporting WWTP management and control of solids separation problems. This procedure involved experimental and/or field validation of the aspects (or criteria) previously proposed, which include: (1) reliability of communication and data gathering; (2) consistency and completeness of decision trees; (3) accuracy of ES diagnosis; (4) usefulness of the expert control plans; (5) accuracy of case and episode retrieval; (6) usefulness of the retrieved episodes; and (7) global validation of the overall system.

In the final section, a specific case study concerning a real episode of filamentous bulking is provided in order to illustrate the overall performance, usability and usefulness of the KBDSS.



		EXPERIMENTAL VALIDATION				FIELD VALIDATION						
		09/04	10/04	11/04	12/04	01/05	02/05	03/05	04/05	05/05	06/05	07/05
	Reliability of communication and data gathering	🟫	🟫	🟫	🟫							
EXPERT SYSTEM	Consistency and completeness of decision trees	🟦	🟦	🟦	🟦			🟦			🟦	
	Accuracy of Expert System diagnosis	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦
	Usefulness of the expert control plans					🟦	🟦	🟦	🟦	🟦	🟦	🟦
EPISODE-BASED REASONING SYSTEM	Adequacy of the episode library	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
	Accuracy of case and episode retrieval					🟩	🟩	🟩	🟩	🟩	🟩	🟩
	Usefulness of the retrieved episodes								🟩	🟩	🟩	🟩
	Global validation of the overall system					🟥	🟥	🟥	🟥	🟥	🟥	🟥

Figure 8.1 Temporal distribution of the different sub-stages of the validation process

8.2 Reliability of Communication and Data Gathering

The validation of the communication and data gathering module consisted mainly in the verification of the reliable transferability of data between the environment where the system operates (the activated sludge process) and the KBDSS. This process was continuously preformed during the first four months of validation by means of an experimental procedure that validated the following components and functions:

- ➢ The ActiveXlink bridge (communication tool between the Excel worksheet and the KBDSS)
- ➢ Data filtration and handling of missing data
- ➢ Discretisation of data
- ➢ Calculation of new data (trends, means, etc.)
- ➢ Compilation of the required data into the specific vectors

The experimental validation was performed during the first phase of the validation procedure in order to ensure that the other modules (ES and EBRS) were provided with the most appropriate data to perform their reasoning processes. During the experimental procedure, a historical library of data from the Girona WWTP was used to verify the different components and functions listed above. Once all these key components were validated, improvement actions were carried out to solve the identified errors. Some examples include:

- ⇒ the modification of the discretisation procedure was necessary due to an error found in the discretisation of some of the quantitative values by using fuzzy sets

- ⇒ the calculation of the average MCRT was included instead of the daily value of MCRT in order to provide more useful information to the system

⇒ apart from calculating the SVI trend, the calculation of effluent-TSS was also included as a monitoring parameter of solids separation problems

The proper functioning of data gathering module is basic for the proper operation of the whole system, since it is the basis of the suggested KBDSS architecture. Once the data gathering module was validated, the ES and EBRS modules were ready to be validated too.

8.3 Consistency and Completeness of Decision Trees

The core of the reasoning process of the Expert System is the set of decision trees which allow the system to reach conclusions about problem diagnosis (or prediction), cause identification and proper control strategy. In order to avoid the first cause of errors or inefficiency listed above (i.e. the lack of system specifications and poor understanding of the problem, together with semantic and syntactic errors introduced during the implementation), the set of decision trees developed in the different modules of diagnosis were verified both for consistency and for completeness.

8.3.1 Checking for consistency

Checking for consistency in the knowledge base of the ES implies checking for syntactic and semantic errors that the developers often introduce during the development phase. During the validation of decision trees, the base of rules was checked for the existence of the following types of erroneous rules:

- Redundant rules: e.g.

  IF *SVI is high AND Sludge Settleability is 2* ***THEN*** *Filamentous Bulking is True*

  IF *Sludge Settleability is 2 AND SVI is high* ***THEN*** *Filamentous Bulking is True*

- Conflicting rules: e.g.

  IF *SVI is low AND Sludge Settleability is 5* ***THEN*** *Dispersed Growth is True*

  IF *SVI is low AND Sludge Settleability is 5* ***THEN*** *Dispersed Growth is False*

- Subsumed rules: one rule is subsumed by another if it has more constraints in the premise while having identical solutions. In this example, rule 1 is subsumed by rule 2 because the former has one more constraint than the latter.

  1. ***IF*** *SVI is high AND Sludge Settleability is 2 AND Floc Characteristics is F* ***THEN*** *Filamentous Bulking is True*

  2. ***IF*** *Sludge Settleability is 2 AND SVI is high* ***THEN*** *Filamentous Bulking is True*

- Unnecessary IF conditions: similar to subsumed rules.

  IF *SVI is high AND V30 is high AND Sludge Settleability is 2* ***THEN*** *Filamentous Bulking is True*

  IF *SVI is high AND V30 is not high AND Sludge Settleability is 2* ***THEN*** *Filamentous Bulking is True*

The premises of these rules are identical except for one, which appears to be contradictory. In this case the contradictory rule is unnecessary (the conclusion is independent of the value of V30), and therefore the two rules can be collapsed into the single rule:

__IF__ SVI is high AND Sludge Settleability is 2 __THEN__ Filamentous Bulking is True

- Circular rules: these rules lead to an infinite loop of useless rule firings. Forward-chaining systems rarely check if rule conclusions have been derived previously.

  __IF__ SVI is high AND Sludge Settleability is 2 __THEN__ Filamentous Bulking is True

  __IF__ Filamentous Bulking is True __THEN__ SVI is high AND Sludge Settleability is 2

8.3.2 Checking for completeness

When checking for completeness of the rule base, the following types of errors have been checked:

- Dead-end rules: rules that have actions that do not affect any conclusions and are not used by other rules to generate any other conclusion.

  __IF__ DO is 2.3 __THEN__ conclude that the DO level is normal (None of the rules use the premise “IF the DO level is normal THEN…”)

- Unreachable rules: rules with a premise that will never be matched.

  __IF__ V30 is very low AND Floc Characteristics is F __THEN__ Filamentous Bulking is not True

- Missing rules: rules characterised by facts that are not used in the inference process, and conclusions that do not affect any other rule or procedure.

There exist tools to perform automatically the verification process of decision trees by comparing premises and conclusions of each rule individually with all the others, comparing premises of one rule with all the others, comparing conclusions of one rule with all the others, and comparing all the premises with all the conclusions. However, we have preferred to do it manually and in co-operation with the expert.

The validation of decision trees was performed experimentally, during the 4-month period before the KBDSS was globally validated. As in the case of the validation of communication and data gathering, validating the decision trees was extremely important to ensure the correct operation of the ES and EBRS. In addition, once the field validation started, the decision trees were continuously validated, particularly when a problem of accuracy or adequacy of the KBDSS (especially of the ES) was detected. This is why in Figure 8.1 the validation of decision trees sporadically continues throughout the field validation period.

8.4 Accuracy of Expert System Diagnosis

Different approaches exist for validating expert systems. Among them, the techniques of face validation and the use of a set of historical test cases (in which the results are known) were combined to check for the accuracy and adequacy of the knowledge base of the Expert System. The criteria of the plant manager and even our expert criteria on the domain were used as reference standards. The methodology used to discover inaccuracy and inadequacy in the ES knowledge base was based on the proposal of Comas (2000). It includes the following processes:

1. Validation of the Expert System with historical real cases (historical test cases with known solutions). This permits the validation of the accuracy and adequacy of the ES in diagnosing and/or predicting solids separation problems, identifying the possible cause and suggesting an appropriate control method.

2. Arrangement of meetings between one or more experts to discuss in a qualitative way each conclusion reached by the system (face validation). During these meetings, the system diagnosis was compared with the real state of the activated sludge system that, according to the expert criteria, the facility was undergoing during the evaluated day. The reasoning path followed by the ES was also compared and discussed with the expert of the facility. A questionnaire administered to the expert helped to collect information to support this process. The main topics in the questionnaire were:

    i. was the analysed situation correctly identified?
    ii. did the inference followed by the system correspond to the expert heuristics?
    iii. were the decision variables and their values appropriate?
    iv. did the system detect the cause properly?
    v. do you think that the solution proposed is credible enough?
    vi. what would you suggest to improve the inference process of the expert system?
    vii. evaluate the overall performance of the expert system as very good, good, fair or poor.

The next steps focused on the detection of concrete incorrect paths in each one of the decision trees:

3. Identification and registration of which problems were well diagnosed, which were not and which were not detected. In a similar way, identification of the causes detected and the control plan proposed. The trends of the key variables (SVI and TSS) and the prediction of future problems were also registered.

4. Identification and registration of abnormal qualitative values (very low, low, normal or high) taken by the variables that launched the intermediate alarms of the explored problems (the symptoms).

5. For each of the problems diagnosed by the ES (i.e. for each decision tree explored), the traces followed by the expert system, i.e. the branches explored by the ES, to infer the detected situations were identified (rule trace validation method).

6. With the information derived from sections 4 and 5, the symptoms and the traces followed by the ES were underlined on a paper representation of the decision tree under exploration. This representation makes it easier for the plant manager to identify correct or incorrect inference paths. Red branches of Figure 8.2 show the symptoms and the rules traced by the ES when diagnosing a situation of filamentous bulking.

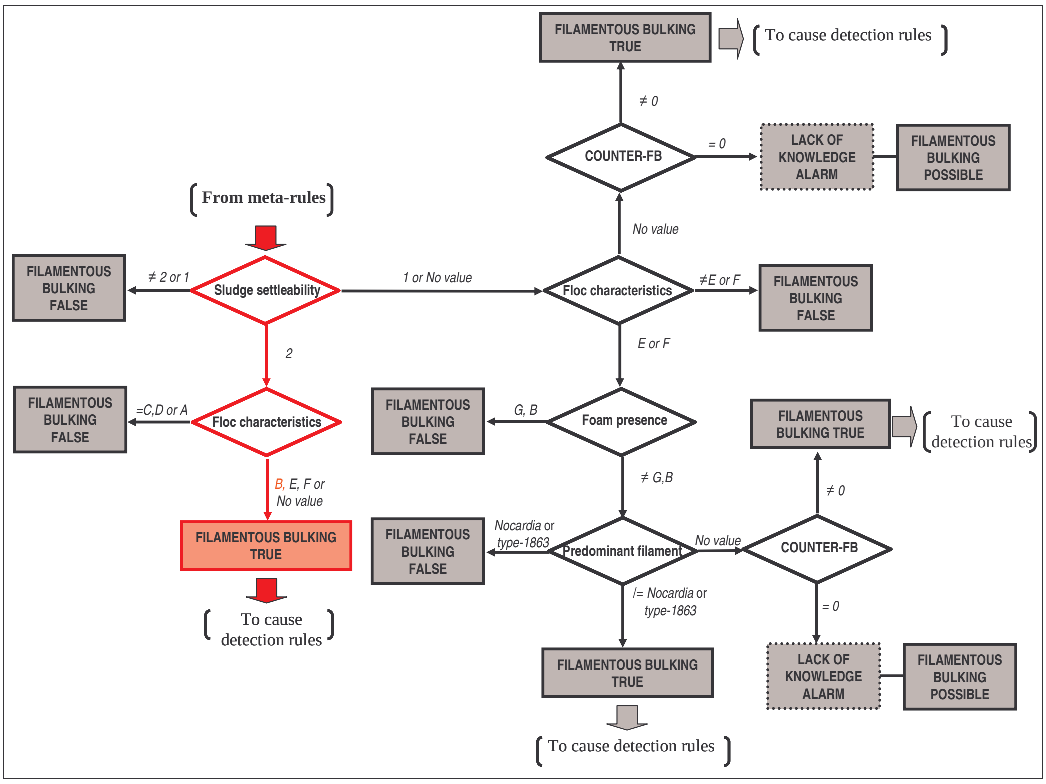


Figure 8.2 Filamentous bulking detection through the filamentous bulking diagnosis decision tree

7. Comparison of these representations with the expert's criteria (the expert compares the traces with his/her problem-solving strategies). This comparison enabled us to discover which of the rules were correctly fired and which were not and in which sequence, as well as to find errors of commission or omission. These errors can be in the symptoms that alarm for a possible problem, in the reasoning process followed by the expert system and/or in the conclusions reached.

8. The Expert System validation permitted the refinement and further expansion of the system so as to better reflect the experts' solution procedures. As a result, during this period, new rules, procedures or facts were added whilst others were modified or deleted (rule refinement, reformulation and revision). The modifications made to the knowledge base during the expert system validation included:

- ⇒ changes in meta-rules
- ⇒ addition of some causes of problems previously not considered
- ⇒ adaptation of the general actions proposed to the specifications of the studied plant
- ⇒ changes in the modalities of each variable (low/normal/high)
- ⇒ addition of some problem-solving strategy previously not considered
- ⇒ modification of the reasoning path of some problems

⇒ detection of inference paths with missing values for some variables

⇒ discovery of some situations not covered by the expert system

9. Finally, the experts were also consulted for the validation of the system's adequacy in order to detect areas of the domain knowledge not yet covered by the system. This led, for example, to the addition of new possible causes, the consideration of other control methods, etc.

As with the rest of modules (data gathering and EBRS), the modifications resulting from this validation procedure were already included in the developed system described in chapter 5.

Following the above methodology, the procedure used to validate the reasoning results of the ES concerning the problem diagnosis, the identification of possible causes and the prediction of future solids separation problems, was divided into two main phases: (1) validation of the diagnosis and prediction of solids separation problems with real data from a historical library of experiences (experimental validation), and (2) validation of the diagnosis of problems, the identification of the related causes, and the predictions of the ES by running the system in real time (field validation). Both phases are depicted in the following sections.

8.4.1 First phase of the ES diagnosis validation

During the first phase of validation (from September 2004 to January 2005), a database was built with the historical experiences gathered at the Girona WWTP from January 2004 to September 2004 (8-months test period). This historical database included both a set of the most relevant variables from the process and a description of the situation experienced by the facility during these specific days (mainly problem diagnosis; the identification of the possible cause was only reported in very specific situations). This historical database was build from the required historical data collected in the corresponding on-line, analytical, microscopic and macroscopic worksheets at the Girona WWTP. The real situation of the plant for each of the days comprised between January and September 2004 was acquired from a "operational diary" used by the plant operators to register all the problematic situations experienced in the plant.

It is worth noting that during this preliminary validation, the information retrieved from the operational diary was used only as a reference since the diagnosis recorded there concerning the solids separation problems situations could not be demonstrated to be absolutely correct. On the other hand, since the possible cause of such problematic situations was only registered for some specific days (72% of missing values), the identification of the possible cause was not evaluated during the first phase of validation.

Once the historical database was built, the ES was launched in order to obtain an expert diagnosis for each of the days comprised between January and September 2004. The different diagnoses were reached according to the data from the process (involving both quantitative and qualitative data) registered in the historical database. These results were then compared with the real situations reported in the operational diary. Tables 8.1 and 8.2 summarize the results obtained from the first validation procedure. Table 8.1 shows the validation results obtained from the diagnosis of solids separation problems while Table 8.2 depicts the results obtained from validating the prediction of future problematic situations.

Table 8.1 Results of the first validation phase of the Expert System concerning problem diagnosis

EXPERT SYSTEM DIAGNOSIS			Percentage (%)
Situations correctly diagnosed	**180**		**65.5**
Situations incorrectly diagnosed	Lack of data	Inefficiency of the system	
	72	23	
Total incorrectly diagnosed situations	**95**		**34.5**

Table 8.2 Results of the first validation phase of the Expert System concerning the prediction of future events

EXPERT SYSTEM PREDICTION			Percentage (%)
Type of situations	**Episodes properly diagnosed**	**Episodes diagnosed in advance**	
Filamentous bulking	4	4	
Non-filamentous bulking	4	2	
Biological foaming	2	2	
Rising sludge	1	0	
Normal state	5	0	
Total episodes	16	**8**	**50.0**

Among the different diagnosis obtained, the Expert System dealt with different types of solids separation problems including filamentous bulking, non-filamentous bulking, biological foaming and rising sludge, in addition to periods of normal situation (i.e., when no solids separation problem occurred). The performance of the ES throughout the 8-months period resulted in a total of 275 diagnosed situations corresponding to 11 different episodes of solids separation problems including 4 episodes of filamentous bulking, 4 episodes of non-filamentous bulking, 2 episodes of biological foaming and 1 episode of rising sludge, together with 5 episodes where no solids separation problems where diagnosed (normal state). Each one of the 275 diagnosed situations corresponds to the different diagnosis made in each one of the 244 days during the 8-months test period. The resulting 275 diagnosis were obtained given that on 31 specific days, two situations were simultaneously diagnosed due to the overlapping episodes of filamentous and non-filamentous bulking. This explains the 275 (244+31) diagnosed situations.

Regarding the results obtained in this first phase of validation, out of the 275 diagnosed situations, 180 were correctly detected, representing **65.5%** efficiency in diagnosing solids separation problems. On the other hand, 34.5% of the diagnosed situations were incorrectly detected, mainly due to a lack of data. Among the 180 correctly detected diagnoses, the dynamic ES identified 16 different episodes of variable length, 50% of which were detected in advance.

As commented before, during this first phase the **diagnosis of the possible causes** could not be validated due to a lack of registration of the historical causes.

The results obtained throughout the first phase of validation enabled us to reach several conclusions which resulted in refinements and improvements to the final Expert System prototype. The following conclusions were reached and the corresponding ES improvements implemented:

⇒ A substantial number of situations were incorrectly diagnosed. The main cause for this inefficiency was a lack of data that occurred on several days, especially at weekends, when for example microscopic and macroscopic observations were not performed.

In order to solve this setback, the existing decision trees were extended by means of dynamic reasoning, which enabled the system to reach a conclusion about the problem diagnosis even when not all the required data were available (see chapter 5). The following rule is an example of the additional dynamic reasoning incorporated:

IF filamentous bulking= possible and %-certainty of filamentous bulking >50.0 and fb-counter ≠ 0
THEN conclude that filamentous bulking= TRUE

In situations where the lack of macroscopic and microscopic data (especially at weekends) impeded the ES to diagnose whether the process was suffering from a solids separation problem, the dynamic reasoning was applied to make a decision about the current situation of the process. By using information about the closest past diagnosis of the process (registered by specific counters) and the percent certainty in reaching a conclusion derived from the different fuzzy sets, a possible situation was turned into a true or false conclusion. The dynamic reasoning applied follows the logic reasoning of an expert, i.e. in spite of missing the value for the abundance of filaments, if the process was suffering from a filamentous bulking problem the day before, and if today's SVI and V30 values indicate 85% certainty of filamentous bulking, then it is almost true that the process is still suffering from the same problematic situation.

⇒ A substantial fraction of incorrect diagnoses was caused by an inefficient reasoning process. The reported incorrect diagnoses were reviewed and, after considering the possibility of error in the registered diagnosed obtained from the operational diary, all the necessary modifications were implemented in the reasoning process to improve these inadequacies. These modifications included: (1) changes in meta-rules; (2) changes in the modalities (fuzzy sets) of each variable (very low/low/normal/high); (3) modification of the reasoning path of some problems, and (4) detection of inference paths with lack of values for some variables.

⇒ The percentage of predicted episodes was not satisfactory. Half of the detected episodes were not predicted or detected in advance, in particular episodes of non-filamentous bulking, biological foaming or rising sludge. On the other hand, the results of validation showed that the tendency towards a normal situation (recovery of the process) had not been considered by the system. Once the results of the validation were analyzed, the prediction of future improvements in the process (signs of a recovery of the system from a solids separation problem) was considered as important as the prediction of future separation problems. As a result: (1) a set of heuristic rules was developed to detect any improvement on the process once a solids separation problem was initiated, and (2) the reasoning processes used to predict solids separation problem, especially those concerning biological foaming and rising sludge, were also revised and improved. All the resulting modifications were already included in the final prototype described in chapter 5.

The first phase of the ES validation (experimental validation), together with the corresponding modifications, was carried out during the first 4 months of the validation process (September 2004-January 2005) dedicated to the validation of the ES diagnosis (see Figure 8.1).

8.4.2 Second phase of the ES diagnosis validation

The second phase of validation was carried out once the modifications resulting from the first phase of validation were applied. The ES diagnoses were validated from January 2005 to August 2005. The real-time operation of the KBDSS implied that during this 7-month period a field validation of the ES diagnosis was performed. The validation process involved a direct evaluation (face validation) of the results obtained by the ES in comparison to the real situation of the process (real diagnoses of the experts on the process, the diagnoses of the plant operators). In addition to the validation of diagnoses of solids separation problems and of the prediction of future situations, during this second phase of the ES diagnosis, the identification of the possible causes could also be validated.

This process of validation (as well as the validation of the EBRS and the whole KBDSS) was carried out by means of periodical meetings with plant operators. The meetings were usually performed 2-3 times per week in order to validate the specific performance of the ES in properly diagnosing and/or predicting the problematic situations and their corresponding causes. Different plant operators, including the plant manager, lab technicians, process controllers and process engineers, participated in these periodical meetings. The following questions were used as an outline of the validation process:

- Was the analysed situation correctly identified?
- Did the inference followed by the ES match the expert heuristics?
- Were the decision variables and their values appropriate?
- Did the system detect the cause properly?

Regarding the prediction of future episodes, every time a new episode was identified, the results of the ES corresponding to the days previous to the episode were checked in order to validate its predictions. Table 8.3 summarizes the results of the second phase of the ES diagnosis validation concerning the diagnosis of solids separation problems and the identification of possible causes. Table 8.4 depicts the results of the ES validation regarding the prediction of future problematic situations or improvements to the process.

Table 8.3 Results of the second validation phase of the ES concerning problem diagnosis and cause identification

EXPERT SYSTEM DIAGNOSIS			Percentage (%)
Situations correctly diagnosed	**219**		**97.3**
Situations incorrectly diagnosed	Lack of data 0	Inefficiency of the system 6	
Total incorrectly diagnosed situations	**6**		**2.7**
Causes correctly identified	**98**		**43.5**
Causes not identified	Lack of data 119	Inefficiency of the system 0	
Total incorrectly identified causes	**127**		**56.5**

Table 8.4 Results of the second validation phase of the ES concerning the prediction of future events

EXPERT SYSTEM PREDICTION			Percentage (%)
Type of situations	**Episodes properly diagnosed**	**Episodes diagnosed in advance**	
Filamentous bulking	4	3	
Non-filamentous bulking	1	0	
Biological foaming	1	1	
Normal state	4	4	
Total episodes	10	**8**	**80.0**

The validation of the ES diagnoses throughout the 7-months validation period resulted in a significant improvement in the efficiency of the system regarding problem diagnosis and prediction of future episodes of solids separation problems. During this period, 219 situations were correctly detected, representing 97.3% of efficiency. Notably, the inefficiency of diagnosis as a consequence of lack of data was totally suppressed thanks to the addition of dynamic reasoning to the decision trees, as commented above. This implies that the ES can now reach a diagnosis even when it does not have at its disposal all the required data from the process. On the other hand, the cause of the incorrect diagnoses due to an inefficiency of the system was revised, and once identified, the corresponding rules were modified. Once all these modifications were applied, it can be stated that approximately **100%** of the situations were correctly detected.

As during the first phase of validation, the prediction capability of the ES was validated by considering its efficiency in predicting future episodes of solids separation problems. First of all, the information provided by the calculated trends of the SVI and the TSS seems to be very useful in helping plant operators to study the daily evolution of the process. On the other hand, regarding the prediction of future episodes or events, an efficiency of 80% was obtained, given that from the 219 situations correctly identified, which correspond to 10 different episodes, 8 episodes were properly predicted. Also, 3 of the 4 identified episodes of filamentous bulking were properly predicted at least 1 day in advance (the episode not identified corresponded to an episode of filamentous bulking that appeared at the same time as an episode of biological foaming). The episode of biological foaming was also correctly predicted. Finally, the improvement of the process (trend to normal state) was efficiently predicted in all the situations. On the other hand, the prediction of the viscous bulking episode was not specifically predicted (just a general deterioration on the sludge settleability was predicted). This can be explained by the fact that the proliferation of *Zoogloea* sp. or the overabundance of EPS can not be as evidently predicted as for example the presence of filamentous bacteria, given that as in episodes of biological foaming (and in general as in every solids separation problem), not all the possible effects and causes are sufficiently understood.

Finally, the validation of the ES diagnoses regarding the detection of the possible causes of solids separation problems resulted in an efficiency of just **43.3 %** of correctly identified causes. The origin of this 56.7% of unidentified causes (meaning not identified causes rather than incorrectly identified causes) is certainly the lack of data that hinders the verification of all the possible causes. An example of this fact is for example the lack of data regarding the percent of soluble COD, the sulphide concentration or the ratios COD/P and COD/N which due to their low frequency of measurement delayed the proper identification of the cause. In such occasions, therefore, it can be concluded that an increase in the frequency of measurement of such variables will absolutely increase the efficiency of the system in diagnosing the possible causes of solids separation problems. This was therefore recommended to plant operators in order to improve the diagnosis of the possible causes and the prediction of future problems based on the trends of such variables.

It should be noted, therefore, that in spite of our attempt to select the most generally measured variables (based on the study of chapter 3) to facilitate the reasoning process of the system, in the end the set of input variables used by the system to identify the possible cause (see Table 5.2) can be scarcely modified. Its measurement is strictly necessary (other variables can not be measured instead) whenever the possible cause of a problematic episode needs to be identified or when the efficiency of an specific control method is being monitored.

In order to overcome the limitations found in the identification of possible causes, a set of improvements were carried out. The following are some examples:

⇒ Whenever a cause can not be properly identified due to a lack of data, the user is informed about which are the specific variables that need to be measured to identify the cause. In addition, according to the predominant filament, the system will provide the user with a list of probable causes.

⇒ As soon as a cause is identified and a specific control plan proposed, the system recommends monitoring daily the evolution of the specific variable to study the effects of the control strategy. For those cases in which this daily measurement can not be performed, the reasoning process has been modified in order to continue with the specific control plan instead of suggesting non-specific control methods. Numerous examples can be observed in most decision trees designed to define a specific control strategy (see Annex).

Once again, all the improvement measures derived from the second phase of validation of the ES are already incorporated in the final prototype described in chapter 5.

8.5 Usefulness of the Expert Control Plans

Once the first phase of the ES diagnosis validation was completed, and the accuracy in diagnosing the problem and the possible cause was improved, the functionality of the ES in suggesting complete control methods to solve solids separation problems was ready to be validated. Since no historical data was available about the control strategies applied in the past, historical cases could not be used to validate the control plans suggested by the ES (experimental validation). Instead, a field validation procedure, stretching from January 2005 to August 2005 was followed, whereby the system was run in real time and the usefulness of the control plans that it suggested were validated. The collaboration of plant operators, and especially of the plant manager, was also crucial during this process. Hence, the periodic meetings performed during the validation of ES dignosis were also used to validate the suggested control plans.

Figure 8.3 depicts the different control plans (in %) suggested by the ES for the 6 different episodes of solids separation problems experienced during the 7-month period of validation. Notice that non-specific control methods were directly suggested only when the cause was not properly identified. On such occasions, a list of possible non-specific methods such as chlorination or addition of coagulants/flocculants was proposed (it was up to the user to decide which of these possibilities to follow).

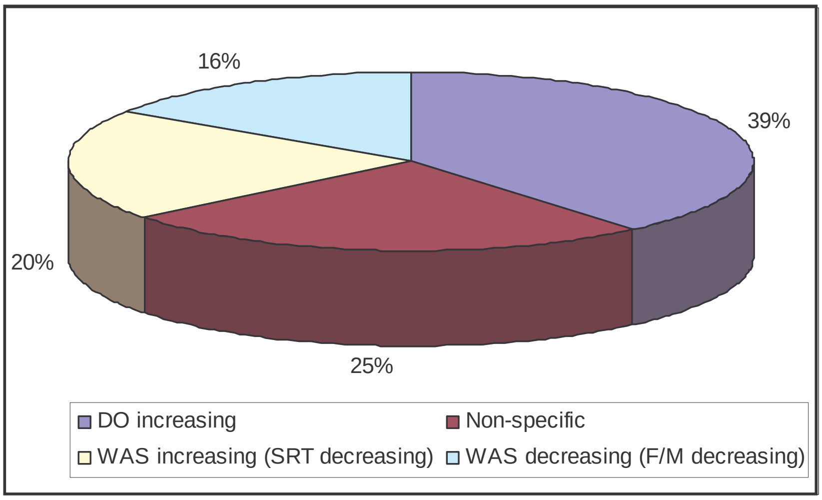


Figure 8.3 Different control plans (in %) suggested during the 7 months of validation

As observed in Figure 8.3, an increase in the DO concentration was the control plan most frequently suggested by the Expert System. This can be explained by the fact that the facility continuously suffers from high loadings which usually result in high oxygen demand (i.e. low DO cause). On the other hand, the use of non-specific control methods (especially in episodes of filamentous bulking caused by Type 021N) was suggested whenever the specific cause could not be identified due to a lack of data. It needs to be mentioned though, that as soon as the percent of soluble COD was available, a readily biodegradable substrate was identified as the possible cause of the filamentous bulking episodes caused by Type 021N. The use of a selector was then suggested as a possible control plan. When the plant operator rejected this option (due to limitations of the facility), the ES automatically suggested alternative actions such as decreasing at source the concentration of soluble COD or the use of non-specific control methods. Accordingly, the percent of occasions when the ES suggested the use of a selector has not been included in Figure 8.3 because in fact the suggested control strategy was turned into a non-specific control method. Finally, 36% of the suggested control actions corresponded to a modification of the WAS flow rate (20% of control plans involved an increase of WAS to decrease sludge age, and 16% involved a decrease of WAS to decrease the F/M ratio).

The moment these control plans were suggested, the plant operators were asked to validate their usefulness. Out of the 151 suggested control plans (no control plan was suggested during the episodes diagnosed as normal state), **91.5%** were considered as very useful for plant operators, especially those concerning the manipulation of WAS, when a target WAS flow was recommended by the system, to avoid sudden increases in sludge age or to thwart the effects of a high organic load. The 8.5% useless control plans corresponded to occasions for which the suggested method was not considered feasible enough to be applied, i.e. whenever a use of selectors was suggested or when an increase of DO set-point over 2.5 ppm was recommended.

Apart from validating the usefulness of the suggested control plans, their applicability was also validated by analysing the percentage of the plans that were directly applied. It was found that **57%** of the control methods eventually applied corresponded to the control method suggested by the ES. The remaining 43% of the occasions, in which the applied control action was different from the one suggested by the ES, can be explained by the existence of some particularities and limitations of the facility which hindered the direct application of the control plan suggested by the ES. For example, sometimes the required DO set-point could not be reached due to some limitations of the aeration system. Another example is found in episodes when an increase of the MLSS (by decreasing the purge flow) was suggested in order to counteract high organic loads. On such occasions, it was found that a high concentration of MLSS could have resulted in sludge settleability problems in secondary settlers.

Based on the results obtained during the second phase of the ES validation regarding the suggestion of control strategies, a set of conclusions were identified. Moreover, given that one of the main objectives of the KBDSS is to provide efficient and useful control strategies by applying an intelligent reasoning process, a set of improvement actions were also identified. The following are some examples:

⇒ The proposal of complete control plans including the exact required doses of chemicals or the exact control parameters such as the target WAS flow or DO set-point was evaluated as being very positive by plant operators. Likewise, the suggested duration of the control plan was also qualified as being very useful.

⇒ Given the dynamism observed in the evolution of all the episodes of solids separation problems, the necessity of a continuously dynamic reasoning process was confirmed to be very useful and necessary in order to update the suggested control strategies according to the evolution of the process. Monitoring some key variables such as the SVI, the effluent TSS, the floc characteristics or the identified cause (COD/N or COD/P ratios, F/M, DO concentration, etc.) whenever a control strategy was being applied, was confirmed to be extremely necessary in order to continuously update the control strategies and increase their usefulness. The most precise and specific the suggested control strategy was (e.g., the exact WAS flow rate or DO set-point is provided), the more useful it was considered.

⇒ For most control methods, alternative or complementary actions must be suggested in case a specific control action could not be directly applied. For example, in the specific control method suggested for the solution of low DO concentrations, it has been found that the suggestion of increasing the DO set-point is not feasible enough because some facilities, such as the Girona WWTP, can show some limitations in supplying the required air to reach the targeted set-point, especially when the origin of this low DO is a extremely high organic load. Accordingly, the plan to increase the DO was completed by including other possibilities of action such as the manipulation of WAS (see Figure 7.21).

⇒ In order to provide the user with options other than the suggested control plan, access to non-specific control plans were also provided in case the user rejected the specific control method suggested, or in case the specific control method turned out not to be effective. For instance, whenever the use of a selector is suggested to control low F/M filaments or to counteract readily biodegradable substrates, a non-specific control method is also suggested, since the use of selectors can be considered a long-term control strategy and, sometimes it can not be directly adopted by most facilities.

As with the rest of the validation phases, the partial conclusions and the related improvement measures derived from the second validation phase were applied in parallel to the validation process. Thus, at the end of the validation phase all the corrective actions had already been applied, resulting in an increase in system efficiency and usefulness.

8.6 Adequacy of the Episode Library

An error in the EBRS may be basically due to (1) a small episode-library (insufficient adequacy, i.e., insufficient representation of the WWTP domain), (2) an inefficient retrieval mechanism (the most similar case and its corresponding episode are not retrieved), or (3) the low usefulness of the retrieved episodes (i.e., the information stored in the retrieved episodes is not useful enough to the user).

Based on our experience with previous applications of CBR in the WWTP domain (Roda *et al.*, 2001 and Martínez *et al.* 2005b), we took into account the fact that the adequacy of the EBRS as it relates to the representation of the domain within the episode library grows with time, i.e. as the number of retrieved episodes of solids separation problems increases, the adequacy of the system increases too.

At the beginning of the field validation phase, a complete library of episodes (with all the necessary attributes including the applied control strategies and lessons) could not be built due to the lack of information related to the exact control strategies applied in the past in the Girona WWTP. As a consequence, a experimental validation could not be properly performed. Even so, an initial seed of incomplete episodes (they lack the identified cause, the applied control strategy, the comments and lesson) was built with the historical data of the facility from 2004 in order to initially perform a first phase of validation concerning the accuracy of case and episode retrieval (see section 8.7). Once the real-time operation of the KBDSS progressed, the initial seed could be renovated with complete episodes. At the end of the field validation period, the episode library included 6 complete episodes of solids separation problems and 151 single cases.

Consequently, we assumed that at the beginning of the field validation phase, the adequacy of the episode library was very low, but that it gradually increased as new episodes were incorporated to the episode library. Actually, at the end of the field validation phase, the adequacy of the episode library increased significantly. Nevertheless, the adequacy can not be considered as being absolute, given that episodes of dispersed growth or pin-point floc (never experienced in the facility) are not represented. Regarding the representation of rising sludge episodes, 3 historical episodes of rising sludge could be reconstructed by studying historical data and with the support of plant operators.

According to the characteristics of this phase of validation, in Figure 8.1, the validation period set aside for the adequacy of the episode library is represented throughout the 11 months of validation.

8.7 Accuracy of Case and Episode Retrieval

The accuracy of case retrieval was validated on the one hand by carefully analysing the “real” similarity of the retrieved cases and episodes, i.e. by checking that the retrieved case and episode were indeed similar to the current situation; and, on the other hand, by validating the usefulness of the retrieved control plans. Both validation processes involved our expert knowledge of the domain and the expert knowledge of plant operators. Other methods of validation requiring a complex validation model (which can predict the performance of EBRS under certain pre-selected conditions) were avoided because of the extensive effort required to derive the mathematical method.

The validation of the accuracy of the EBRS in retrieving the most similar cases and episodes from the episode library was performed in two main phases. During the first phase of validation (from January 2005 to April 2005) an initial episode library was used for the validation of the system accuracy. This initial episode library was compiled from historical data from the Girona WWTP corresponding to the whole year 2004. It consisted in 326 days divided into 21 different episodes: 6 episodes of filamentous bulking, 3 episodes of non-filamentous bulking, 2 episodes of biological foaming, 3 episodes of rising sludge, and 7 episodes of normal situation. During 2004 the facility never experienced episodes of pin-point floc or dispersed growth, so initially these problems were not represented in the initial episode library. Also, since a historical compilation of the control methods applied in 2004 was not possible, the initial episode library could only be used to test the accuracy of case and episode retrieval.

The validation procedure was performed by examining, for every day between January 2005 and April 2005, which historical cases and episodes were retrieved from the initial episode library. This validation phase was carried out according to our expert knowledge of the domain. Thus, for every retrieval process, the most similar case, and its corresponding episode, chosen by the EBRS were analysed according to our expert criteria in order to check if it really was a similar experience.

The process of validation consisted in rating the retrieved most similar case and episode as valid or not valid. The retrieved case and episode were considered **valid** whenever: (1) the description of the problem (diagnosed problem and identified cause(s)) coincided with the current case, and (2) the most relevant data, i.e. SVI, effluent TSS, predominant filament (in situations of filamentous bulking or biological foaming) or floc characteristics were effectively similar to the data available for the current case. Likewise, the retrieved case and episode were considered **not valid** when none of the mentioned characteristics fitted the current case. In those occasions when, as a result, a problem during the retrieval process, no conclusions about the most similar case or episode were properly reached, the retrieval process was also considered as not valid. It needs to be mentioned that during this validation process plant operators also were consulted frequently.

If during the validation process no valid cases and episodes were retrieved, it was considered that either the overall accuracy of the system was under the pre-established minimum value or that some other error of accuracy must have been present in the system. In such circumstances, the EBRS had to be corrected, changed, altered or extended.

Table 8.3 shows an example of a valid retrieved case and episode. In the example, the current day, 02/21/05, was identified by the ES as the first day of an episode of filamentous bulking caused by low DO.

The case 05/26/04 was retrieved by the EBRS, together with its corresponding episode, the episode FB-3. The episode FB-3 is an episode of filamentous bulking caused by low DO with a total length of 33 days. The problem and the cause in effect matched the current situation. During the validation of the information gathered from the retrieved case, we noted the high similarity between the SVI and the e-TSS values; other key variables, such as the predominant filament or the DO concentration, were also similar. These led us to consider the retrieval process a valid one.

Table 8.3 Example of valid case and episode

	CASE-STUDY	RETRIEVED CASE
episode-id.	FB-8	FB-3
case-id.	02/21/05	05/26/04
temporal-id.	1	2
Case-situation-description		
SVI	392	354
V30	333	410
e-TSS	67.6	52
i-sulphide	-	-
COD/N	5.7	7.1
COD/P	65.4	69.2
SRT	2.9	3.0
F/M	1.4	1.2
i-pH	7.7	7.5
DO	1.1	0.9
COD soluble	-	-
T	14.3	24.3
Sludge set.	2	2
Pred.fil.	*S.natans*	*S.natans*
Foam.pres.	N	N
Floc char.	B	E
Pred. Group	1	2
Zoog.pres.	NO	NO
%COD rem.	89.7	87.3
W-Flow	989	848
Air-Flow	128080	132450
In-flow	46308	42125
case-problem-diagnosis	Filamentous bulking	Filamentous bulking
case-cause-diagnosis	Low DO	Low DO
case-parallel-problems	none	none
case-control-strategy	-	-
case-comments	-	-
%similarity	-	97.1

RETRIEVED EPISODE	
episode-id.	FB-3
episode-length	33
episode-description	$C_1, C_2, \ldots, C_{33}$
episode-diagnosis	Filamentous bulking
episode-lesson	-
initial-case	05/25/05
final-case	06/27/04

Throughout the 3 months of the first validation phase, 87 situations were diagnosed. These situations included 2 episodes of filamentous bulking, 1 episode of biological foaming and 1 episode of normal situation (the retrieval process was not validated for the normal situations). In each case, the validation criteria stated above were used to determine the accuracy of case retrieval. When the first validation phase was finished, a set of results were obtained. These can be summarized as follows:

- During the first episode of filamentous bulking (an episode of filamentous bulking caused by Type 021N), 88% of the retrieved cases and episodes corresponded to filamentous bulking caused by Type 021N. In most of the retrieved cases, however, and ultimate cause was not identified (mainly to the lack of soluble COD and sulphide data). In these circumstances, the retrieval process was considered valid, but the need to measure and register the missing data was noted in order to improve the usefulness of the retrieved cases. The remaining 12% of not valid cases or episodes corresponded to processes that retrieved historical cases of filamentous bulking originated by different types of filaments or originated by other causes. It also included situations where no conclusions were properly reached.

- During the second episode of filamentous bulking (an episode caused by *S. natans* and encouraged mostly by a low DO), 77% of the retrieved cases and episodes were considered valid because they corresponded to situations of filamentous bulking caused by *S. natans* (or occasionally *H. hydrossis*) whose originating cause was a low DO concentration (mainly related to a extremely high F/M ratio). The same criteria as in the previous episode was followed to determine invalid retrievals.

- Finally, during the episode of biological foaming, caused by Type 1863 and also originated by a low DO concentration, 68% of the retrieved cases and episodes were considered valid enough due to their similarity with the characteristics of the current cases. The rest of not valid processes did not reach a final conclusion or even retrieved cases and episodes corresponding to other problems such as filamentous bulking.

Considering these partial results, the global validation process of the first phase of validation resulted in **77.6%** of valid retrieval processes in front of 22.4% of not valid or inefficient retrievals.

According to these preliminary results, a set of conclusions were reached, which resulted in refinements and improvements of the final EBRS. For example, for every retrieval process where the retrieved cases and episodes could not be considered as valid, all variables of the current and the most similar retrieved case were compared, and the trace followed by the new case through the hierarchical memory was checked, in an attempt to locate the problem.

The main goal of the reported actions was to improve the retrieval process and subsequently to increase the efficiency of the system. The set of conclusions and the corresponding corrective actions taken to refine the EBRS prototype were the following:

⇒ After accepting that the validation of the Eixample distance by Núñez *et al.* (2004) warranted that this similarity measurement was the most suitable or accurate for our domain, the assignment of the relevant weights was identified as the key factor to improve the retrieval mechanism of the EBRS. It needs to be remembered that during the development of the EBRS a leave-one-out cross validation analysis was performed in order to determine the importance of assigning relevant weights to the different variables (Martínez *et al.*, 2005c). Thus, the set of assigned weights were revised and adjusted in order to improve the efficiency of the retrieval process. For example, the weight of the predominant filament and floc characteristics variables was increased from 3 to 4 due to their maximum importance in characterizing the different episodes of solids separation problems (the possibility of giving a weight = 5 to these variables was rejected due to the subjective nature of such variables). The final set of applied weights is presented in Figure 5.23.

⇒ Also related to the *Eixample* distance, the maximum and minimum values of each considered variable (*upperval* and *lowerval*) were adjusted to eliminate the errors observed when applying a very wide interval, i.e. when the distance between the lowerval and the upperval is excessively large, the distance between the attributes becomes insignificant. The criteria used to define those limits were based on the statistical analysis of data from the Girona WWTP. In general, the upperval limit corresponds to the double of the maximum value observed for the specific attribute. The lowerval limit corresponds to the half of the minimum value observed for the specific attribute. The final set of maximum and minimum values (upperval and lowerval limits) considered for each of the quantitative variables used in the similarity calculation is presented in Table 8.4[3].

Table 8.4 Set of upperval and lowerval limits considered for each one of the quantitative variables

LIMIT	VALUE	LIMIT	VALUE
SVI upperval	500	SVI lowerval	25
V30 upperval	1000	V30 lowerval	0
e-TSS upperval	200	e-TSS lowerval	0
COD-rem. upperval	100	COD-rem. lowerval	0
i-S^{2-} upperval	200	i-S^{2-}lowerval	0
COD/N upperval	80	COD/N lowerval	0
COD/P upperval	400	COD/P lowerval	0
SRT upperval	50	SRT lowerval	0
F/M upperval	5.0	F/M lowerval	0
i-pH upperval	14	i-pH lowerval	0
DO upperval	6	DO lowerval	0
%CODs upperval	100	CODs lowerval	0
T upperval	30	T lowerval	0
WAS-flow upperval	5000	WAS-flow lowerval	0
air-flow upperval	250000	air-flow lowerval	0
inflow upperval	90000	inflow lowerval	10000

⇒ In order to avoid the retrieval of cases and episodes which did not match the diagnosis of the current case (e.g., retrieval of non-filamentous bulking episodes when the current situation corresponds to filamentous bulking), the use of a discrimination tree for the retrieval process was proposed (see Figure 6.4).

⇒ Finally, during this first phase of the validation process, it was proven that the bigger the initial episode library, the more accurate the retrieval process was; in other words, the greater the number of different episodes of solids separation problems (including different causes and different predominant filaments) represented in the episode library, the bigger the probability of retrieving similar and useful cases and episodes. This was proven, for example, in the studied episode of biological foaming, where due to the lack of sufficient historical episodes of biological foaming, the retrieved episodes were not adequately similar. During the first 3 months of the KBDSS field validation, 3 different episodes of solids separation problems were registered into the episode library and these will assist plant operators in the resolution of future episodes.

---

[3] The values presented in Table 8.4 correspond to the final limits defined after the validation phase was concluded and once the comments of the expert reviewers were considered. The initial values used during the validation phase are slightly different: SVI upperval was 700, SRT upperval was 100, DO upperval was 10, %CODs upperval was 800 and WAS flow upperval was 30000.

Therefore, by continuously running the EBRS, new experiences are retained and the adequacy of the system increases. In facilities where a previous and complete historical library can be developed (initial seed) the adequacy of the system can be very high at the beginning of the KBDSS implementation.

According to the preliminary validation results, these corrective actions were implemented several times in an iterative process, with constant expert collaboration, until an acceptable accuracy and adequacy of the retrieval process was reached. Thus, the validation of the retrieval process continued for a second phase that included the next four months (April to August 2004), during which the validation of the usefulness of the retrieved plans was performed in parallel. During these months of validation, the conclusion was confirmed that as the system was used and specific corrective actions were applied, the number of valid retrieved cases and episodes increased significantly until the percent of valid retrievals exceeded **97%**.

8.8 Usefulness of the Retrieved Episodes

Once the validation process concerning the retrieval of cases and episodes resulted in an acceptable percentage of valid results, and once the episode library was completed with an initial set of complete cases (87 cases and 3 episodes of solids separation problems), which included the information related to the applied control method, specific comments, and the valuable piece of knowledge –the lesson–, a preliminary validation of the EBRS suggested control plan could be performed. This preliminary validation involved the evaluation of the usefulness of the retrieved cases and episodes in supporting plant operators when deciding which control strategy to carry out to solve an episode of solids separation problems. Nevertheless, the results obtained during this preliminary validation phase can not be considered definitive given that the episode library is still not adequate enough to help in the solution of any solids separation problem appearing in the process.

During this validation period, extended from April 2005 to August 2005, the episode library that included all the experiences stored during the first phase of validation (from January 2005 to April 2005) was used to validate the usefulness of the retrieved control strategies. During this second period of validation, 2 episodes of filamentous bulking and 1 episode of non-filamentous bulking happened at the Girona WWTP. The usefulness of the retrieved episodes in each one of the 3 new episodes was preliminarily validated. However, a global percentage of usefulness was not calculated due to the considerations stated above about the limitations of the current episode library. The partial results obtained from each of the 3 new episodes are summarized next:

- During the first episode of filamentous bulking (an episode of filamentous bulking caused by Type 021N), 86% of the retrieved cases corresponded to episode FB-7, an episode of filamentous bulking caused by Type 021N. During that episode, a non-specific control method was applied in an attempt to avoid the loss of solids with the effluent and the sudden variations in the organic load. The main control action applied was the manipulation of WAS. The retrieval of historical experiences was considered very useful for plant operators. For example, it could be noted that the manipulation of WAS was not enough to completely eliminate Type 021N.The remaining 24% of the retrieved cases corresponded to episode FB-8, an episode of filamentous bulking caused by *S. natans*. The information stored for this episode was also considered very useful for plant operators because for some days of the current episode *S. natans* was, together with Type 021N, one of the predominant filaments, mainly due to a parallel problem with the DO concentration. On such occasions, the retrieved control plans that were applied to increase the DO concentration proved useful for plant operators.

- During the second episode of filamentous bulking (an episode of filamentous bulking also caused by Type 021N), 67% of the retrieved cases corresponded to the recently stored episode of filamentous bulking caused by Type 021N (episode FB-9) which mainly provided plant operators with information about an inefficient control strategy where different flocculants were applied to decrease the presence of Type 021N. The rest of cases belonged again to the episode FB-7 commented above.
- During the episode of non-filamentous bulking, the lack of historical cases regarding such a problem, or at least containing information about the applied control strategy, made it impossible for the user to receive the necessary information to help them make a decision based on historical experience. Finally, once the plant manager decided which control strategy to carry out (helped by the control strategy suggested by the ES), the new experience of non-filamentous bulking was added to the episode library. The next time that an episode of non-filamentous bulking appears in the process, a historical episode will be available.

Even though a global evaluation can not be stated from this validation phase, a preliminary conclusion is the confirmation by plant operators that the information provided by the retrieval of complete episodes is indeed much more useful than the information provided by a unique case. The evolution of the process (stored in the retrieved episodes) whenever a control strategy was applied, was considered very useful to decide the most suitable control strategy to be applied. The usefulness but also the usability of the EBRS impinges directly (specially once the adequacy and accuracy of the system is completed) on the proposal and application of the most efficient control plans, which on most occasions imply a reduction of the response time of the process, with the consequent environmental and economical savings.

8.9 Global Validation of the Overall System

Apart from the partial validation of each one of the components of the KBDSS, including the data gathering and communication module, the ES module and the EBRS module, a general validation of the overall KBDSS also had to be performed. The validation of the overall KBDSS was carried out by means of a field validation at the Girona WWTP. Since the adequacy, accuracy and usefulness of the system had already been validated, in this phase only the usability of the system was validated, where usability was understood as the proper communication between the system and the user.

The validation of the overall system was based especially on the consistency of the results provided by the KBDSS. In the present section, a summary of the results obtained by the KBDSS throughout the field validation period (from January to August 2005) is presented. To complement the validation study, the performance of the overall KBDSS is illustrated with a real case-study. Thus, a general description of the results obtained from the real-time operation of the system at the Girona WWTP during a real episode of filamentous bulking are presented.

8.9.1 General results of the KBDSS

During the field validation period, 6 different episodes of solids separation problems and 4 episodes of normal state were identified. Figure 8.4 depicts the temporal distribution of the episodes throughout the period of validation, based on the percent certainty concluded by the system in diagnosing each type of problem. Table 8.5 summarizes the results obtained by the KBDSS regarding the type of solids separation problem diagnosed, the main cause of such problematic situations, the suggested expert strategy, the suggested strategy based on past experiences and finally the control strategy applied by the plant manager.

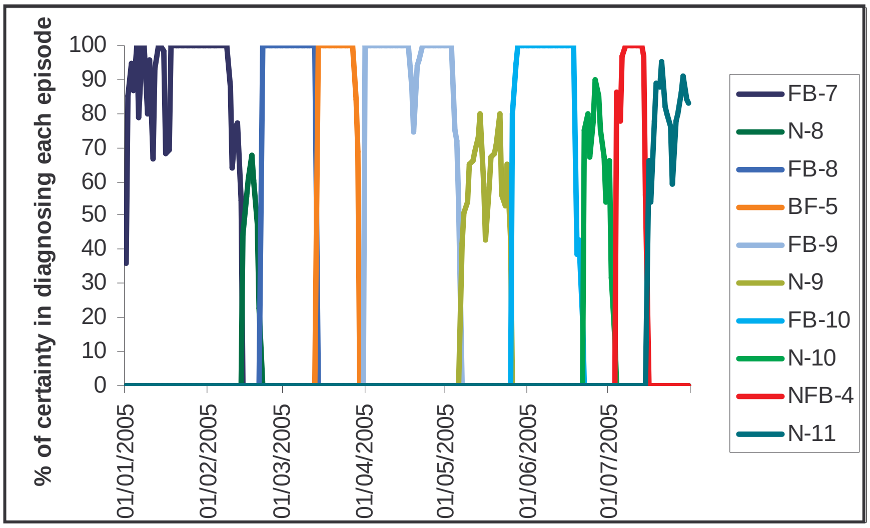


Figure 8.4 Temporal distribution of the different episodes identified by the KBDSS

Table 8.5 Summary of the results obtained by the KBDSS throughout the field validation period

EPISODE	DIAGNOSIS	CAUSE	SUGGESTED CONTROL STRATEGY		APPLIED CONTROL STRATEGY
			ES	EBRS	
FB-7	Filamentous bulking	Not identified but possibly high readily biodegradable substrate	Non-specific control methods (use of flocculants/coagulants)	-	Manipulation of WAS and addition of ferric sulphate
FB-8	Filamentous bulking	Low DO and high F/M	DO increasing and/or WAS decreasing	-	DO increasing, manipulation of WAS and addition of flocculants
BF-5	Biological foaming	Low DO and high F/M	DO increasing and/or WAS decreasing	-	Addition of defoaming agents and flocculants
FB-9	Filamentous bulking	High readily biodegradable substrate	Non-specific control methods (use of flocculants/coagulants)	Manipulation of WAS	Addition of flocculants and WAS manipulation
FB-10	Filamentous bulking	High readily biodegradable substrate	Non-specific control methods (use of flocculants/coagulants)	Addition of flocculants	Addition of flocculants
NFB-4	Non-filamentous bulking	High F/M and high readily biodegradable substrate	Non-specific control methods (use of flocculants/coagulants)	-	Addition of flocculants and WAS manipulation

-EPISODE FB-7: This episode of filamentous bulking lasted 40 days (from 01/03/05 to 02/13/05) and originated in the overabundance of Type 021N. The cause could not be properly identified because key variables referring to the possible causes, such as percent of readily biodegradable substrate or sulphide concentration, were not measured. As a consequence, the system suggested the periodic measurement of such variables, a suggestion that was taken up by plant operators. From the occurrence of this episode, the sulphide concentration began to be measured sporadically (2-3 times per month) and the percent of soluble COD was measured at least 1-2 times per week. Since the cause could not be properly identified, the system recommended the application of non-specific control methods: chlorination or flocculants/coagulants addition. Only on specific days when a high SRT was detected the system suggested to increase the WAS flow rate. The EBRS could not be used because the episode library did not yet include any historical experiences with reported control strategies for this type of problem. Finally, the plant manager decided to tackle this problematic situation on the one hand, by stopping the WAS flow rate for 15 days in order to avoid an excessive decrease in MLSS (a lot of solids were already being washed out by the effluent) and on the other hand, by applying ferric sulphide during the pre-treatment to decrease the high organic load entering the process.

-EPISODE FB-8: This episode of filamentous bulking was originated by an excessive abundance of *S. natans* mainly as a consequence of a high F/M ratio which eventually led to a low DO concentration (see Figure 8.5). The episode lasted 21 days (from 02/21/05 to 03/13/05). Once the possible causes of the episode were identified, the KBDSS recommended the use of specific control methods to restore the process based on increasing DO and decreasing WAS (to decrease F/M ratio) (see example in Figure 7.21). As in the last episode, no experience could be retrieved from the episode library. The strategy followed by the plant manager was two-folded. On the one hand, the DO set-point was increased up to 2.5 ppm in order to ensure that enough oxygen was available to microorganisms and the WAS flow rate was decreased to increase the MLSS concentration. On the other hand, different types of flocculants were tested to accelerate the improvement of sludge settleability.

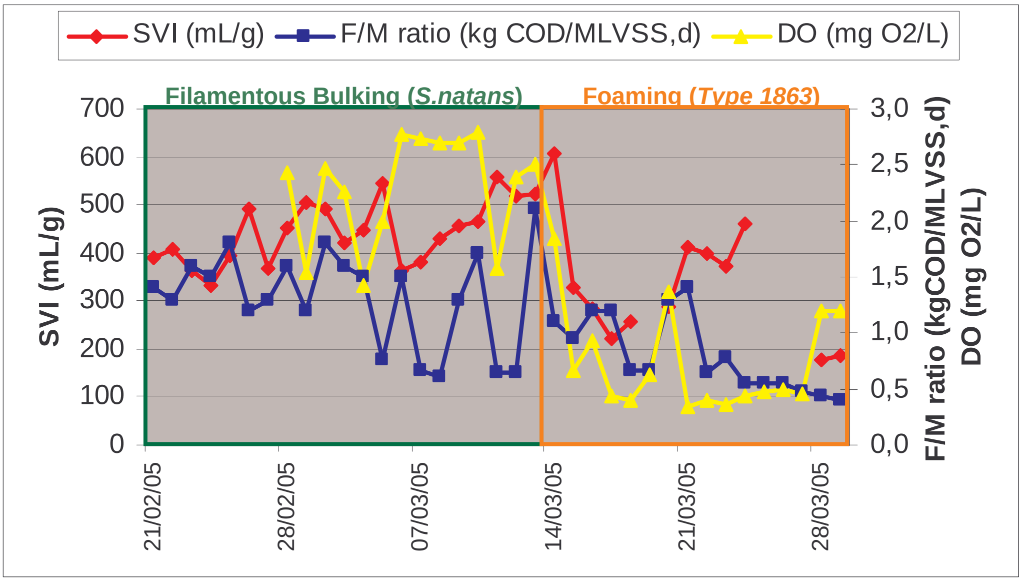


Figure 8.5 Evolution of SVI, F/M ratio and DO concentration throughout episodes FB-8 and BF-5

-EPISODE BF-5: The next episode of solids separation problems concerned an episode of biological foaming originated by an overabundance of Type 1863. This episode started right after the episode of filamentous bulking (it lasted from 03/14/04 to 03/29/05) due to a change in the predominant filament. As in the previous episode of filamentous bulking, the main cause of this problematic situation was the high F/M ratio, which eventually resulted in a deficiency of oxygen (see Figure 8.5). Accordingly, the KBDSS suggested to increase DO concentration by increasing the DO set-point and decreasing the WAS flow rate in order to decrease the F/M ratio. The control strategy finally applied was the addition of defoaming agents to control the production of foams, and the addition of flocculants to improve the sludge settleability.

-EPISODE FB-9: The next episode of filamentous bulking lasted 36 days in the process (from 04/01/05 to 05/06/05). The origin of the episode was the proliferation of Type 021N and the possible cause the presence of a readily biodegradable substrate. The control strategy suggested on the base of expert criteria was to use selectors to counteract the effects of the readily biodegradable substrate. This option was ruled out by the plant manager, and the KBDSS suggested then the use of non-specific control methods (chlorination or use of coagulants/flocculants) to restore the process. On the other hand, the strategy suggested by the KBDSS on the base of past experiences (since April 2005 the episode library could be used to retrieve historical experiences), included the manipulation of the WAS to avoid an excessive washout of microorganisms by the effluent. The control strategy eventually applied by plant operators included the addition of flocculants and the manipulation of the WAS.

-EPISODE FB-10: (See case-study)

-EPISODE NFB-4: The next episode involving a solids separation problem was an episode of non-filamentous bulking originated by an excessive production of EPS. It persisted during 12 days on the process (from 07/04/05 to 07/15/05). The high F/M ratio and the high percent of readily biodegradable substrate were again identified as the possible causes of the episode. The control strategy suggested by the KBDSS was based on expert criteria (no historical experiences regarding non-filamentous bulking were stored in the current episode library) and suggested on the one hand, the manipulation of WAS to equilibrate the F/M ratio and on the other hand, the addition of flocculants to facilitate sludge settleability. The plant manager finally decided to follow both strategies.

8.9.2 Case study

The presented case study includes a complete description of a real episode of filamentous bulking at the Girona WWTP together with the results obtained by the KBDSS and the effects of its conclusions on the evolution of the process.

8.9.2.1 Description of the episode

The problematic situation corresponds to an episode of filamentous bulking caused by the filament Type 021N (FB-10). The total length of the episode was 27 days; it started on May 26th, 2005 and finished on June 21st, 2005. Table 8.6 summarizes the most relevant variables that characterize the episode.

Table 8.6 Most relevant variables during the episode FB-10 ((1) and (2) see Figures 5.2 and 5.3)

DATE	SVI	TSS	%COD$_s$	COD/N	COD/P	F/M	SRT	DO	Sludge settleability (1)	Floc characteristics(2)
05/26/05	185	31.2	50.5	7.8	70.9	1.2	3.1	0.4	2	B
05/27/05	196	34.8				1.1	3.1	1.7	2	B
05/28/05	218					0.6	3.5	1.9		
05/29/05	355	12.4				0.6	4.2	1.8		
05/30/05	370	17				1.0	4.2	2.1	2	F
05/31/05	423	56	50.9			1.0	4.2	1.5	2	F
06/01/05	458	16				1.1	5.5	2.0	2	F
06/02/05	390	20.4	55.9	8.7	68.6	1.0	5.5	2.1	2	F
06/03/05	451	25.2				0.9	6.1	2.4	2	F
06/04/05	400					0.8	6.2	2.2		
06/05/05	373	15.2				0.7	5.9	2.1		
06/06/05	462	23.2		7.3	61.3	0.9	6.3	1.8	2	F
06/07/05	231	45.6				0.8	6.4	2.3	1	F
06/08/05	287	31.6	55.9	7.4	59.7	0.8	5.8	2.4	2	F
06/09/05	356	24.8	64.5			1.1	5.3	2.4	2	F
06/10/05	335	4.8				1.1	4.8	2.2	2	F
06/11/05	488					0.7	4.1	2.2		
06/12/05	438	13.6				0.7	4.1	2.2		
06/13/05	359	27.6		7.7	73.8	0.8	4.5	1.4	2	F
06/14/05	300	10.4	44.8			0.4	5.1	2.3	2	F
06/15/05	228	18.8		7.7	71.2	0.7	5.3	1.7	2	F
06/16/05	228	19.6	47.2			0.8	5.9	1.4	1	F
06/17/05	382	20.4				0.8	6.2	1.3	2	F
06/18/05	368					0.5	6.5	1.4		
06/19/05	154	9.2				0.5	7.1	2.3		
06/20/05	157	8		8.6	69.2	0.8	7.4	1.1	1	E
06/21/05	140	8.4	51.0			0.9	7.9	1.0	1	E

8.9.2.2 Results of the dynamic KBDSS

The main results obtained by the dynamic KBDSS during the whole episode are summarized in Table 8.7, which includes both the results of the dynamic ES (the diagnosis of the problem, the percent of certainty in the diagnosis, the identification of the cause and the suggested control strategy) and the results of the dynamic EBRS (the retrieved case, the retrieved episode and a summary of the applied control strategy).

Table 8.7 Summary of the results obtained by the dynamic KBDSS during the episode FB-10 (N.I.: not identified cause)

	DYNAMIC EXPERT SYSTEM				DYNAMIC EPISODE-BASED REASONING SYSTEM		
DATE	**Problem diagnosis**	**% cert.**	**Cause(s) diagnosed**	**Suggested control strategy**	**Retrieved case**	**Retrieved episode**	**Retrieved control strategy**
05/26/05	F.Bulking	80.0	rbCOD	Use of selectors + coag./floc. addition	04/06/05	FB-9	Flocculant addition + WAS manip.
05/27/05	F.Bulking	94.6	N.I.	Coag./floc. addition	01/14/05	FB-7	WAS manipulation
05/28/05	F.Bulking	100	N.I.	Coag./floc. addition	01/16/05	FB-7	WAS manipulation
05/29/05	F.Bulking	100	N.I.	Coag./floc. addition	01/19/05	FB-7	WAS manipulation
05/30/05	F.Bulking	100	N.I.	Coag./floc. addition	01/23/05	FB-7	WAS manipulation
05/31/05	F.Bulking	100	rbCOD	Coag./floc. addition	04/12/05	FB-9	Flocculant addition + WAS manip.
06/01/05	F.Bulking	100	N.I.	Coag./floc. addition	01/19/05	FB-7	WAS manipulation
06/02/05	F.Bulking	100	rbCOD	Coag./floc. addition	04/11/05	FB-9	Flocculant addition + WAS manip.
06/03/05	F.Bulking	100	N.I.	Coag./floc. addition	04/15/05	FB-9	Flocculant addition + WAS manip.
06/04/05	F.Bulking	100	N.I.	Coag./floc. addition	01/16/05	FB-7	WAS manipulation
06/05/05	F.Bulking	100	N.I.	Coag./floc. addition	01/23/05	FB-7	WAS manipulation
06/06/05	F.Bulking	100	N.I.	Coag./floc. addition	02/01/05	FB-7	WAS manipulation
06/07/05	F.Bulking	100	N.I.	Coag./floc. addition	04/20/05	FB-9	Flocculant addition
06/08/05	F.Bulking	100	rbCOD	Coag./floc. addition	04/15/05	FB-9	Flocculant addition + WAS manip.
06/09/05	F.Bulking	100	rbCOD	Coag./floc. addition	04/18/05	FB-9	Flocculant addition + WAS manip.
06/10/05	F.Bulking	100	N.I.	Coag./floc. addition	04/23/05	FB-9	Flocculant addition
06/11/05	F.Bulking	100	N.I.	Coag./floc. addition	04/23/05	FB-9	Flocculant addition
06/12/05	F.Bulking	100	N.I.	Coag./floc. addition	02/01/05	FB-7	WAS manipulation
06/13/05	F.Bulking	100	N.I.	Coag./floc. addition	04/20/05	FB-9	Flocculant addition + WAS manip.
06/14/05	F.Bulking	100	rbCOD	Coag./floc. addition	04/25/05	FB-9	Flocculant addition + WAS manip.
06/15/05	F.Bulking	100	N.I.	Change coag./floc. addition	05/02/05	FB-9	Flocculant addition
06/16/05	F.Bulking	100	rbCOD	Coag./floc. addition	05/05/05	FB-9	Flocculant addition
06/17/05	F.Bulking	100	N.I.	Coag./floc. addition	04/28/05	FB-9	Flocculant addition
06/18/05	F.Bulking	100	N.I.	Coag./floc. addition	05/07/05	FB-9	Flocculant addition
06/19/05	F.Bulking	38.6	N.I.	Coag./floc. addition	05/08/05	FB-9	Flocculant addition
06/20/05	F.Bulking	42.6	N.I.	Coag./floc. addition	05/09/05	FB-9	Flocculant addition
06/21/05	F.Bulking	20.0	rbCOD	Coag./floc. addition	05/12/05	FB-9	Flocculant addition

In order to fully understand the usability and usefulness of the overall system, together with its global field performance, a detailed description of the results obtained by the system and of the evolution of the problematic situation is described next:

⇒ From **May 18th to May 25th** (before the episode FB-10 started) the ES concluded that the plant was in a *normal situation* (most relevant variables presented normal values). Nevertheless, on May 23rd and on May 24th the ES predicted that the process was at risk of filamentous bulking (see Figure 8.6), the SVI trend was increasing and the value of floc characteristics described a common abundance of filaments per floc (category B). Subsequently, an increase on the percent of soluble COD was identified as the possible origin of the filamentous bulking risk.

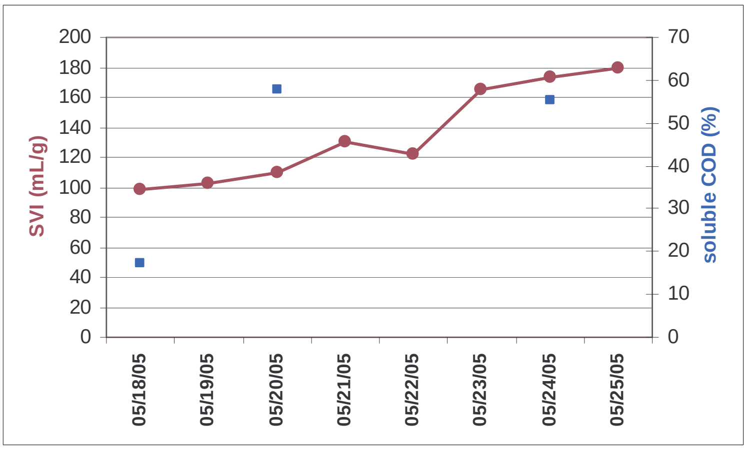


Figure 8.6 Evolution of SVI and percent of soluble COD during the previous period to episode FB-10

⇒ On **May 26th** the ES finally identified (with 80% certainty) the beginning of an episode of filamentous bulking originated by Type 021N (see Figure 8.7). Once the problem was identified, deficiency of nutrients, low F/M ratio and septic conditions were ruled out as possible causes and finally the presence of a high readily biodegradable substrate was identified as the potential cause of the filamentous bulking situation, and more concretely as the potential source of the excessive proliferation of Type 021N (see Figure 8.8).

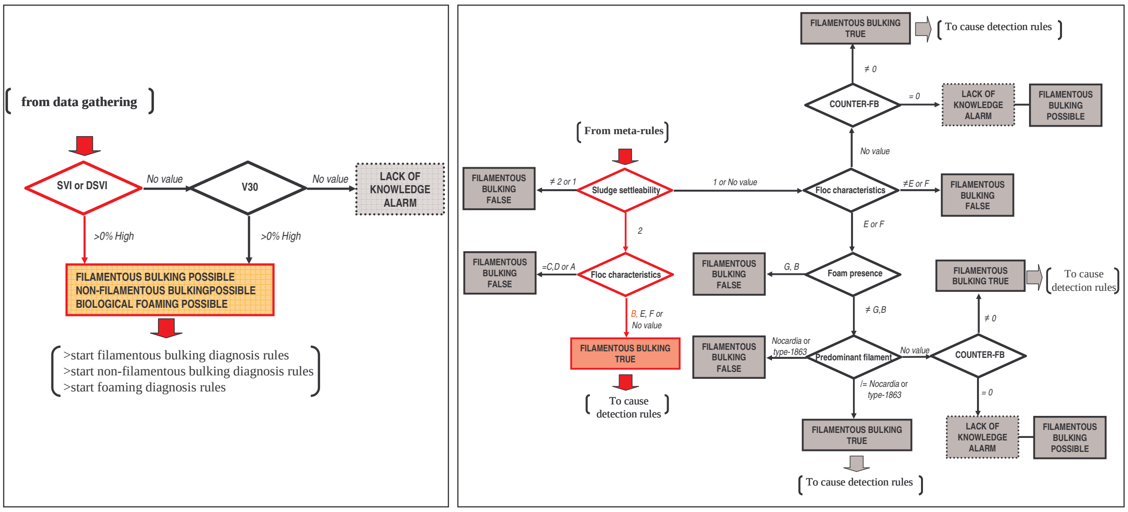


Figure 8.7 Reasoning process followed by the ES to diagnose the problem: meta-rules and filamentous bulking diagnosis decision trees

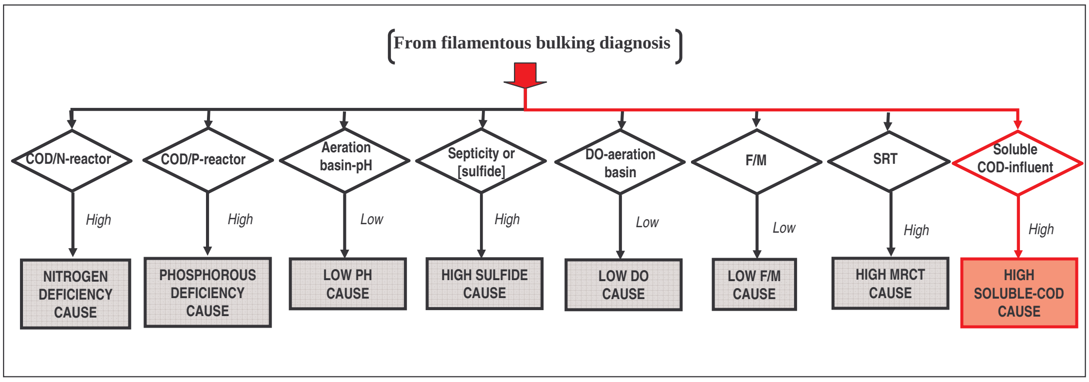
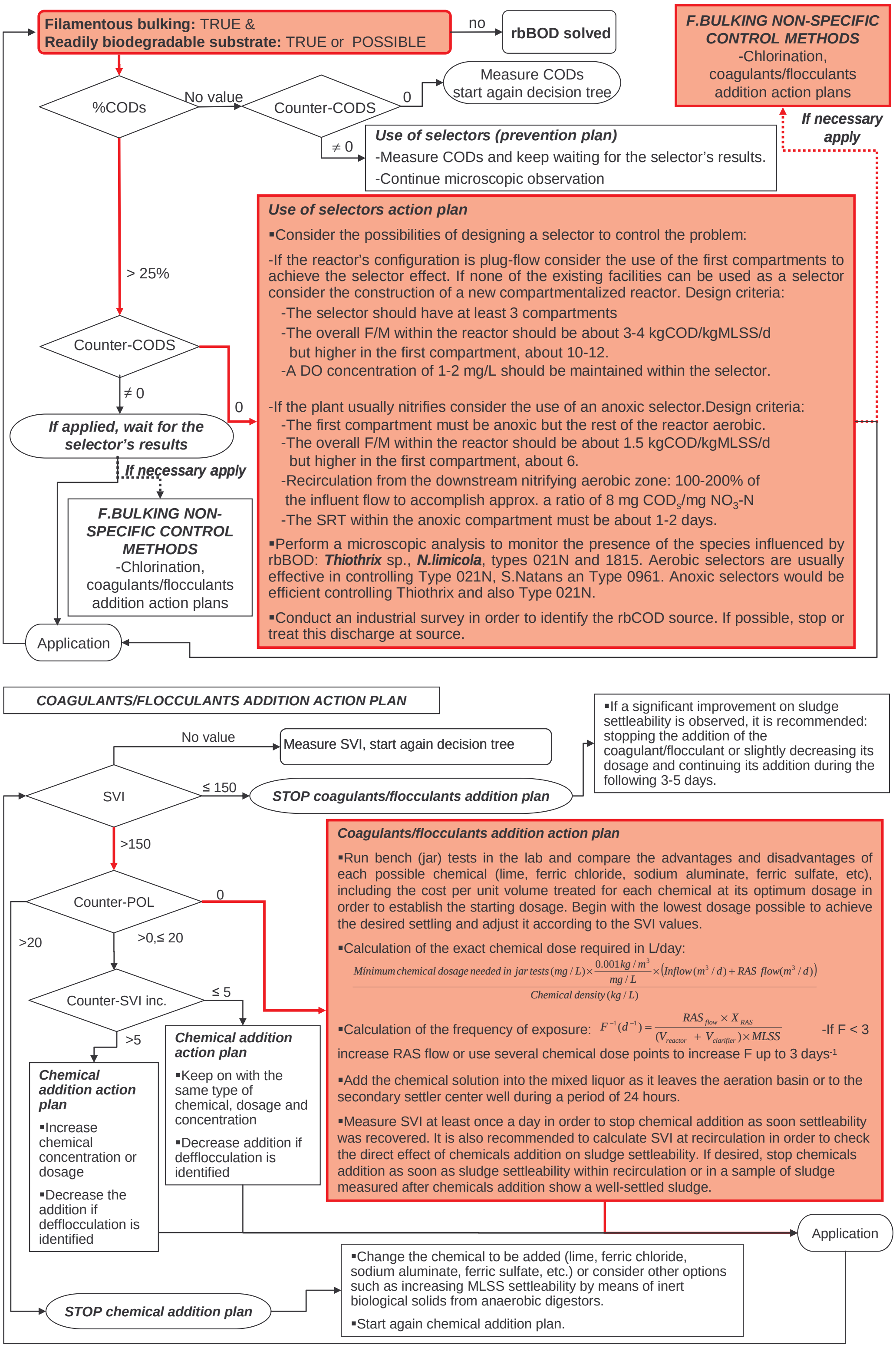


Figure 8.8 Reasoning process followed by the ES to diagnose the cause: filamentous bulking causes decision tree

According to these results, the ES suggested the *use of selectors action plan* as the most appropriate strategy to solve the situation. The suggested plan included a list of possible actions such as the design and installation of a selector, the identification of the source of the readily biodegradable substrate and the use of non-specific control methods (Figure 8.9).



Use of selectors (prevention plan)
-Measure CODs and keep waiting for the selector's results.
-Continue microscopic observation

Use of selectors action plan

▪Consider the possibilities of designing a selector to control the problem:

-If the reactor's configuration is plug-flow consider the use of the first compartments to achieve the selector effect. If none of the existing facilities can be used as a selector consider the construction of a new compartmentalized reactor. Design criteria:
- -The selector should have at least 3 compartments
- -The overall F/M within the reactor should be about 3-4 kgCOD/kgMLSS/d but higher in the first compartment, about 10-12.
- -A DO concentration of 1-2 mg/L should be maintained within the selector.

-If the plant usually nitrifies consider the use of an anoxic selector.Design criteria:
- -The first compartment must be anoxic but the rest of the reactor aerobic.
- -The overall F/M within the reactor should be about 1.5 kgCOD/kgMLSS/d but higher in the first compartment, about 6.
- -Recirculation from the downstream nitrifying aerobic zone: 100-200% of the influent flow to accomplish approx. a ratio of 8 mg COD_S/mg NO_3-N
- -The SRT within the anoxic compartment must be about 1-2 days.

▪Perform a microscopic analysis to monitor the presence of the species influenced by rbBOD: ***Thiothrix*** sp., ***N.limicola***, types 021N and 1815. Aerobic selectors are usually effective in controlling Type 021N, S.Natans an Type 0961. Anoxic selectors would be efficient controlling Thiothrix and also Type 021N.

▪Conduct an industrial survey in order to identify the rbCOD source. If possible, stop or treat this discharge at source.

▪If a significant improvement on sludge settleability is observed, it is recommended: stopping the addition of the coagulant/flocculant or slightly decreasing its dosage and continuing its addition during the following 3-5 days.

Coagulants/flocculants addition action plan

▪Run bench (jar) tests in the lab and compare the advantages and disadvantages of each possible chemical (lime, ferric chloride, sodium aluminate, ferric sulfate, etc), including the cost per unit volume treated for each chemical at its optimum dosage in order to establish the starting dosage. Begin with the lowest dosage possible to achieve the desired settling and adjust it according to the SVI values.

▪Calculation of the exact chemical dose required in L/day:

$$\frac{Mínimum\ chemical\ dosage\ needed\ in\ jar\ tests\,(mg/L)\times\frac{0.001\,kg/m^3}{mg/L}\times\left(Inflow\,(m^3/d)+RAS\ flow(m^3/d)\right)}{Chemical\ density\,(kg/L)}$$

▪Calculation of the frequency of exposure: $F^{-1}(d^{-1})=\frac{RAS_{flow}\times X_{RAS}}{(V_{reactor}+V_{clarifier})\times MLSS}$ -If F < 3 increase RAS flow or use several chemical dose points to increase F up to 3 days^{-1}

▪Add the chemical solution into the mixed liquor as it leaves the aeration basin or to the secondary settler center well during a period of 24 hours.

▪Measure SVI at least once a day in order to stop chemical addition as soon settleability was recovered. It is also recommended to calculate SVI at recirculation in order to check the direct effect of chemicals addition on sludge settleability. If desired, stop chemicals addition as soon as sludge settleability within recirculation or in a sample of sludge measured after chemicals addition show a well-settled sludge.

Chemical addition action plan

▪Increase chemical concentration or dosage

▪Decrease the addition if defflocculation is identified

Chemical addition action plan

▪Keep on with the same type of chemical, dosage and concentration

▪Decrease addition if defflocculation is identified

▪Change the chemical to be added (lime, ferric chloride, sodium aluminate, ferric sulfate, etc.) or consider other options such as increasing MLSS settleability by means of inert biological solids from anaerobic digestors.

▪Start again chemical addition plan.

Figure 8.9 Reasoning process followed by the ES to suggest a specific control plan: (top) use of selectors action plan and (bottom) use of coagulants/flocculants control plan

On the other hand, the working cycle of the EBRS retrieved the case 04/06/05 (a case of filamentous bulking originated by Type 021N and caused by a high readily biodegradable substrate) and the whole episode FB-9 as the most similar historical experiences from the episode library. From this past experience, apart from the value of the most relevant attributes that define each case, the control strategy and the related comments were retrieved. Besides, the information retrieved in each one of the cases belonging to episode FB-9 was also available for the plant manager. Table 8.8 shows the control action, the corresponding comments and the global episode lesson retrieved from the historical data base.

Table 8.8 Control strategy and comments from case 04/06/05 and lesson from episode FB-9

Case-control-strategy	ZETAG-7689 (0.3% concentrated) is added at a dosage of 300L/h to facilitate sludge settleability. Besides, for 10 hours 98 m^3/h (980 m^3/d) of secondary sludge is recirculated up to the primary clarifiers to facilitate sludge removal and to decrease the high oxygen demand.
Case-comments	A lot of *Spirochaetes* have appeared as a consequence of the oxygen deficiency. Sludge settleability does not improve.
Episode-lesson	Two different flocculants were applied during the episode: ZETAG-7689 (0.3% conc.) at a dosage of 300L/h and ZETAG-63 (0.3% conc.) at a dosage of 600 L/h does not seem to provide a significant improvement on the sludge settleability. In the end, the problematic episode ended as a consequence of the filaments competition. The low DO throughout the last 8 days of the episode, made *S. natans* to be more predominant than Type 021N. The increase of the DO set-point up to 2.5 ppm together with the occurrence of a rainy period made *S. natans* disappear, and the sludge settleability recovered slightly. The recirculation of sludge to primary settlers (WAS increase) during several days of the episode, avoided an excessive loss of solids.

Once the plant manager received the results from both systems, she decided that the best strategy to follow was the addition of flocculants. The use of selectors was at the moment ruled out due to plant limitations, but it will be considered for the future enlargement of the plant. Besides, thanks to the experience gained from past episodes of filamentous bulking also caused by a possible readily biodegradable substrate, the plant manager decided to carry out a local survey to study which industrial facility from the surrounding area could be the source of the high concentration of readily biodegradable substrates.

The selected control strategy was initiated by applying the flocculant ZETAG-63, at a dosage of 600 L/h, but at a higher concentration (0.6%) than previous experiences. This new information and the corresponding comments were then added to the episode library.

⇒ On **May 27th**, the dynamic ES identified the current situation as the second day of the initiated episode FB-10. This time, the percent of certainty increased up to 94.6%. Now, the SVI was higher than the day before. Regarding the cause identification, the ES could not reach a definitive conclusion because the value of soluble COD was not available. Hence, the system recommended the user to measure the value of soluble COD in order to verify the possible cause.

Meanwhile, it was proposed (see Decision tree at the bottom of Figure 8.8) to keep on with the addition of the same flocculant that was applied on the previous day (same type of flocculant, same dosage and same concentration), because no significant increasing trend in SVI was observed.

The results of the EBRS involved the retrieval of case 01/04/05 (a case of filamentous bulking caused by type 021N with no identified cause) and the whole episode of FB-7 as the most similar historical experiences. Table 8.9 depicts the applied control action, the corresponding comments and the global episode lesson retrieved from the historical data base.

Table 8.9 Control strategy and comments from case 01/04/05 and lesson from episode FB-7

Case-control-strategy	WAS flow rate was stopped in order to increase the MLSS concentration and to slightly decrease the F/M ratio.
Case-comments	A low oxygen demand is observed as a consequence of the low MLSS concentration. The loss of solids with the effluent makes this situation worse.
Episode-lesson	The manipulation of WAS was not enough to control the proliferation of filaments. During the present episode, stopping the WAS flow rate for 15 days made the sludge become very old; as a consequence, the proliferation of filaments increased during the following days. Once again, a final episode of heavy storm caused a decrease on the organic load (and consequently a lower % of readily biodegradable COD). This fact resulted in a decrease on the population of type 021N. The sludge settleability improved in the following 3-5 days, although Type 021N remained present in the sludge (floc characteristics=B, common abundance of filaments).

According to these results, the plant manager decided to keep on applying the same non-specific control method (the addition of ZETAG-63 (0.6% concentrated) at 600 L/h.

⇒ On **May 28th and 29th**, the system did not have qualitative data available (qualitative variables were not measured on weekends), but despite this lack of information, the dynamic reasoning (see Figure 8.10) allowed the KBDSS to diagnose a filamentous bulking situation, identified as days 3 and 4 of episode FB-10. Regarding the evolution of the problem, the system also identified an *increase* in the SVI trend and a decrease in the loss of solids on day 3, as well as a stationary trend in the TSS on day 4. On the other hand, since the value of soluble COD was not available, the cause of the situation could not be properly identified. Nevertheless, the system suggested it as the most probable cause and recommended again its measurement to perform a proper verification. According to these results the ES recommended to keep on with the application of a non-specific control method (same type, dosage and concentration of flocculant).

Finally, the results of the EBRS on day 3 and 4, resulted in the retrieval of cases 01/16/05 and 01/19/05 of episode FB-7, where the applied control strategy continued with WAS manipulation to increase the MLSS concentration. Considering these results, the plant manager decided to keep on with the addition of flocculants, as suggested by the ES. The manipulation of WAS was rejected due to the inefficient results obtained in the past. The flocculant ZETAG-63 (0.6% concentrated) at 600 L/h was applied on days 2 and 3 of the episode (same type, dosage and concentration as suggested by the ES). These new experiences were also added to the historical database.

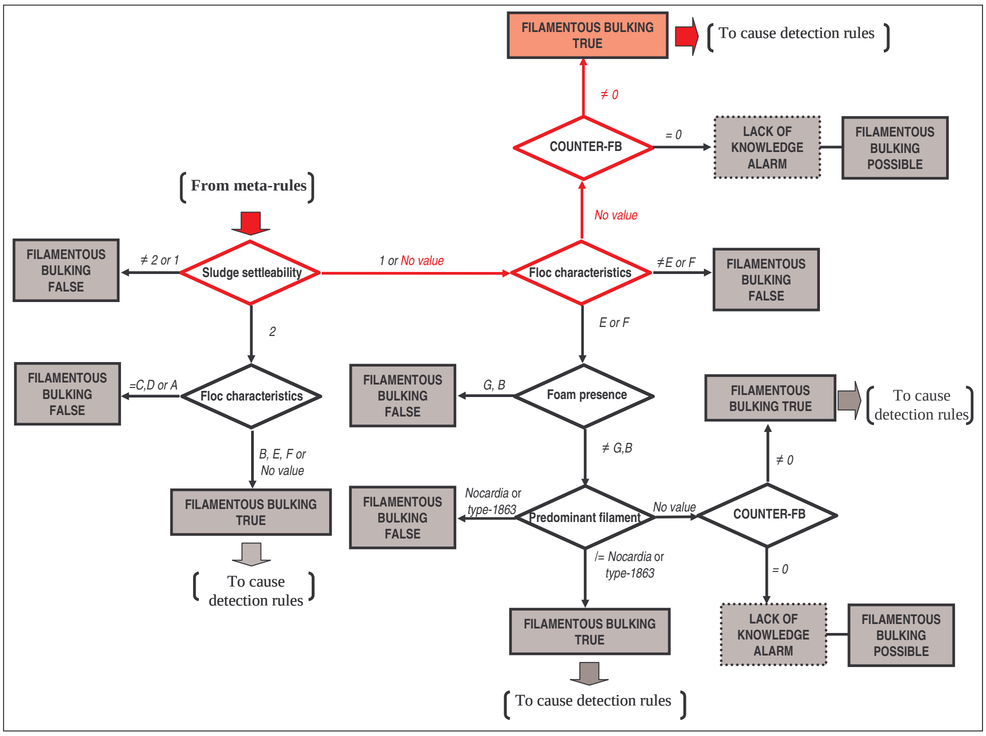


Figure 8.10 Reasoning process followed by the ES to diagnose a filamentous bulking situation

⇒ From **May 30th to June 14th**, the ES identified the same situation of filamentous bulking with 100% certainty. The high percentage of readily biodegradable substrate was still identified as the possible cause of the proliferation of Type 021N and the application of non-specific control methods was suggested as the most appropriate control strategy. Likewise, the results of the EBRS allowed the plant operator to revise the control actions applied on the past experiences of the similar episodes FB-7 and FB-9 (lessons of the episodes described above).

According to these results, the plant manager decided to keep on applying the same control strategy regarding flocculants addition. Following the *flocculants addition action plan* suggested by the ES and considering that the value of SVI trend did not show an increasing trend for more than 5 days, the ES did not suggest increasing the flocculant dose or concentration. On the other hand, on June 7th the system identified a high loss of solids within the effluent (TSS = 45.6 mg/L) as a possible consequence of a flocculant overdose, and it suggested decreasing the dosage or concentration of the applied flocculant (see the decision tree at the bottom of Figure 8.8). Accordingly, on June 7th the plant operator decided to stop the flocculant addition at night (from 10 pm to 8 am), when the influent flow and the organic load were lower. This strategy was followed from June 7th to June 14th. On the other hand, according to the specific experience retrieved from episode FB-9, during some days of the present episode, part of the secondary sludge was recirculated up to primary settlers with the intention of increasing the potential of the facility to eliminate more sludge (i.e., increasing the WAS). This action helped to counteract the effects of the inefficient sludge settleability on the days when a concentration of MLSS higher than 1300-

1500 mL/g could have resulted in an overwhelming loss of solids with the effluent. In Figure 8.11 the evolution of WAS and the effluent TSS throughout the whole episode are represented. In general, after an increase in the WAS flow rate (WAS flow rate > 1200 m^3/d indicates recirculation to primary settlers) it can be observed how the effluent TSS decreases. All these new experiences were also retained in the episode library.

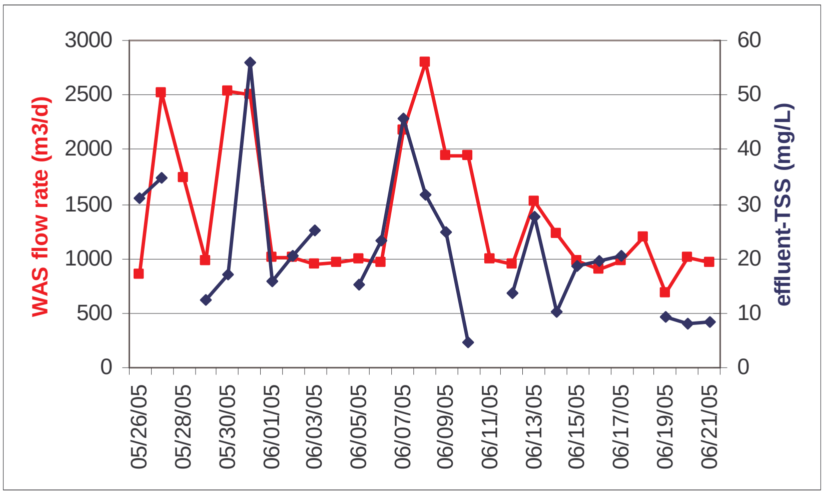


Figure 8.11 Evolution of WAS flow rate and effluent TSS during the current episode

⇒ On **June 15th**, the ES kept on identifying the same situation of filamentous bulking originated by Type 021N and caused by a readily biodegradable substrate. Nevertheless, according to the *flocculants addition control plan*, the limit of days during which the applied strategy should have produced a significant improvement in sludge settleability had been exceeded (June 15th was the day 21 of the filamentous bulking episode). Consequently, the ES suggested to discontinue the initiated plan (see decision tree at the bottom of Figure 7.7), and either change the type of coagulant/flocculant or consider other non-specific control strategies.

Given this new information, the plant manager decided to change the type of flocculant. The addition of ZETAG-63 was stopped and the flocculant ZETAG-7689 was used instead. The information retrieved from past experiences, made the user decide to initiate the treatment with a higher dosage: 600 L/h (0.6% concentrated). In order to avoid deflocculation, the addition of flocculant was also stopped at night.

⇒ On **June 16th, 17th and 18th**, the same situation was still identified but a decreasing trend was detected in SVI, meaning a slight improvement on the sludge settleability, while the trend of the effluent TSS was stationary, without exceeding the limit of 35 mg/L. Given this positive results, the suggested control plan was to keep on with the same type of flocculant and the same dosage.

Given these results, the plant manager decided to continue with the same control strategy. This time, also according to the information retrieved from the past cases, the plant operator decided not to recirculate the secondary sludge, given that the sludge settleability was improving, and no risk of loss of solids was identified.

⇒ On **June 19th, 20th and 21st**, the ES system finally detected a significant improvement in the process. Filamentous bulking was still identified as the current situation, but now the percent of certainty decreased to 38.6%, 42.6% and 20.0% respectively, given that the SVI values had decreased considerably. Actually, the ES identified an improvement in the process (trend to normal situation) given the decreasing trend of both SVI and TSS (see Figure 8.12). Nevertheless, according to the reasoning process depicted in the decision tree of Figure 8.10, the presence of a very common abundance of filaments (category E of floc characteristics) led the ES to decide that the process was still suffering from the same problematic situation (days 25, 26 and 27 of the episode FB-10).

In accordance with these results, on June 19th and 20th the ES recommended keeping on with the same control strategy (same type of flocculant and same dosage). On June 21st though, given that the SVI value had significantly recovered (<150 mg/L), the system suggested either to stop the *flocculants addition control plan* by stopping the addition of the flocculant or to keep on adding it for the following 3-5 days at a lower dosage (see decision tree at the bottom of Figure 8.8).

Given all this new information, together with the experiences retrieved by the EBRS, the plant manager decided to keep on applying the same dosage and concentration of ZETAG-7689 also during June 21st in order to ensure the improvement of sludge settleability.

⇒ On **June 22nd**, the system finally diagnosed a normal situation and consequently the end of episode FB-10. The SVI value was now 118 mg/L and the qualitative variables (sludge settleability = 1, meaning a proper sludge settleability, and Floc characteristics = B, meaning a common abundance of filaments) characterized a significant improvement in the sludge settleability. According to theses results, the plant operator decided to keep on applying ZETAG-7689 at 600 L/h during the following 4 days and a minor dosage (400 L/h) during the next 3 days. The preventive control strategy was stopped on June 29th. Figure 8.12 shows the evolution of SVI and effluent TSS throughout the whole episode. Figures 8.13, 8.14 and 8.15 show some pictures of the floc characteristics at different stages of the episode.

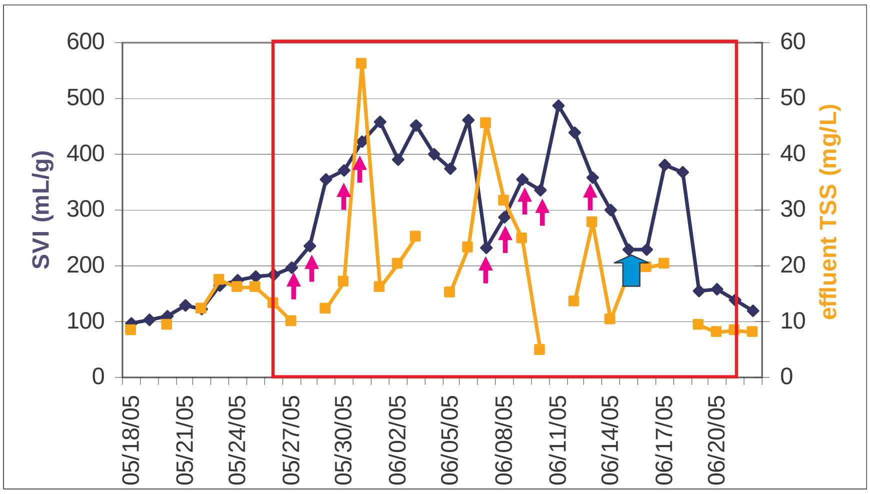


Figure 8.12 Evolution of SVI and effluent TSS throughout the whole episode FB-10. The red box marks the beginning and end of the episode, the pink arrows mark WAS manipulation and the blue arrow indicates the day the applied flocculant was changed.

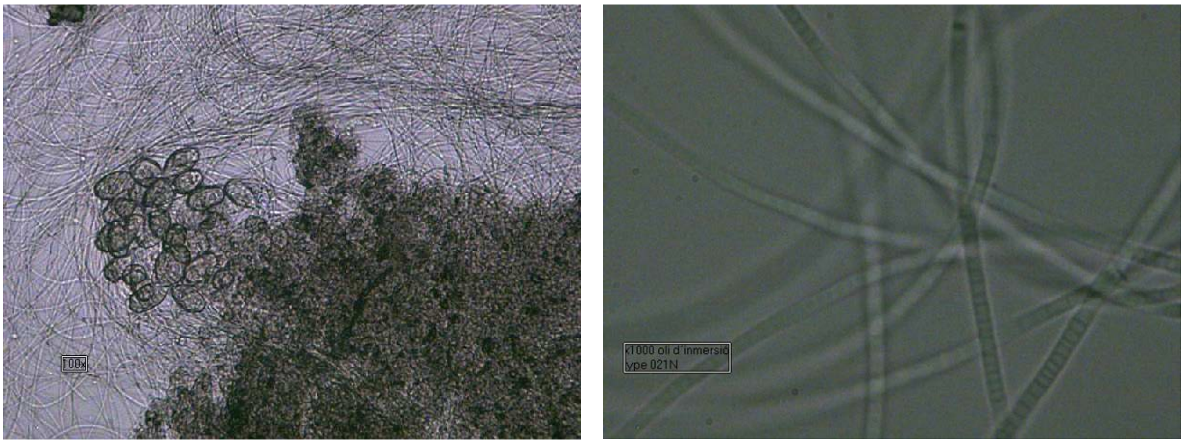


Figure 8.13 Phase contrast micrographs of floc characteristics at the beginning of the episode: (left) Excessive abundance of filaments (Floc characteristics = F (more than 20 filaments per floc)) on May 31st (100x magnification) and (right) detailed micrograph of the predominant filament: Type 021N (1000x magnification).

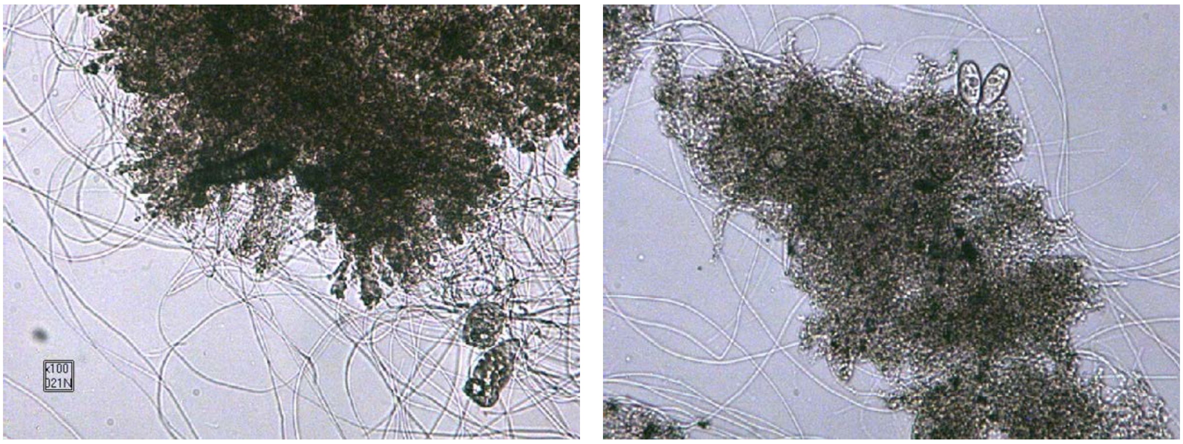


Figure 8.14 Phase contrast micrographs of floc characteristics at the final stage of the episode: (left and right). Very common abundance of filaments (Floc characteristics = E (5-20 filaments per floc)) on June 20th. The effects of the flocculant, which gives the floc a compact consistency, can be observed. Thus, despite the high number of filaments, the floc settles more easily (100x magnification).

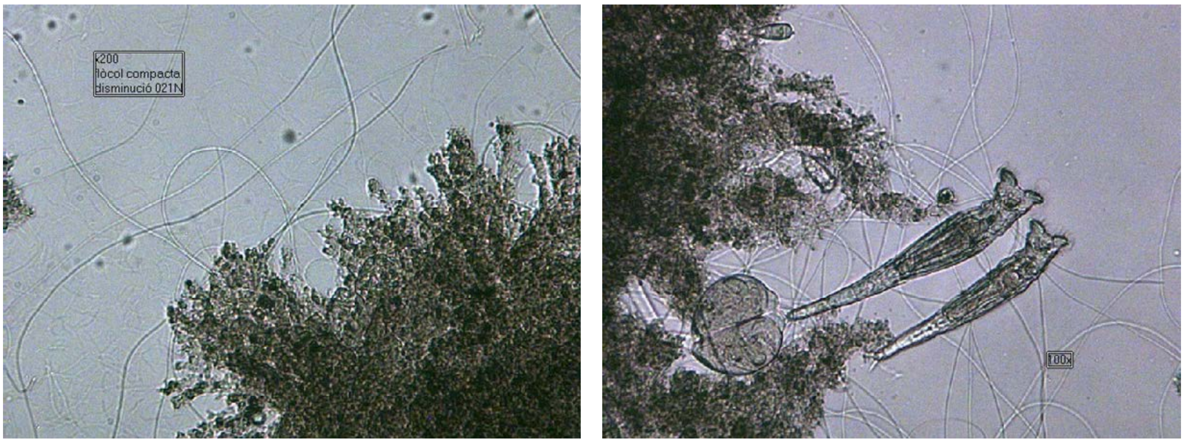


Figure 8.15 Phase contrast micrographs of floc characteristics at the end of the episode: (left and right) Common abundance of filaments (Floc characteristics = B (<5 filaments per floc)) on June 23th. The compactness of the floc and the fewer number of filaments can also be observed. Some stalked ciliates and rotifers are observed as a consequence of the high sludge age of the last days of the episode (100x magnification).

8.9.2.3 Registration of the new episode

Once the system identified the end of the episode, it was time to retain the new experience in the episode library as episode FB-10. The partial experiences (case-control-strategy and case-comments) gained from the daily operation throughout the episode, were registered every day. Now, with the conclusion of the episode, the general lesson completes the piece of knowledge provided by this new experience. The lesson registered in the depicted case study, is showed as an example of the important experience provided by the new experience.

The lesson registered in episode FB-10 is the following:

Two different flocculants have been applied throughout the episode: ZETAG-63 and ZETAG-7689. At the end, it seems that the flocculant ZETAG-7689 (0.6% concentrated), at a dosage of 600 L/h, was effective after 7 days of application. Although Type 021N did not disappear completely, at least sludge settleability improved considerably. After SVI decreased down to 150 mg/L, the flocculant was still applied during an extra week (the last 3 days the dosage was decreased to 400 L/h). On the other hand, in order to prevent the deflocculation of sludge as a consequence of flocculant overdose, during the whole episode, the addition of flocculant was stopped at nights (from 1 am to 8 am) when the organic load and the influent flow were lower. Besides, in order to avoid the loss of solids with the effluent, part of the secondary sludge was periodically recirculated to primary settlers to increase WAS flow. Finally, it seems that the cause of the extreme proliferation of Type 021N is the high percent of readily biodegradable substrate. An industrial survey is being performed in order to identify the source of this type of substrate. The slaughterhouse is suspected to be a major source of the readily biodegradable substrate.

8.10 Discussion

The validation process has been considered as a multi-stage period covering the entire life cycle of the KBDSS and including both the development and implementation phases. It has been extended during a period of 11 months (4 months of experimental validation and 7 months of field validation), whereby the system was continuously upgraded, according to the partial results obtained from the validation of the following aspects: reliability of communication and data gathering; consistency and completeness of decision trees; accuracy of ES diagnosis; usefulness of the expert control plans; adequacy of the episode library; accuracy of case and episode retrieval; usefulness of the retrieved episodes and global validation of the overall system.

Regarding the accuracy of the ES in diagnosing and predicting solids separation problems of microbiological origin, once the specific modifications derived from the experimental and field validation phase were applied, an accuracy of almost 100% in diagnosing solids separation problems and about 80% of accuracy in predicting future problematic events were obtained.

The validation of the ES diagnosis regarding the identification of the possible causes of solids separation problems resulted in an efficiency of just 43.3 % of correctly identified causes. The origin of this 56.7% of unidentified causes (meaning not identified causes rather than incorrectly identified causes) is certainly the lack of data that hinders the verification of all the possible causes (apart from the existence of open question related to the possible causes affecting solids separation problems) . In such occasions, it can be concluded that an increase in the frequency of measurement of such variables will absolutely increase the efficiency of the system in diagnosing the possible causes of solids separation problems. This was therefore recommended to plant operators in order to improve the KBDSS performance in identifying the possible causes and the prediction of future problems based on the trends of such variables.

In spite of our attempt to select the most commonly measured variables (based on the results presented in chapter 3), in the end the set of input variables used by the system to identify the possible cause can not be modified, their measurement is strictly necessary and no other variables can be measured instead. In order to overcome these limitations and whenever a possible cause cannot be properly checked, a set of complementary actions such as the hypothesis of probable causes (according to the predominant filament) or the suggestion of preventive control strategy are provided.

The validation of the control plans suggested by the ES resulted in 91.5% of usefulness of the suggested control plans, from which 57% exactly correspond to the control strategies finally applied. In particular, the proposal of complete control plans and the continuous adjustment of the suggested strategies, where the exact required doses of chemicals or the exact control parameters such as the target WAS flow or DO set-point were included, was evaluated as being very useful for plant operators.

As expected, at the beginning of the field validation phase, the adequacy of the episode library was very low, but it increased gradually as new episodes were incorporated to the episode library. Nevertheless, at the end of the validation phase, the adequacy could not be considered as being complete, given that for example, episodes of dispersed growth or pin-point floc (never experienced in the facility) were not represented. Regarding the representation of rising sludge episodes, 3 historical episodes of rising sludge could be reconstructed by studying historical data and with the support of plant operators.

Regarding the accuracy of case and episode retrieval and once the specific modifications resulted from the first phase of validation were applied, an accuracy of 97% in retrieving similar episodes and cases was obtained.

A general result from the usefulness of the retrieved episodes could not be represented (mainly due to the inexistence of an adequate episode library to critically validate the usefulness of the stored episodes). Just a partial usefulness of 86% and 67% were obtained during the 2 episodes of filamentous bulking that could be validated.

It can be confirmed that, according to plant operators, the information provided by the retrieval of complete episodes is indeed much more useful than the information provided by a unique case. The evolution of the process (stored in the retrieved episodes) after a control strategy is applied was considered very useful to decide the most suitable control strategy to follow. The usefulness of the EBRS, but also its usability, directly affects (specially once the validation of the adequacy and accuracy of the system is completed) the proposal and application of the most efficient control plans, which on most occasions imply a reduction in the response time of the process, with the consequent environmental and economical savings.

Validation of Decision Support Systems is a challenging issue in the domain of Artificial Intelligence. The complexity of such applications, especially when applied within environmental domains, makes finding a standard methodology to validate any DSS difficult. Further research on this issue should be addressed to the development of standard methodologies such as sensitivity analysis to determine how sensitive the results of a DSS are to missing data, erroneous data, misclassification of the diagnosis, etc. This validation would assess how robust the system is and how sensitive results are to changes in the assumptions.

In general, it was confirmed the difficulties in validating a KBDSS, where the use of supporting tools such as simulators is generally not possible. As a result, during the day-to-day validation, some constraints were found, among them: all the type of solids separation problems initially considered did not occur during the phase of validation (e.g. episodes of dispersed growth or pin-point floc never happened during the field validation) and thus the system could not be completely validated in diagnosing, predicting or suggesting solutions for such type of situations; a complete implication of the user was not always obtained, specially during the registration of the information requested to retain the new experiences in the EBRS; at the beginning of the validation process, the user was not confident enough to directly apply the control strategies suggested by the system. According to all these facts and limitations, the presented results obtained from the validation phase should not be stated as definitive conclusions but as the preliminary results which has allowed us to improve the feasibility of the KBDSS.

CHAPTER 9

CONCLUSIONS

The most important conclusions arising from the present thesis are the following:

- A Knowledge-Based Decision Support System to support plant operators in handling solids separation problems of microbiological origin was designed, developed, implemented and validated in a full-scale WWTP.

- In order to evaluate the extent of solids separation problems in activated sludge systems a literature review, regarding the characteristics of the problems and the available methods to control the occurrence of these undesired situations, was carried out. The following are the most relevant conclusions derived from this revision:

    - Solids separation problems, including filamentous bulking, non-filamentous bulking, biological foaming, dispersed growth, pin-point floc and rising sludge are among the most common and problematic operational problems occurring in conventional activated sludge systems.

    - Despite the extensive research on the causes and control of solids separation problems, general guidelines to irrefutably prevent and manage these undesired situations have not been published yet. Likewise, feasible mathematical models capable of simulating and predicting the occurrence of solids separation problems are still not available.

    - Whenever solids separation problems occur, experts tend to combine both “literature” and “experiential” knowledge to face these situations.

- The conclusions reached during the specific study in Catalonia have contributed to the design of a viable and realistic KBDSS. First of all, the necessity of a decision support tool to handle solids separation problems was confirmed after considering the existing difficulties in efficiently managing these kind of situations. Secondly, according to the results of the study, the most commonly measured variables were selected as the input information for the suggested KBDSS and only the most reliable control methods were considered as possible solutions. The following set of conlusions were obtained from this local study:

    - In Catalonia, filamentous bulking and biological foaming are the two types of solids separation problems that most commonly affect the correct operation of conventional activated sludge systems.

    - Among the 13 studied WWTPs in Catalonia, the predominant filamentous bacteria involved in episodes of filamentous bulking and foaming are (in order): nocardioforms, Type 021N, *M. parvicella*, Type 0041 and *N. limicola*.

    - The probable causes of solids separation problems in the studied plants of Catalonia are: sudden variations in the organic load, variations in the DO concentration, low temperatures, nutrient deficiency and high percentages of readily biodegradable substrate (in episodes of bulking and foaming), and the arrival of toxics or a consequence of overchlorination (in episodes of deflocculation).

  - Also in Catalonia, the most common parameters generally used to monitor sludge settleability are the V30 test and the calculation of the SVI. Microscopic analysis is also commonly used as a monitoring tool mainly to study floc characteristics and abundance of filaments. Its main shortages are, firstly, the low frequency of measurement, mainly attributed to lack of time or skills of plant operators, and secondly the lack of standardisation in registering such information. Moreover, some of the analytical data and operational parameters necessary to monitor the possible cause of solids separation problems should be measured at a higher frequency. Nowadays, some key variables such as sulphide concentration or soluble COD are not generally being measured at any of the studied plants.

  - Non-specific control methods such as chlorination or the addition of coagulants or flocculants and the control of DO or manipulation of WAS flow were the most common control strategies applied in the studied plants. The selection of the control method usually depended on the cost of their application and on the characteristics or limitations of the facility. In general, plant operators tend to apply those plans that in the past helped them to restore a similar problematic situation.

- The use of the Domino Model was crucial during the conceptual design of the dynamic KBDSS, as it facilitates: (1) an intuitive visualization and a direct characterization and organization of the different knowledge bases and reasoning processes necessary to develop the KBDSS; (2) the possibility of structuring the suggested control plans into a well-organized and scheduled set of actions, and (3) the use of any suitable technique that would allow the system to accomplish the pre-defined objectives and reach an appropriate conclusion.

- The use of the Domino Model also permitted to structure the development of the system according to 3 distinct cycles of operation defined to accomplish different objectives: (1) identification and prediction of solids separation problems; (2) identification of the possible causes, and (3) proposal of a complete control strategy to restore the process from the identified problematic situations. The beginning of the first cycle is determined by the acquisition of a new set of data that characterize the new situation of the process (Situation Beliefs). The following cycles are dynamically chained in such a way that the decision reached in the previous cycle becomes the beginning of the next cycle (by being part of the new situation beliefs base). Likewise, different cycles may run in parallel whenever different objectives need to be pursued. The cyclic process is repeated until the pre-defined goals are accomplished.

- With the aim of developing an intelligent system, capable of using intelligent reasoning, an Expert System, based on expert knowledge, and a Case-Based Reasoning System, based on experiential knowledge, were chosen as the knowledge-based tools to compose the KBDSS.

- By incorporating dynamic reasoning, the Expert System is capable of adapting its reasoning process to the evolving necessities of the process. The system is aware of the number of days that the same type of problem or the same cause has been diagnosed, of the number of days that a specific strategy has been applied, as well as the response of the process to a specific control action. Moreover, the use of dynamic reasoning allows the system to make diagnoses even when not all the necessary information is available. Regarding the control methods, complete control strategies –including a set of scheduled actions and the expected deadline of effectiveness–, are now provided. The suggested control strategies are continuously updated. Now the decision is taken daily about whether to continue on with the same control strategy, to modify it (even suggesting a different control strategy) or to suggest a preventive control plan whenever the lack of information prevents the system from confirming the diagnosis.

- Eight decision trees have been developed to diagnose any type of solids separation problems, three decision trees have been developed to predict future solids separation problems, eight more to identify the possible causes and finally thirty-nine decision trees have been developed to decide the control strategy to be applied.

- In order to facilitate the acquisition of qualitative knowledge, including microscopic and macroscopic observations of the activated sludge, both a microscopic and a macroscopic worksheet have been developed.

- The use of qualitative knowledge during the reasoning process gives the final conclusion a certain degree of uncertainty as a consequence of the subjectivity related to the measurement of such variables.

- The development of fuzzy sets for each of the quantitative variables used in the reasoning process allows the system to inform the user about the percentage of certainty related to some of the conclusions reached by the ES.

- The quantity of required information has been restricted to the set of variables needed by the reasoning process to reach a conclusion. Hence, just by checking the value of one quantitative variable (i.e. SVI or V30) and a qualitative variable (i.e. sludge settleability or floc characteristics) the system is capable of reaching a diagnosis about the type of solids separation problem occurring in the process. Regarding the identification of the possible cause, from 5 to 8 different variables (depending on the type of problem) must be checked to reach a conclusion. Whenever a possible cause cannot be properly checked, a set of complementary actions such as a list of probable causes or a preventive control strategy are provided. Finally, once a conclusion concerning the problem and cause diagnosis is reached, just by checking 2 or 3 different variables (including trends, counters or other analytical variables) the type of control strategy to apply can be reached.

- Because of the complexity of solids separation problems means that there are still a lot of open questions left, related to their possible causes and effects. Thus, the efficiency of the ES in diagnosing and predicting problematic situations and their potential causes is directly affected by this fact.

- The limitations of the Expert System related to the staticity of the knowledge base and to the systematic repetition of errors during the decision making process (inablility to learn from new experiences) were overcome by the incorporation of a Case-Based Reasoning technique into the overall system.

- A new temporal approach for classical CBRS was developed: the Episode-Based Reasoning System. The new EBRS goes one step further and considers the basic pieces of knowledge of the system, the cases, at a larger and undefined length. Instead of dealing with single cases, the new temporal approach deals with episodes, defined as a temporal sequence of cases. This temporal approach definitively allowed the system to manage solids separation problems, characterized by their variable temporal definition.

- A set of 22 attributes (including quantitative and qualitative variables) were chosen to be used during the similarity calculation. These variables were selected according to their relevance in characterising a specific solids separation problem.

- An exponential weight-sensitive algorithm (*Eixample* distance) has been selected as the similarity algorithm due to the importance that the weight assigned to each attribute has during similarity calculation. Likewise, a relevance matrix has been built to assign different weights to each one of the selected attributes, according to the context (the type of solids separation problem and its cause). In this way, the retrieval of similar episodes (sharing the same context) is favoured.

- In order to streamline the searching process, a discriminating tree was developed. The resulting heuristic rules use the diagnosis reached by the Expert System, regarding the type of solids separation problem that the facility is currently facing, and uses it to discriminate and finally select the set of episodes that the system must explore during the searching mechanism.

- In order to simplify the retrieval process, the similarity calculation between the current situation and each one of the historical experiences was at the same scale (i.e., case to case).

- The usefulness of the retrieved information contained in the historical experiences has been improved by the retrieval of complete episodes (instead of single cases). This allows the user not only to consult the applied control strategy, but also to study the whole effect of its application on the process, as well as to learn from the final lesson registered by the expert.

- The reasoning process steps of the EBRS concerning episode reuse, adaptation and evaluation are not automatic. Given the complexity of the domain, the automatic adaptation, reuse and evaluation of the control strategy would be difficult to perform considering the limited amount of available information. As a consequence, these tasks are currently left to the final user, with the subsequent risk related to human errors (registration of subjective evaluations, errors in adapting the suggested control strategy, etc.).

- EBRSs cannot cope with situations that have never happened before in the process given that the episode library does not include reference cases to be compared to. This fact is overcome by the complementary task of the ES. When the EBRS is not capable of retrieving a similar historical experience to handle the current problematic situation, the suggestion of the ES provides the user with a control strategy based on expert knowledge. Once the user applies the selected control strategy, the new experience immediately becomes part of the episode library and will be available for future use.

- The Girona WWTP was an appropriate candidate facility to implement the developed KBDSS. The frequent occurrence of solids separation problems, mainly filamentous bulking, non-filamentous bulking and biological foaming, the complex identification of the possible causes, and the lack of registration of historical experiences (especially the applied control strategies and the results of their application) are some of the characteristics that supports the use of the KBDSS in such facility as a decision support tool to handle solids separation problems.

- The developed KBDSS were implemented in the G2-Gensym environment, an object-oriented shell that facilitated the implementation of the system.

- The implementation of the KBDSS has been done using a three-layer architecture which include: (i) data gathering, (ii) reasoning process (where the ES and EBRS are integrated), and (iii) decision support. The use of a three-layer architecture enabled the system to: (1) have the appropriate modularity and independence to guarantee the re-design and transferability of the system to other facilities; (2) to deal with any kind of data gathered from the process (quantitative and qualitative); (3) to include an array of different kinds of knowledge (numerical, heuristic, experiential, and predictive), and (4) to integrate the advantages of different intelligent technologies.

- The system has been implemented to be routinely executed once a day, although it could also be launched manually at any time, whenever an alarm symptom is detected. According to the slow dynamics of the domain, the daily execution of the system is adequate enough to fulfil the main objectives of the system.

- The implementation of the reasoning process of the ES was performed by means of 5 different modules including: meta-diagnosis, diagnosis of solids separation problems, prediction of future solids separation problems, identification of the possible cause(s) and proposal of the control strategy. The implementation of the EBRS was performed by means of 5 different modules including: current case definition; case-retrieval; episode-retrieval; review and retention. All these modules consist of several heuristic rules and procedures. The modular implementation of both systems facilitates their future modification and transferability to other WWTPs.

- A menu-based interface was developed in the G2 environment to provide a simple and feasible way for users to run the KBDSS and to know the state and evolution of the process over time at any time. In order to promote and assist the interaction of the user, several user-friendly interfaces were developed. They can provide the results of both the ES and the EBRS, explanations about the conclusions reached and the deductive processes followed, the retrieval of certain quantitative or qualitative values and the access to trend charts, graphics or readouts of the most important values, which facilitates the understanding of the large amounts of information generated by the process. Furthermore, the user, at any time, can calibrate or customize the system by modifying some key parameters such as the considered fuzzy sets, the limits of effectiveness, etc.

- The maintenance of the KBDSS is in general a tough task. For example, whenever the activated sludge process undergoes some modifications (i.e. enlargement of the facility, integration of new sensors, etc.) most of the modules should be modified. Regarding the maintenance of the episode library, it directly depends on the user implication, which is in charge of registering new experiences.

- The use and interaction of both ES and EBRS techniques improves the reasoning capabilities (deeper knowledge) and upgrades the decision making reliability of the KBDSS. The system is able to reason both based on expert knowledge from literature and experts (ES) and based on experiential knowledge accumulated through years of operation in the facility (EBRS), which is crucial for diagnosing, predicting and solving abnormal or problematic situations. This intelligent feature of the system allows it to overcome the limitations of classical control in this kind of processes.

- The validation process has been considered as a multi-phase period covering the entire life cycle of the KBDSS. It extended for a period of 11 months (4 months of experimental validation and 7 months of field validation), during which the system was continuously upgraded, according to the partial results obtained from the validation of the following aspects: reliability of communication and data gathering; consistency and completeness of decision trees; accuracy of ES diagnosis; usefulness of the expert control plans; adequacy of the episode library; accuracy of case and episode retrieval; usefulness of the retrieved episodes, and global validation of the overall system.

- During the field validation phase, 6 different episodes of solids separation problems were identified, including 4 episodes of filamentous bulking, 1 episode of biological foaming and 1 episode of non-filamentous bulking, together with 5 episodes of normal state where no solids separation problems occurred. The length of the filamentous bulking episodes varied from 21 to 40 days, the episode of biological foaming lasted 16 days and the episode of non-filamentous bulking persisted for 12 days. On 73 days (32% of the validation period) the facility did not suffer from any solids separation problem.

- Regarding the accuracy of the ES in diagnosing and predicting solids separation problems, an accuracy of almost 100% in diagnosing solids separation problems and 80% accuracy in predicting future problematic events were obtained. The validation of the ES diagnoses regarding the identification of the possible causes of solids separation problems resulted in an efficiency of just 43.3% of correctly identified causes. The origin of the remaining 56.7% of unidentified causes is the lack of data that hinders the verification of all the possible causes. An increase in the frequency of measurement of such variables is recommended in order to improve the efficiency of the system in diagnosing the possible causes and the prediction of future problems based on the trends of such variables. The validation of the control plans suggested by the ES resulted in 91.5% of usefulness of the suggested control plans, from which 57% exactly correspond to the control strategies finally applied. In particular, the proposal of complete control plans and the continuous adjustment of the suggested strategies (where the exact required doses of chemicals or the exact control parameters such as the target WAS flow or DO set-point were included) was evaluated as being very useful for plant operators.

- As expected, at the beginning of the field validation phase, the adequacy of the episode library was very low, but it increased gradually as new episodes were incorporated to the episode library. Nevertheless, at the end of the validation phase, the adequacy could not be considered as being complete, given that for example, episodes of dispersed growth or pin-point floc (never experienced in the facility) were not represented. Regarding the accuracy of case and episode retrieval, and once the specific modifications resulting from the first phase of validation were applied, an accuracy of 97% was obtained in retrieving similar episodes and cases. A general evaluation of the usefulness of the retrieved episodes could not be done mainly due to the lack of an adequate episode library to critically validate the usefulness of the stored episodes. Only partial usefulness scores of 86% and 67% were obtained for the 2 episodes of filamentous bulking that could be validated.

- It can be concluded that, according to plant operators, the information provided by the retrieval of complete episodes is indeed much more useful than the information provided by a unique case. The evolution of the process (stored in the retrieved episodes) once a control strategy is applied was considered very useful to decide the most suitable control strategy to follow. The usefulness of the

EBRS, but also its usability, directly affects (specially once the validation of the adequacy and accuracy of the system is completed) the proposal and application of the most efficient control plans, which on most occasions imply a reduction in the response time of the process, with the consequent environmental and economical savings.

- Validation of Decision Support Systems is a challenging issue in the domain of Artificial Intelligence. The complexity of such applications, especially when applied within environmental domains, makes finding a standard methodology to validate any DSS difficult. In our approach, some constraints were found during the validation of the KBDSS, among them: all the type of solids separation problems initially considered did not occur during the phase of validation (e.g. episodes of dispersed growth or pin-point floc never happened during the field validation) and thus the system could not be completely validated in diagnosing, predicting or suggesting solutions for such type of situations; a complete implication of the user was not always obtained, specially during the registration of the information requested to introduce the new experiences in the EBRS; at the beginning of the validation process, the user was not confident enough to directly apply the control strategies suggested by the system. According to all these facts and limitations, the presented results, obtained from the validation phase, should not be stated as definitive conclusions but as the preliminary results which has allowed us to improve the feasibility of the KBDSS.

As a final conclusion, it can be stated that the Knowledge-Based Decision Support System developed and implemented in the Girona WWTP has demonstrated a high usefulness in handling solids separation problems of microbiological origin in activated sludge systems.

REFERENCES

Aamodt A. and Plaza E., Case-based reasoning: fundamental issues, methodological variations and system approaches, *AI Comm.*, **7(1)**,39-59 (1994).

Allen, J., Mantaining knowledge about temporal intervals, *Communications of the ACM*, **26(11)**, 832-843 (1983).

Al-Shahwani, S.M. and Horan, N.J., The use of protozoa to indicate changes in the performance of activated sludge plants. *Water Res.*, **25**, 633-638 (1991).

ATV Working Group, 2.6.1. Report, prevention and control of bulking sludge and scum, *Korr. Abwass.*, **36**, 165 (1989).

Andreadakis, A.D., Physical and chemical properties of activated sludge floc, *Water Res.*, **27(12)**, 1707-1714 (1993).

Andreasen, K. and Nielsen, P.H., Application of microautoradiography to the study of substrate uptake by filamentous microorganisms in activated sludge. *Appl. Environm. Microbiol.*, **63**, 3662-3668 (1997).

Ansari, S.A., Farrah, S.R. and Chadhry, G.R., Presence of human immuno deficiency virus in wastewater and their detection by the polymerase chain reaction. *Appl. Environm. Microbiol.*, **58**, 3984-3990 (1992).

Ashraf Islam, K., Newell, B. and Lant, P., Advanced process control for biological nutrient removal, *Water Sci. Technol.*, **39(6)**, 97-103 (1999).

Avesani P., Ricci F. and Perini A., Interactive case-based planning for forest fire fighting. *Appl. Intell.*, **13(1)**, 41-58 (2000).

Baeza, J.A., Gabriel, D. and Lafuente, J., Improving the nitrogen removal efficiency of an A2/O based WWTP by using an on-line knowledge based expert system, *Water Res*. **36(8)**, 2109-2123 (2002).

Barbunski, K., Koscielniak, H., Influence of substrate loading intensity on foc size in activated sludge process. *Water Res.*, **29 (7)**, 1703-1710 (1995).

Barnett, M. W., Patry, G. G., Takács, I. and Bruce R.A., A model-based expert system for activated sludge separation: expert system design, *Research Symposium VI: Innovative Concepts, Water Environment Federation 66th Annual Conference & Exhibition*, 1 - 10, California (1993).

Beck, M.B., Latten, A. and Tong, R.M., Modelling and operational control of the activated sludge process in wastewater treatment. *Professional Paper PP-78-10*, International Institute for Applied Systems Analysis, Luxemburg (1978).

Beebe, R.D., Jenkins, D. and Daigger, G.T., Activated Sludge Bulking Control at the Santa Jose/Santa Clara, California Water Pollution Control Plant, presented at *55th Annual Conference of the Water Polution control Federation*, Sant Louis, MO, USA (1982).

Belanche Ll., Valdés J.J., Comas J., R.-Roda I. and Poch M., A soft computing techniques study in wastewater treatment plants. *Artif. Intell. Eng.*, **14(4)**, 307-317 (2000).

Bellmann, R. and Zadeh, L.A., Decision-making in a fuzzy environment. *Manage. Sci*. **17**, 141-164 (1970).

Beltramini L. and Motard R.L., Expert systems in chemical engineering. *Chem. Biochem. Eng. Q.,* **2(3),** 119-126 (1988).

Bergh, S.G. and Olsson, G. Knowledge based diagnosis of solids-liquid separation problems. *Water Sci. Technol.*, **33(2)**, 219-226 (1996).

Berthouex, P. M., M. Lai and A. Darjatmoko. A statistics-based information and expert system for plant control and improvement. *Proc. of 5th National Conf. on Microcomputers in Civil Engineering*, 146 - 150, Orlando (1987).

Blackall, L.L., Harbers, A.E., Greenfield, P.F. and Hayward, A.C. Actinomycete scum problems in Australian activated sludge plants. *Water Sci.Technol.*, **20**, 493-495 (1988).

Blackall, L.L., Harbers, A.E., Greenfield, P.F. and Hayward, A.C. Activated sludge foams: effects of environmental variables on organism growth and foam formation. *Environ.Technol.*, **12**, 241-248 (1991).

Blackbeard, J.R., Ekama, G.A. and Marais, G.v.R., A survey of bulking and foaming in activated sludge plants in South Africa, *Water Pollut. Control*, **85(1)**, 90-100 (1986).

Bongards, M. Improving the efficiency of a wastewater treatment plant by fuzzy control and neural networks. *Water Sci.Technol.*, **43(11)**, 189-196 (2001).

Borenstein D., Towards a practical method to validate Decision Support Systems, *Decis. Support Syst.*, **23**, 227-239 (1998).

Branting L.K., Hastings J.D., and Lockwood, J.A. Integrating cases and models for prediction in biological systems, *AI Applications*, **11(1),** 29-48 (1997).

Brdjanovic, D., van Loosdrecht, M.C.M., Versteeg, P., Hooijmans, C.M., Alaerts, G.J. and Heijnen, J.J., Modelling COD, N and P removal in a full-scale WWTP Haarlem Waarderpolder, *Water Res.*, **34**, 846-858 (2000).

Capodaglio, A.G., Jones, H.V., Novotny, V. and Feng, X., Sludge bulking analysis and forecasting: application of system identification and artificial neural computing technologies, *Water Res.*, **25(10)**, 1217-1224 (1991).

Caravelli, A., Contreras, E.M., Giannuzzi, L. and Zaritzky, N., Modelling of chlorine effect on floc forming and filamentous micro-organisms of activated sludges, *Water Res.*, **37**, 2097-2105 (2003).

Casey, T.G., Ekama, G.A., Wentzel, M.C. and Marais, G.v.R., Filamentous organisms bulking in nutrient removal activated sludge systems. Paper 1: a historical overview of causes and control. *Water SA*, **21(3)** (1995).

Cenens, C., Smets, I.Y. and van Impe, J.F., Modelling the competition between floc-forming and filamentous bacteria in activated sludge wastewater treatment systems-I. Evaluation of mathematical models based on kinetic selection theory, *Water Res.*, **34(9)**, 2525-2534 (2000a).

Cenens, C., Smets, I.Y. and van Impe, J.F., Modelling the competition between floc-forming and filamentous bacteria in activated sludge wastewater treatment systems-II. A prototype mathematical model based on kinetic selection and filamentous backbone theory, *Water Res.*, **34(9)**, 2535-2541 (2000b).

Cenens, C., van Beurden, K.P., Jenné, R. and van Impe, J.F., On the Development of a Novel Image Análisis Technique to Distinguish Between Flocs and Filaments in Activated Sludge Images. *Water Sci.Technol.*, **46 (1-2)**, 381-387 (2002).

Chan, W.T. and Koe, L.C.C. A knowledge based framework for the diagnosis of sludge bulking in the activated sludge process. *Water Sci. Technol.* **23**, 847-855 (1991).

Chen, W. C., Chang, N-B., and Shieh, W. K. Advanced hybrid fuzzy-neural controller for industrial wastewater treatment. *J. Environ. Eng.*, **127(11)**, 1048-1059 (2001).

Chiesa, S.C and Irvine, R.L., Growth and control of filamentous microbes in activated sludge: an integrated hypothesis, *Water Res.*, **19**, 471-479 (1985).

Chudoba, J., Control of activated aludge filamentous bulking-VI. Formulation of basic principles. *Water Res.*, **19**, 1017-1022 (1985).

Chudoba, J., Grau, P. and Ottova, V., Control of Activated Sludge Filamentous Bulking II. Selection of Microorganisms by Means of a Selector. *Water Res.*, **7(10)**, 1389-1406 (1973).

Comas, J., Development, implementation and evaluation of an activated sludge supervisory system for the Granollers WWTP, PhD Thesis, Universitat de Girona (2000).

Comas, J., Cortés, C., Rodríguez-Roda, I., Sànchez-Marrè, M., Cortés, U., Ekster, A. and Poch, M. Improvements of the Decision Support system at the Granollers WWTP. *Proc. of the 75th annual conference WEFTEC'02*, Session 88 (2002).

Comas, J., Rodríguez-Roda, I., Sànchez-Marrè, M., Cortés, U., Freixó, A., Arráez, J. and Poch, M. A knowledge-based approach to the deflocculation problem: integrating on-line, off-line, and heuristic information. *Water Res.*, **37(10)**, 2377-2387 (2003).

Cortés, C., Supervisory Systems in wastewater treatment plants: systematise their implementation, PhD Thesis, Universitat de Girona (2002).

Cortés, U., Sànchez-Marrè, M., Cecaronni, L., R.-Roda, I. and Poch, M., Artificial Intelligence and Environmental Decision Support Systems*, Appl. Intell.*, **13(1)**, 77-91 (2000).

Cortés, U., Martínez, M., Comas, J., Sànchez-Marrè, M., Poch, M and Rodríguez-Roda, I., A conceptual model to facilitate knowledge sharing in waste water treatment plants, *AI Com.*, **16(3)**, 193-207 (2003).

Costerton, J.W., Irvin, R.T. and Cheng K.-J. The bacterial glycocalyx in nature and disease. *Ann. Rev. Microbiol.*, **35**, 299-324 (1981).

Côte M., Grandjean B.P.A., Lessard P. and Thibault J., Dynamic modelling of the activated sludge process: improving prediction using neural networks, *Water Res.*, **29(4)**, 995-1004 (1995).

Curds, C.R., The ecology and role of protozoa in aerobic sewage treatment processes. *Annu. Rev. Microbiol.*, **36**, 27-46 (1982).

Da Motta, M., Pons, M.N. and Roche, N., Study of filamentous bacteria by image analysis and relation with settleability. *Water Sci. Technol.*, **46 (1-2)**, 363-369 (2002).

De Clercq, J., Devisscher, M., Boonen, I., Vanrolleghem, P.A., and Defrancq, J., A new one-dimensional clarifier model-verification using full-scale experimental data, *Water Sci. Technol.*, **47 (12)**, 105-112 (2003).

De los Reyes, F.L., Ritter, W. and Raskin, L. Group-specific small sub-unit rRNA hybridisation probes to characterize filamentous foaming in activated sludge systems. *Appl. Env. Microb*, **63**, 1107-1117 (1997).

Dick, R.I. and Vesilind P.A., The Sludge Volume Index- what is it?, *J. Water Pollut. Contr. Fed.*, **41**, 1285-1291 (1969).

Echeverria, E., Seco, A. and Ferrer, J., Study of the factors affecting activated sludge settling in domestic wastewater treatment plants, *Water Sci. Technol.*, **25**, 273-279 (1992).

Eikelboom, D.H., Filamentous organisms observed in bulking activated sludge, *Water Res.,* **9**, 365-388 (1975).

Eikelboom, D.H., Scum Formation in Carroussel Treatment Plants, *Proceedings of the IAWPRC Specialised Seminar on Interactions of Wastewater, Biomass and Reactor Configurations in Biological Treatment Plants, paper No.24.* August 21-23, Cophenagen*,* Denmark (1991).

Eikelboom, D.H., The *Microthrix parvicella* puzzle, *Water Sci. Technol.*, **29 (7)**, 271-279 (1994).

Eikelboom, D.H. and van Buijsen, H.J.J., *Microscopic sludge investigation Manual*, 2nd ed, TNO Res. Inst. for Env. Hygiene, Delft, The Netherlands (1983).

Eikelboom D.H., Andreadakis, A. and Andreasen, K., Survey of filamentous populations in nutrient removal plants in four european countries, *Water Sci. Technol.*, **37 (4-5)**, 281-289, (1998).

Ekama, G.A. and Marais, G.v.R, Two improved sludge settleabilitty parameters, *IMIESA*, **9(6)**, 20 (1984).

Ekama, G.A., Wentzel, M.C., Casey, T.G., and Marais, G.v.R., Filamentous organisms bulking in nutrient removal sctivated sludge systems. Paper 6: Review, evaluation and consolidation of results. *Water SA*, **22(2)**, 147-152 (1995).

Ekama, G.A., Barnard, F.W., Günthert, F.W., Krebs, P., McCorquodale, J.A., Parker, D.S. and Wahlberg, E.J., *Secondary Settling Tanks: Theory, Modelling, Design and Operation*, Scientific and Technical Report No.6, IAWQ, London (1997).

Ekster, A. and Jenkins, D., Nocardioform foam control in a BNR plant, *Proc. of the 72th annual conference WEFTEC'99*, 1-10 (1999).

Esteban, G., Tellez, C. and Bautista, L.M. Dynamics of ciliated protozoa communities in activated sludge processes, *Water Res.*, **25**, 967-972 (1991).

Fiter, M., Güell, D., Comas, J., Colprim, J., Poch, M. and Rodríguez-Roda, I., Energy saving in a Wastewater Treatment Process: an Application of Fuzzy Logic Control, *Environ. Technol.*, **26**, 1263-1270 (2005a).

Fiter, M., Güell, D., Freixó, A., Comas, J., Colprim, J., Poch, M. and Rodríguez-Roda, I., Fuzzy Control: gaining ground in the activated sludge process, Proc. of the 2nd IWA Conference on Instrumentation, Control and Automation (ICA'05), 1185-1187, Busan, Korea (2005b).

Fiter, M., Control basat en lògica difusa per sistemes de fangs activats. Disseny, implementació i validació en EDAR real (in catalan), PhD Thesis (in process of revision), Universitat de Girona (2006).

Fox, J. and Das, J., *Safe and Sound: Artificial Intelligence in Hazardous Applications*, AAAI Press/ The MIT Press (2000).

Furukawa S., Tokimori K., Hirotsuji J. and Shiono S., New operational support system for high nitrogen removal in oxidation ditch process, *Water Sci. Technol.*, **37(12)**, 63-68 (1998).

Gabb, D.M.D., Still, D.A., Ekama, G.A., Jenkins, D., Wentzel, M.C. and Marais, G.v.R. Development and Full Scale Evaluation of Preventative and Remedial Methods for Control of Activated Sludge bulking. *Final Report to the Water Research Commission*, University of Cape Town, South Africa (1989).

Gall, R. and Patry, G., Knowledge-based system for the diagnosis of an activated sludge plant, *In Dynamic Modelling and Expert Systems in Wastewater Engineering*, G. Patry and D. Chapman editors, Chelsea, MI., Lewis Publishers (1989).

Gernaey, K.V., van Loosdrecht, M.C.M, Henze, M., Lind, M. and Jorgensen, S.B., Activated wludge wastewater treatment plant modelling and simulation: state of the art, *Environ. Model. Softw.*, **19**, 763-783 (2004).

Gensym, *G2 Classic 5.1 Rev.7.* Gensym Corporation, Cambridge, USA (2000).

Gerardi, M.H. *Settleability Problems and Loss of Solids in the Activated Sludge Process*, John Wiley & Sons, Hoboken, New Jersey (2002).

Geselbracht, J. J., Brill, E.D and Pfeffer, J.T., Incorporating a rule-based model of judgement into a wastewater treatment plant design optimization model. *Water Resources Center, Research Report No. 199*, University of Illinois at Urbana-Champaign (1986).

González, A.J. and Douglas, D.D., *The Engineering of Knowledge-Based Systems. Theory and Practice*, Prentice Hall Int.Editions, New Jersey (1993).

Grabinska-Loniewska, A., Denitrification unit biocenosis, *Water Res.*, **25(12)**, 1565-1573 (2002).

Grau, P., Criteria for nutrient-balanced operation of activated sludge process, *Water Sci. Technol.*, **24(3-4)**, 251-258 (1991).

Grady, C.P.L., Daigger, G.T. and Lim., H.C., *Biological Wastewater Treatment*, 2nd ed., Marcel Dekker, New York (1999).

Grijspeerdt, K. and Verstraete, W., Image analysis to estimate the settleability and concentration of activated sludge, *Water Res.*, **31(5)**, 1126-1134 (1997).

Gujer, W, Henze, M., Mino, T. and van Loosdrecht, M.C.M., Activated Sludge Model No.3, ASM2D, *Water Sci. Technol.*, **39(1)**, 183-193 (1999).

Henze, M., Grady Jr, C.P.L, Gujer, W, Marais, G.v.R and Matsuo, T., Activated Sludge Model No.1, *IAWPRC Scientific and Technical Report no.1*. London: IAWPRC (1987).

Henze, M., Gujer, W, Mino, T., Matsuo, T. and Wentzel, M.C., Activated Sludge Model No.2, *IAWQ Scientific and Technical Report no.3*. London: IAWQ (1995).

Henze, M., Gujer, W, Mino, T., Matsuo, T., Wentzel, M.C., Marais, G.v.R and van Loosdrecht, M.C.M., Activated Sludge Model No.2D, ASM2D, *Water Sci. Technol.*, **39(1)**, 165-182 (1999).

Hermanowicz, S., Dynamic changes in populations of the activated sludge community: effect of dissolved oxygen variations, *Water Sci. Technol.*, **19**, 889-895 (1987).

Hernandez, M., Jenkins, D. and Beaman, B., Mass and viability estimations of *Nocardia* in activated sludge and anaerobic digesters using conventional stains and immunofluorescent methods, *Water Sci. Technol.*, **29**, 249-259 (1994).

Hertz J., Krogh, A. and Palmer R.G., *Introduction to the Theory of Neural Computation*, Addison-Wesley, Redwood City, 1991.

Hecht-Nielsen R. *Neurocomputing*. Addison-Wesley, Reading, MA, 1990.

Higgins, M.J. and Novak, J.T. Characterization of Extracellular Protein and its Role in Bioflocculation, *J. Env. Eng.*, **123**, 479-485 (1997).

Hong Y.T., Rosen M.R. and Bhamidimarri R. Analysis of a municipal wastewater treatment plant using a neural network-based pattern analysis. *Water Res.*, **37(7)**, 1608-1618 (2003).

Horan, N.J., *Biological Wastewater Treatment Systems. Theory and Operation*, John Wiley & Sons, Chichester (1990).

Huang, Y.L, Sundar, G. and Fan, L.T., Min-cyanide: an expert system for cyanide waste minimization in electroplating plants. *Environ. Prog.* **10(2)**, 89-95 (1991).

Hudson, A. D., Sanders, D. A., Golding, H., Tewkesbury, G. E. and H. Cawte, Aspects of an expert design system for the wastewater treatment industry, *J. Syst. Architect*, **43**, 59-65 (1997).

Hunt K.J., Sbarbaro D., Zbikowski R. and Gawthrop P.J., Neural Networks for Control Systems – A Survey, *Automatica*, **28 (6)**, 1083-1112 (1992).

Jackson P, *Introduction to Expert Systems*, Addison-Wesley, 3rd ed., Essex: England, (1999).

Jaczynski, M., A framework for the management of past experiences with time-extended situations. *Procc. of the 6th Int. Conference on Information and Knowledge Management (CIKM'97)*, pages 32-39, Las Vegas, Nevada, USA (1997).

Jaere, M., Aamodt, A. and Shalle, P., Representing temporal knowledge for Case-Based Reasoning. *Proc. of the 6th European Conference on Case-Based Reasoning, ECCBR 2002*, pages 174-188, Aberdeen, Scotland, UK, September (2002).

Jenkins, D., Richard M. G. and Daigger G. T. *Manual on the Causes and Control of Activated Sludge Bulking and Foaming*, Pretoria, SA: Water Research Commission (1986).

Jenkins, D., Richard M. G. and Daigger G. T. *Manual on the Causes and Control of Activated Sludge Bulking, Foaming & other Solids Separation Problems*, 3rd ed., Lewis Publishers, Chelsea (2003).

Jenné, R., Banadda, E.N., Smets, I.Y and Van Impe, J.F., Monitoring activated sludge settling properties using image analysis, *Water Sci.Technol.*, **50(7)**, 281-287 (2004).

Jeppson, U., Alex, J., Pons, M.N., Spanjers, H. and Vanrolleghem, P.A., Status and future trends of ICA in wastewater treatment-A european perspective, *Water Sci. Technol.*, **45(4-5)**, 485-495 (2002).

Johnston, D. M. Diagnosis of wastewater treatment processes. *Proc. of Computer Applications in Water Resources*, ASCE, 31 - 38, New York (1985).

Jones E. and Roydhouse A. Retrieving structured spatial information from large databases: a progress report. *Proc. of IJCAI Workshop on Artificial Intelligence and the Environment*, 49-57, Montréal (1995).

Kappeler, J and Brodmann, R. Low F/M bulking and scumming: towards a better understanding by modelling, *Water Sci. Technol.*, **31(2)**, 225-234 (1995).

Kappeler, J and Gujer, W., Development of a mathematical model for “aerobic bulking”, *Water Res.*, **28(2)**, 303-310 (1994).

Klein, M.R. and Methlie, L.B., *Knowledge-Based Decision Support Systems with Applications in Business*, John Wiley & sons, 2nd ed., West Sussex, England (1995).

Kolodner J., *Case-Based Reasoning*. Morgan Kaufmann (1993)

Knoop, S. and Kunst, S., Influence of Temperature and Sludge Loading on Activated Sludge Settling, especially on *Microthrix parvicella*, *Water Sci.Technol.*, **37(4-5)**, 27, (1998).

Kocianova, E., Foot, R.J. and Forster, C.F., Physicochemical Aspects of Activated Sludge in Relation to Stable Foam Formation, *J. Inst. Water Environ. Manage.*, **6(3)**, 342-350 (1992).

Kraslawski A., Koiranen T. and Nystrom L., Case-Based Reasoning System For Mixing Equipment Selection, *Comput. Chem. Eng.*, **19(Suppl.)**, 821-826, (1995).

Kreft, J-U., Picioreanu, C., Wimpanny, J.W.T. and van Loosdrecht, M.C.M., Individual-based modelling of biofilms, *Microbiol.*, **147**, 2897-2912 (2001).

Krichten, D.J., Wilson, K.D and Tracy, K.D., Expert Systems guide biological phosphorus removal, *Water Environ. Tech.*, 60-64 (1991).

Krovvidy, S. and Wee, W.G., Wastewater treatment systems from Case-Based Reasoning, *Mach. Learn.*, **10**, 341-363, 1993.

Ladiges, G. and Kayser, R. On-line and offline expert systems for the operation of wastewater treatment plants. *Water Sci. Technol.*, **28(11-12)**, 315-323 (1993).

Lakay, M.T., Hulsman, A., Ketley, D., Warburton, C., de Villiers, M., Casey, T.G., Wentzel, M.C. and Ekama, G.A., Filamentous organisms bulking in nutrient removal activated sludge systems. Paper 7: exploratory experimental investigations. *Water SA*, **25(4)**, 383-396 (1999).

Lapointe, J., Marcos, J. B., Veillette, M., Laflamme, G. and Dumontier, M., Bioexpert – an Expert System for wastewater treatment process diagnosis, *Comput. Chem. Eng.*, **13(6)**, 619-630 (1989).

Lau, A.O., Strom, P.F. and Jenkins, D., The competitive growth of floc-forming and filamentous bacteria: a model for activated sludge bulking, *J. Water Polln. Control Fedn.*, **56(1)**, 52-61 (1984).

Lee, S-E., Koopman, B.L., Bode, H. and Jenkins, D., Evaluation of alternative sludge settleability indices, *Water Res.*, **17**, 1421 (1983).

Lee, N., Nielsen, P.H., Andreasen, K.H., Juretschko, S., Nielsen, J.L., Schleifer, K.H. and Wagner, M. Combination of Fluorescent In Situ Hybridization and Microautoradiography- a new tool for structure-function analyses in microbial ecology. *Appl. Environ. Microbiol.* **65**, 1289-1297 (1999).

Lemmer, H. and Kroppenstedt, R.M., Chemotaxonomy and physiology of some actinomycetes isolated from scumming activated sludge, *Appl. Microbiol.*, **5**, 124-135 (1984)

Lewis, G.D., Austin, F.J., Loutit, M.W. and Shaples, K. Enterovirus removal from sewage: the effectiveness of four different treatment plants, *Water Res.*, **20**, 1291-1297 (1986).

Li, D-H., and Ganczarkzyk, J.J., Stroboscopic determination of settling velocity, size and porosity of activated sludge flocs, *Water Res.*, **21(3)**, 257-262 (1987).

Li, D-H., and Ganczarkzyk, J.J., Size distribution of activated sludge flocs, *J. Water Pollut. Contr. Fed.*, **63(5)**, 806-814(1991).

Liao, T.W. and Zhang, Z. Similarity measures for retrieval in case-based reasoning systems, *Appl. Artif. Intell*, **12**, 267-288 (1998).

Liao, B.Q., Allen D.G., Droppo, I.G., Leppard, G.G and Liss, S.N., Surface Properties of Sludge and their Role in Bioflucculation and Settleability, *Water Res.*, **35**, 339-350 (2001).

Ma, J. and Knight, B., A Framework for Historical Case-Based Reasoning. *In Procc. of 5th Int. Conference on Case-Based Reasoning (ICCBR'2003)*, 246-260, LNCS2689, 2003.

Madoni, P. Estimates of Ciliated Protozoa Biomass in Activated Sludge and Biofilm. *Bioresource Technol.*, **48**, 245-249 (1994).

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.*, **28**, 67-75 (1994).

Maeda, K., A knowledge-based system for the wastewater treatment plant, *Future Gener. Comp. Sy.*, **5**, 29-32, (1989).

Magara, Y. and Nambu, S., Biochemical and physical properties of an activated sludge on settling characteristics, *Water Res.*, **10(1)**, 71-77 (1976).

Mamais, D., Jenkins, D. and Pitt, P. A rapid physical chemical method for the determination of readily biodegradable soluble COD in municipal wastewater, *Water Res.*, **27**, 195-197 (1993).

Manesis S.A., Sapidis D.J. and King R.E., Intelligent control of wastewater treatment plants, *Artif. Intell. Eng.*, **12**, 275-281 (1998).

Martin, F.J. and Plaza, E., Ceaseless Case-Based Reasoning, *Procc. of the 7th European Conference, ECCBR 2004*, pages 285-301, Madrid, Spain, September (2004).

Martínez, M., *Feasibility testing of a conceptual model to develop knowledge based decision support systems to manage filamentous bulking in a WWTP*. Master's thesis, University of Girona (2002).

Martínez, M., Cortés, U., Bonmatí, A., Poch, M. and Rodríguez-Roda, I., Conceptualisation of a decision support system to solve separation problems, *Proc. of the 9th IWA Specialised Conference on Design, Operation and Economics of Large Wastewater Treatment Plants*, pp. 131-134, September 1-4, Praha, Czech Republic (2003).

Martínez, M., Rodríguez-Roda, I., Poch, M., Cortés, C. and Comas, J., Dynamic reasoning to solve complex problems in activated sludge processes: a step further in Decision Support Systems, *Water Sci. Technol*, (in press) (2005a).

Martínez, M., Sànchez-Marrè, M., Comas, J. and Rodríguez-Roda, I. Case-Based Reasoning, a promising tool to face solids separation problems in the activated sludge process. *Water Sci. Technol*, (in press) (2005b).

Martínez, M., Mérida-Campos, C., Sànchez-Marrè, M., Comas J. and Rodríguez-Roda, I. Improving the efficiency of Case-Based Reasoning to deal with activated sludge solids separation problems. *Environ. Technol.* (accepted) (2005c).

Martins, A.M.P., Pagilla, K., Heijnen, J.J. and Van Loosdrecht, M.C.M. (2004). Filamentous bulking sludge-a critical review. *Water Res.*, **38**, 793-817 (2004).

Matsui, S. and Yamamoto, R., Measuring the lenghts of filamentous microbes in activated sludges using a video color TV with a microscope, *Water Sci. Technol.*, **16(10)**, 11-69 (1983).

Mérida-Campos, C. and Rollón, E., Using a relevance model for performing feature weighting. *Proc. of the Congrés Català d'Intel•ligència Artificial (CCIA'03)*, Mallorca (2003).

Metcalf & Eddy, Inc., *Wastewater Engineering: Treatment and Reuse*, 4rd ed. revised by George Tchnobanoglous , Franklin L. Burton, H. David Stensel, McGraw-Hill, New York (2003).

Mizzaro, S., How many relevances in information retrieval?, *Interact. Comput*, **10(3)**, 303-320 (1998).

Mohlman, F.W., The Sludge Index, *Sewage Works J.*, **6**, 119 (1934)

Mudrack, K. and Kunst, S., *Biology of Sewage Treatment and Water Pollution Control*, John Wiley & Sons, New York (1986).

Musvoto, E.V., Lakay, M.T., Casey, T.G., Wentzel, M.C. and Ekama, G.A., Filamentous Organisms Bulking in Nutrient Removal Activated Sludge Systems. Paper 8: The effect of nitrate and nitrite. *Water SA*, **25(4)**, 397-407 (1999).

Nakhaeizadeh, G., Learning Prediction of Time Series: A Theoretical and Empirical Com-parison of CBR with Some Other Approaches. *Procc. of the Workshop on Case-Based Reasoning,* pages 67-71, AAAI-94. Seattle, Washington (1994).

Neethling, J.B., Asce, A.M., Chung, Y.C. and Jenkins, D., Activated sludge chlorine reactions during bulking control, *J. Environ. Eng.*, **113**, 134-146 (1987).

Ng W.J., Ong S.L., Hossain F. An algorithm approach for system-specific modelling of activated sludge bulking in an SBR, *Environ. Modell. Softw.*, **15**, 199-210 (2000).

Nuñez, H., Sànchez-Marré, M., Cortés, U., Comas, J., Martínez, M., Rodríguez-Roda, I. and Poch, M., A comparative study on the use of similarity measures in case-based reasoning to improve the classification of environmental systems situations, *Environ. Modell. Softw.*, **19(9)**, 809-819 (2004).

Ohtsuki T., Kawazoe T. and Masui T., Intelligent Control System Based on Blackboard Concept for Wastewater Treatment Processes, *Water Sci. Technol.*, **37(12)**, 77-85 (1998).

Okubo, T., K. Kubo, M. Hosomi i A. Murakami. A knowledge-based decision support system for selecting small-scale wastewater treatment processes. *Water Sci. Technol.*, **30(2)**, 175-184 (1994).

Olsson, G., Aspergren, H. and Nielsen, M.K. Operation and control of wastewater treatment-a Scandinavian perspective over 20 years. *Water Sci. Technol.*, **37(12)**, 1-13 (1998).

Olsson, G., Newell, B., Rosen, C. and Ingildsen, P., Application of Information Technology to Decision Support in Treatment Plant Operation, *Water Sci. Technol.*, **47(12)**, 35-42 (2003).

Olsson, G., Instrumentation, Control and Automation in the Water Industry-State-of-the-Art and New Challenges, *Proc. of the 2nd IWA Conference on Instrumentation, Control and Automation*, 19-31, May 29-June 2, Busan, Korea (2005).

Ozgur, N.H. and Stenstrom, M.K. Development of a knowledge-based expert system for process control of nitrification in the activated sludge process. *J. Environ. Eng.-ASCE*, **120**, 87-107 (1994).

Palm, J.C., Jenkins, D. and Parker, D.S., Relationship Between Organic Loading, Dissolved Oxygen Concentration and Sludge Settleability in the completely-mixed Activated Sludge Process, *J. Water Pollut. Contr. Fed.*, **52**, 2484-2506 (1980).

Parker, S.C. and Parker, D.G., *An Expert System for Managing an Activated Sludge Wastewater Treatment Plant*, Arkansas Water Resources Research Center, Publication No. 141, University of Arkansas, Fayetteville, AR (1989).

Patry, G. and Chapman, D., *Dynamic modelling and expert systems in wastewater engineering*, Lewis Publishers, Chelsea (1989).

Pitman, A.R., Bulking and Foaming in BNR Plants in Johannesburg: Problems and Solutions, *Water Sci.Technol.*, **34(3-4)**, 291-298 (1996).

Pitt, P and Jenkins, D., Causes and Control of *Nocardia* in Activated Sludge, *J. Water Pollut. Contr. Fed.*, **62**, 143-150 (1990).

Poch, M., Comas. J., Rodríguez-Roda, I., Sànchez-Marrè, M. and Cortés, U. Designing and building real environmental decision support systems, *Environ. Modell. Softw*, **19(9)**, 857-873 (2004).

Poole, J.E.P. A Study of the Relationships between the Mixed Liquor Microfauna and Plant Performance for a Variety of Activated Sludge Sewage Treatment Works. *Water Res.*, **18**, 281-287 (1984).

Prescott, L.P., Harley, J.P and Klein, D.A, *Microbiology*, Wm.C.Brown Publishers, USA (1990).

Pretorius, W.A. and Laubscher, C.J.P, Control of Biological Scum in Activated Sludge Plants by Means of Selective Flotation, *Water Sci.Technol.*, **19(5-6)**, 1003-1011 (1987).

Pujol, R., Vachon, A. and Martin, G., *Technical Guide to Activated Sludge Bulking*, Anthony:CEMAGREF (in French) (1990).

Pujol, R., Duchene, P., Schetrite, S. and Canler, J.P., Biological Foams in Activated Sludge Plants:characterization and situation, *Water Res.*, **25**, 1399-1404 (1991).

Puñal A., Roca E. and Lema J. M. An expert system for monitoring and diagnosis of anaerobic wastewater treatment plants. *Water Res.*,**36(10)**, 2656-2666 (2002).

Rammacher, J. and Hansen, J. Control of nutrient removal based on simple measurement values. *Water Sci. Technol.*, **41(1)**, 29-32 (2000).

Rao M., Yang H. And Yang H., Integrated distributed intelligent system architecture for incidents and diagnosis, *Comput. Ind.*, **37**, 143-151 (1998).

Richard, M.G., *Activated Sludge Microbiology*, Water Pollut. Contr. Fed., Alexandria, USA (1989).

Robusté, J., Modelització i identificació del procés de fangs activats, PhD Thesis (in catalan), Universitat Autònoma de Barcelona (1990).

R-Roda, I., Desenvolupament d'un protocol per l'aplicació de Sistemes Basats en el Coneixement a la gestió d'estacions depuradores d'aigües residuals urbanes, PhD Thesis (in catalan), Universitat de Girona (1998).

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, *Chem. Eng. Prog.*, **95(6)**, 39-47, (1999).

Rodríguez-Roda, I., Sànchez-Marrè, M., Comas, J., Baeza, J., Colprim, J., Lafuente, J., Cortés, U., and Poch M., A hybrid supervisory system to support wastewater treatment plant operation, *Water Sci.. Technol.*, **45(4-5)**, 289, (2002).

Rosenberg, M. and Doyle, R.J., Microbial cell surface hydrophobicity: history, measurement, and significance. In Doyle, R.J and Rosenberg, M. (Ed.), *Microbial Cell Surface Hydrophobicity*, 1-38, ASM Publications, Washington D.C (1990).

Sadalgekar, V. V., Mahajan, B. A. and Shaligram, A. M., Evaluation of sludge settleability by floc characteristics, *J. Water Pollut. Contr. Fed.*, **60(10)**, 1862--1863 (1988).

Salvadó, H., Garcia, M.P and Amigo, J.M. Capability of ciliated protozoa as indicators of effluent quality in activated sludge plants, *Water Res.*, **29**, 1041-1050 (1995).

Sànchez-Marrè, M., DAI-DEPUR: an integrated supervisory multi-level architecture for wastewater treatment plants, PhD Thesis, Universitat Politècnica de Catalunya (1996).

Sànchez-Marrè, M., Cortés, U., Lafuente, J., R.-Roda, I., and Poch M., DAI-DEPUR: an integrated and distributed architecture for Wastewater Treatment Plants Supervision, *Artif. Intell. Eng.*, **10(3)**, 275-285 (1996).

Sànchez-Marrè, M., Cortés, U., R.-Roda, I., Poch, M. and Lafuente, J., Sustainable case learning for continuous domains, *Environ. Modell. Softw.*, **14(5)**, 349-357 (1999).

Sànchez-Marrè, M., Cortés, U., R.-Roda, I. and Poch, M. Using meta-cases to improve accuracy in hierarchical case retrieval. *Computación y Sistemas*, **4(1)**, 53-63 (2000).

Sànchez-Marrè, M., Cortés, U., Martínez, M., Comas, J. and Rodríguez-Roda, I., An approach for temporal Case-Based Reasoning: Episode-Based Reasoning, *Procc. of the 6th International Conference, ICCBR 2005* pages 465-476, Chicago,IL,USA, August (2005).

Sanin, F.C. and Vesilind, P.A., Bioflocculation of activated sludge: the role of calcium ions and extracellular polymers. *Environ Technol*, **21**, 1405-1412 (2000).

Sawyer, C.N and Bradney, L. Rising of activated sludge in final settling tanks, *Sewage Works J.*, **17**, 1191-1209 (1945).

Schamber, L. Relevance and information behaviour, *Annu. Rev. Info. Sci.*, M.E. Williams ed. American Society for Information Science, 29, 3-48. Medford, NJ.(1994).

Schank, R., *Dynamic Memory: a Theory of Learning in Computers & People*. Cambridge University Press (1982).

Serra, P., Desenvolupament d'un sistema basat en el coneixement per al control I supervisió de plantes depuradores d'aigües residuals urbanes, PhD Thesis (in catalan), Universitat Autònoma de Barcelona (1993).

Serra, P., Sánchez, M., Lafuente, J., Cortés, U., and Poch, M., ISCWAP: A knowledge-based system for supervising activated sludge processes, *Comput. Chem. Eng.*, **21(2)**, 211-221 (1997).

Seviour, E.M., Williams, C., DeGrey, B., Soddell, J.A., Seviour, R.J and Lindrea K.C., Studies on filamentous bacteria from Australian activated sludge plants, *Water Res.*, **28**, 2335-2342 (1994).

Seviour, R.J. and Blackall, L.L. *The Microbiology of Activated Sludge,* Edited by R.J.Seviour and L.L.Blackall, Kluwer Academic Publishers, The Netherlands (1999).

Sezgin, M., Jenkins, D. and Parker, D.S., A unified theory of filamentous activated sludge bulking, *J. Water Polln. Control Fedn.*, **50**, 362-381 (1978).

Simon, J., Wiese, J. and Steinmetz, H., Results of field tests and possible applications for in-situ sludge volume probes, *Proc. of the 2nd IWA Conference on Instrumentation, Control and Automation*, 305-316, May 29-June 2, Busan, Korea (2005).

Snaidr, J., Amann, R., Huber, I., Ludwig, W. and Schleifer, K.-H. Phylogenetic analysis and in situ identification of bacteria in activated sludge. *Appl. Env. Microb*, **63(7)**, 2884-2896 (1997).

Soddell, J.A., Seviour, R.J, Seviour, E.M and Stratton, H.M., Foaming and foam control in activated sludge systems, In *Prevention and Control of Bulking Activated Sludge*, 115-132, Jenkins, D., Ramadori, R. and Cingolani, L. eds, Luigi Bazzucchi center, Perugia, Italy (1993).

Starkey, J.E. and Karr, J.E., Effect of Low Dissolved Oxygen Concentration on Effluent Turbidity. *J. Water Pollut. Contr. Fed.*, **56(7)**, 837-843 (1984).

Stephanopoulos G. and Han C., Intelligent Systems in process Engineering: a review. *Comput. Chem. Eng.*, **20(6-7)**, 743-791 (1996).

Steyer JP, Rolland D., Bouvier JC., Moletta R., Hybrid fuzzy neural network for diagnosis-application to the anaerobic treatment of wine distillery wastewater in a fluidized bed reactor. *Water Sci. Technol.*, **36(6-7)**, 209-217 (1997).

Stobbe, G., *Über das Verhalten von Belebtschramm in aufsteigender Wasserbewegung*, Veröffentlichungen des Institutes für Siedlungwasserwirtschaft der Technischen Hochschule Hannover,18 (1964).

Surma J. and Brauschweig B. Case-based retrieval in process engineering: supporting design by reusing flowsheets. *Eng. Appl. Artif. Intel.*, **9(4)**, 385-391 (1996).

Takács I., Patry G.G. and Nolasco D., A dynamic model of the clarification-thickening process, *Water Res.*, **25 (10)**, 1263-1271 (1991).

Takács, I. and Fleit, E., Modelling of the micromorphology of the activated sludge floc: low DO, low F/M bulking, *Water Sci.Technol.*, **31(2)**, 235-243 (1995).

Tang, R., Shaw Jr., W. M., and Vevea, J.L., Towards the identification of the optimal number of relevance categories, *J. Am. Soc. Inform*, **50(3)**, 254-264 (1999).

Tanzio, M., Validate your engineering software!, *Chem. Eng. Prog.*, **97(7)**, 64-73 (2001)

Tsai, M-W. Wentzel, M.C and Ekama, G.A., The Effect of Residual Ammonia Concentration Under Aerobic Conditions on the Growth of *Microthrix parvicella* in Biological Nutrient Removal plants, *Water Res.*, **37**, 3009-3015 (2003).

Van der Waarde, J., Krooneman, J., Geurkink, B., van der Werf, A., Eikelboom, D., Beimfohr, C., Snaidr, J., Levantesi, C. and Tandoi, V., Molecular monitoring of bulking sludge in industrial wastewater treatment plants, *Water Sci.Technol.*, **46(1-2)**, 551-558 (2002).

Vanderhasselt, A., Aspegren, H., Vanrollegherm, P and Verstraete, W., Settling characterisation using on-line sensors at a full-scale wastewater treatment plant, *Water SA*, **25 (4)**, 453- 458 (1999).

van Niekerk, N., Kawahigashi, J., reichlin, D., Malea, A. and Jenkins, D., Foaming in anaerobic digesters-a survey and laboratory investigation, *J. Water Poll. Contr. Fed.*, **59**, 149-153 (1987).

Vanrolleghem, P.A., De Clercq, B., De Clercq, J., Devisscher, M., Kinnear, D.J. and Nopens, I., New measurement techniques for secondary settlers: a review, *Proc. of the 2nd IWA Conference on Instrumentation, Control and Automation*, 305-316, May 29-June 2, Busan, Korea (2005).

van Veldhuizen, H.M., van Loosdrecht, M.C.M and Heijnen, J.J , Modelling biological phosphorous and nitrogen removal in a full scale activated sludge process, *Water Res.*, **33**, 3459-3468 (1999).

Vitasovic C., Specific challenges for automation of wastewater facilities. WEFTEC 2001 Seminar #105 *Implementing Instrumentation, Control, and Information Technology in Activated Sludge Systems*, Water Environment Federation, Alexandria (USA) (2001).

Vouros G.A., Pantelakis I.S. and Lekkas T.D. Knowledge representation in an activated sludge plant diagnosis system. *Expert Syst.*, **17(5)**, 226-240 (2000).

Wagner, F., Study of the causes and prevention of sludge bulking in Germany, Chapter 2 in *Bulking of Activated Sludge: Preventative and Remedial Methods*, Eds. B.Chambers and E.J. Tomlinson, Ellis Horwood Ltd., Chichester, England (1982).

Wagner, M., Amann, R., Kämpfer, P., Assmus, B., Hartmann, A., Hutzler, P., Springer, N. and Scheleifer, K.-H. Identification and in situ detection of Gram-negative filamentous bacteria in activated sludge, *System. Appl. Microb.*, **17**, 405-417 (1994).

Walker, B.A., Smith, K.D. and Lenhart, J.E., Optimize control room communications, *Chem. Eng. Prog.*, **97(10)**, 54-59 (2001)

Wanner, J. *Activated Sludge Bulking and Foaming Control*, Technomic Publishing Co., Lancaster, PA (1994).

Wanner, J. and Grau, P., Identification of filamentous microorganisms from activated sludge: a compromise between wishes, needs and possibilities, *Water Res.*, **23**, 883-891 (1989).

WEF, Water Environment Federation, *Activated Sludge. WEF Manual of Practice No.OM-9*, second ed., Water Environment Federation, Alexandria (2002).

Wen, C. and Vassiliadis, C.A., Applying hybrid artificial intelligence techniques in wastewater treatment. engineering, *Eng. Appl. Artif. Intel*, **11**, 685-705 (1998).

White, M.J.D, *Settling of Activated Sludge*, Tech. Rept TR11, Water Res. Centre, Stevenage, England (1975).

White, D.C., Pinkart, H.C. and Ringelberg, D.B., *Manual of Environmental Microbiology*, C.J. Hurst Ed., ASM Press, Washington D.C (1997).

Wiese, J., Schmitt, A., Stahl, A., Hansen, J. and Schmitt, T.G., Experience management for wastewater treatment. GI-Workshopwoche "Lehren-Lernen-Wissen-Adaptivität", *Proceedings of the Workshop on Knowledge and Experience Management, FGWM03*, Oct. 6-8, Karlsruhe, Germany (2003).

Wilén, B.-M., Nielsen, J.L., Keiding, K., and Nielsen, P.H., Influence of microbial activity on the stability of activated sludge flocs, *Colloids Surf. B: Biointerfaces*, **18**, 145-156 (2000).

Wilén, B.-M., Keiding, K. and Nielsen, P.H., Influence flocculation of activated sludge flocs by stimulation of the aerobic biological activity, *Water Res.*, **38**, 3909-3919 (2004).

Wilderer, P.A., Bungartz, H-J., Lemmer, H., Wagner, M., Keller, J. and Wuertz, S., Modern scientific methods and their potential in wastewater science and technology. *Water Res.*, **36**, 370-393 (2002).

Wilson, D.R. and Martínez, T.R., Improved heterogeneous distance functions, *J. Artificial Intellig. Res.*, **6**, 1-34 (1997).

Woombs, M. and Laybourn-Parry, J. Seasonal species composition, density and role of nematodes in activated sludge effluent treatment works. *Water Res.*, **21**, 459-468 (1987).

Yang, C. T. and J. J. Kao, An expert-system for selecting and sequencing waste-water treatment processes, *WaterSci.Technol.*, **34(3-4)**, 347-353 (1996).

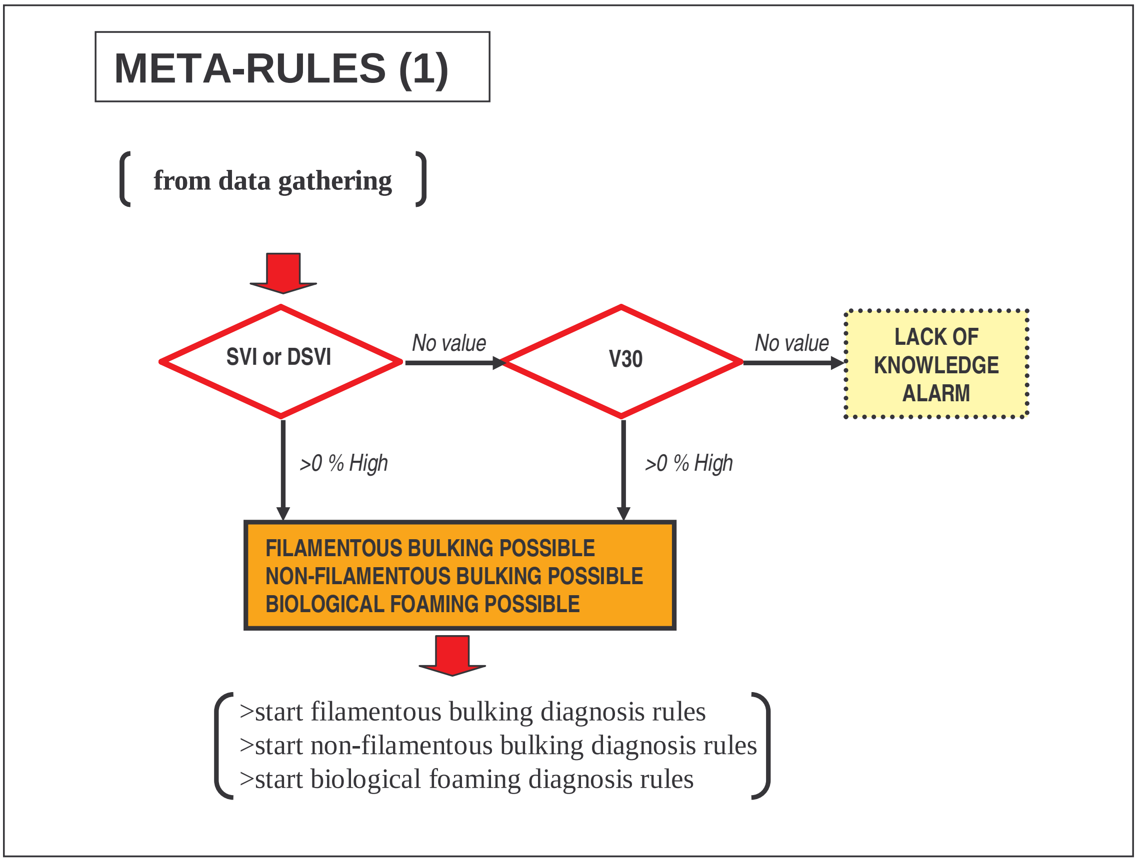
Zadeh, L., Fuzzy sets. *Inform. Control*, **8**, 338-353 (1965).

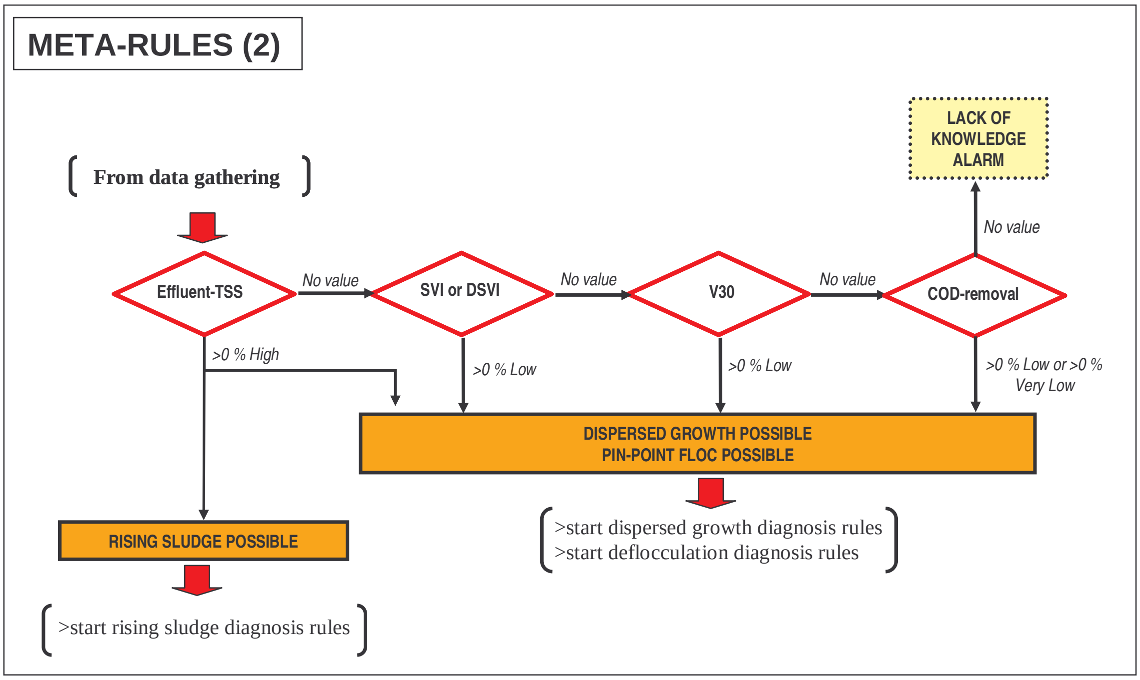
Zhao H., Hao O.J. and McAvoy T.J., Approaches to Modeling Nutrient Dynamics: ASM2, Simplified Model and Neural Nets, *Water Sci. Technol.*, **39(1)**, 227-234 (1999).

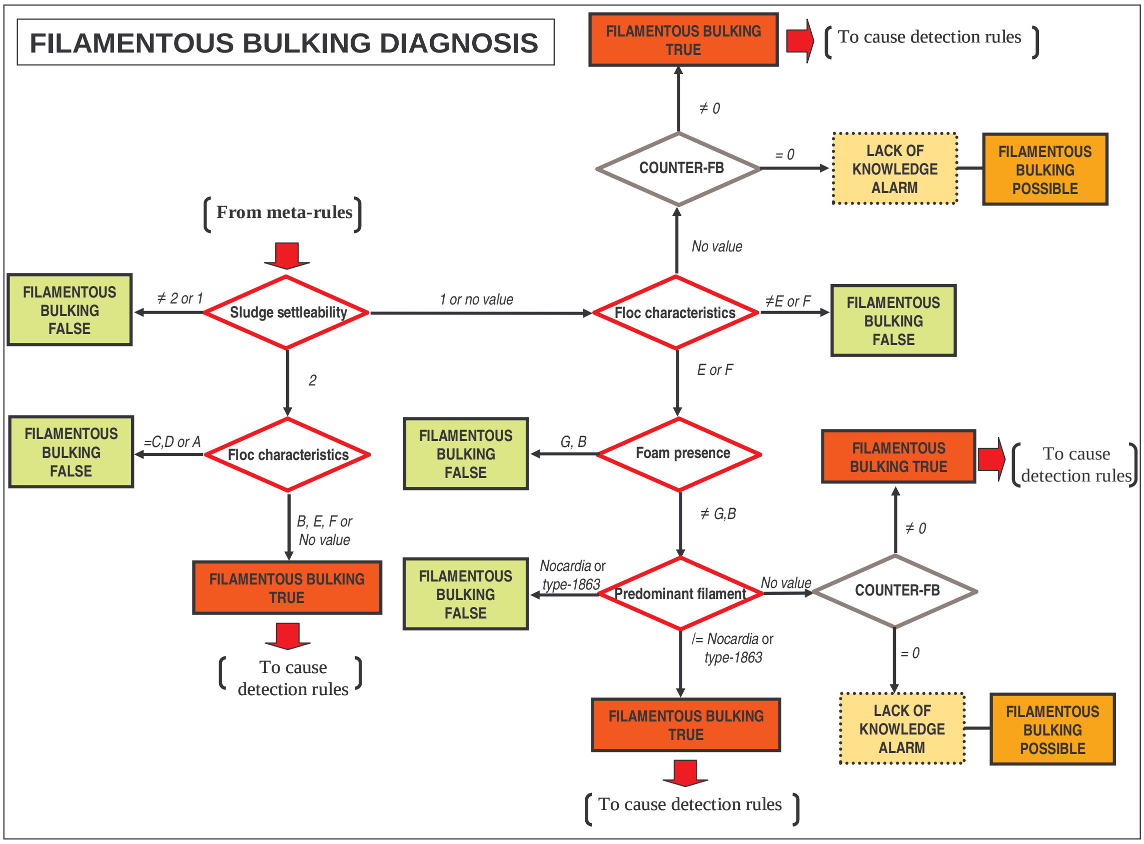
Zhu, X. X. and Simpson, A.R., Expert system for water treatment plant operation. *J. Environ. Eng.*, **122(9)**, 822 – 829 (1996).

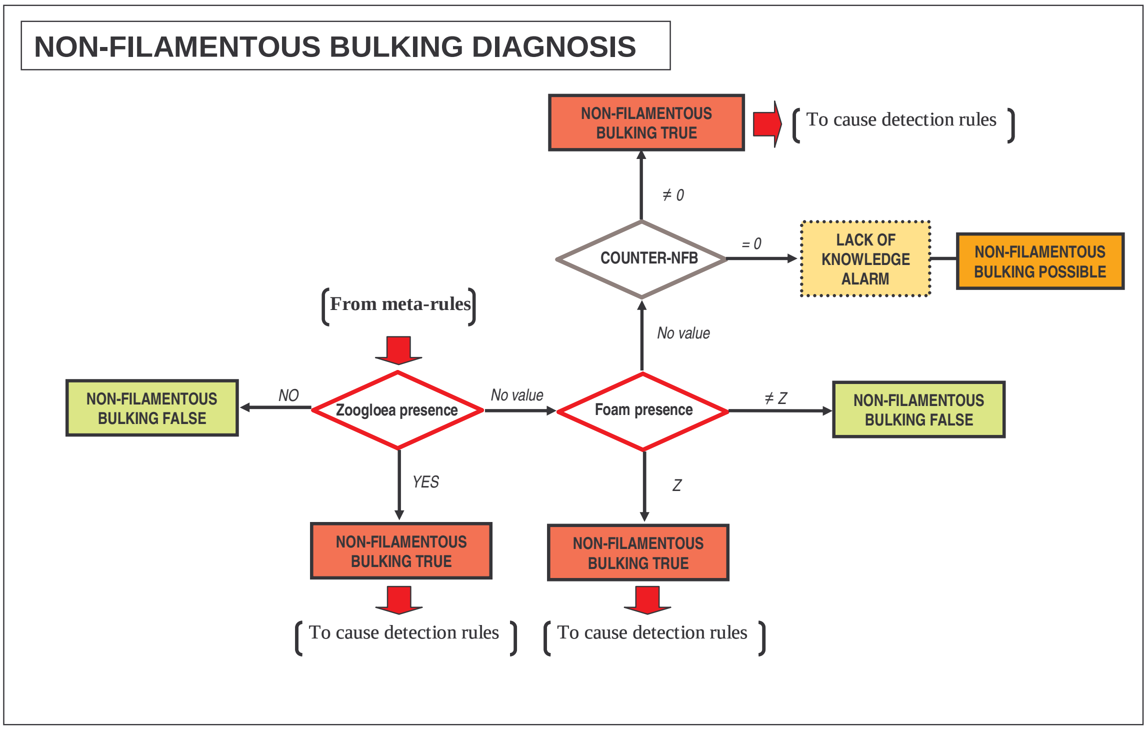
ANNEX

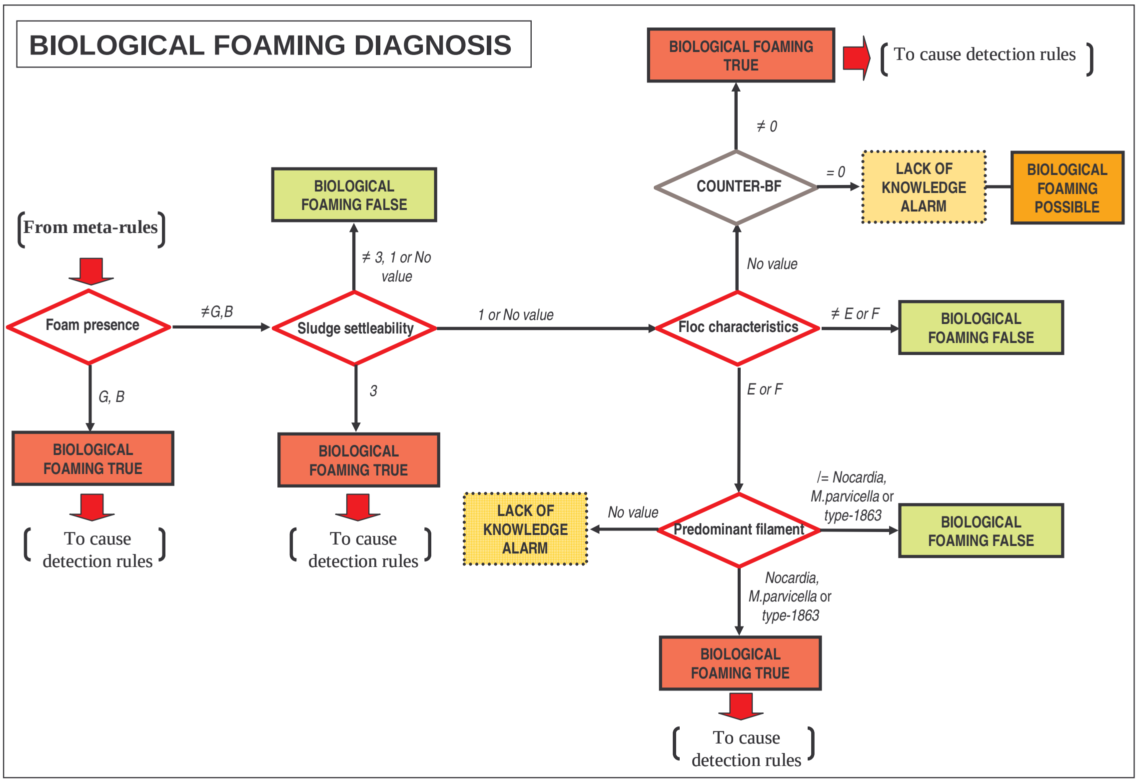
Decision trees to diagnose the different solids separation problems

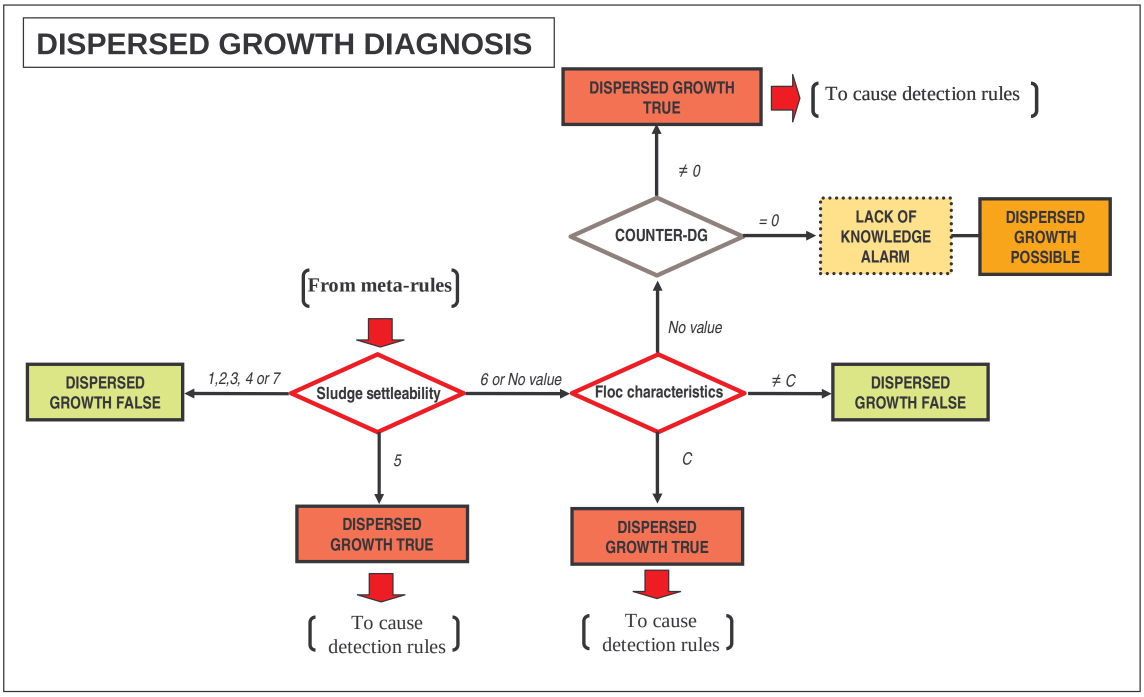


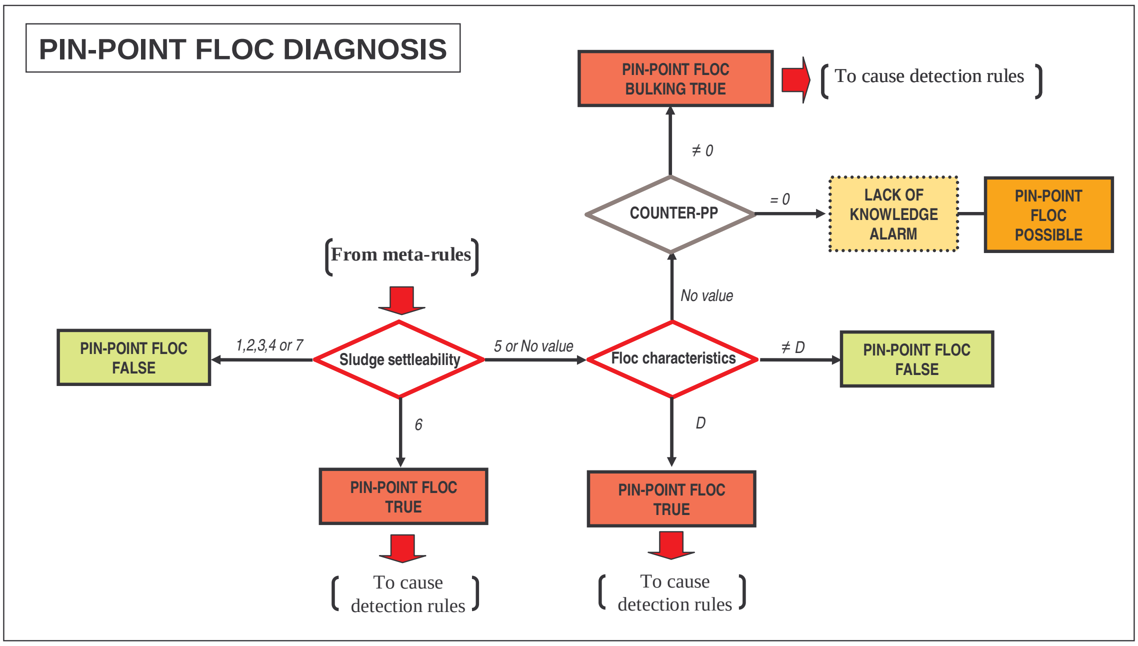



FILAMENTOUS BULKING DIAGNOSIS
FILAMENTOUS BULKING TRUE
To cause detection rules
≠ 0
COUNTER-FB
= 0
LACK OF KNOWLEDGE ALARM
FILAMENTOUS BULKING POSSIBLE
From meta-rules
No value
FILAMENTOUS BULKING FALSE
≠ 2 or 1
Sludge settleability
1 or no value
Floc characteristics
≠E or F
FILAMENTOUS BULKING FALSE
2
E or F
FILAMENTOUS BULKING FALSE
=C,D or A
Floc characteristics
FILAMENTOUS BULKING FALSE
G, B
Foam presence
FILAMENTOUS BULKING TRUE
To cause detection rules
B, E, F or No value
≠ G,B
≠ 0
FILAMENTOUS BULKING TRUE
FILAMENTOUS BULKING FALSE
Nocardia or type-1863
Predominant filament
No value
COUNTER-FB
To cause detection rules
/= Nocardia or type-1863
= 0
FILAMENTOUS BULKING TRUE
LACK OF KNOWLEDGE ALARM
FILAMENTOUS BULKING POSSIBLE
To cause detection rules

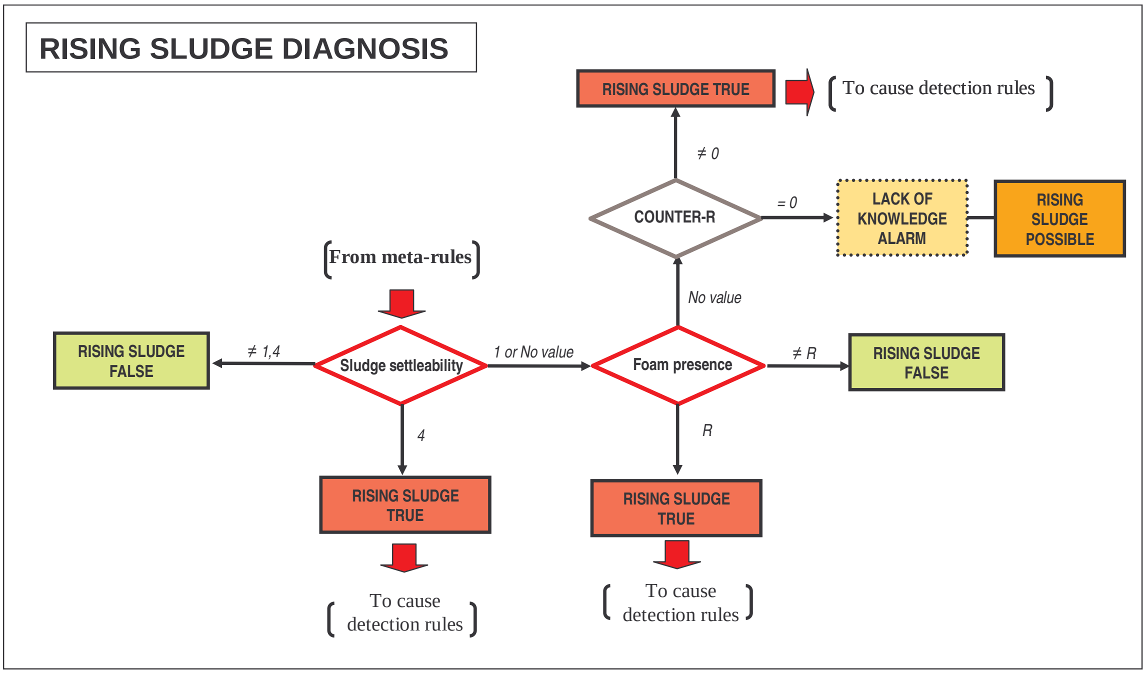

NON-FILAMENTOUS BULKING DIAGNOSIS
NON-FILAMENTOUS BULKING TRUE
To cause detection rules
≠ 0
COUNTER-NFB
= 0
LACK OF KNOWLEDGE ALARM
NON-FILAMENTOUS BULKING POSSIBLE
From meta-rules
No value
NON-FILAMENTOUS BULKING FALSE
NO
Zoogloea presence
No value
Foam presence
≠ Z
NON-FILAMENTOUS BULKING FALSE
YES
Z
NON-FILAMENTOUS BULKING TRUE
NON-FILAMENTOUS BULKING TRUE
To cause detection rules
To cause detection rules


BIOLOGICAL FOAMING DIAGNOSIS
BIOLOGICAL FOAMING TRUE
To cause detection rules
≠ 0
COUNTER-BF
= 0
LACK OF KNOWLEDGE ALARM
BIOLOGICAL FOAMING POSSIBLE
BIOLOGICAL FOAMING FALSE
From meta-rules
≠ 3, 1 or No value
No value
Foam presence
≠G,B
Sludge settleability
1 or No value
Floc characteristics
≠ E or F
BIOLOGICAL FOAMING FALSE
G, B
3
E or F
BIOLOGICAL FOAMING TRUE
BIOLOGICAL FOAMING TRUE
To cause detection rules
To cause detection rules
LACK OF KNOWLEDGE ALARM
No value
Predominant filament
/= Nocardia, M.parvicella or type-1863
BIOLOGICAL FOAMING FALSE
Nocardia, M.parvicella or type-1863
BIOLOGICAL FOAMING TRUE
To cause detection rules


DISPERSED GROWTH DIAGNOSIS
DISPERSED GROWTH TRUE
To cause detection rules
≠ 0
COUNTER-DG
= 0
LACK OF KNOWLEDGE ALARM
DISPERSED GROWTH POSSIBLE
From meta-rules
No value
DISPERSED GROWTH FALSE
1,2,3, 4 or 7
Sludge settleability
6 or No value
Floc characteristics
≠ C
DISPERSED GROWTH FALSE
5
C
DISPERSED GROWTH TRUE
DISPERSED GROWTH TRUE
To cause detection rules
To cause detection rules

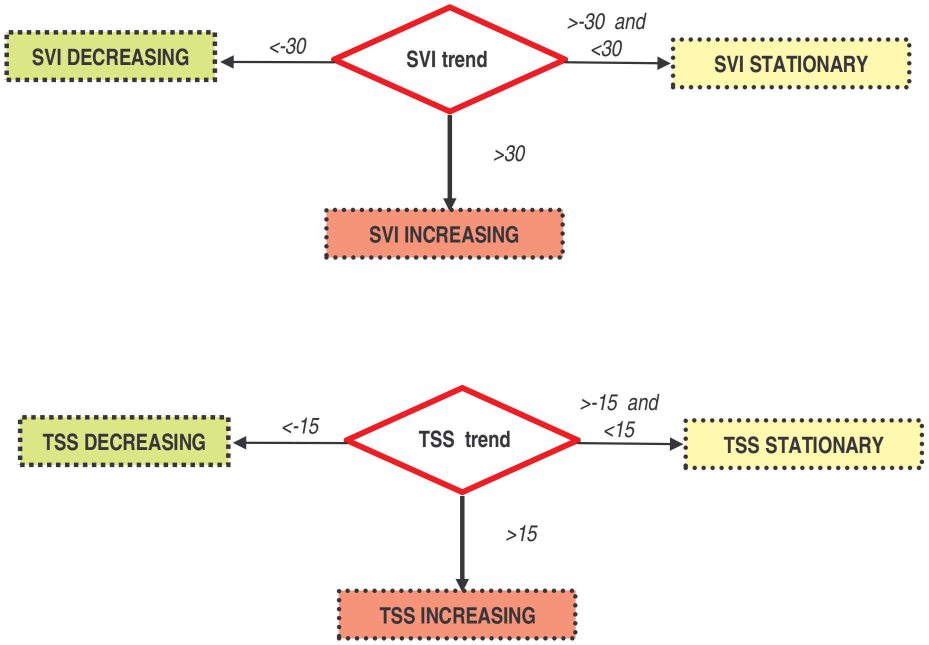


PIN-POINT FLOC DIAGNOSIS

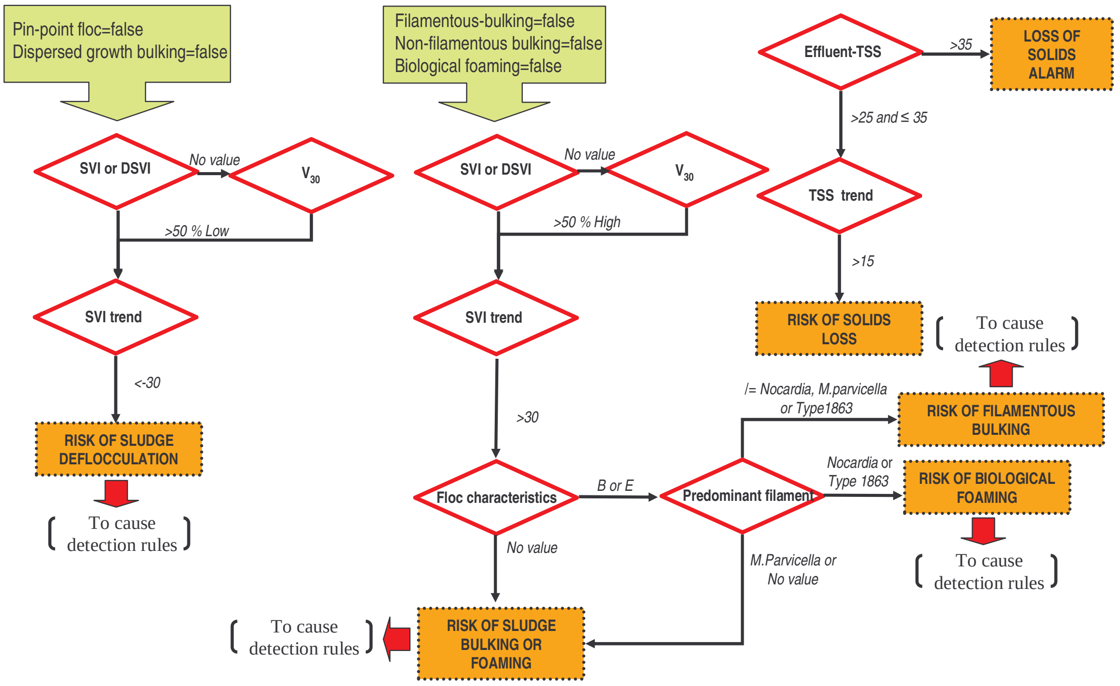


RISING SLUDGE DIAGNOSIS

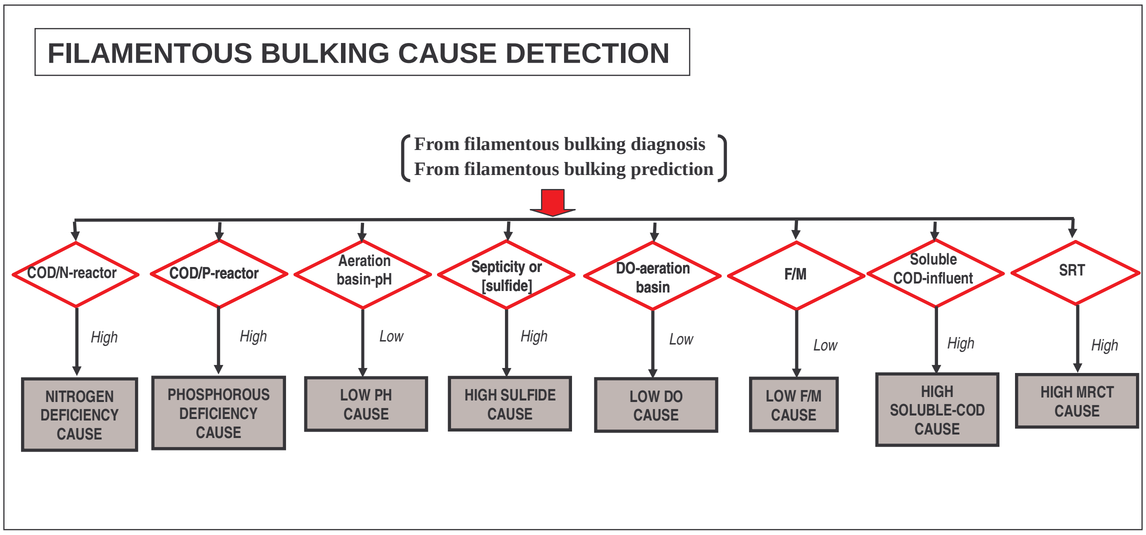
TRENDS CALCULATIONS

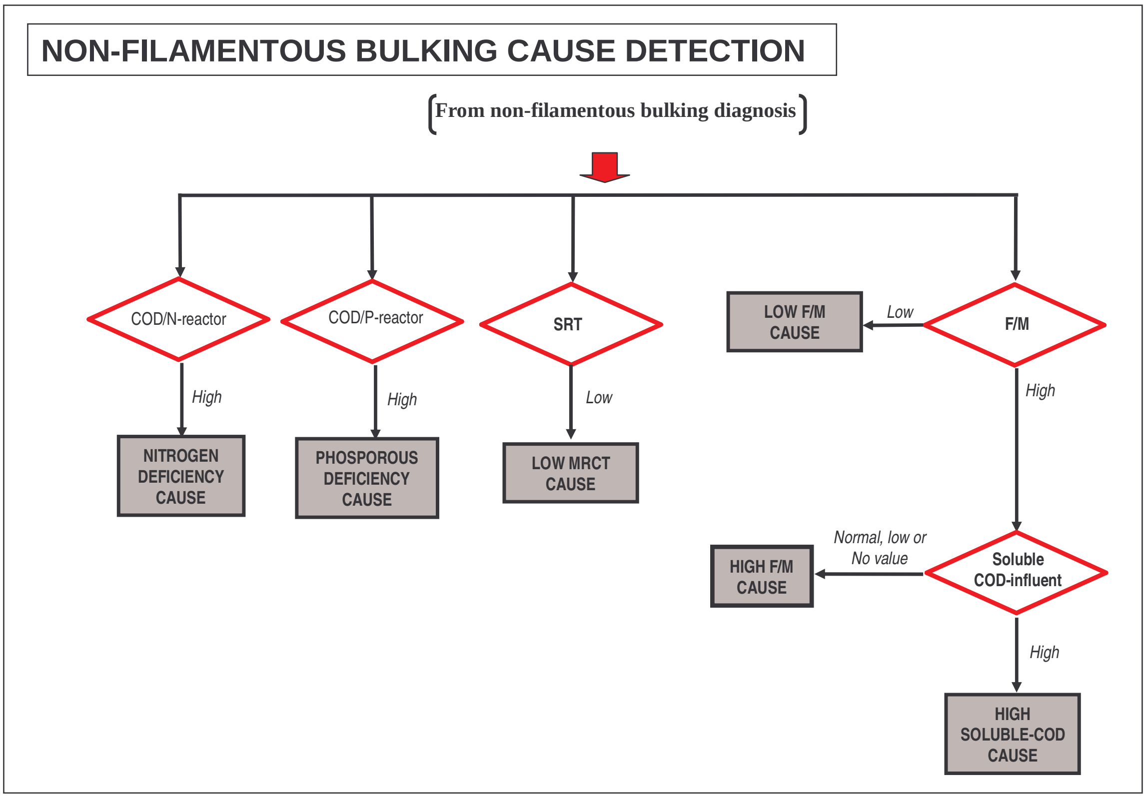


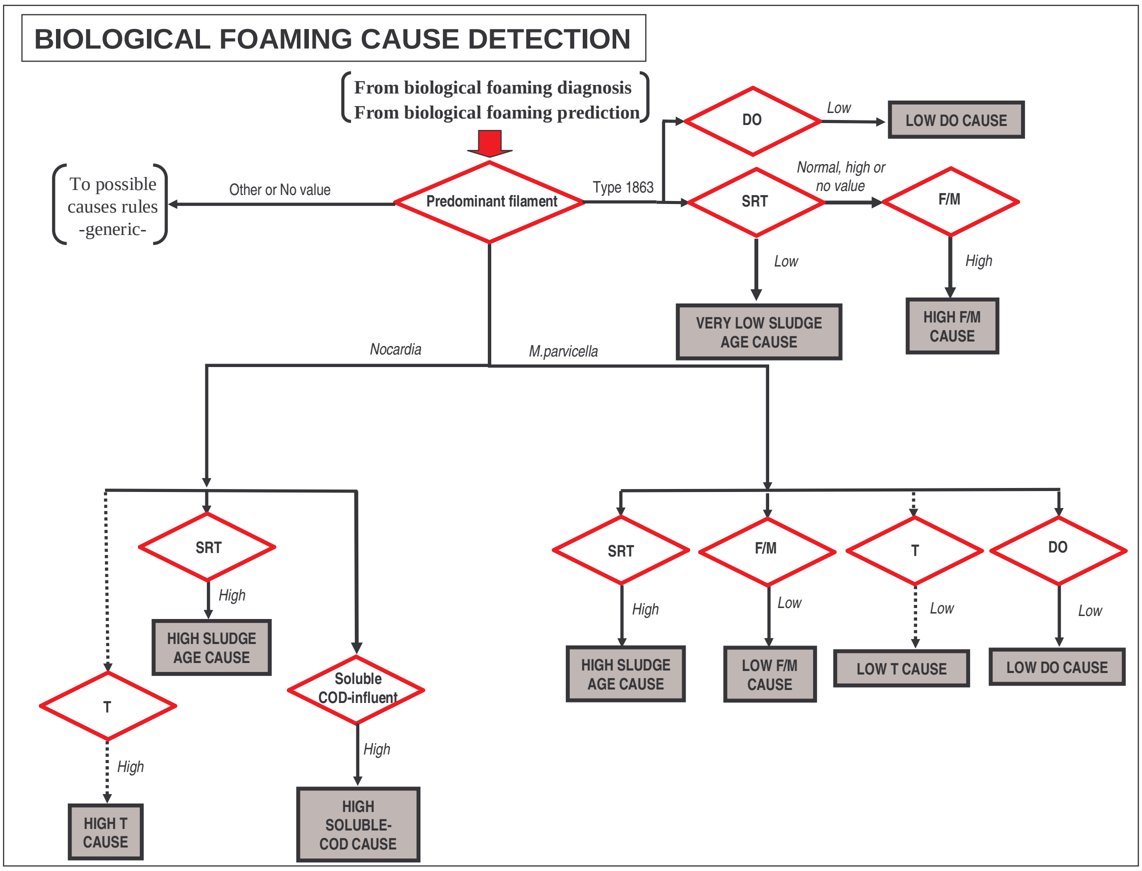
PREDICTION OF RISKY SITUATIONS AND ALARMS

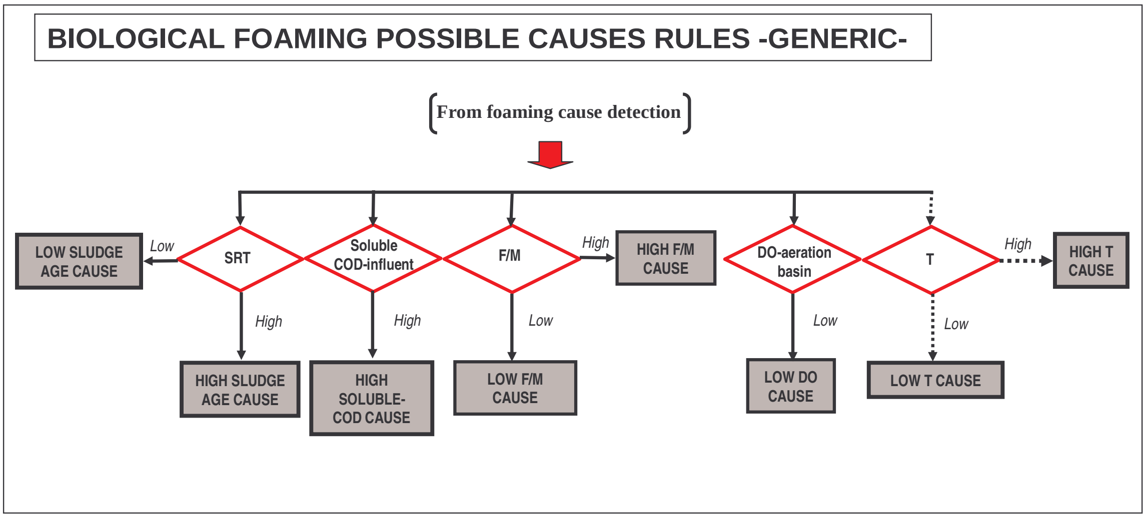


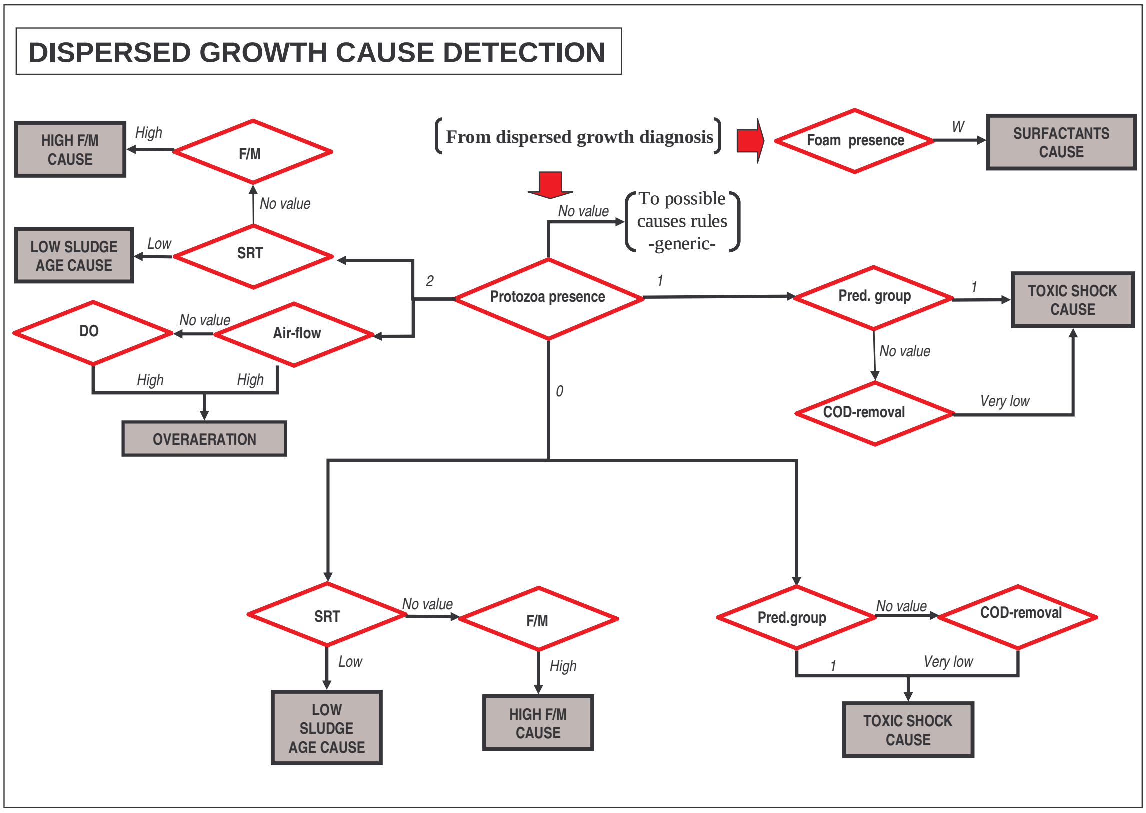
Decision trees to identify the possible causes



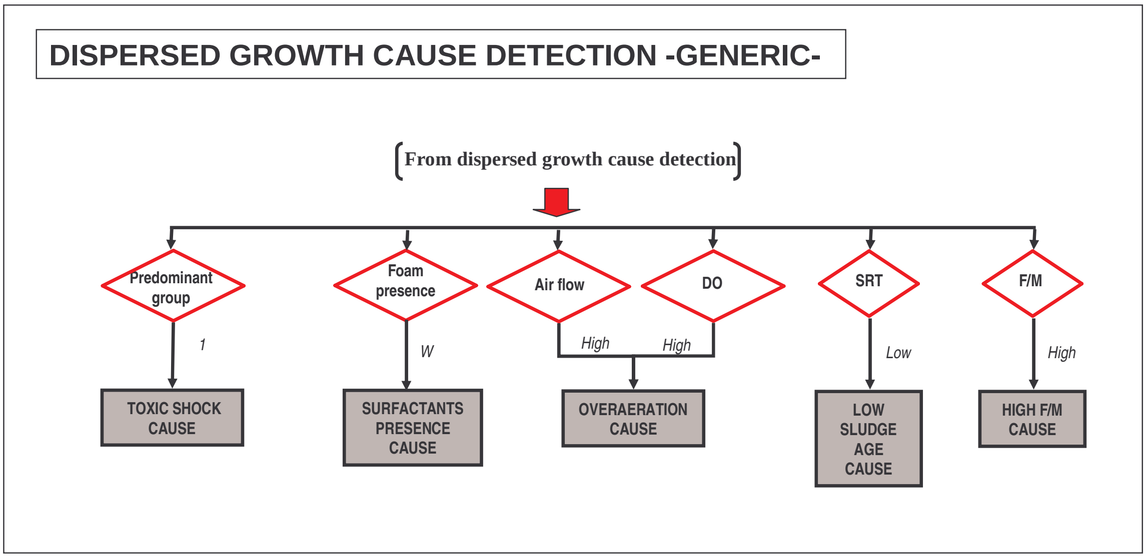



BIOLOGICAL FOAMING CAUSE DETECTION
From biological foaming diagnosis
From biological foaming prediction
To possible causes rules -generic-
Other or No value
Predominant filament
Type 1863
DO
Low
LOW DO CAUSE
Normal, high or no value
SRT
F/M
Low
High
VERY LOW SLUDGE AGE CAUSE
HIGH F/M CAUSE
Nocardia
M.parvicella
SRT
High
HIGH SLUDGE AGE CAUSE
T
High
HIGH T CAUSE
Soluble COD-influent
High
HIGH SOLUBLE-COD CAUSE
SRT
High
HIGH SLUDGE AGE CAUSE
F/M
Low
LOW F/M CAUSE
T
Low
LOW T CAUSE
DO
Low
LOW DO CAUSE


BIOLOGICAL FOAMING POSSIBLE CAUSES RULES -GENERIC-
From foaming cause detection
LOW SLUDGE AGE CAUSE
Low
SRT
High
HIGH SLUDGE AGE CAUSE
Soluble COD-influent
High
HIGH SOLUBLE-COD CAUSE
F/M
High
HIGH F/M CAUSE
Low
LOW F/M CAUSE
DO-aeration basin
Low
LOW DO CAUSE
T
High
HIGH T CAUSE
Low
LOW T CAUSE

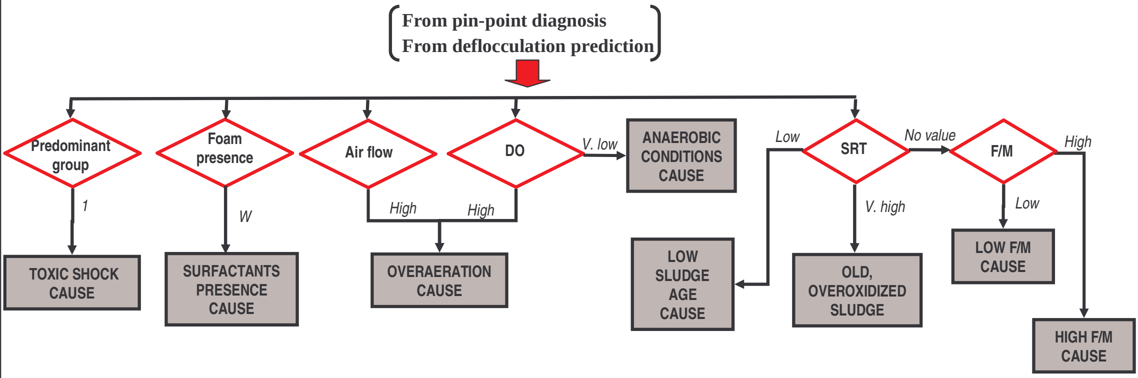


DISPERSED GROWTH CAUSE DETECTION



DISPERSED GROWTH CAUSE DETECTION -GENERIC-

PIN-POINT CAUSE DETECTION



From pin-point diagnosis
From deflocculation prediction

Predominant group
1
TOXIC SHOCK CAUSE

Foam presence
W
SURFACTANTS PRESENCE CAUSE

Air flow
High
DO
High
OVERAERATION CAUSE

DO
V. low
ANAEROBIC CONDITIONS CAUSE

SRT
Low
LOW SLUDGE AGE CAUSE
V. high
OLD, OVEROXIDIZED SLUDGE
No value

F/M
Low
LOW F/M CAUSE
High
HIGH F/M CAUSE

RISING SLUDGE CAUSE DETECTION

From rising sludge diagnosis

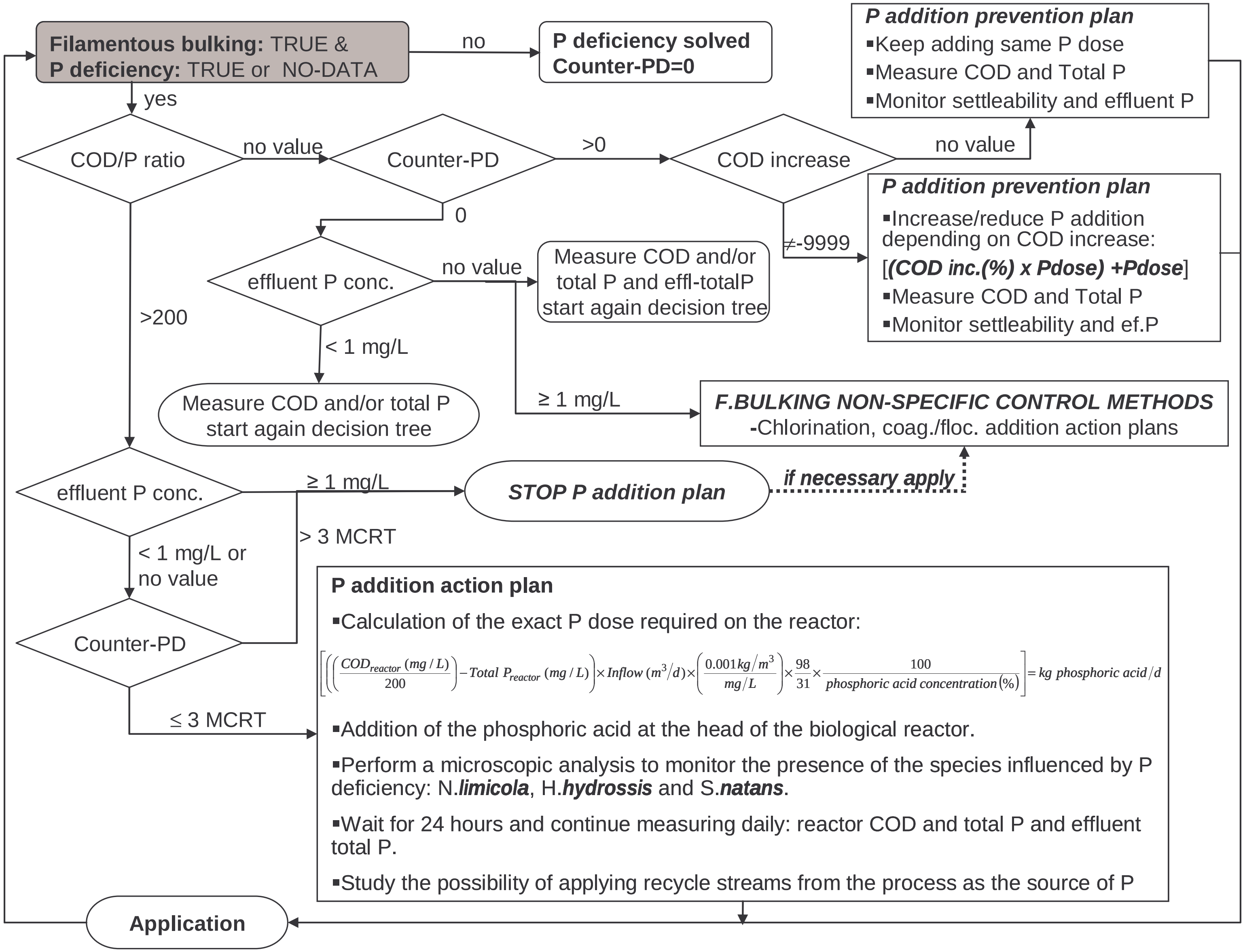
FAVORABLE CONDITIONS TO DENITRIFICATION IN CLARIFIERS CAUSE:

-High SRT

-High NO_3^- concentration

-High COD concentration

-Low DO concentration



P addition prevention plan

- Keep adding same P dose
- Measure COD and Total P
- Monitor settleability and effluent P

P addition prevention plan

- Increase/reduce P addition depending on COD increase:

[***(COD inc.(%) x Pdose) +Pdose***]

- Measure COD and Total P
- Monitor settleability and ef.P

F.BULKING NON-SPECIFIC CONTROL METHODS

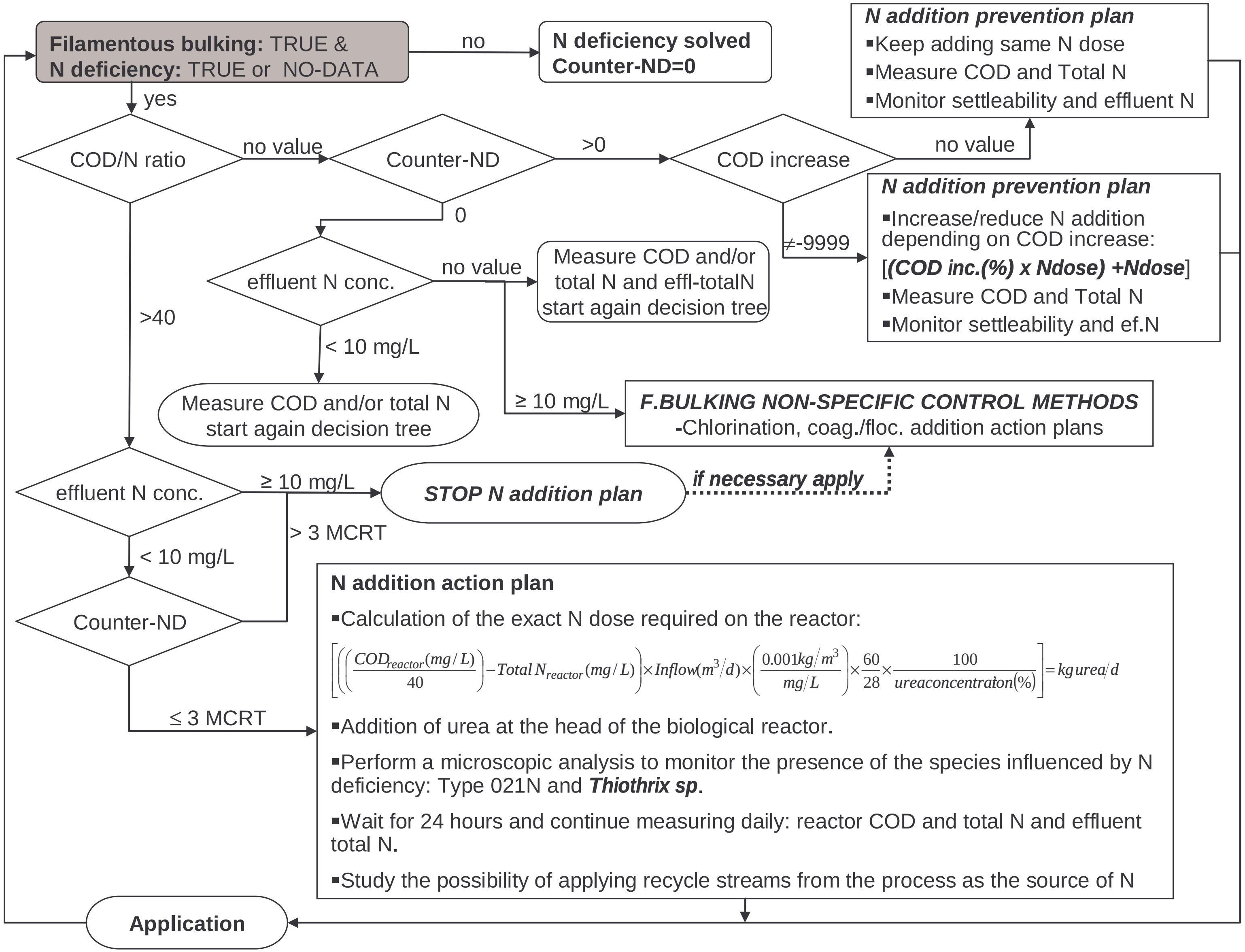
-Chlorination, coag./floc. addition action plans

P addition action plan

- Calculation of the exact P dose required on the reactor:

$$\left[\left(\left(\frac{COD_{reactor}\,(mg/L)}{200}\right)-Total\ P_{reactor}\,(mg/L)\right)\times Inflow\,(m^3/d)\times\left(\frac{0.001kg/m^3}{mg/L}\right)\times\frac{98}{31}\times\frac{100}{phosphoric\ acid\ concentration\,(\%)}\right]=kg\ phosphoric\ acid/d$$

- Addition of the phosphoric acid at the head of the biological reactor.
- Perform a microscopic analysis to monitor the presence of the species influenced by P deficiency: N.***limicola***, H.***hydrossis*** and S.***natans***.
- Wait for 24 hours and continue measuring daily: reactor COD and total P and effluent total P.
- Study the possibility of applying recycle streams from the process as the source of P

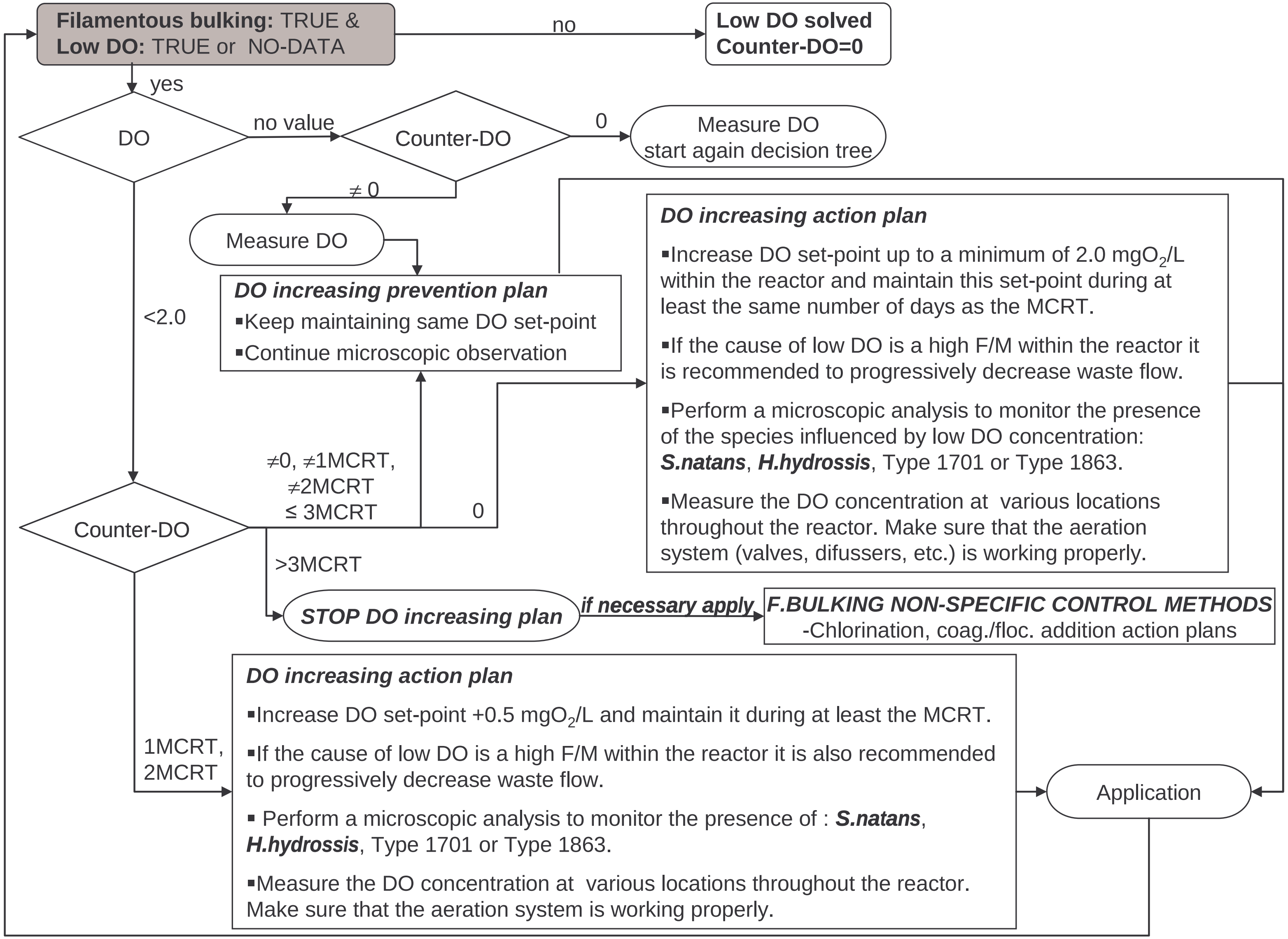


N addition action plan

- Calculation of the exact N dose required on the reactor:

$$\left[\left(\left(\frac{COD_{reactor}(mg/L)}{40}\right)-Total\,N_{reactor}(mg/L)\right)\times Inflow(m^3/d)\times\left(\frac{0.001kg/m^3}{mg/L}\right)\times\frac{60}{28}\times\frac{100}{ureaconcentraton(\%)}\right]=kg\,urea/d$$

- Addition of urea at the head of the biological reactor.
- Perform a microscopic analysis to monitor the presence of the species influenced by N deficiency: Type 021N and ***Thiothrix sp***.
- Wait for 24 hours and continue measuring daily: reactor COD and total N and effluent total N.
- Study the possibility of applying recycle streams from the process as the source of N



DO increasing prevention plan

- Keep maintaining same DO set-point
- Continue microscopic observation

DO increasing action plan

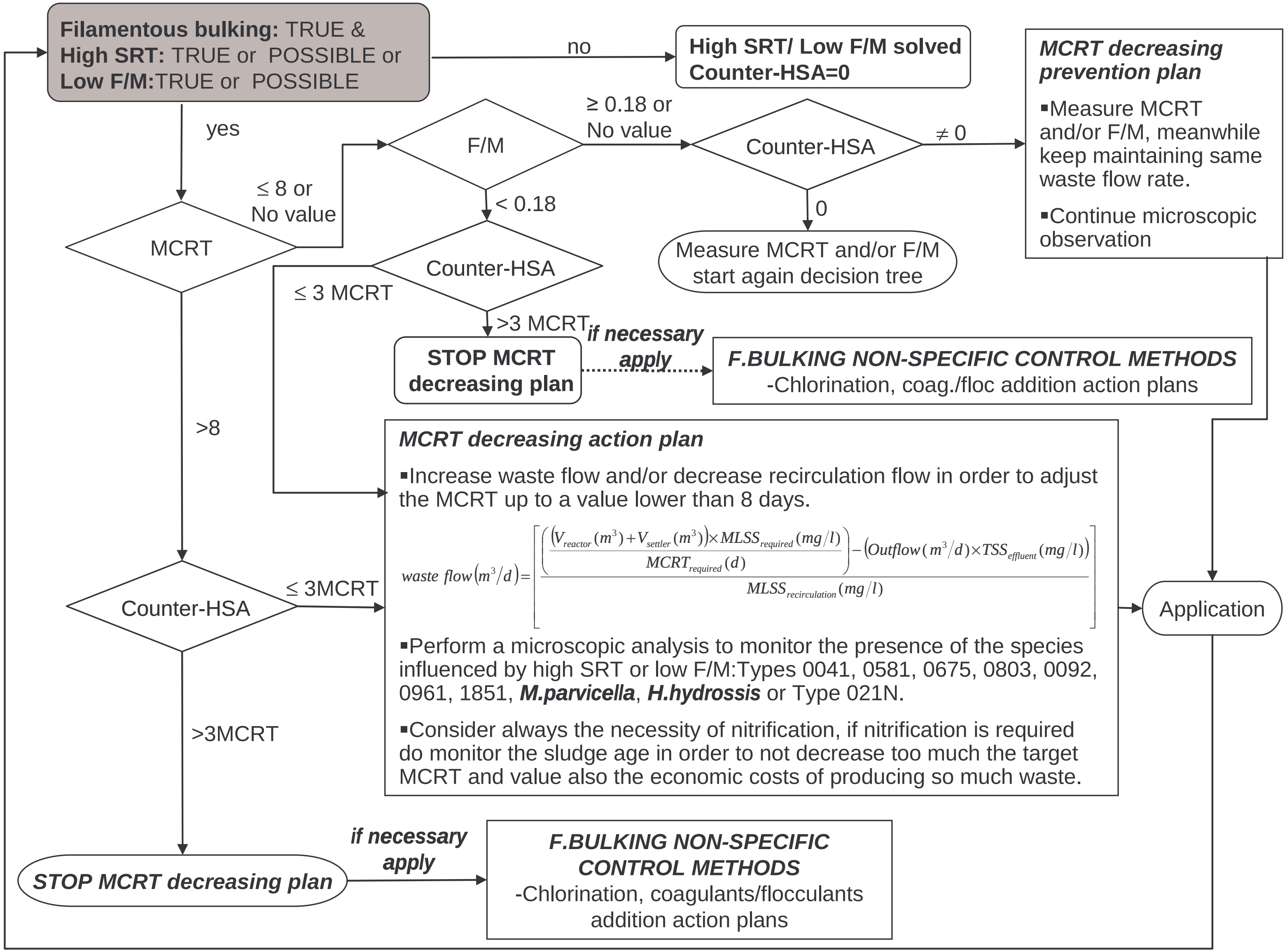
- Increase DO set-point up to a minimum of 2.0 mgO_2/L within the reactor and maintain this set-point during at least the same number of days as the MCRT.
- If the cause of low DO is a high F/M within the reactor it is recommended to progressively decrease waste flow.
- Perform a microscopic analysis to monitor the presence of the species influenced by low DO concentration: ***S.natans***, ***H.hydrossis***, Type 1701 or Type 1863.
- Measure the DO concentration at various locations throughout the reactor. Make sure that the aeration system (valves, difussers, etc.) is working properly.

F.BULKING NON-SPECIFIC CONTROL METHODS

-Chlorination, coag./floc. addition action plans

DO increasing action plan

- Increase DO set-point +0.5 mgO_2/L and maintain it during at least the MCRT.
- If the cause of low DO is a high F/M within the reactor it is also recommended to progressively decrease waste flow.
- Perform a microscopic analysis to monitor the presence of : ***S.natans***, ***H.hydrossis***, Type 1701 or Type 1863.
- Measure the DO concentration at various locations throughout the reactor. Make sure that the aeration system is working properly.



MCRT decreasing prevention plan

- ▪Measure MCRT and/or F/M, meanwhile keep maintaining same waste flow rate.
- ▪Continue microscopic observation

F.BULKING NON-SPECIFIC CONTROL METHODS

-Chlorination, coag./floc addition action plans

MCRT decreasing action plan

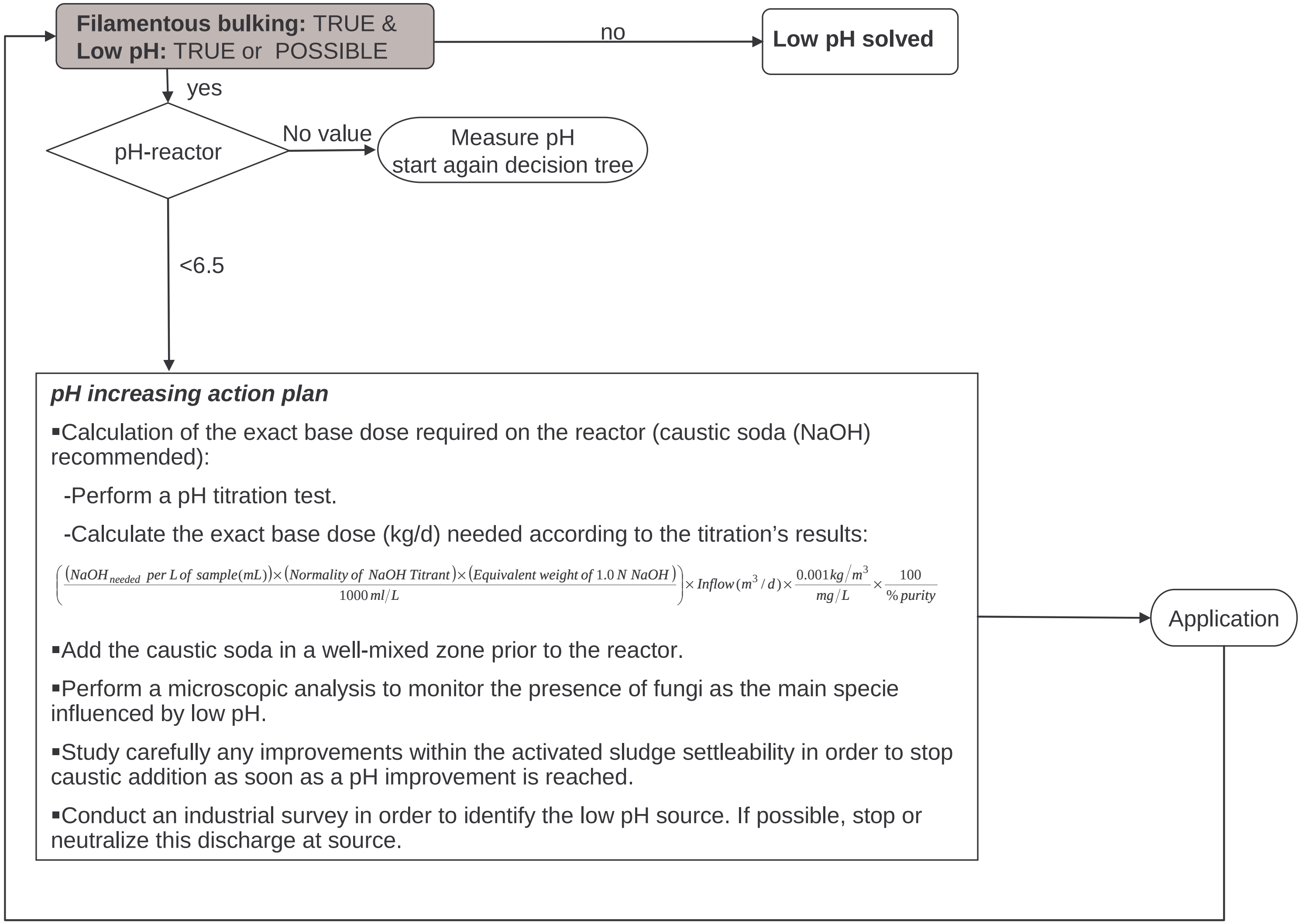
- ▪Increase waste flow and/or decrease recirculation flow in order to adjust the MCRT up to a value lower than 8 days.

$$waste\ flow(m^3/d)=\left[\frac{\left(\frac{\left(V_{reactor}(m^3)+V_{settler}(m^3)\right)\times MLSS_{required}(mg/l)}{MCRT_{required}(d)}\right)-\left(Outflow(m^3/d)\times TSS_{effluent}(mg/l)\right)}{MLSS_{recirculation}(mg/l)}\right]$$

- ▪Perform a microscopic analysis to monitor the presence of the species influenced by high SRT or low F/M:Types 0041, 0581, 0675, 0803, 0092, 0961, 1851, ***M.parvicella***, ***H.hydrossis*** or Type 021N.
- ▪Consider always the necessity of nitrification, if nitrification is required do monitor the sludge age in order to not decrease too much the target MCRT and value also the economic costs of producing so much waste.

F.BULKING NON-SPECIFIC CONTROL METHODS

-Chlorination, coagulants/flocculants addition action plans



pH increasing action plan

▪Calculation of the exact base dose required on the reactor (caustic soda (NaOH) recommended):

-Perform a pH titration test.

-Calculate the exact base dose (kg/d) needed according to the titration’s results:

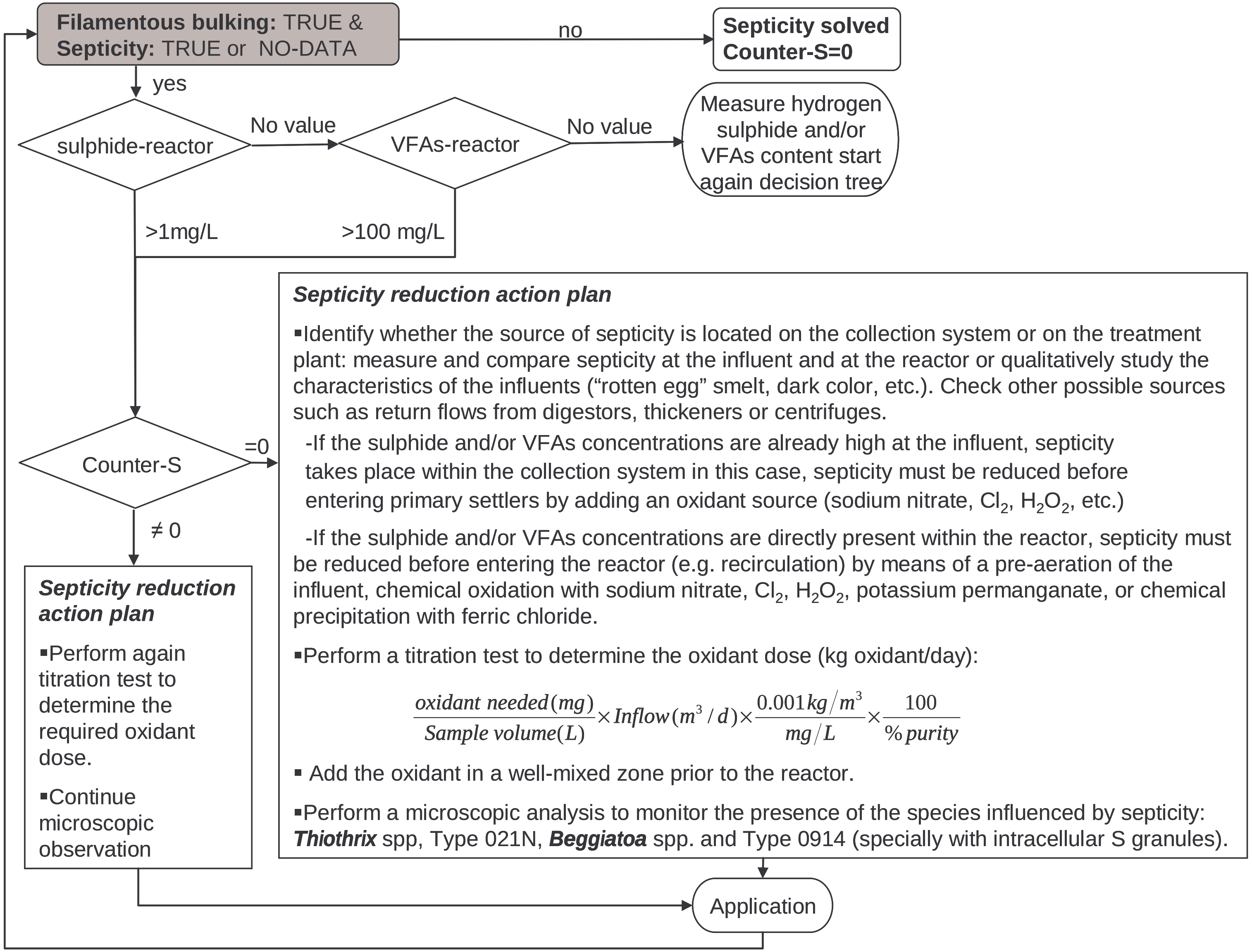
$$\left(\frac{(NaOH_{needed}\ per\ L\ of\ sample(mL))\times(Normality\ of\ NaOH\ Titrant)\times(Equivalent\ weight\ of\ 1.0\ N\ NaOH)}{1000\,ml/L}\right)\times Inflow\,(m^3/d)\times\frac{0.001\,kg/m^3}{mg/L}\times\frac{100}{\%\,purity}$$

▪Add the caustic soda in a well-mixed zone prior to the reactor.

▪Perform a microscopic analysis to monitor the presence of fungi as the main specie influenced by low pH.

▪Study carefully any improvements within the activated sludge settleability in order to stop caustic addition as soon as a pH improvement is reached.

▪Conduct an industrial survey in order to identify the low pH source. If possible, stop or neutralize this discharge at source.



Septicity reduction action plan

- Identify whether the source of septicity is located on the collection system or on the treatment plant: measure and compare septicity at the influent and at the reactor or qualitatively study the characteristics of the influents (“rotten egg” smelt, dark color, etc.). Check other possible sources such as return flows from digestors, thickeners or centrifuges.

-If the sulphide and/or VFAs concentrations are already high at the influent, septicity takes place within the collection system in this case, septicity must be reduced before entering primary settlers by adding an oxidant source (sodium nitrate, Cl_2, H_2O_2, etc.)

-If the sulphide and/or VFAs concentrations are directly present within the reactor, septicity must be reduced before entering the reactor (e.g. recirculation) by means of a pre-aeration of the influent, chemical oxidation with sodium nitrate, Cl_2, H_2O_2, potassium permanganate, or chemical precipitation with ferric chloride.

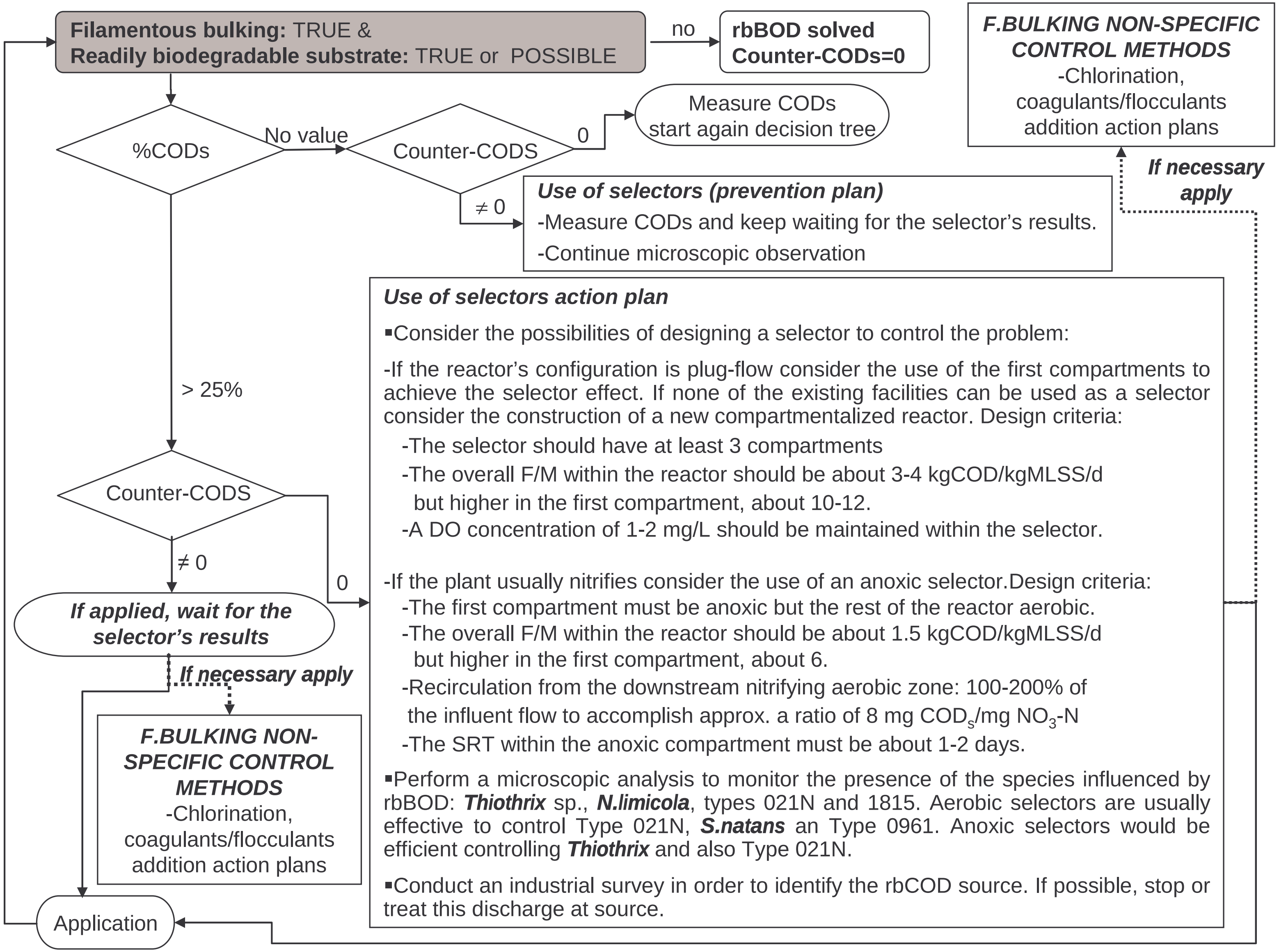
- Perform a titration test to determine the oxidant dose (kg oxidant/day):

$$\frac{oxidant\ needed(mg)}{Sample\ volume(L)}\times Inflow(m^3/d)\times\frac{0.001kg/m^3}{mg/L}\times\frac{100}{\%\,purity}$$

- Add the oxidant in a well-mixed zone prior to the reactor.
- Perform a microscopic analysis to monitor the presence of the species influenced by septicity: ***Thiothrix*** spp, Type 021N, ***Beggiatoa*** spp. and Type 0914 (specially with intracellular S granules).

Septicity reduction action plan

- Perform again titration test to determine the required oxidant dose.
- Continue microscopic observation



F.BULKING NON-SPECIFIC CONTROL METHODS
-Chlorination, coagulants/flocculants addition action plans

Use of selectors (prevention plan)
-Measure CODs and keep waiting for the selector's results.
-Continue microscopic observation

Use of selectors action plan

▪Consider the possibilities of designing a selector to control the problem:

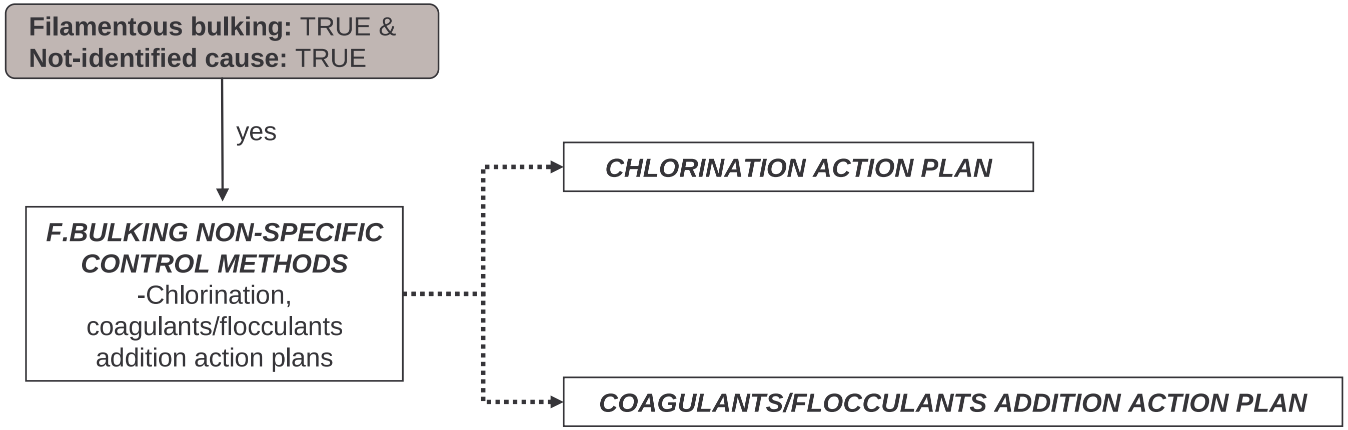
-If the reactor's configuration is plug-flow consider the use of the first compartments to achieve the selector effect. If none of the existing facilities can be used as a selector consider the construction of a new compartmentalized reactor. Design criteria:
  -The selector should have at least 3 compartments
  -The overall F/M within the reactor should be about 3-4 kgCOD/kgMLSS/d but higher in the first compartment, about 10-12.
  -A DO concentration of 1-2 mg/L should be maintained within the selector.

-If the plant usually nitrifies consider the use of an anoxic selector.Design criteria:
  -The first compartment must be anoxic but the rest of the reactor aerobic.
  -The overall F/M within the reactor should be about 1.5 kgCOD/kgMLSS/d but higher in the first compartment, about 6.
  -Recirculation from the downstream nitrifying aerobic zone: 100-200% of the influent flow to accomplish approx. a ratio of 8 mg COD_s/mg NO_3-N
  -The SRT within the anoxic compartment must be about 1-2 days.

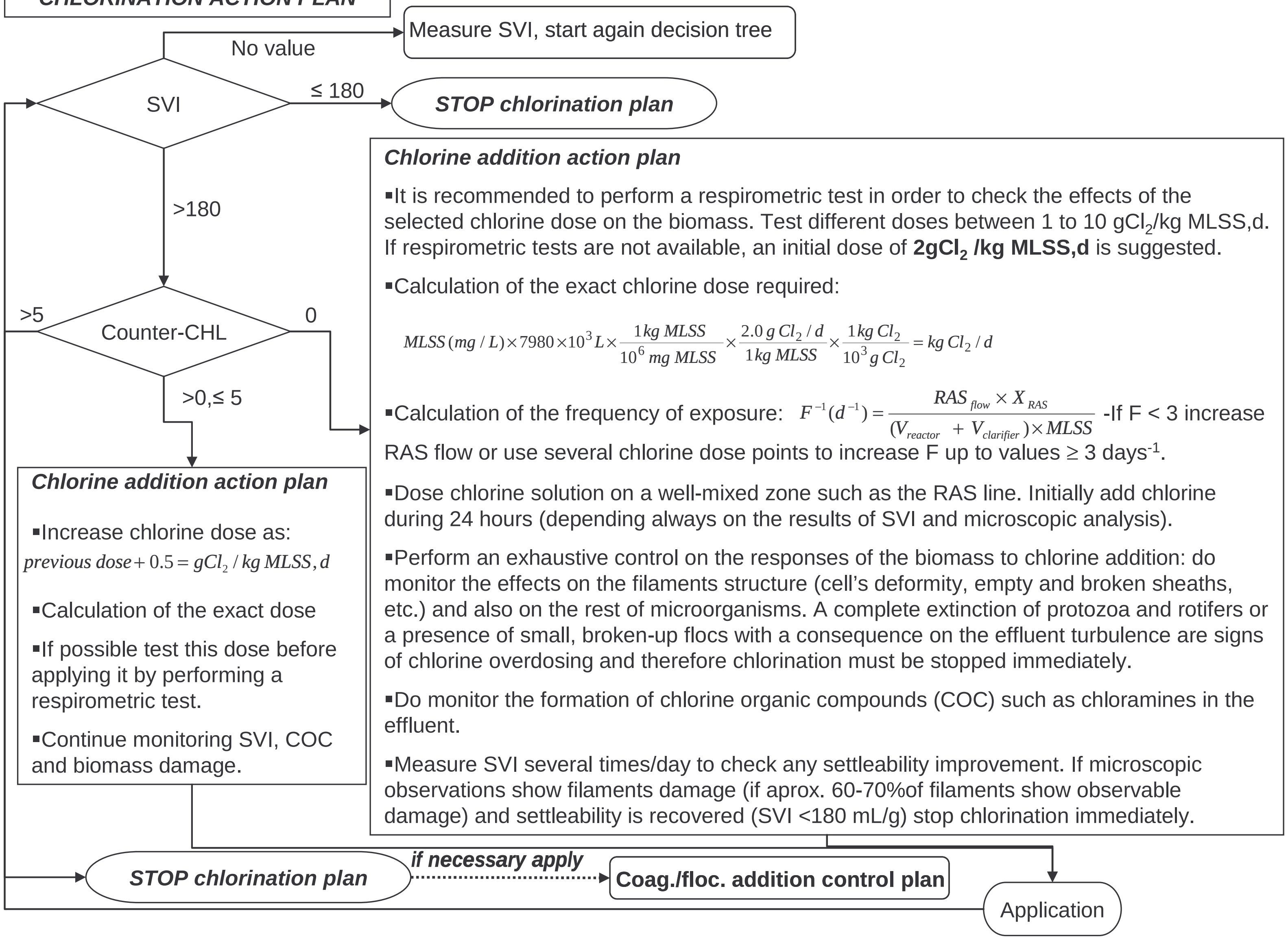
▪Perform a microscopic analysis to monitor the presence of the species influenced by rbBOD: ***Thiothrix*** sp., ***N.limicola***, types 021N and 1815. Aerobic selectors are usually effective to control Type 021N, ***S.natans*** an Type 0961. Anoxic selectors would be efficient controlling ***Thiothrix*** and also Type 021N.

▪Conduct an industrial survey in order to identify the rbCOD source. If possible, stop or treat this discharge at source.

F.BULKING NON-SPECIFIC CONTROL METHODS
-Chlorination, coagulants/flocculants addition action plans


Filamentous bulking: TRUE &
Not-identified cause: TRUE
yes
F.BULKING NON-SPECIFIC CONTROL METHODS
-Chlorination, coagulants/flocculants addition action plans
CHLORINATION ACTION PLAN
COAGULANTS/FLOCCULANTS ADDITION ACTION PLAN

CHLORINATION ACTION PLAN



Chlorine addition action plan

- Increase chlorine dose as:

$$previous\ dose + 0.5 = gCl_2/kg\ MLSS, d$$

- Calculation of the exact dose
- If possible test this dose before applying it by performing a respirometric test.
- Continue monitoring SVI, COC and biomass damage.

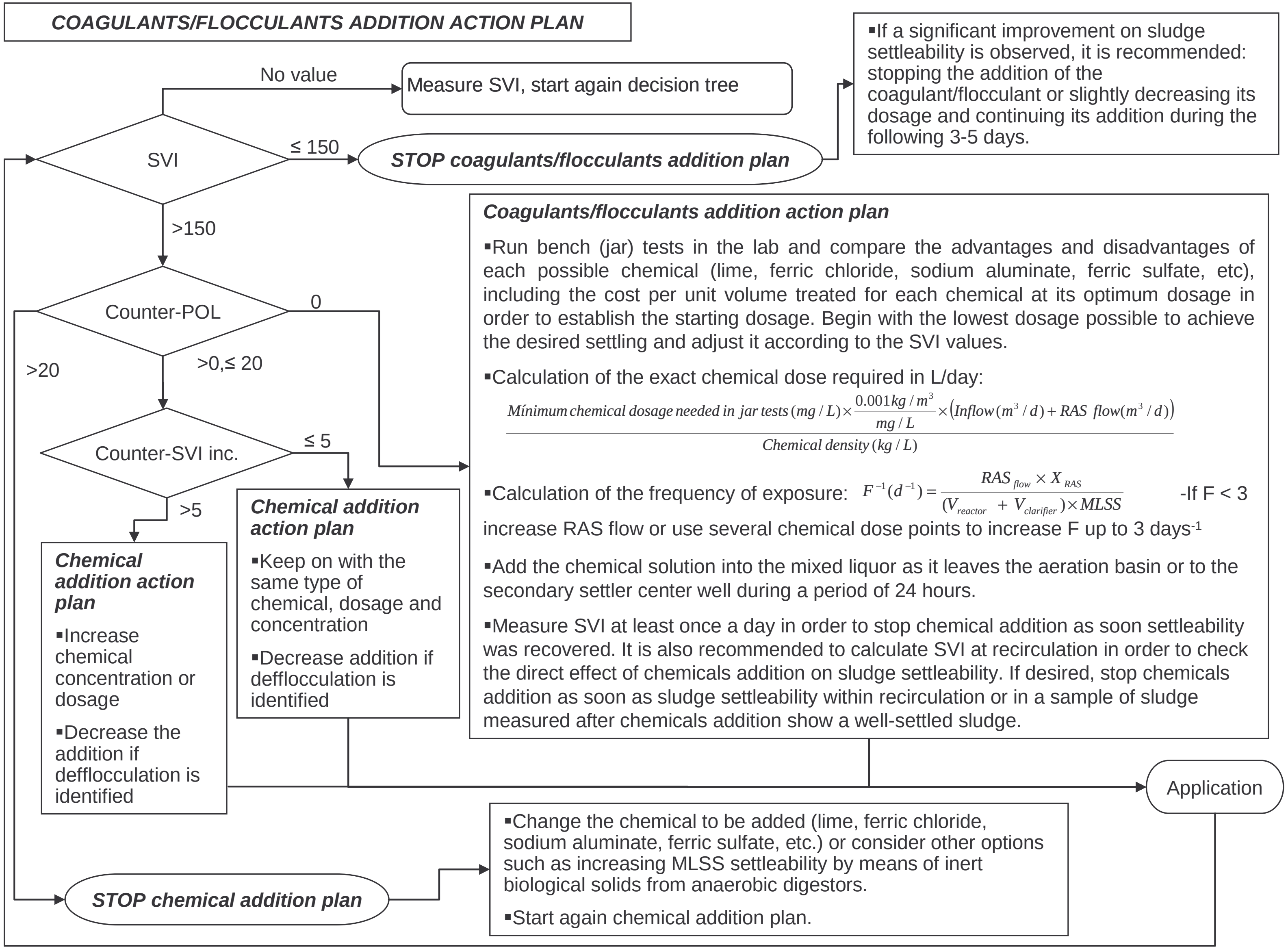
Chlorine addition action plan

- It is recommended to perform a respirometric test in order to check the effects of the selected chlorine dose on the biomass. Test different doses between 1 to 10 gCl_2/kg MLSS,d. If respirometric tests are not available, an initial dose of **2gCl_2 /kg MLSS,d** is suggested.
- Calculation of the exact chlorine dose required:

$$MLSS\,(mg/L)\times 7980\times 10^3\,L\times\frac{1\,kg\,MLSS}{10^6\,mg\,MLSS}\times\frac{2.0\,g\,Cl_2/d}{1\,kg\,MLSS}\times\frac{1\,kg\,Cl_2}{10^3\,g\,Cl_2}=kg\,Cl_2/d$$

- Calculation of the frequency of exposure: $F^{-1}(d^{-1})=\frac{RAS_{flow}\times X_{RAS}}{(V_{reactor}+V_{clarifier})\times MLSS}$ -If F < 3 increase RAS flow or use several chlorine dose points to increase F up to values ≥ 3 $days^{-1}$.
- Dose chlorine solution on a well-mixed zone such as the RAS line. Initially add chlorine during 24 hours (depending always on the results of SVI and microscopic analysis).
- Perform an exhaustive control on the responses of the biomass to chlorine addition: do monitor the effects on the filaments structure (cell's deformity, empty and broken sheaths, etc.) and also on the rest of microorganisms. A complete extinction of protozoa and rotifers or a presence of small, broken-up flocs with a consequence on the effluent turbulence are signs of chlorine overdosing and therefore chlorination must be stopped immediately.
- Do monitor the formation of chlorine organic compounds (COC) such as chloramines in the effluent.
- Measure SVI several times/day to check any settleability improvement. If microscopic observations show filaments damage (if aprox. 60-70%of filaments show observable damage) and settleability is recovered (SVI <180 mL/g) stop chlorination immediately.

COAGULANTS/FLOCCULANTS ADDITION ACTION PLAN



- If a significant improvement on sludge settleability is observed, it is recommended: stopping the addition of the coagulant/flocculant or slightly decreasing its dosage and continuing its addition during the following 3-5 days.

Coagulants/flocculants addition action plan

- Run bench (jar) tests in the lab and compare the advantages and disadvantages of each possible chemical (lime, ferric chloride, sodium aluminate, ferric sulfate, etc), including the cost per unit volume treated for each chemical at its optimum dosage in order to establish the starting dosage. Begin with the lowest dosage possible to achieve the desired settling and adjust it according to the SVI values.
- Calculation of the exact chemical dose required in L/day:

$$\frac{Mínimum\ chemical\ dosage\ needed\ in\ jar\ tests\,(mg/L)\times\frac{0.001 kg/m^3}{mg/L}\times\left(Inflow\,(m^3/d)+RAS\ \ flow(m^3/d)\right)}{Chemical\ density\,(kg/L)}$$

- Calculation of the frequency of exposure: $F^{-1}(d^{-1})=\frac{RAS_{flow}\times X_{RAS}}{(V_{reactor}+V_{clarifier})\times MLSS}$ -If F < 3 increase RAS flow or use several chemical dose points to increase F up to 3 days^{-1}
- Add the chemical solution into the mixed liquor as it leaves the aeration basin or to the secondary settler center well during a period of 24 hours.
- Measure SVI at least once a day in order to stop chemical addition as soon settleability was recovered. It is also recommended to calculate SVI at recirculation in order to check the direct effect of chemicals addition on sludge settleability. If desired, stop chemicals addition as soon as sludge settleability within recirculation or in a sample of sludge measured after chemicals addition show a well-settled sludge.

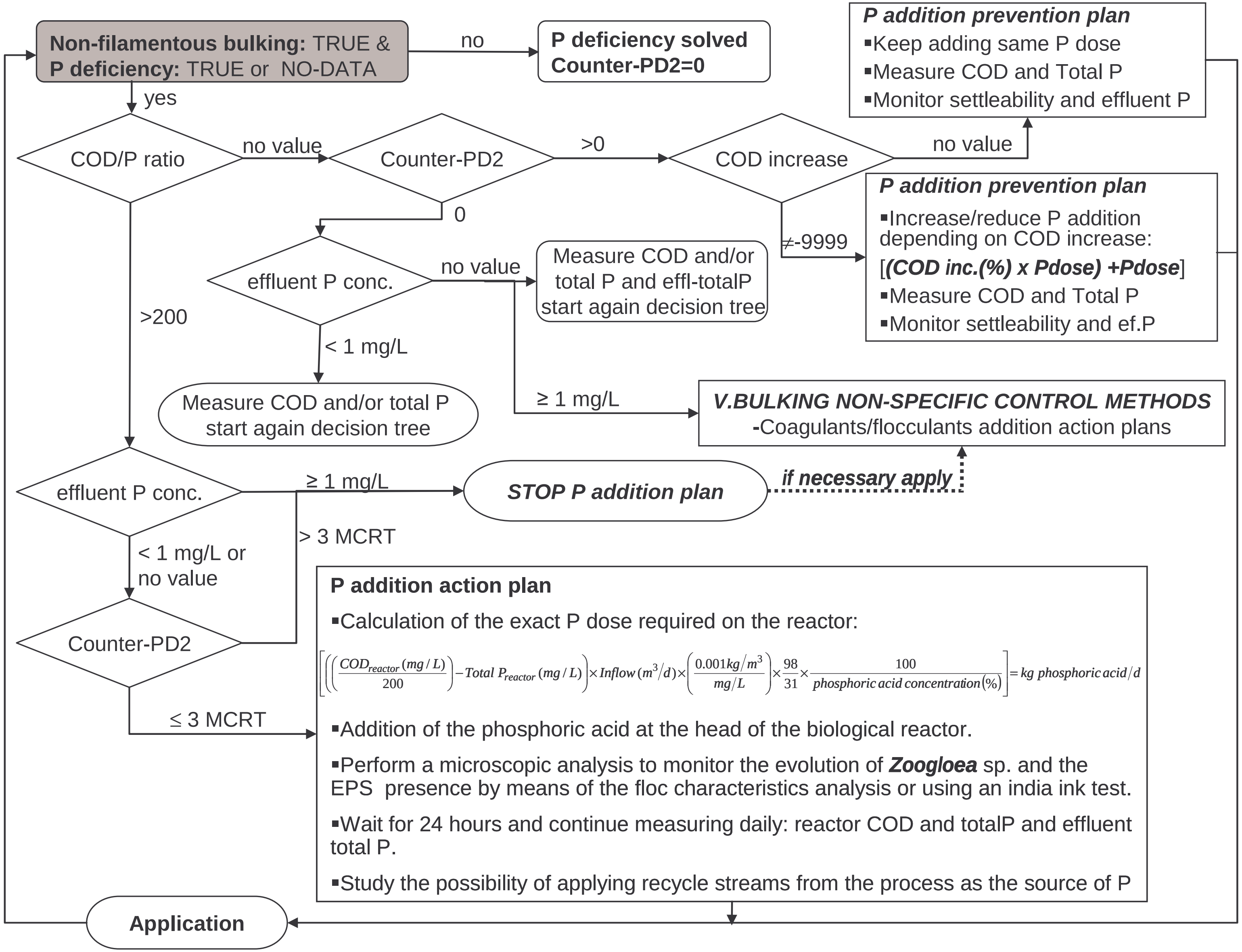
Chemical addition action plan

- Keep on with the same type of chemical, dosage and concentration
- Decrease addition if deflocculation is identified

Chemical addition action plan

- Increase chemical concentration or dosage
- Decrease the addition if deflocculation is identified

- Change the chemical to be added (lime, ferric chloride, sodium aluminate, ferric sulfate, etc.) or consider other options such as increasing MLSS settleability by means of inert biological solids from anaerobic digestors.
- Start again chemical addition plan.



P addition prevention plan

- Keep adding same P dose
- Measure COD and Total P
- Monitor settleability and effluent P

P addition prevention plan

- Increase/reduce P addition depending on COD increase:
[***(COD inc.(%) x Pdose) +Pdose***]
- Measure COD and Total P
- Monitor settleability and ef.P

Measure COD and/or total P and effl-totalP start again decision tree

Measure COD and/or total P start again decision tree

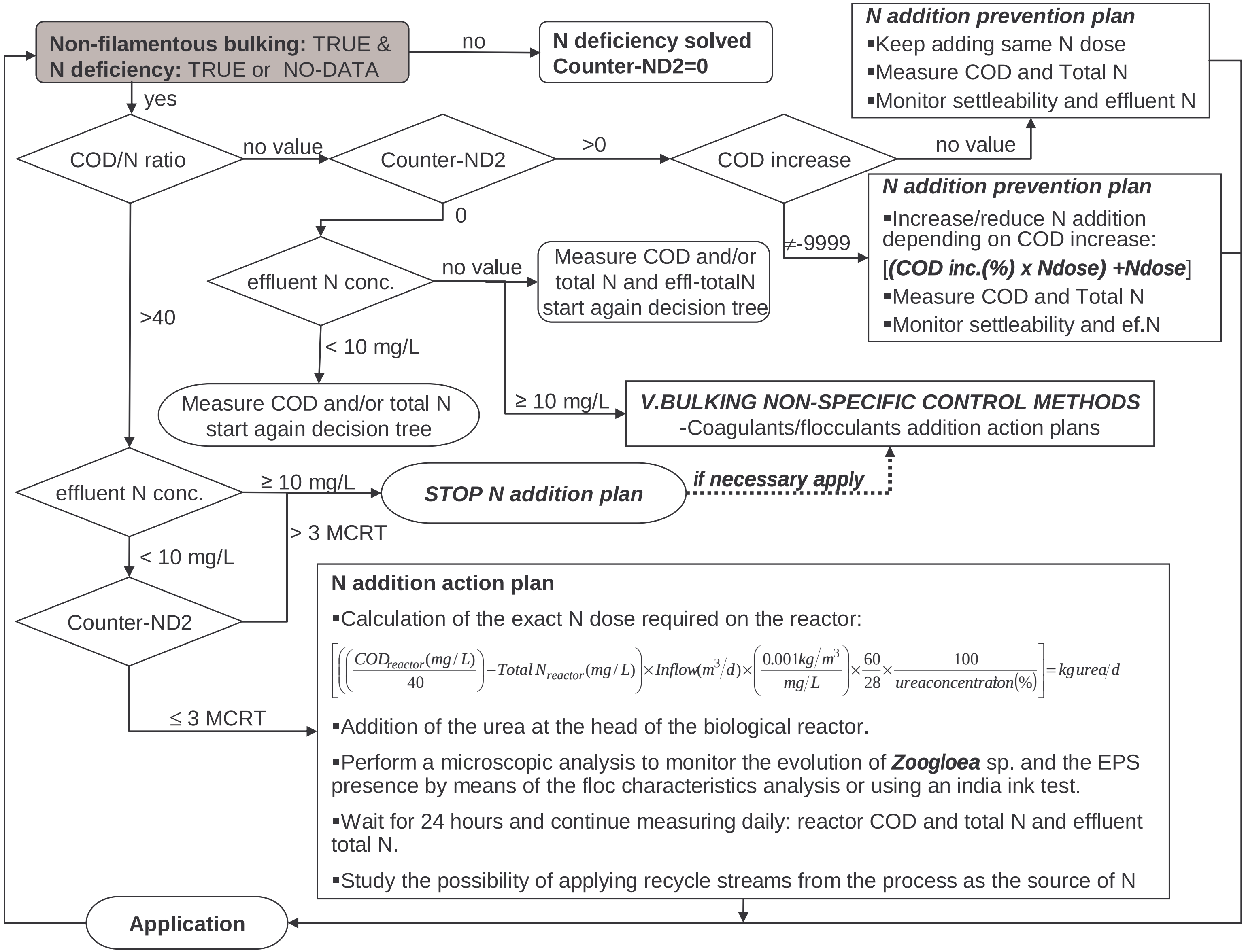
V.BULKING NON-SPECIFIC CONTROL METHODS
-Coagulants/flocculants addition action plans

P addition action plan

- Calculation of the exact P dose required on the reactor:

$$\left[\left(\left(\frac{COD_{reactor}(mg/L)}{200}\right)-Total\ P_{reactor}(mg/L)\right)\times Inflow\,(m^3/d)\times\left(\frac{0.001kg/m^3}{mg/L}\right)\times\frac{98}{31}\times\frac{100}{phosphoric\ acid\ concentration(\%)}\right]=kg\ phosphoric\ acid/d$$

- Addition of the phosphoric acid at the head of the biological reactor.
- Perform a microscopic analysis to monitor the evolution of ***Zoogloea*** sp. and the EPS presence by means of the floc characteristics analysis or using an india ink test.
- Wait for 24 hours and continue measuring daily: reactor COD and totalP and effluent total P.
- Study the possibility of applying recycle streams from the process as the source of P

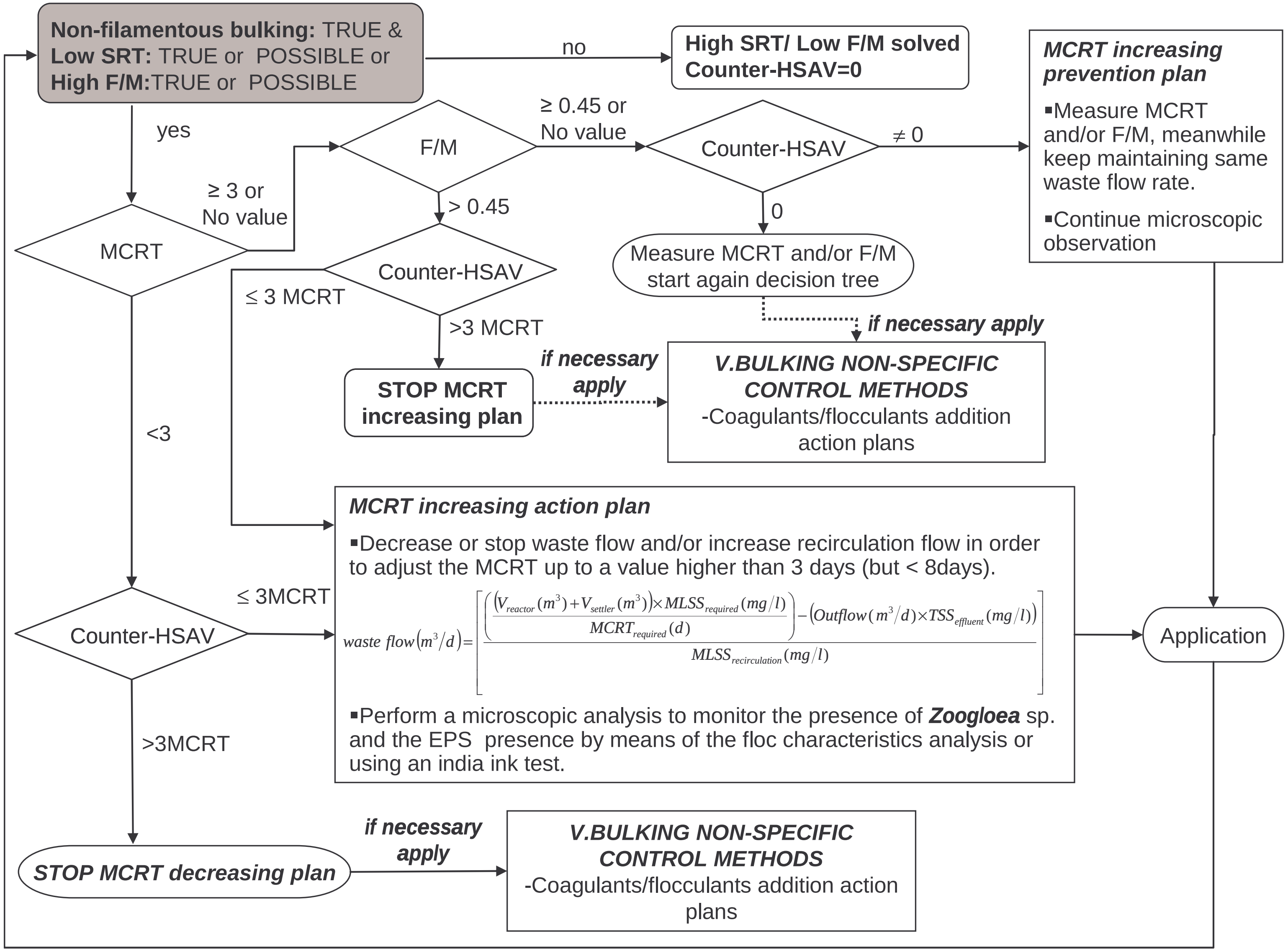


N addition action plan

- Calculation of the exact N dose required on the reactor:

$$\left[\left(\left(\frac{COD_{reactor}(mg/L)}{40}\right)-Total\,N_{reactor}(mg/L)\right)\times Inflow(m^3/d)\times\left(\frac{0.001kg/m^3}{mg/L}\right)\times\frac{60}{28}\times\frac{100}{ureaconcentration(\%)}\right]=kg\,urea/d$$

- Addition of the urea at the head of the biological reactor.
- Perform a microscopic analysis to monitor the evolution of ***Zoogloea*** sp. and the EPS presence by means of the floc characteristics analysis or using an india ink test.
- Wait for 24 hours and continue measuring daily: reactor COD and total N and effluent total N.
- Study the possibility of applying recycle streams from the process as the source of N



MCRT increasing prevention plan

- Measure MCRT and/or F/M, meanwhile keep maintaining same waste flow rate.
- Continue microscopic observation

V.BULKING NON-SPECIFIC CONTROL METHODS

-Coagulants/flocculants addition action plans

MCRT increasing action plan

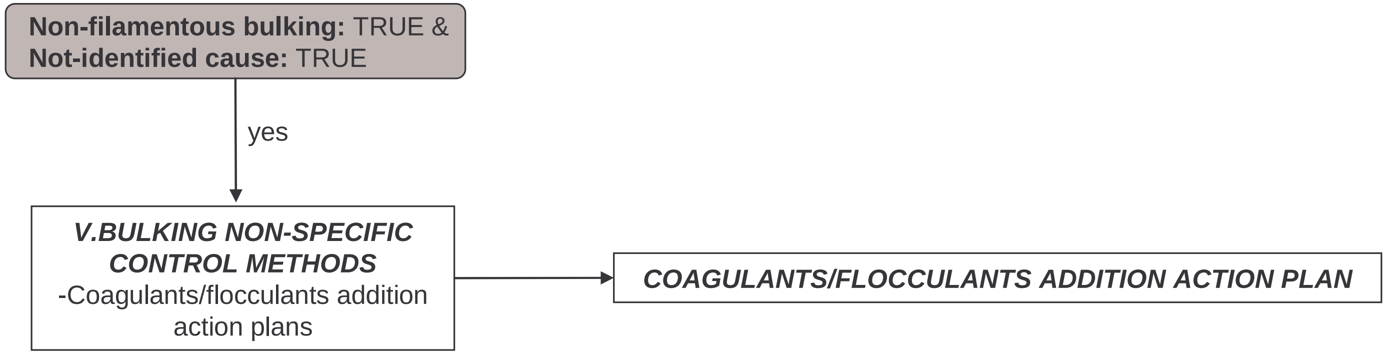
- Decrease or stop waste flow and/or increase recirculation flow in order to adjust the MCRT up to a value higher than 3 days (but < 8days).

$$waste\ flow\left(m^3/d\right)=\left[\frac{\left(\frac{\left(V_{reactor}(m^3)+V_{settler}(m^3)\right)\times MLSS_{required}(mg/l)}{MCRT_{required}(d)}\right)-\left(Outflow(m^3/d)\times TSS_{effluent}(mg/l)\right)}{MLSS_{recirculation}(mg/l)}\right]$$

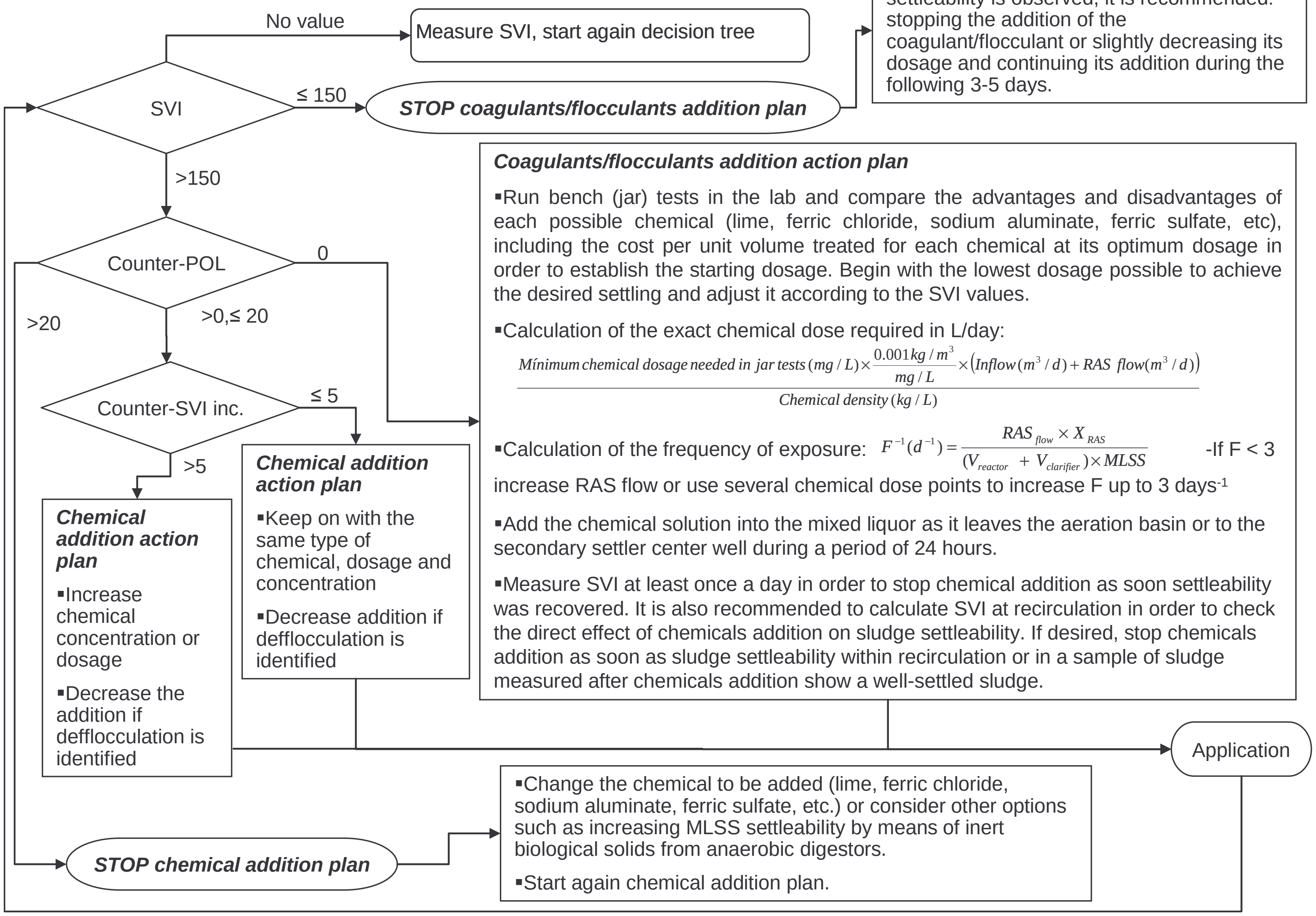
- Perform a microscopic analysis to monitor the presence of ***Zoogloea*** sp. and the EPS presence by means of the floc characteristics analysis or using an india ink test.

V.BULKING NON-SPECIFIC CONTROL METHODS

-Coagulants/flocculants addition action plans


Non-filamentous bulking: TRUE &
Not-identified cause: TRUE
yes
V.BULKING NON-SPECIFIC CONTROL METHODS
-Coagulants/flocculants addition action plans
COAGULANTS/FLOCCULANTS ADDITION ACTION PLAN

COAGULANTS/FLOCCULANTS ADDITION ACTION PLAN



- If a significant improvement on sludge settleability is observed, it is recommended: stopping the addition of the coagulant/flocculant or slightly decreasing its dosage and continuing its addition during the following 3-5 days.

Coagulants/flocculants addition action plan

- Run bench (jar) tests in the lab and compare the advantages and disadvantages of each possible chemical (lime, ferric chloride, sodium aluminate, ferric sulfate, etc), including the cost per unit volume treated for each chemical at its optimum dosage in order to establish the starting dosage. Begin with the lowest dosage possible to achieve the desired settling and adjust it according to the SVI values.
- Calculation of the exact chemical dose required in L/day:

$$\frac{\text{Mínimum chemical dosage needed in jar tests}\,(mg/L)\times\frac{0.001\,kg/m^3}{mg/L}\times\left(Inflow\,(m^3/d)+RAS\ \ flow(m^3/d)\right)}{\text{Chemical density}\,(kg/L)}$$

- Calculation of the frequency of exposure: $F^{-1}(d^{-1})=\frac{RAS_{flow}\times X_{RAS}}{(V_{reactor}+V_{clarifier})\times MLSS}$ -If F < 3 increase RAS flow or use several chemical dose points to increase F up to 3 days^{-1}
- Add the chemical solution into the mixed liquor as it leaves the aeration basin or to the secondary settler center well during a period of 24 hours.
- Measure SVI at least once a day in order to stop chemical addition as soon settleability was recovered. It is also recommended to calculate SVI at recirculation in order to check the direct effect of chemicals addition on sludge settleability. If desired, stop chemicals addition as soon as sludge settleability within recirculation or in a sample of sludge measured after chemicals addition show a well-settled sludge.

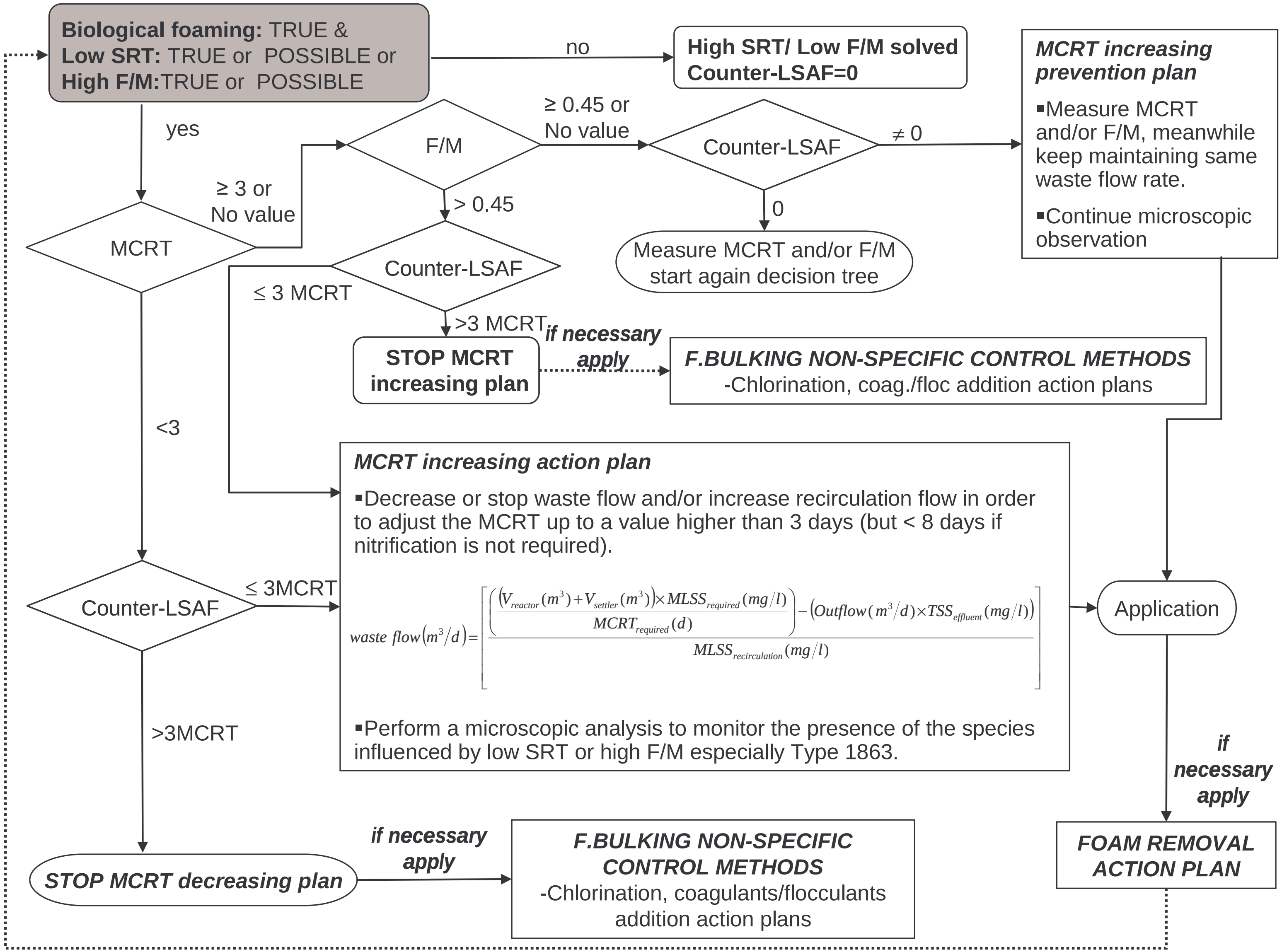
Chemical addition action plan

- Keep on with the same type of chemical, dosage and concentration
- Decrease addition if deflocculation is identified

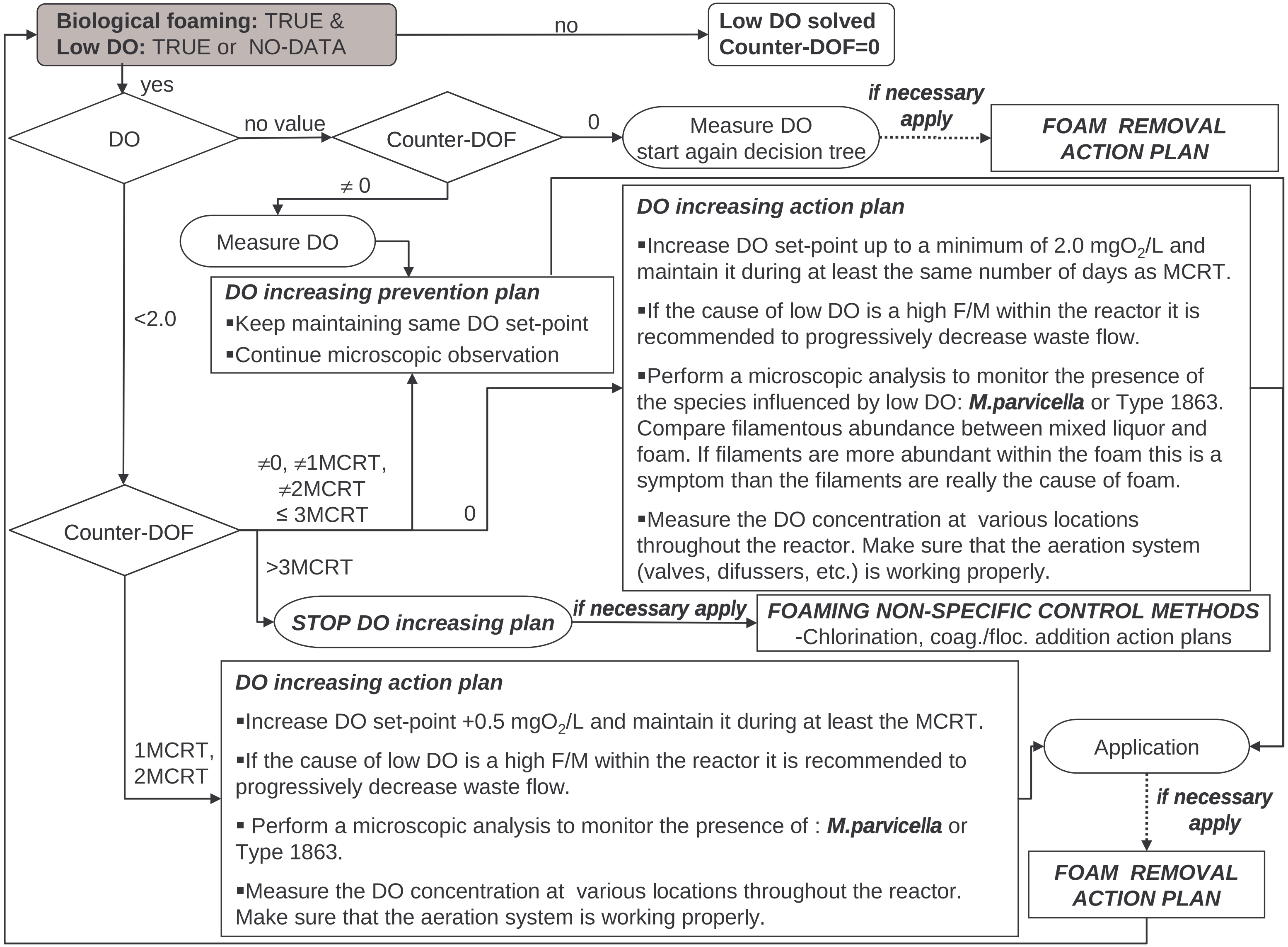
Chemical addition action plan

- Increase chemical concentration or dosage
- Decrease the addition if deflocculation is identified

- Change the chemical to be added (lime, ferric chloride, sodium aluminate, ferric sulfate, etc.) or consider other options such as increasing MLSS settleability by means of inert biological solids from anaerobic digestors.
- Start again chemical addition plan.



$$waste\ flow(m^3/d) = \left[\frac{\left(\frac{\left(V_{reactor}(m^3) + V_{settler}(m^3)\right) \times MLSS_{required}(mg/l)}{MCRT_{required}(d)} \right) - \left(Outflow(m^3/d) \times TSS_{effluent}(mg/l) \right)}{MLSS_{recirculation}(mg/l)} \right]$$



DO increasing prevention plan

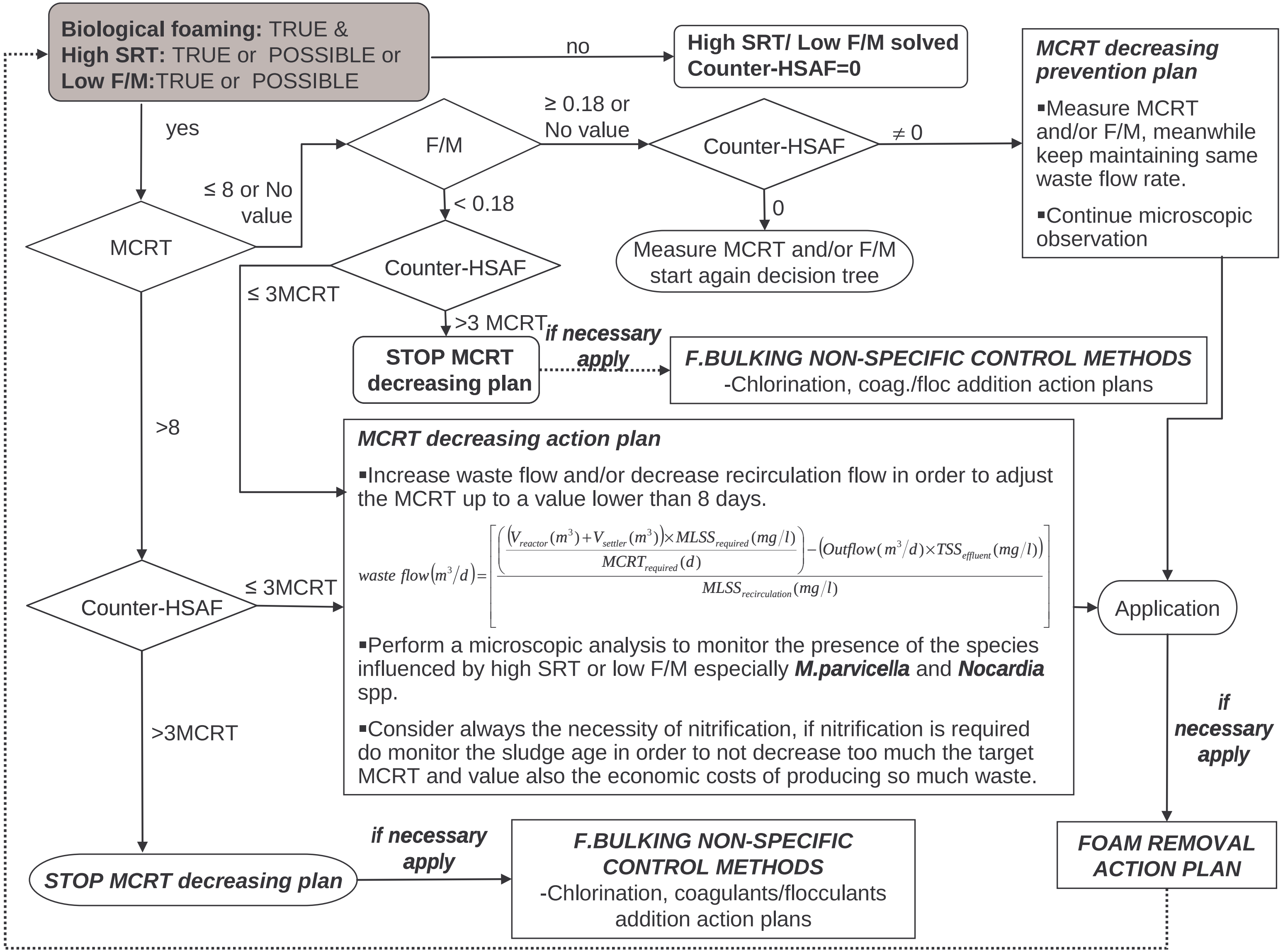
- Keep maintaining same DO set-point
- Continue microscopic observation

DO increasing action plan

- Increase DO set-point up to a minimum of 2.0 mgO_2/L and maintain it during at least the same number of days as MCRT.
- If the cause of low DO is a high F/M within the reactor it is recommended to progressively decrease waste flow.
- Perform a microscopic analysis to monitor the presence of the species influenced by low DO: ***M.parvicella*** or Type 1863. Compare filamentous abundance between mixed liquor and foam. If filaments are more abundant within the foam this is a symptom than the filaments are really the cause of foam.
- Measure the DO concentration at various locations throughout the reactor. Make sure that the aeration system (valves, difussers, etc.) is working properly.

DO increasing action plan

- Increase DO set-point +0.5 mgO_2/L and maintain it during at least the MCRT.
- If the cause of low DO is a high F/M within the reactor it is recommended to progressively decrease waste flow.
- Perform a microscopic analysis to monitor the presence of : ***M.parvicella*** or Type 1863.
- Measure the DO concentration at various locations throughout the reactor. Make sure that the aeration system is working properly.



MCRT decreasing prevention plan

- Measure MCRT and/or F/M, meanwhile keep maintaining same waste flow rate.
- Continue microscopic observation

F.BULKING NON-SPECIFIC CONTROL METHODS
-Chlorination, coag./floc addition action plans

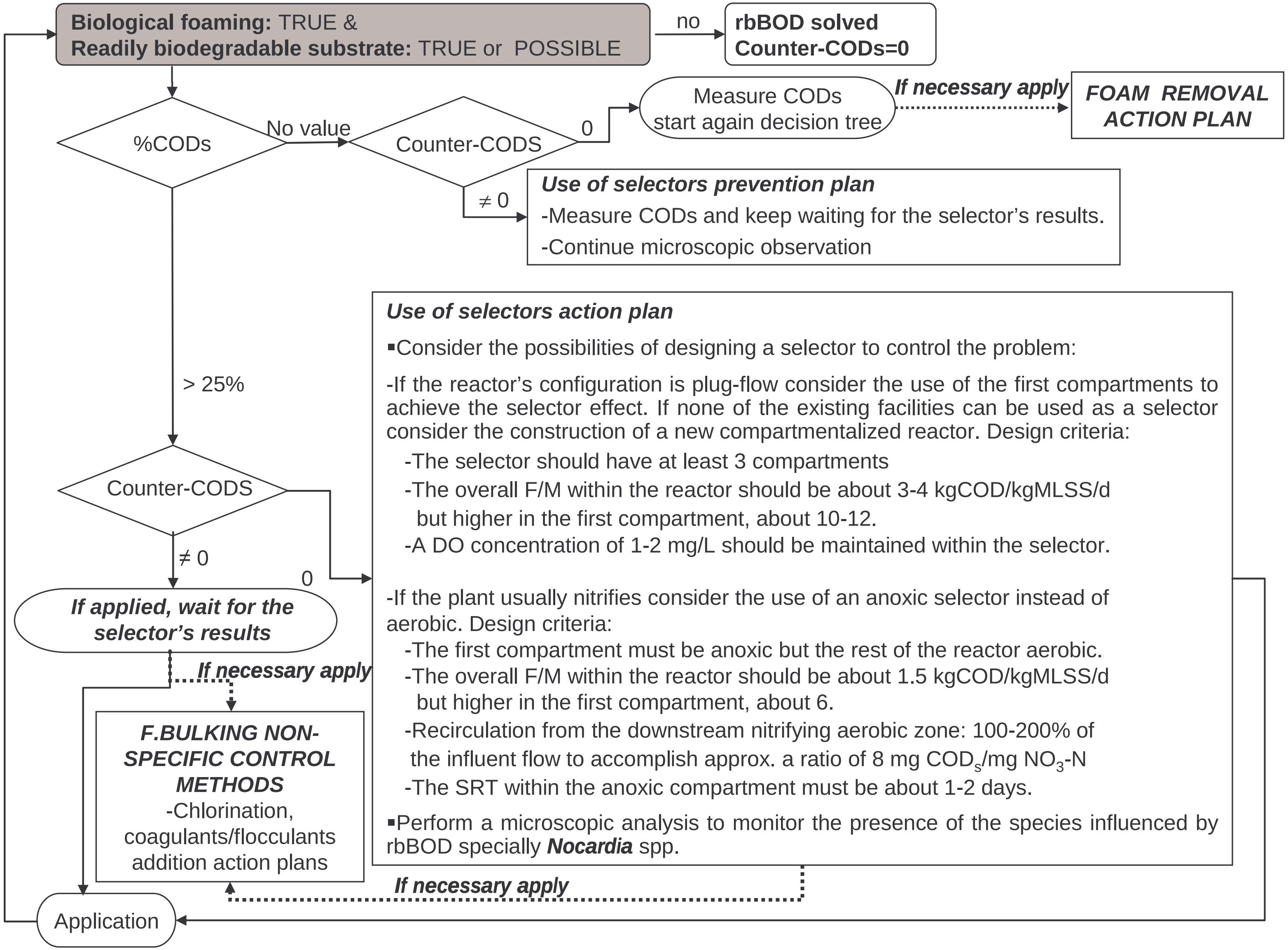
MCRT decreasing action plan

- Increase waste flow and/or decrease recirculation flow in order to adjust the MCRT up to a value lower than 8 days.

$$waste\ flow\left(m^3/d\right)=\left[\frac{\left(\frac{\left(V_{reactor}(m^3)+V_{settler}(m^3)\right)\times MLSS_{required}(mg/l)}{MCRT_{required}(d)}\right)-\left(Outflow(m^3/d)\times TSS_{effluent}(mg/l)\right)}{MLSS_{recirculation}(mg/l)}\right]$$

- Perform a microscopic analysis to monitor the presence of the species influenced by high SRT or low F/M especially ***M.parvicella*** and ***Nocardia*** spp.
- Consider always the necessity of nitrification, if nitrification is required do monitor the sludge age in order to not decrease too much the target MCRT and value also the economic costs of producing so much waste.

F.BULKING NON-SPECIFIC CONTROL METHODS
-Chlorination, coagulants/flocculants addition action plans



Use of selectors prevention plan

-Measure CODs and keep waiting for the selector's results.

-Continue microscopic observation

Use of selectors action plan

▪Consider the possibilities of designing a selector to control the problem:

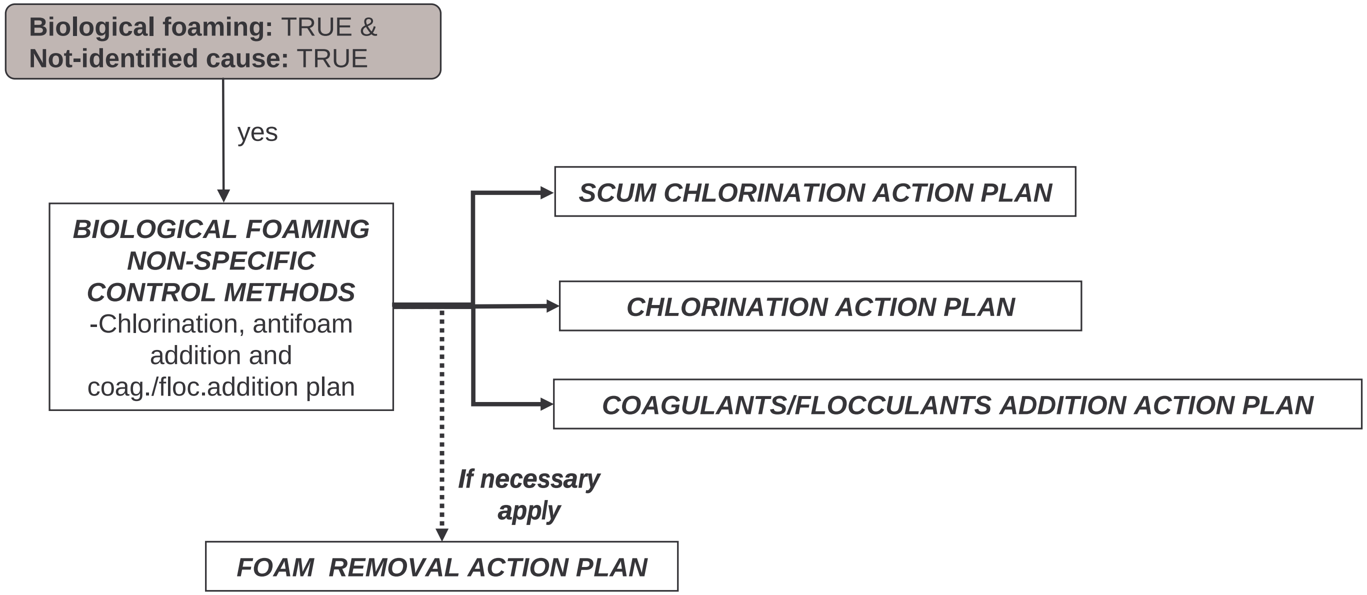
-If the reactor's configuration is plug-flow consider the use of the first compartments to achieve the selector effect. If none of the existing facilities can be used as a selector consider the construction of a new compartmentalized reactor. Design criteria:

- -The selector should have at least 3 compartments
- -The overall F/M within the reactor should be about 3-4 kgCOD/kgMLSS/d but higher in the first compartment, about 10-12.
- -A DO concentration of 1-2 mg/L should be maintained within the selector.

-If the plant usually nitrifies consider the use of an anoxic selector instead of aerobic. Design criteria:

- -The first compartment must be anoxic but the rest of the reactor aerobic.
- -The overall F/M within the reactor should be about 1.5 kgCOD/kgMLSS/d but higher in the first compartment, about 6.
- -Recirculation from the downstream nitrifying aerobic zone: 100-200% of the influent flow to accomplish approx. a ratio of 8 mg COD_s/mg NO_3-N
- -The SRT within the anoxic compartment must be about 1-2 days.

▪Perform a microscopic analysis to monitor the presence of the species influenced by rbBOD specially ***Nocardia*** spp.

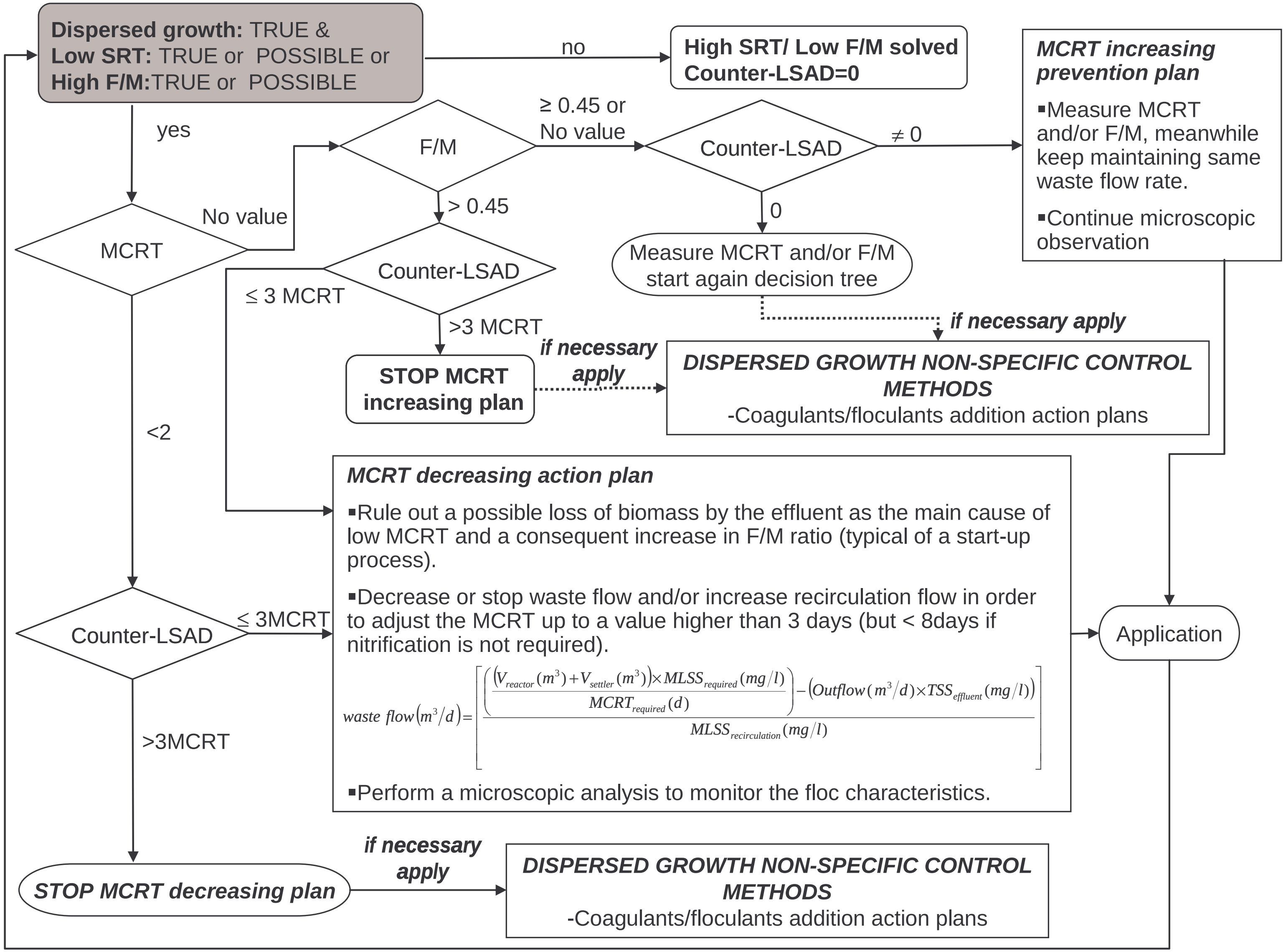

Biological foaming: TRUE &
Not-identified cause: TRUE
yes
BIOLOGICAL FOAMING NON-SPECIFIC CONTROL METHODS
-Chlorination, antifoam addition and coag./floc.addition plan
SCUM CHLORINATION ACTION PLAN
CHLORINATION ACTION PLAN
COAGULANTS/FLOCCULANTS ADDITION ACTION PLAN
If necessary apply
FOAM REMOVAL ACTION PLAN

FOAM REMOVAL ACTION PLAN

- Avoid and/or separate the possible fats, oils or greases present at the influent.
- Spray foam with effluent water through a bib sprayer or lawn sprinkler or treat the foam with a defoaming agent such as the polyglycol-based defoaming agent.
- Vacuum or rake foam from the surface of the aeration tank and/or secondary settlers. The removed foam must be treated and disposed properly in order to avoid the return of foam-producing filaments into the activated sludge process.

SCUM CHLORINATION ACTION PLAN

- Spray foam with a 10-15% sodium hypochlorite solution. The solution should remain in contact with the foam for 2-3 hours.
- Collapse the foam with a spray of effluent water.



MCRT increasing prevention plan

- Measure MCRT and/or F/M, meanwhile keep maintaining same waste flow rate.
- Continue microscopic observation

DISPERSED GROWTH NON-SPECIFIC CONTROL METHODS

-Coagulants/floculants addition action plans

MCRT decreasing action plan

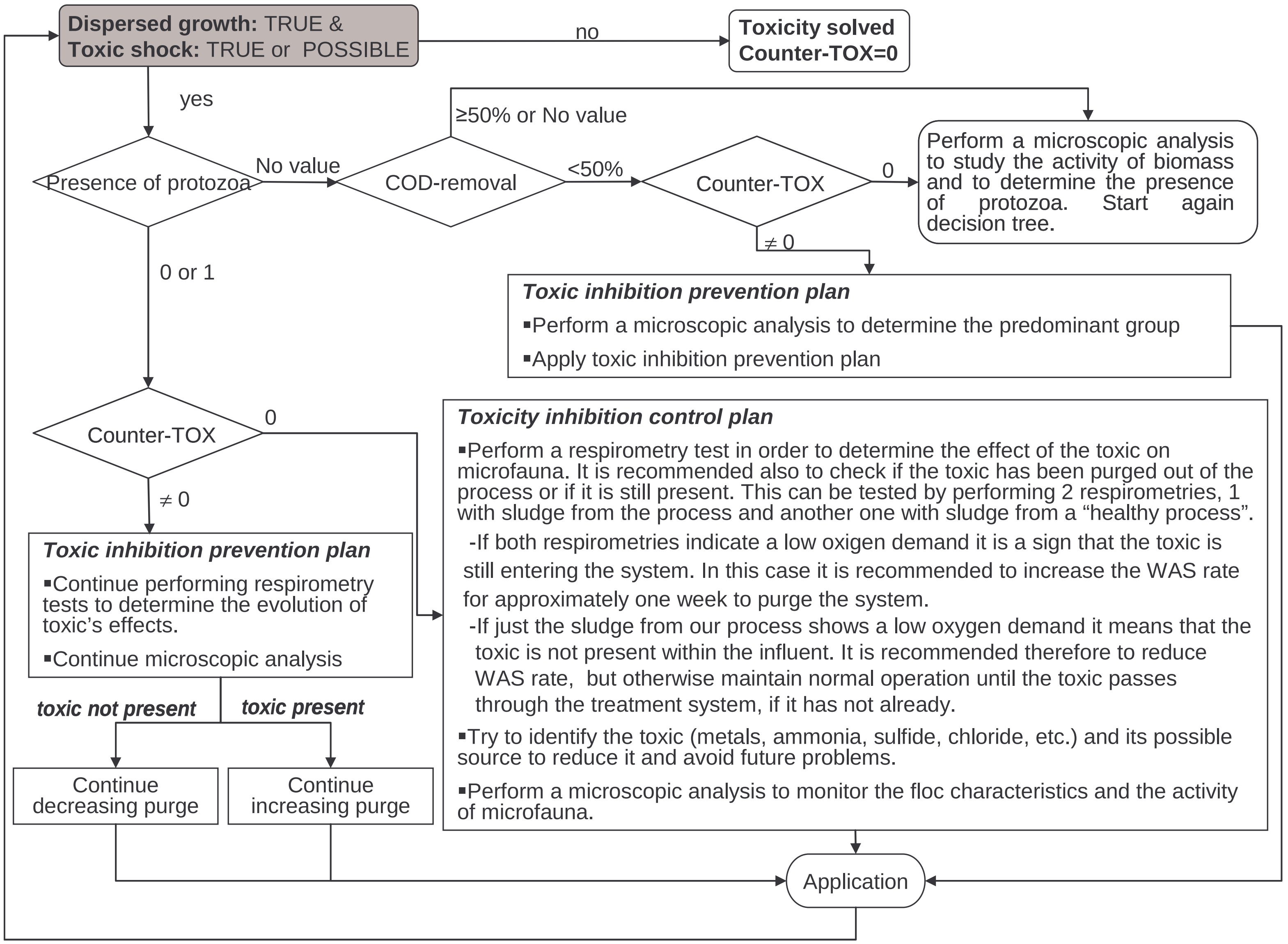
- Rule out a possible loss of biomass by the effluent as the main cause of low MCRT and a consequent increase in F/M ratio (typical of a start-up process).
- Decrease or stop waste flow and/or increase recirculation flow in order to adjust the MCRT up to a value higher than 3 days (but < 8days if nitrification is not required).

$$waste\ flow\left(m^3/d\right)=\left[\frac{\left(\frac{\left(V_{reactor}\left(m^3\right)+V_{settler}\left(m^3\right)\right)\times MLSS_{required}\left(mg/l\right)}{MCRT_{required}\left(d\right)}\right)-\left(Outflow\left(m^3/d\right)\times TSS_{effluent}\left(mg/l\right)\right)}{MLSS_{recirculation}\left(mg/l\right)}\right]$$

- Perform a microscopic analysis to monitor the floc characteristics.

DISPERSED GROWTH NON-SPECIFIC CONTROL METHODS

-Coagulants/floculants addition action plans



Perform a microscopic analysis to study the activity of biomass and to determine the presence of protozoa. Start again decision tree.

Toxic inhibition prevention plan

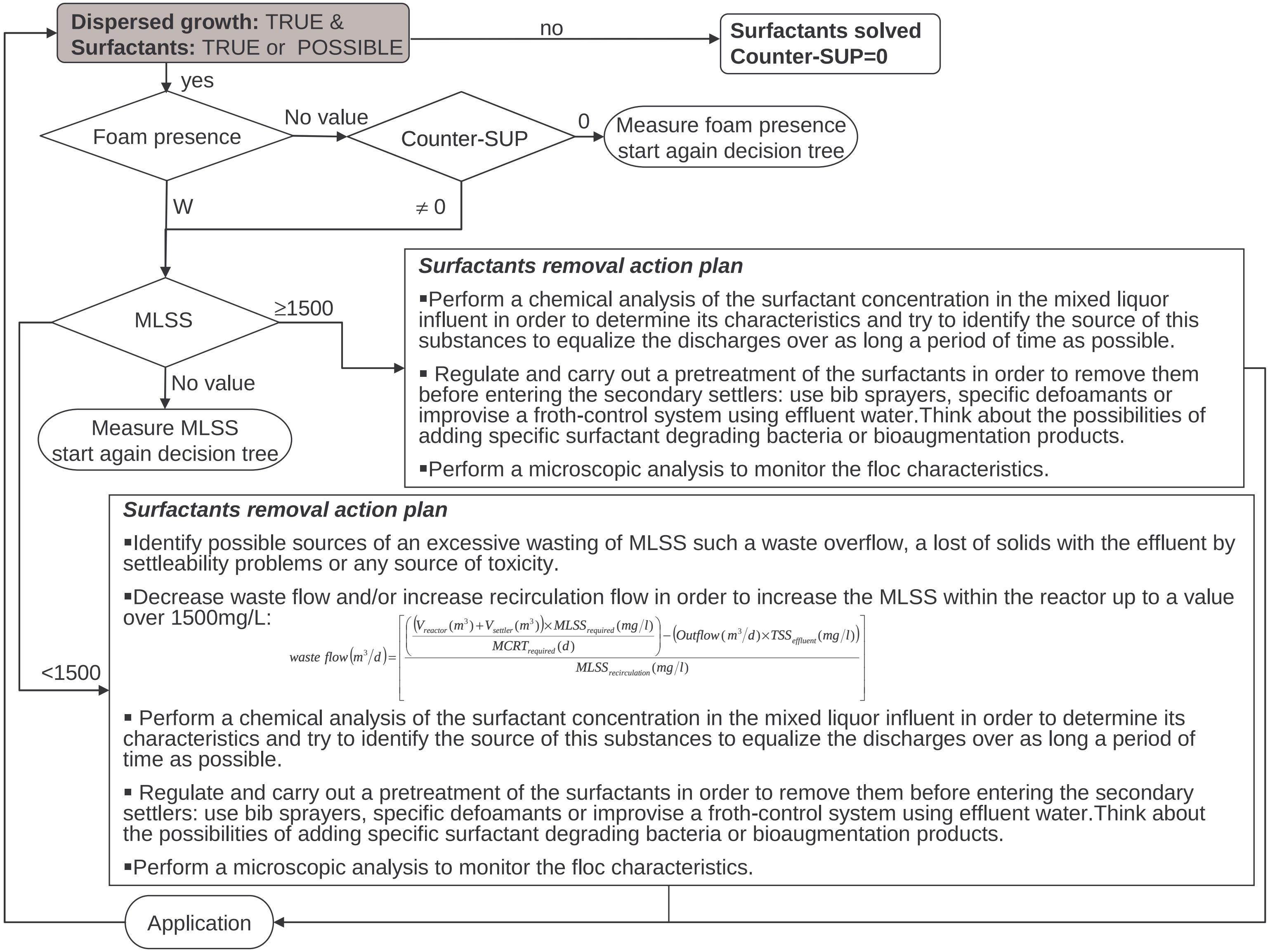
- ▪Perform a microscopic analysis to determine the predominant group
- ▪Apply toxic inhibition prevention plan

Toxicity inhibition control plan

- ▪Perform a respirometry test in order to determine the effect of the toxic on microfauna. It is recommended also to check if the toxic has been purged out of the process or if it is still present. This can be tested by performing 2 respirometries, 1 with sludge from the process and another one with sludge from a "healthy process".
  - -If both respirometries indicate a low oxigen demand it is a sign that the toxic is still entering the system. In this case it is recommended to increase the WAS rate for approximately one week to purge the system.
  - -If just the sludge from our process shows a low oxygen demand it means that the toxic is not present within the influent. It is recommended therefore to reduce WAS rate, but otherwise maintain normal operation until the toxic passes through the treatment system, if it has not already.
- ▪Try to identify the toxic (metals, ammonia, sulfide, chloride, etc.) and its possible source to reduce it and avoid future problems.
- ▪Perform a microscopic analysis to monitor the floc characteristics and the activity of microfauna.

Toxic inhibition prevention plan

- ▪Continue performing respirometry tests to determine the evolution of toxic's effects.
- ▪Continue microscopic analysis



Surfactants removal action plan

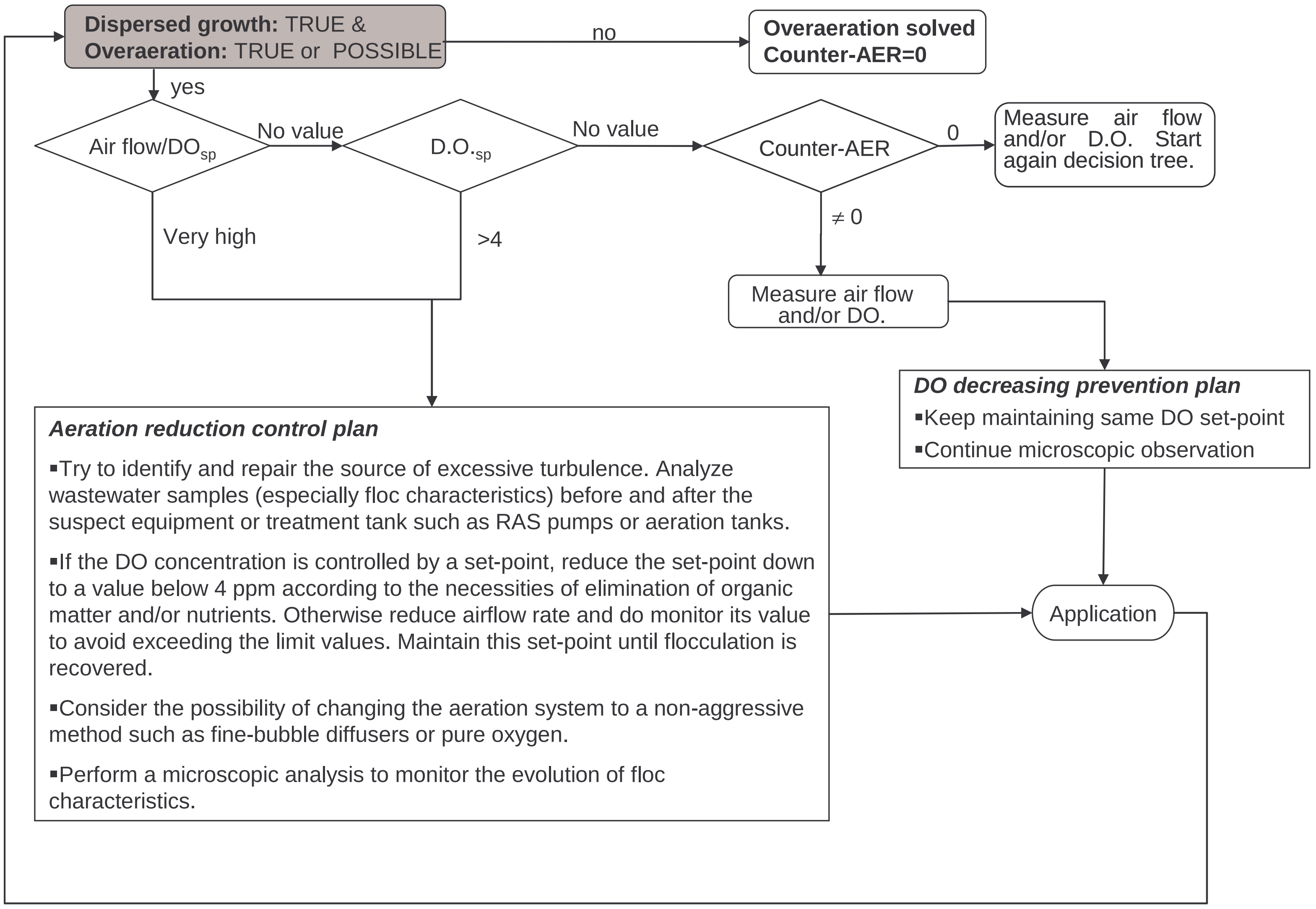
- Perform a chemical analysis of the surfactant concentration in the mixed liquor influent in order to determine its characteristics and try to identify the source of this substances to equalize the discharges over as long a period of time as possible.
- Regulate and carry out a pretreatment of the surfactants in order to remove them before entering the secondary settlers: use bib sprayers, specific defoamants or improvise a froth-control system using effluent water.Think about the possibilities of adding specific surfactant degrading bacteria or bioaugmentation products.
- Perform a microscopic analysis to monitor the floc characteristics.

Surfactants removal action plan

- Identify possible sources of an excessive wasting of MLSS such a waste overflow, a lost of solids with the effluent by settleability problems or any source of toxicity.
- Decrease waste flow and/or increase recirculation flow in order to increase the MLSS within the reactor up to a value over 1500mg/L:

$$waste\ flow\left(m^3/d\right)=\left[\frac{\left(\frac{\left(V_{reactor}(m^3)+V_{settler}(m^3)\right)\times MLSS_{required}(mg/l)}{MCRT_{required}(d)}\right)-\left(Outflow(m^3/d)\times TSS_{effluent}(mg/l)\right)}{MLSS_{recirculation}(mg/l)}\right]$$

- Perform a chemical analysis of the surfactant concentration in the mixed liquor influent in order to determine its characteristics and try to identify the source of this substances to equalize the discharges over as long a period of time as possible.
- Regulate and carry out a pretreatment of the surfactants in order to remove them before entering the secondary settlers: use bib sprayers, specific defoamants or improvise a froth-control system using effluent water.Think about the possibilities of adding specific surfactant degrading bacteria or bioaugmentation products.
- Perform a microscopic analysis to monitor the floc characteristics.

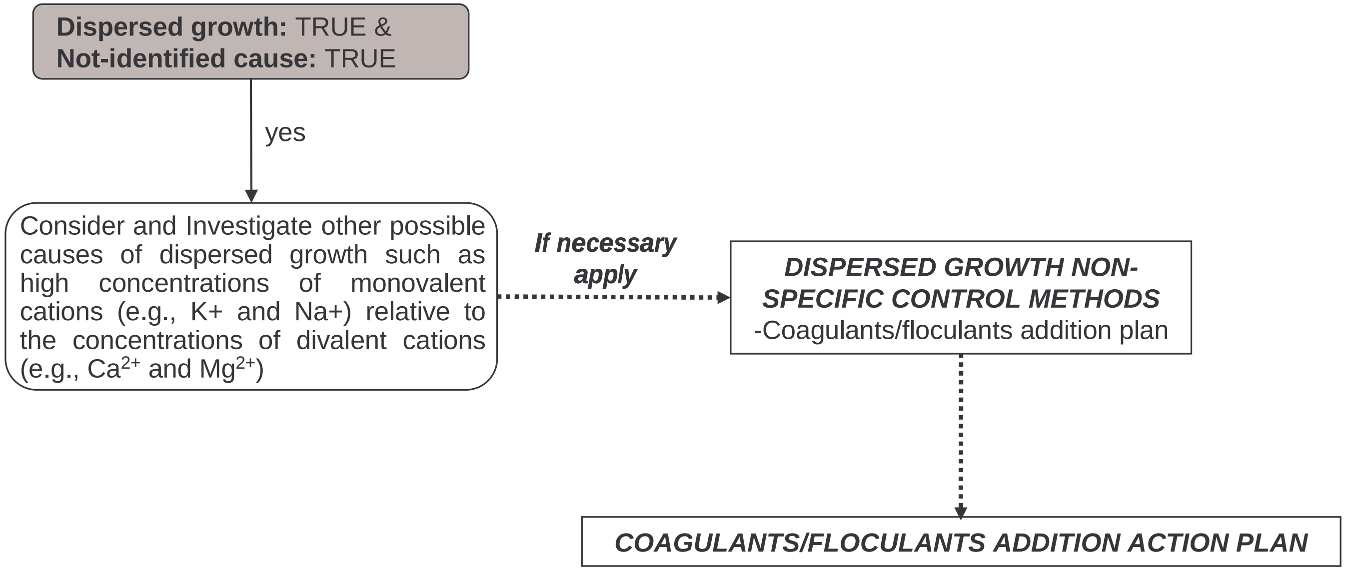


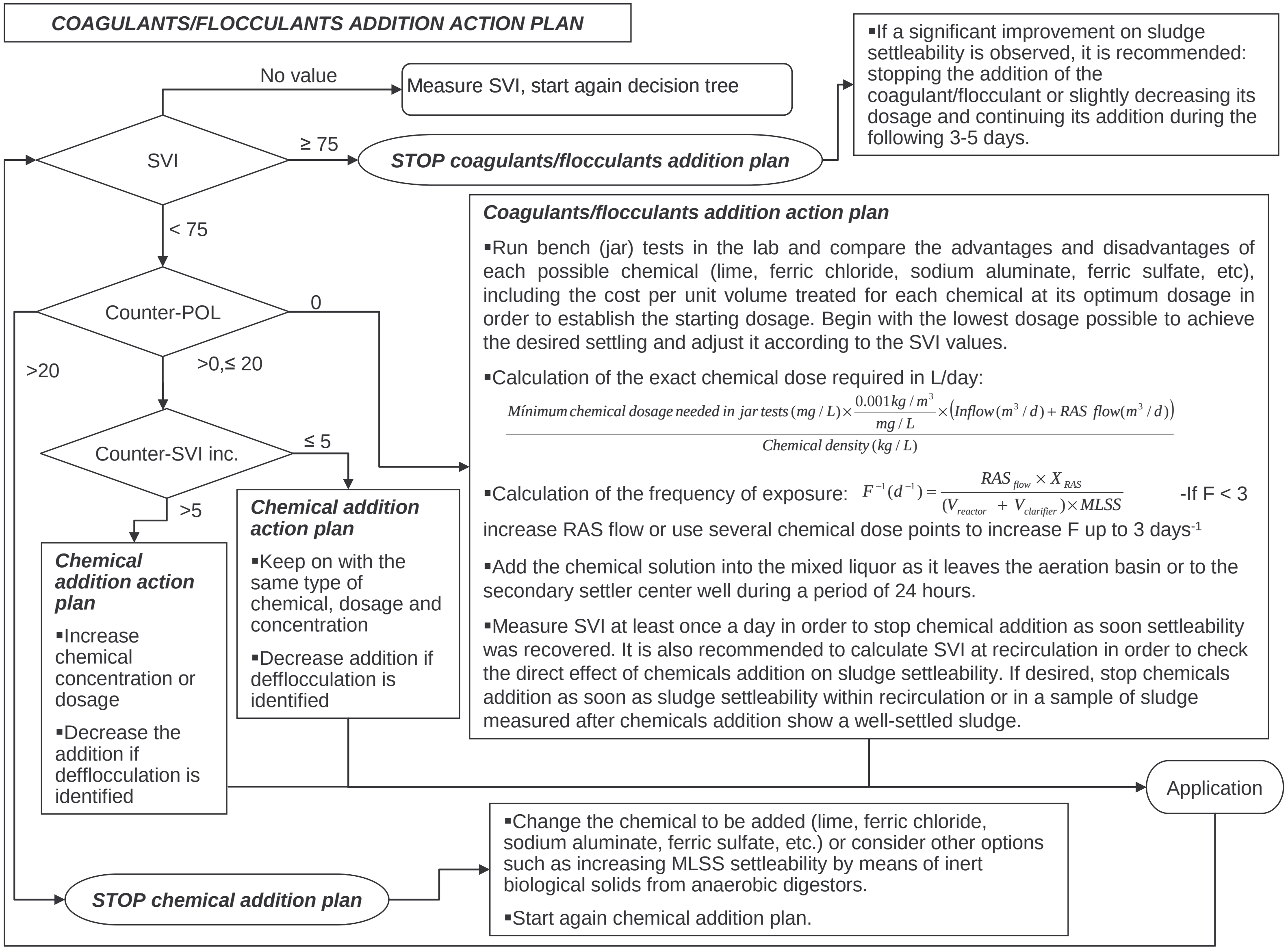
DO decreasing prevention plan

- Keep maintaining same DO set-point
- Continue microscopic observation

Aeration reduction control plan

- Try to identify and repair the source of excessive turbulence. Analyze wastewater samples (especially floc characteristics) before and after the suspect equipment or treatment tank such as RAS pumps or aeration tanks.
- If the DO concentration is controlled by a set-point, reduce the set-point down to a value below 4 ppm according to the necessities of elimination of organic matter and/or nutrients. Otherwise reduce airflow rate and do monitor its value to avoid exceeding the limit values. Maintain this set-point until flocculation is recovered.
- Consider the possibility of changing the aeration system to a non-aggressive method such as fine-bubble diffusers or pure oxygen.
- Perform a microscopic analysis to monitor the evolution of floc characteristics.


Dispersed growth: TRUE &
Not-identified cause: TRUE
yes
Consider and Investigate other possible causes of dispersed growth such as high concentrations of monovalent cations (e.g., K+ and Na+) relative to the concentrations of divalent cations (e.g., Ca^{2+} and Mg^{2+})
If necessary apply
DISPERSED GROWTH NON-SPECIFIC CONTROL METHODS
-Coagulants/floculants addition plan
COAGULANTS/FLOCULANTS ADDITION ACTION PLAN



COAGULANTS/FLOCCULANTS ADDITION ACTION PLAN

Measure SVI, start again decision tree

STOP coagulants/flocculants addition plan

- If a significant improvement on sludge settleability is observed, it is recommended: stopping the addition of the coagulant/flocculant or slightly decreasing its dosage and continuing its addition during the following 3-5 days.

Coagulants/flocculants addition action plan

- Run bench (jar) tests in the lab and compare the advantages and disadvantages of each possible chemical (lime, ferric chloride, sodium aluminate, ferric sulfate, etc), including the cost per unit volume treated for each chemical at its optimum dosage in order to establish the starting dosage. Begin with the lowest dosage possible to achieve the desired settling and adjust it according to the SVI values.
- Calculation of the exact chemical dose required in L/day:

$$\frac{\text{Mínimum chemical dosage needed in jar tests}\,(mg/L)\times\frac{0.001kg/m^3}{mg/L}\times\left(Inflow\,(m^3/d)+RAS\ flow(m^3/d)\right)}{\text{Chemical density}\,(kg/L)}$$

- Calculation of the frequency of exposure: $F^{-1}(d^{-1})=\frac{RAS_{flow}\times X_{RAS}}{(V_{reactor}+V_{clarifier})\times MLSS}$ -If F < 3 increase RAS flow or use several chemical dose points to increase F up to 3 days^{-1}
- Add the chemical solution into the mixed liquor as it leaves the aeration basin or to the secondary settler center well during a period of 24 hours.
- Measure SVI at least once a day in order to stop chemical addition as soon settleability was recovered. It is also recommended to calculate SVI at recirculation in order to check the direct effect of chemicals addition on sludge settleability. If desired, stop chemicals addition as soon as sludge settleability within recirculation or in a sample of sludge measured after chemicals addition show a well-settled sludge.

Chemical addition action plan

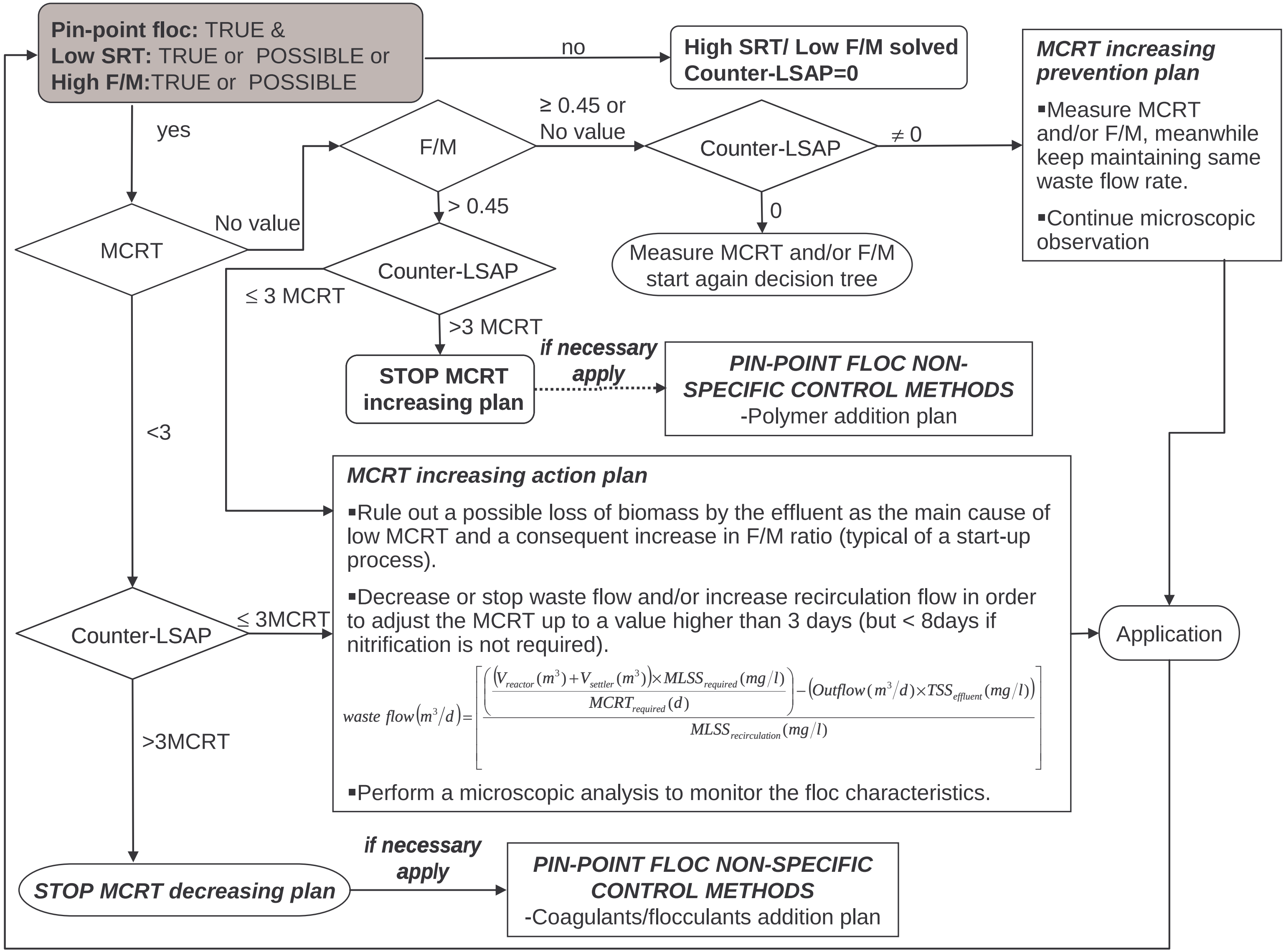
- Keep on with the same type of chemical, dosage and concentration
- Decrease addition if deflocculation is identified

Chemical addition action plan

- Increase chemical concentration or dosage
- Decrease the addition if deflocculation is identified

STOP chemical addition plan

- Change the chemical to be added (lime, ferric chloride, sodium aluminate, ferric sulfate, etc.) or consider other options such as increasing MLSS settleability by means of inert biological solids from anaerobic digestors.
- Start again chemical addition plan.



MCRT increasing prevention plan

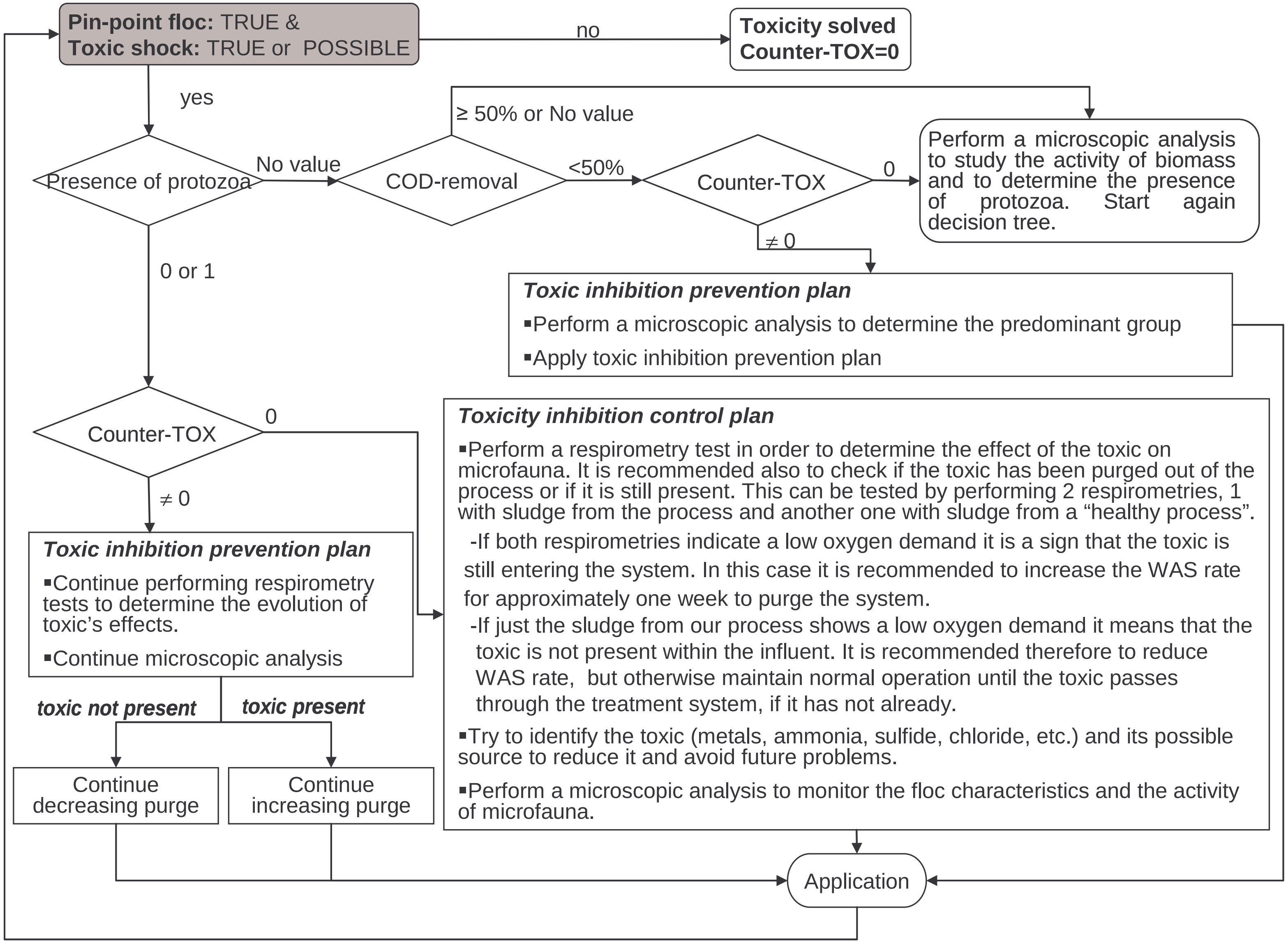
- Measure MCRT and/or F/M, meanwhile keep maintaining same waste flow rate.
- Continue microscopic observation

MCRT increasing action plan

- Rule out a possible loss of biomass by the effluent as the main cause of low MCRT and a consequent increase in F/M ratio (typical of a start-up process).
- Decrease or stop waste flow and/or increase recirculation flow in order to adjust the MCRT up to a value higher than 3 days (but < 8days if nitrification is not required).

$$waste\ flow\left(m^3/d\right)=\left[\frac{\left(\frac{\left(V_{reactor}(m^3)+V_{settler}(m^3)\right)\times MLSS_{required}(mg/l)}{MCRT_{required}(d)}\right)-\left(Outflow(m^3/d)\times TSS_{effluent}(mg/l)\right)}{MLSS_{recirculation}(mg/l)}\right]$$

- Perform a microscopic analysis to monitor the floc characteristics.



Perform a microscopic analysis to study the activity of biomass and to determine the presence of protozoa. Start again decision tree.

Toxic inhibition prevention plan

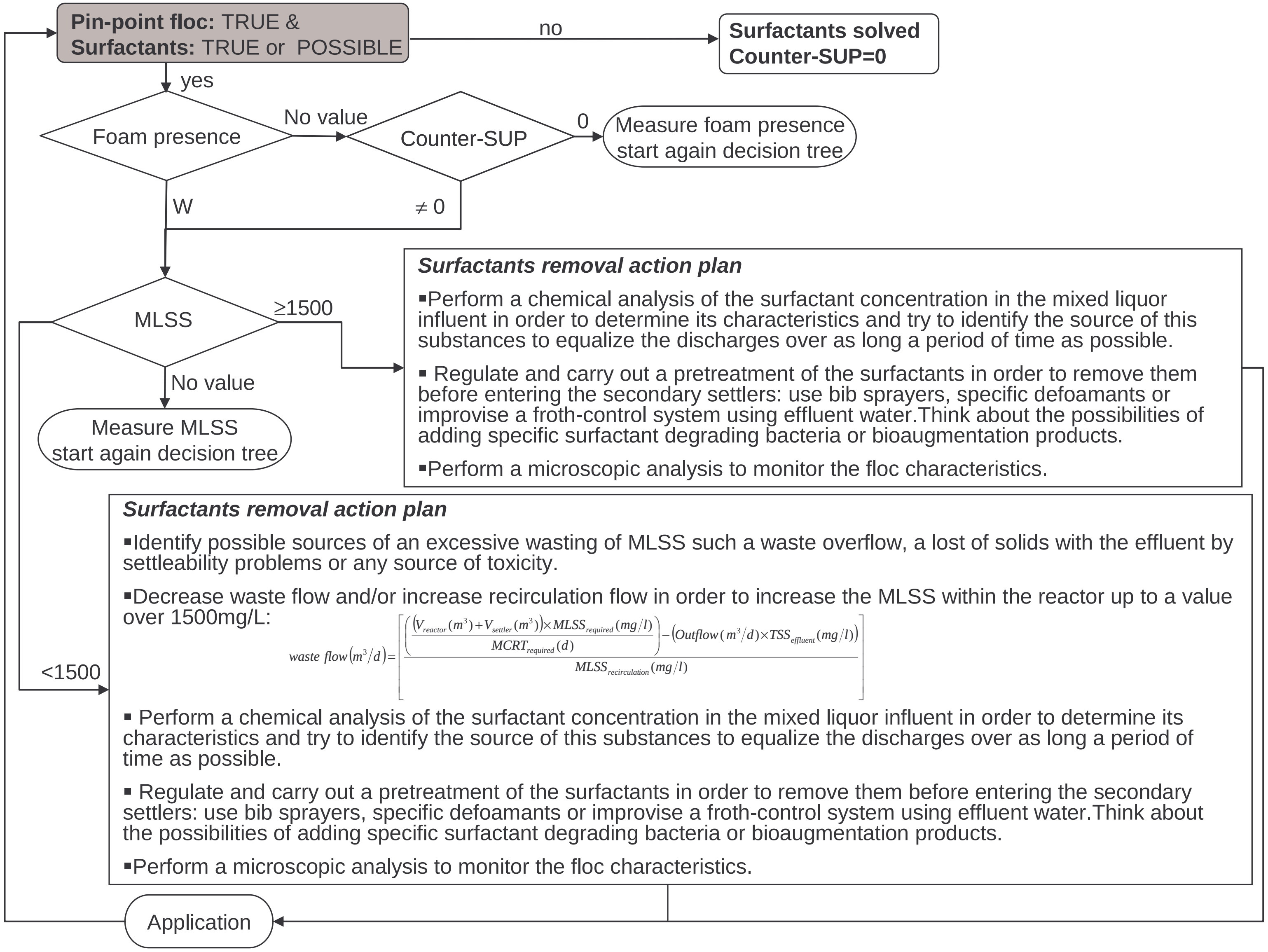
- Perform a microscopic analysis to determine the predominant group
- Apply toxic inhibition prevention plan

Toxicity inhibition control plan

- Perform a respirometry test in order to determine the effect of the toxic on microfauna. It is recommended also to check if the toxic has been purged out of the process or if it is still present. This can be tested by performing 2 respirometries, 1 with sludge from the process and another one with sludge from a "healthy process".
  - If both respirometries indicate a low oxygen demand it is a sign that the toxic is still entering the system. In this case it is recommended to increase the WAS rate for approximately one week to purge the system.
  - If just the sludge from our process shows a low oxygen demand it means that the toxic is not present within the influent. It is recommended therefore to reduce WAS rate, but otherwise maintain normal operation until the toxic passes through the treatment system, if it has not already.
- Try to identify the toxic (metals, ammonia, sulfide, chloride, etc.) and its possible source to reduce it and avoid future problems.
- Perform a microscopic analysis to monitor the floc characteristics and the activity of microfauna.

Toxic inhibition prevention plan

- Continue performing respirometry tests to determine the evolution of toxic's effects.
- Continue microscopic analysis



Surfactants removal action plan

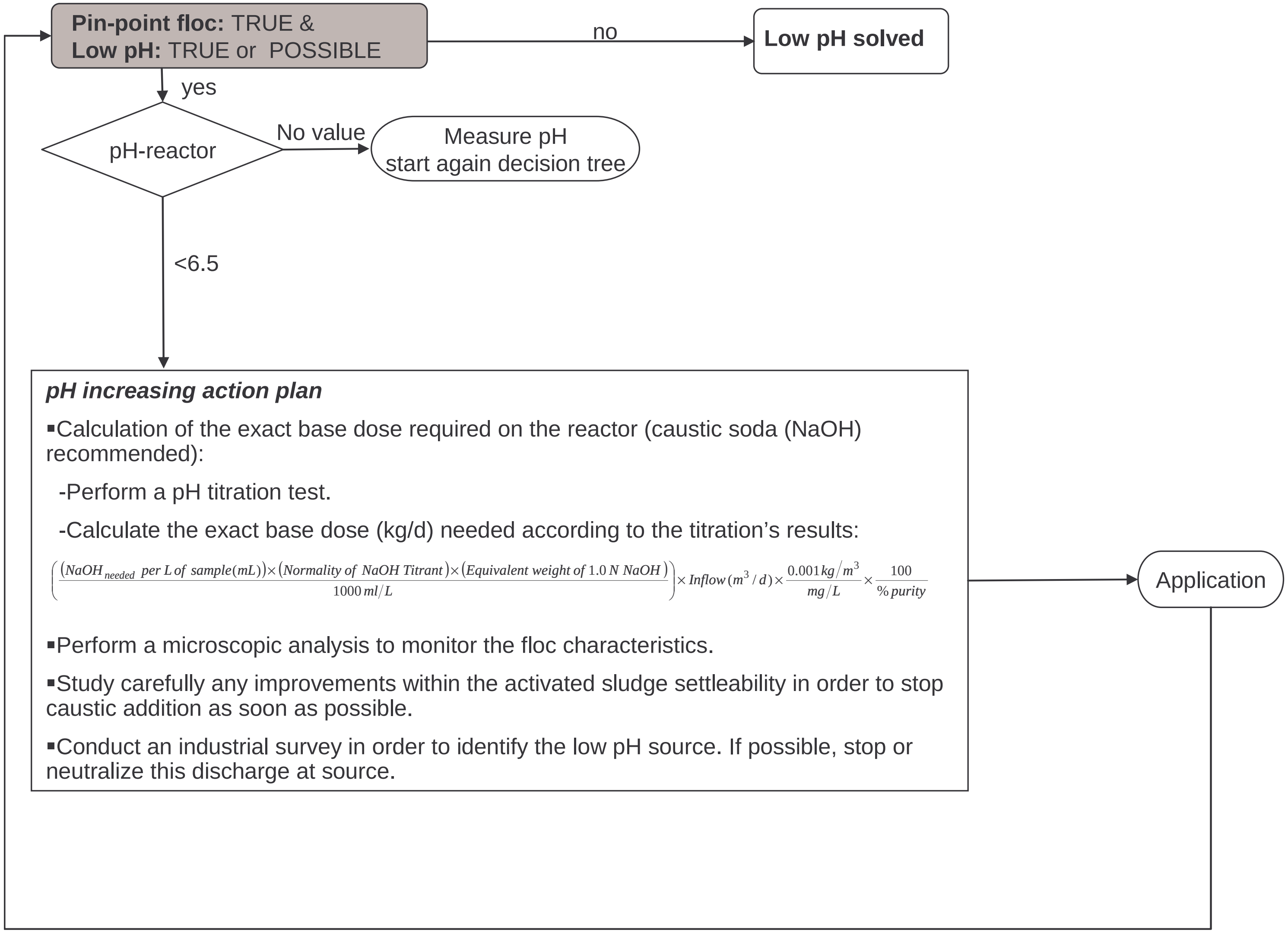
- Perform a chemical analysis of the surfactant concentration in the mixed liquor influent in order to determine its characteristics and try to identify the source of this substances to equalize the discharges over as long a period of time as possible.
- Regulate and carry out a pretreatment of the surfactants in order to remove them before entering the secondary settlers: use bib sprayers, specific defoamants or improvise a froth-control system using effluent water.Think about the possibilities of adding specific surfactant degrading bacteria or bioaugmentation products.
- Perform a microscopic analysis to monitor the floc characteristics.

Surfactants removal action plan

- Identify possible sources of an excessive wasting of MLSS such a waste overflow, a lost of solids with the effluent by settleability problems or any source of toxicity.
- Decrease waste flow and/or increase recirculation flow in order to increase the MLSS within the reactor up to a value over 1500mg/L:

$$waste\ flow\left(m^3/d\right)=\left[\frac{\left(\frac{\left(V_{reactor}\left(m^3\right)+V_{settler}\left(m^3\right)\right)\times MLSS_{required}\left(mg/l\right)}{MCRT_{required}\left(d\right)}\right)-\left(Outflow\left(m^3/d\right)\times TSS_{effluent}\left(mg/l\right)\right)}{MLSS_{recirculation}\left(mg/l\right)}\right]$$

- Perform a chemical analysis of the surfactant concentration in the mixed liquor influent in order to determine its characteristics and try to identify the source of this substances to equalize the discharges over as long a period of time as possible.
- Regulate and carry out a pretreatment of the surfactants in order to remove them before entering the secondary settlers: use bib sprayers, specific defoamants or improvise a froth-control system using effluent water.Think about the possibilities of adding specific surfactant degrading bacteria or bioaugmentation products.
- Perform a microscopic analysis to monitor the floc characteristics.

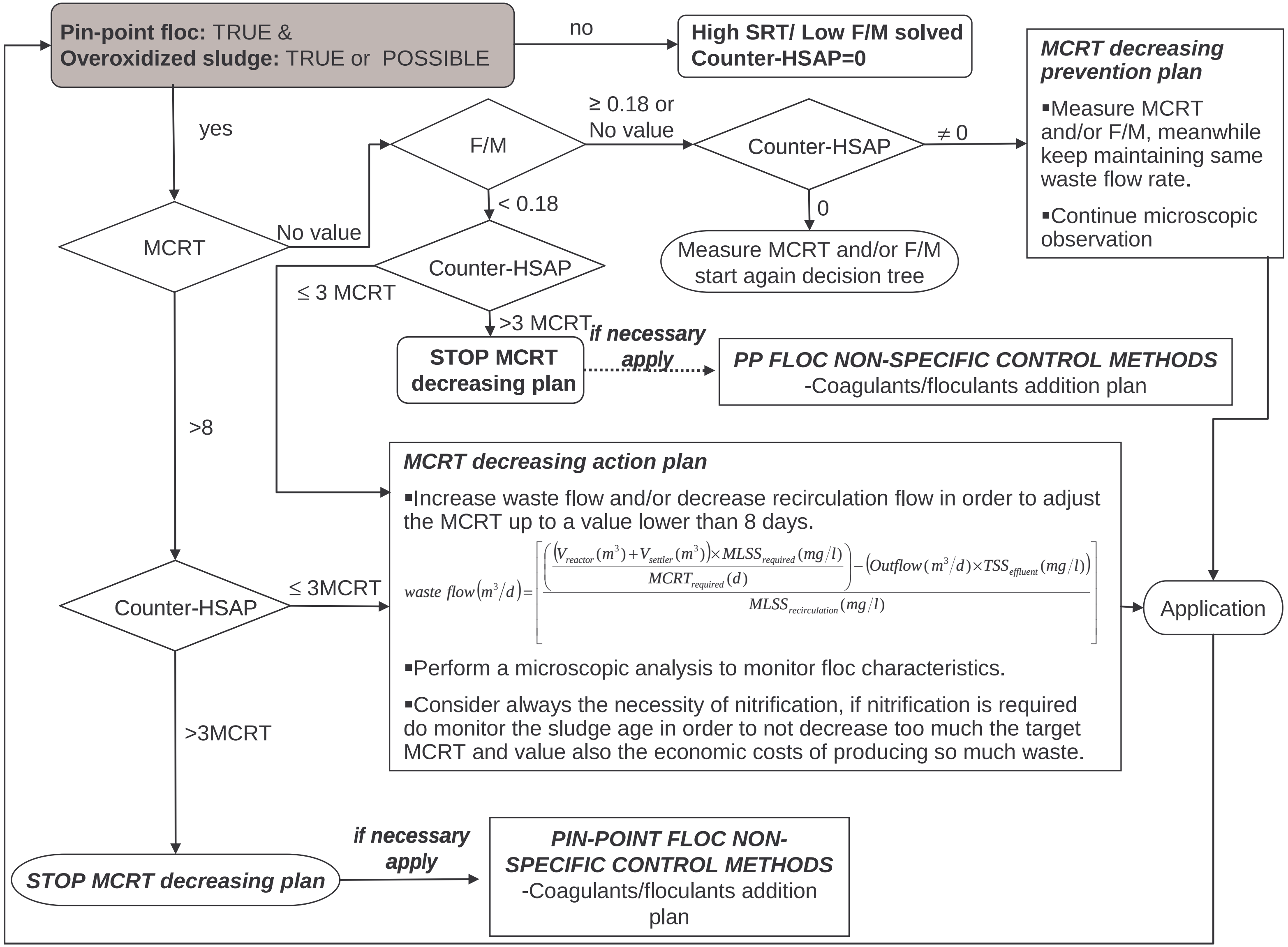


pH increasing action plan

- ▪Calculation of the exact base dose required on the reactor (caustic soda (NaOH) recommended):
  - -Perform a pH titration test.
  - -Calculate the exact base dose (kg/d) needed according to the titration's results:

$$\left(\frac{(\text{NaOH}_{needed}\ \text{per L of sample}(mL))\times(\text{Normality of NaOH Titrant})\times(\text{Equivalent weight of 1.0 N NaOH})}{1000\,ml/L}\right)\times \text{Inflow}\,(m^3/d)\times\frac{0.001\,kg/m^3}{mg/L}\times\frac{100}{\%\,purity}$$

- ▪Perform a microscopic analysis to monitor the floc characteristics.
- ▪Study carefully any improvements within the activated sludge settleability in order to stop caustic addition as soon as possible.
- ▪Conduct an industrial survey in order to identify the low pH source. If possible, stop or neutralize this discharge at source.



MCRT decreasing prevention plan

- Measure MCRT and/or F/M, meanwhile keep maintaining same waste flow rate.
- Continue microscopic observation

Measure MCRT and/or F/M start again decision tree

PP FLOC NON-SPECIFIC CONTROL METHODS
-Coagulants/floculants addition plan

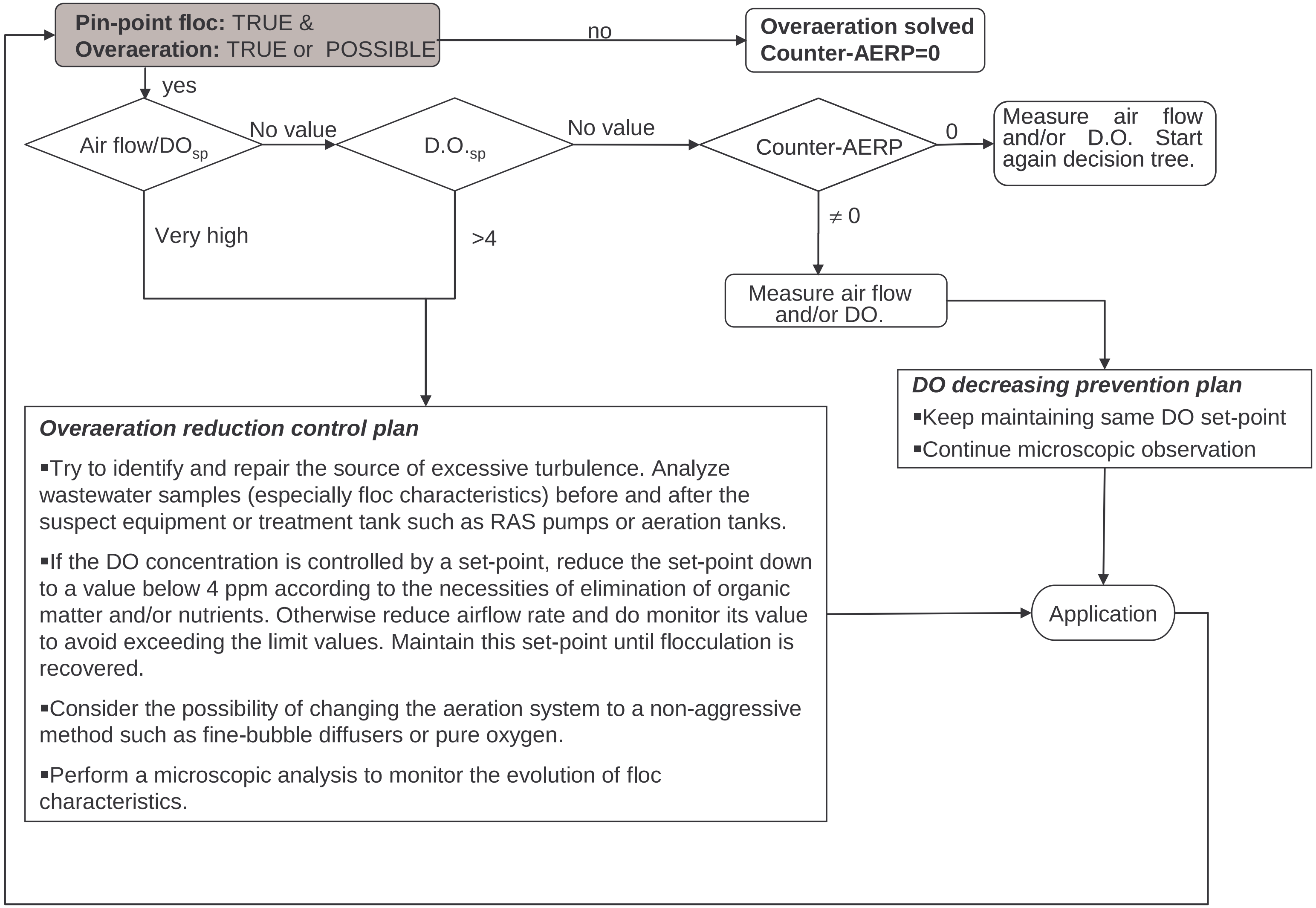
MCRT decreasing action plan

- Increase waste flow and/or decrease recirculation flow in order to adjust the MCRT up to a value lower than 8 days.

$$waste\ flow\left(m^3/d\right)=\left[\frac{\left(\frac{\left(V_{reactor}\left(m^3\right)+V_{settler}\left(m^3\right)\right)\times MLSS_{required}\left(mg/l\right)}{MCRT_{required}\left(d\right)}\right)-\left(Outflow\left(m^3/d\right)\times TSS_{effluent}\left(mg/l\right)\right)}{MLSS_{recirculation}\left(mg/l\right)}\right]$$

- Perform a microscopic analysis to monitor floc characteristics.
- Consider always the necessity of nitrification, if nitrification is required do monitor the sludge age in order to not decrease too much the target MCRT and value also the economic costs of producing so much waste.

PIN-POINT FLOC NON-SPECIFIC CONTROL METHODS
-Coagulants/floculants addition plan

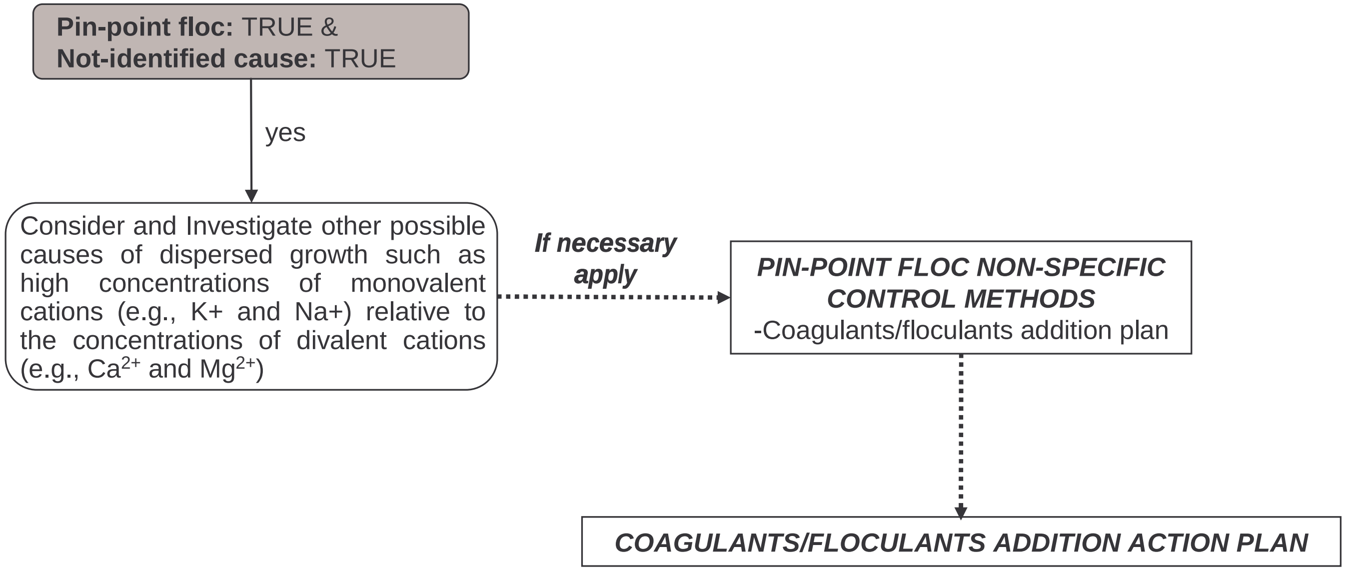


Overaeration reduction control plan

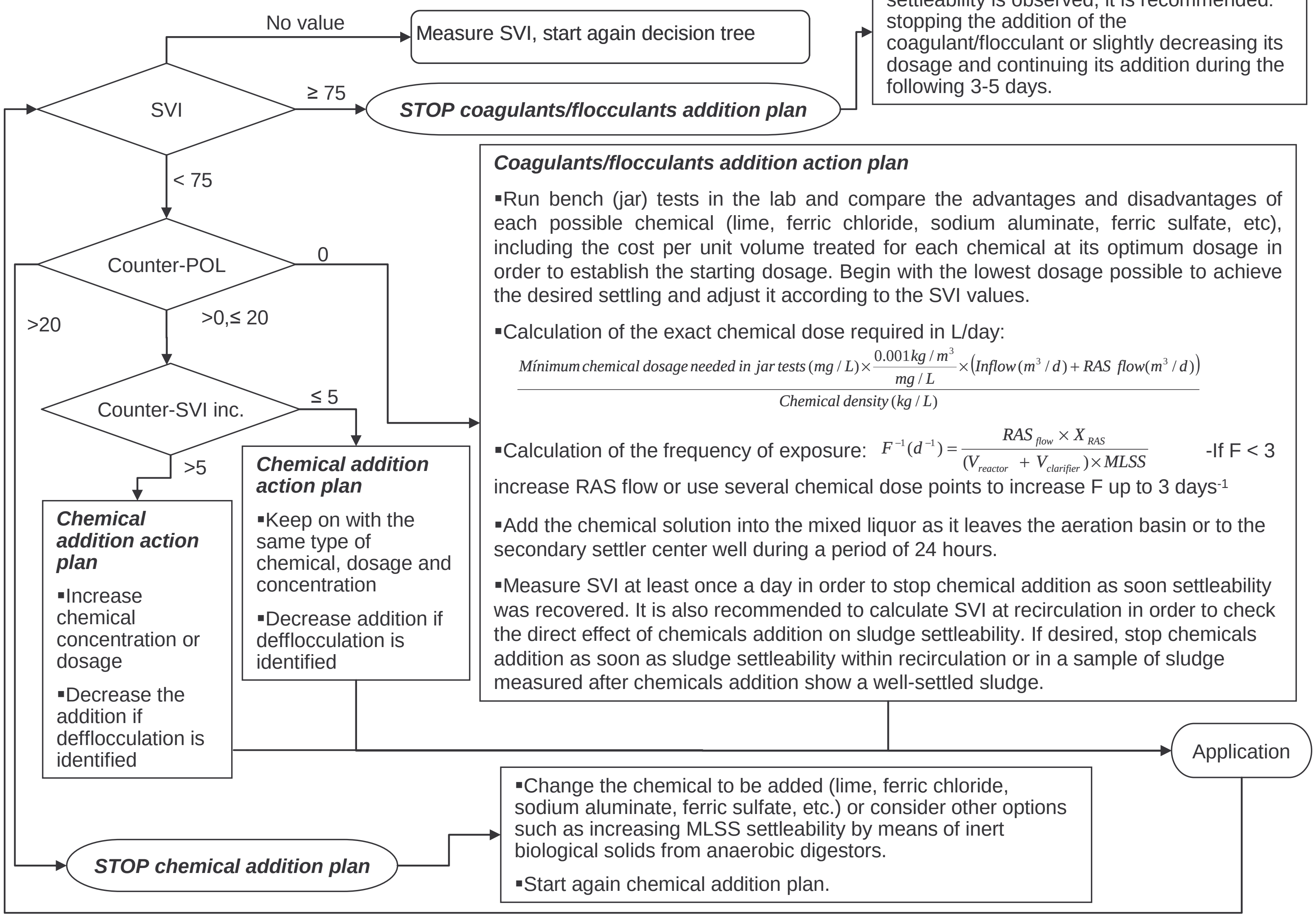
- Try to identify and repair the source of excessive turbulence. Analyze wastewater samples (especially floc characteristics) before and after the suspect equipment or treatment tank such as RAS pumps or aeration tanks.
- If the DO concentration is controlled by a set-point, reduce the set-point down to a value below 4 ppm according to the necessities of elimination of organic matter and/or nutrients. Otherwise reduce airflow rate and do monitor its value to avoid exceeding the limit values. Maintain this set-point until flocculation is recovered.
- Consider the possibility of changing the aeration system to a non-aggressive method such as fine-bubble diffusers or pure oxygen.
- Perform a microscopic analysis to monitor the evolution of floc characteristics.

DO decreasing prevention plan

- Keep maintaining same DO set-point
- Continue microscopic observation


Pin-point floc: TRUE &
Not-identified cause: TRUE
yes
Consider and Investigate other possible causes of dispersed growth such as high concentrations of monovalent cations (e.g., K+ and Na+) relative to the concentrations of divalent cations (e.g., Ca2+ and Mg2+)
If necessary apply
PIN-POINT FLOC NON-SPECIFIC CONTROL METHODS
-Coagulants/floculants addition plan
COAGULANTS/FLOCULANTS ADDITION ACTION PLAN

COAGULANTS/FLOCCULANTS ADDITION ACTION PLAN



- If a significant improvement on sludge settleability is observed, it is recommended: stopping the addition of the coagulant/flocculant or slightly decreasing its dosage and continuing its addition during the following 3-5 days.

Coagulants/flocculants addition action plan

- Run bench (jar) tests in the lab and compare the advantages and disadvantages of each possible chemical (lime, ferric chloride, sodium aluminate, ferric sulfate, etc), including the cost per unit volume treated for each chemical at its optimum dosage in order to establish the starting dosage. Begin with the lowest dosage possible to achieve the desired settling and adjust it according to the SVI values.
- Calculation of the exact chemical dose required in L/day:

$$\frac{\text{Mínimum chemical dosage needed in jar tests}\,(mg/L)\times\frac{0.001\,kg/m^3}{mg/L}\times\left(Inflow\,(m^3/d)+RAS\ \ flow(m^3/d)\right)}{\text{Chemical density}\,(kg/L)}$$

- Calculation of the frequency of exposure: $F^{-1}(d^{-1})=\frac{RAS_{flow}\times X_{RAS}}{(V_{reactor}+V_{clarifier})\times MLSS}$ -If F < 3 increase RAS flow or use several chemical dose points to increase F up to 3 days^{-1}
- Add the chemical solution into the mixed liquor as it leaves the aeration basin or to the secondary settler center well during a period of 24 hours.
- Measure SVI at least once a day in order to stop chemical addition as soon settleability was recovered. It is also recommended to calculate SVI at recirculation in order to check the direct effect of chemicals addition on sludge settleability. If desired, stop chemicals addition as soon as sludge settleability within recirculation or in a sample of sludge measured after chemicals addition show a well-settled sludge.

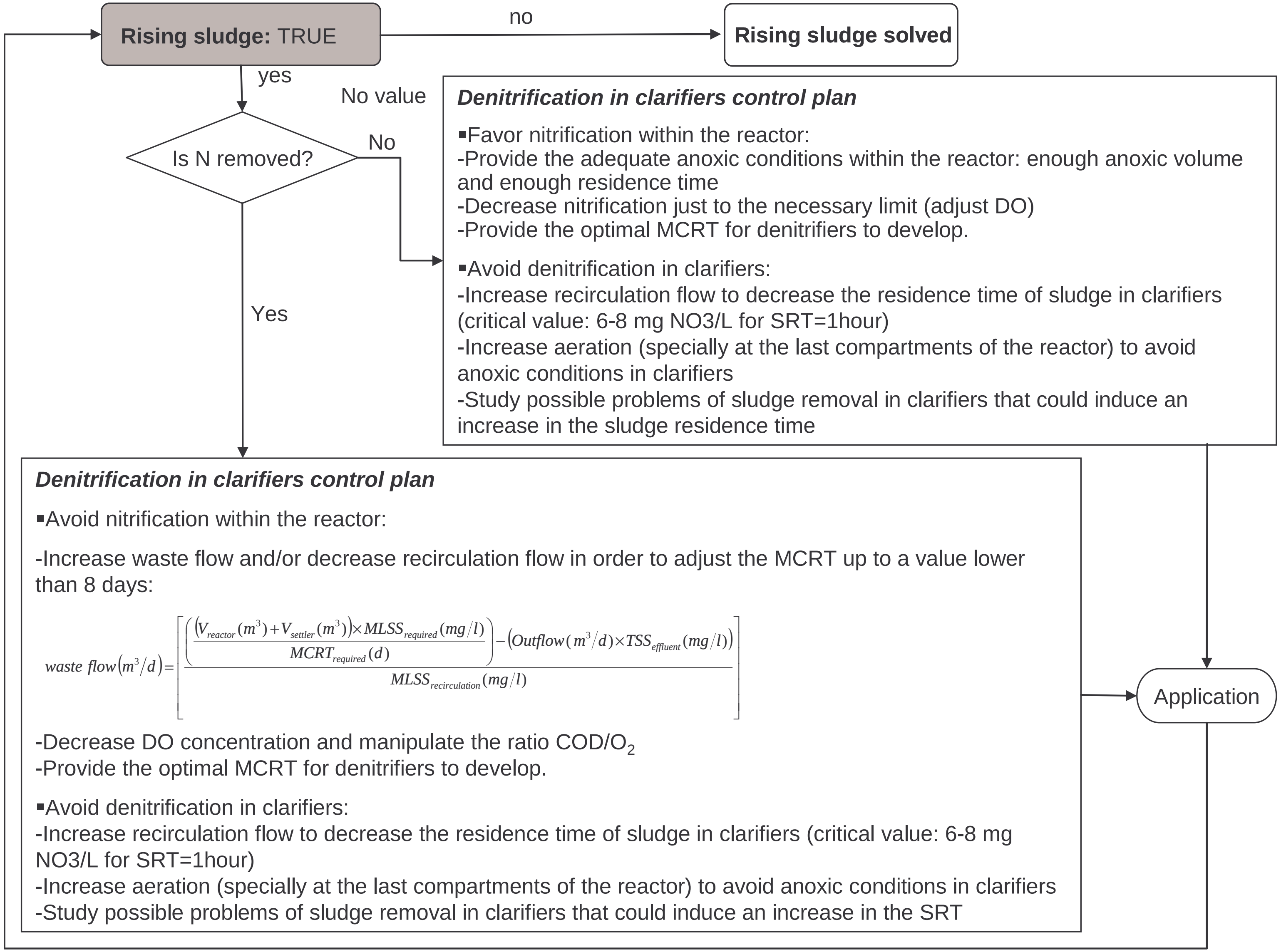
Chemical addition action plan

- Keep on with the same type of chemical, dosage and concentration
- Decrease addition if defflocculation is identified

Chemical addition action plan

- Increase chemical concentration or dosage
- Decrease the addition if defflocculation is identified

- Change the chemical to be added (lime, ferric chloride, sodium aluminate, ferric sulfate, etc.) or consider other options such as increasing MLSS settleability by means of inert biological solids from anaerobic digestors.
- Start again chemical addition plan.



Denitrification in clarifiers control plan

▪Favor nitrification within the reactor:

-Provide the adequate anoxic conditions within the reactor: enough anoxic volume and enough residence time
-Decrease nitrification just to the necessary limit (adjust DO)
-Provide the optimal MCRT for denitrifiers to develop.

▪Avoid denitrification in clarifiers:

-Increase recirculation flow to decrease the residence time of sludge in clarifiers (critical value: 6-8 mg NO3/L for SRT=1hour)
-Increase aeration (specially at the last compartments of the reactor) to avoid anoxic conditions in clarifiers
-Study possible problems of sludge removal in clarifiers that could induce an increase in the sludge residence time

Denitrification in clarifiers control plan

▪Avoid nitrification within the reactor:

-Increase waste flow and/or decrease recirculation flow in order to adjust the MCRT up to a value lower than 8 days:

$$waste\ flow\left(m^3/d\right)=\left[\frac{\left(\frac{\left(V_{reactor}(m^3)+V_{settler}(m^3)\right)\times MLSS_{required}(mg/l)}{MCRT_{required}(d)}\right)-\left(Outflow(m^3/d)\times TSS_{effluent}(mg/l)\right)}{MLSS_{recirculation}(mg/l)}\right]$$

-Decrease DO concentration and manipulate the ratio COD/O_2
-Provide the optimal MCRT for denitrifiers to develop.

▪Avoid denitrification in clarifiers:

-Increase recirculation flow to decrease the residence time of sludge in clarifiers (critical value: 6-8 mg NO3/L for SRT=1hour)
-Increase aeration (specially at the last compartments of the reactor) to avoid anoxic conditions in clarifiers
-Study possible problems of sludge removal in clarifiers that could induce an increase in the SRT