Modelling Aspects of Wastewater Treatment Processes

Ulf Jeppsson

Department of
Industrial Electrical Engineering and Automation (IEA)
Lund Institute of Technology (LTH)
P.O. Box 118
S-221 00 LUND
SWEDEN

ISBN 91-88934-00-4
CODEN:LUTEDX/(TEIE-1010)/1-444/(1996)

Printed in Sweden by Reprocentralen, Lund University
Lund 1996

Abstract

Wastewater treatment processes are inherently dynamic because of the large variations in the influent wastewater flow rate, concentration and composition. Moreover, these variations are to a large extent not possible to control. The adaptive behaviour of the involved microorganisms imposes further difficulties in terms of time-varying process parameters. Mathematical models and computer simulations are essential to describe, predict and control the complicated interactions of the processes. The number of reactions and organism species that are involved in the system may be very large. An accurate description of such systems can therefore result in highly complex models, which may not be very useful from a practical, operational point of view. The thesis contains a thorough discussion on aspects concerning the mathematical modelling of the activated sludge, sedimentation and biofilm processes.

A reduced order dynamic model, describing an activated sludge process performing carbonaceous removal, nitrification and denitrification with reasonable accuracy, is presented. The main objective is to combine knowledge of the process dynamics with mathematical methods for estimation and identification. The identifiability of the model is investigated using both off-line and on-line methods, and its dynamic behaviour is validated by simulations of a recognized model. The information required by the identification algorithms are based on directly measurable real-time data. The simplified model may serve as a tool for predicting the dynamic behaviour of an activated sludge process, since the parameters can be tracked on-line during varying operating conditions. The model is aimed for operation and control purposes as an integral part of a hierarchical control structure.

The main objective of the work on settler modelling is to support and enlighten recent theoretical results. A new settler model is compared to a traditional layer model by means of numerical simulations. Emphasis is put on the numerical solution's ability to approximate the analytical solution of the conservation law written as a non-linear partial differential equation. The new settler model is consistent in this respect. Several problems that occur when integrating a model of the biological reactor with a model of

the settler are also discussed. In particular, the concentrations of the biological components of the particulate material are of importance for an accurate description of the sludge that is recycled to the biological reactor. Two one-dimensional algorithms have been evaluated. The first algorithm is commonly used and some of its inherent problems are discussed. The second algorithm, which is preferred, is a new analytically derived method.

Few attempts have been made to take into account the influence of higher order organisms in biofilm systems when developing or applying mathematical models. This work describes a simplified modelling approach to include some possible effects of higher order organisms on nitrification, based on a proposed hypothesis of their oxygen consumption in the biofilm. Three different models are developed and investigated. Model simulations are validated using data from a laboratory experiment using continuous-flow suspended-carrier biofilm reactors, where the predators were selectively inhibited. The proposed models are capable of reproducing several of the observed effects. They are primarily aimed at capturing the steady-state behaviour of the biofilm but may also prove to be a useful basis for describing the dynamics of the process.

Preface

This thesis contains my work on *Modelling Aspects of Wastewater Treatment Processes* at the Department of Industrial Electrical Engineering and Automation (IEA), Lund Institute of Technology, Lund, Sweden. In order to improve the comprehensibility of the thesis it is divided into five different parts, where Part I and Part V contain a *General Introduction to Mathematical Modelling* and *Conclusions*, respectively. Part II contains my work in the field of *Modelling the Activated Sludge Process*, Part III deals with *Modelling the Settling Process* and Part IV contains some new aspects of importance when *Modelling Biofilm Processes*. Although the different processes are related and in practice normally combined within a wastewater treatment plant, I believe that this division will make it easier for the reader to locate areas related to his/her special interests.

Several parts of the work presented in this thesis have previously been published in or submitted to international journals, or presented at international conferences and workshops. A list of the relevant references in chronological order is given below.

Main Papers

[1] Jeppsson, U. (1993), *On the Verifiability of the Activated Sludge System Dynamics*. Licentiate's thesis, CODEN:LUTEDX/(TEIE-1004)/1-177/(1993), IEA, Lund Inst. of Tech., Lund, Sweden.

[2] Jeppsson, U., Olsson, G. (1993), "Reduced Order Models for On-Line Parameter Identification of the Activated Sludge Process". *Wat. Sci. Tech.*, vol. 28, no. 11-12, pp. 173-183.

[3] Jeppsson, U. (1994), "Simulation and Control of the Activated Sludge Process – a Comparison of Model Complexity". *Proc. IMACS Symposium on Mathematical Modelling*, (I. Troch and F. Breitenecker eds.), Technical University Vienna, Vienna, Austria, vol. 3, pp. 444-451.

[4] Olsson, G., Jeppsson, U. (1994), "Establishing Cause-Effect Relationships in Activated Sludge Plants – what Can Be Controlled". *Proc. 8th Forum Applied Biotechnology*, Med. Fac. Landbouww., University of Gent, Gent, Belgium, vol. 59, pp. 2057-2070.

[5] Diehl, S., Jeppsson, U. (1995), "A Simulation Model of the Reactor–Settler Interaction in Wastewater Treatment". Paper VI in *Conservation Laws with Application to Continuous Sedimentation.* Ph.D. dissertation by S. Diehl (ISBN 91-628-1632-2), Dept of Mathematics, Lund Inst. of Tech., Lund, Sweden.

[6] Jeppsson, U. (1995), "A Simplified Control-Oriented Model of the Activated Sludge Process". *Mathematical Modelling of Systems*, vol. 1, no. 1, pp. 3-16.

[7] Jeppsson, U., Diehl, S. (1995), "Validation of a Robust Dynamic Model of Continuous Sedimentation". *Proc. 9th Forum Applied Biotechnology*, Med. Fac. Landbouww., University of Gent, Gent, Belgium, vol. 60, pp. 2403-2415.

[8] Jeppsson, U., Lee, N., Aspegren, H. (1995), "Modelling Microfauna Influence on Nitrification in Aerobic Biofilm Processes". *Proc. Int. IAWQ Conf. Workshop on Biofilm Structure, Growth and Dynamics*, Delft University of Technology, The Netherlands, pp. 77-85.

[9] Diehl, S., Jeppsson, U. (1996), "A Model of the Settler Coupled to the Biological Reactor". Submitted to *Wat. Res.*

[10] Jeppsson, U., Diehl, S. (1996), "An Evaluation of a Dynamic Model of the Secondary Clarifier". Accepted for presentation at *Water Quality International '96, IAWQ 18th Biennial International Conference*, Singapore, June 23-28.

[11] Jeppsson, U., Diehl, S. (1996), "On the Modelling of the Dynamic Propagation of Biological Components in the Secondary Clarifier". Accepted for presentation at *Water Quality International '96, IAWQ 18th Biennial International Conference,* Singapore, June 23-28.

Subsidiary Papers

[12] Aspegren, H., Andersson, B., Olsson, G., Jeppsson, U. (1990), "Practical Full Scale Experiences of the Dynamics of Biological Nitrogen Removal". *Instrumentation, Control and Automation of Water and Wastewater Treatment and Transport Systems*, (R. Briggs ed.), Pergamon Press, London, England, pp. 283-290.

[13] Jeppsson, U. (1990), "On-Line Estimation of the Nitrosomonas and Nitrobacter Concentrations in a Nitrification Process", (in Swedish). *Technical Report, CODEN:LUTEDX/(TEIE-7025)/1-93/(1990)*, IEA, Lund, Sweden.

[14] Jeppsson, U. (1990), "Expert System for Diagnosis of the Activated Sludge Process", (in Swedish). *Technical Report, CODEN: LUTEDX/(TEIE-7026)/1-83/(1990)*, IEA, Lund, Sweden.

[15] Jeppsson, U. (1994), "A Comparison of Simulation Software for Wastewater Treatment Processes – a COST 682 Program Perspective". *Technical Report, CODEN:LUTEDX/(TEIE-7078)/1-12/(1994)*, IEA, Lund, Sweden.

[16] Olsson, G., Jeppsson, U. (1994), "Modelling, Simulation, and Identification Technologies". *Invited paper, Water Environment Federation (WEF), 67th Annual Conference*, Chicago, Illinois, USA, Oct. 15-19.

[17] Ayesa, E., Carstensen, J., Jeppsson, U., Vanrolleghem, P. (1996), "Identification of the Dynamic Processes in WWTP and Control of WWTP". *COST 682: Environment*, (D. Dochain, P. Vanrolleghem, M. Henze eds.), European Commission, Directorate-General XII: Science, Research and Development, Luxembourg, pp. 89-104.

[18] Carstensen, J., Vanrolleghem, P., Ayesa, E., Jeppsson, U., Urrutikoetxea, A., Vanderhaegen, B. (1996), "Objective Functions for Wastewater Treatment Design and Operation". *COST 682: Environment*, (D. Dochain, P. Vanrolleghem, M. Henze eds.), European Commission, Directorate-General XII: Science, Research and Development, Luxembourg, pp. 105-108.

[19] Vanrolleghem, P., Jeppsson, U. (1996), "Simulators for Modelling of WWTP". *COST 682: Environment*, (D. Dochain, P. Vanrolleghem, M. Henze eds.), European Commission, Directorate-General XII: Science, Research and Development, Luxembourg, pp. 67-78.

[20] Vanrolleghem, P., Jeppsson, U., Carstensen, J., Carlsson, B., Olsson, G. (1996), "Integration of WWT Plant Design and Operation – a Systematic Approach Using Cost Functions". Accepted for presentation at *Water Quality International '96, IAWQ 18th Biennial International Conference*, Singapore, June 23-28.

Comment

For reasons of simplification, the above papers are scarcely referenced explicitly within the thesis although they are all included in the Bibliography to provide a complete description of the used sources. Instead, their main contributions to the thesis are listed below.

Of the main papers, my licentiate's thesis [1] is the basic source for my research on modelling of the activated sludge process and has consequently strongly influenced the contents of Chapters 2, 3 and 4. Chapter 3 further covers material from [4] and parts of Chapter 4 are also discussed in [2], [3] and [6]. Chapters 5 and 6 cover material from [5], [7], [9], [10] and [11], while [8] is the initial work that has been extended in Chapter 8.

The influences of the subsidiary papers are more difficult to define (as they are subsidiary). Chapters 2 and 3 are to some degree influenced by concepts presented in [15], [16], [17], [18], [19] and [20], while [12], [13] and [14] contain some ideas that have been incorporated into Chapter 4.

Acknowledgements

First of all, I would like to express my sincere gratitude to my supervisor Professor Gustaf Olsson for introducing me to the challenges of modern wastewater treatment, for all his support and for the great enthusiasm he has always shown for my work. I cannot imagine having any other supervisor than Gustaf. Furthermore, his careful reading of this thesis has improved its quality considerably. He is also the co-author of several of the papers on which Chapter 4 is based.

I also want to thank my two colleagues Dr Bo Peterson and doctoral student (Teknologie licentiat) Olof Samuelsson for proof-reading the thesis and providing me with valuable comments. In addition to this, Bo has helped me solve various practical problems with regard to the layout of the thesis. At the same time, I would like to thank the rest of the staff and my fellow graduate students at the Department of Industrial Electrical Engineering and Automation for providing the creative and friendly atmosphere which makes work a pleasure.

I have had the privilege of being part of a large, interdisciplinary research group, STAMP, working within the field of wastewater treatment, during the last six years. A large amount of important research has been carried out and presented all over the world. The group is today internationally recognized for its work. I would like to take this opportunity to thank all my friends and colleagues within STAMP (in Lund, Göteborg, Stockholm and Uppsala).

During the last two years I have been involved in a research project with Dr Stefan Diehl (Dept of Mathematics, Lund University). Working with him adds an extra dimension to the pleasures of research. Stefan, who originally developed the new settler model discussed in Chapter 5, has taught me a great deal and I am much indebted to him for his contributions to the papers on which Chapter 6 is based. It is my sincere wish that our collaboration will continue in the future. I also wish to thank him for proof-reading large parts of the thesis, providing me with huge amounts of detailed comments and suggesting important improvements with regard to the content of the thesis.

I would also like to thank Dr Henrik Aspegren (Sjölunda WWTP, Malmö) and Dr (within a few weeks) Natuscka Lee (Dept of Biotechnology, Lund University), who approached me about the idea of trying to model the influence of predators on biofilm systems. As this was a new area to me, I am much indebted to them for explaining the details of the process to me (over and over again). Working with them has been a great experience and I am grateful for their contributions to the original paper on biofilms, which was later expanded into Chapter 8. I would also like to thank them for proof-reading Chapters 7 and 8 and providing me with many useful comments.

I am also grateful to several people abroad who have contributed to my work. I would especially like to thank Dr Eduardo Ayesa (CEIT, San Sebastian, Spain), Dr Jacob Carstensen (Technical University of Denmark, Lyngby, Denmark) and Dr Peter Vanrolleghem (University of Gent, Belgium). They are not only co-authors of some of the papers used in this thesis but also close personal friends. Others who have inspired me and provided me with valuable comments over the years are Professor Emeritus John Andrews (formerly at Rice University, Houston, Texas, USA), Professor Peter Dold (McMaster University, Hamilton, Ontario, Canada) and Professor Gerrit van Straten (Wageningen Agricultural University, Wageningen, The Netherlands).

Finally, I thank my parents and the other members of my immediate family for the love, support and trust that they have always shown towards me.

The work has been supported financially by the Swedish National Board for Industrial and Technical Development (NUTEK) as part of the STAMP project, contract no. 90-4882. This support is gratefully acknowledged.

Lund, Easter Sunday 1996

Ulf Jeppsson

Axioms of Modelling

Three fundamental axioms to consider for everyone working with the development of mathematical models and model simulations are summarized below. The pictures are from Ljung and Glad (1991).

Note! No figure available.

Pygmalion: King of Cyprus, who was also a famous sculptor. He fell in love with one of his sculptures and pleaded with the gods to bring it to life.

Axiom 1: Do not fall in love with your model.

Note! No figure available.

Procrustes' bed

Procrustes: A robber in Greek mythology known for his bed where he tortured the people he caught. He stretched those who were too short until they fitted the bed, and he cut off the legs and the head of those who were too long.

Axiom 2: Do not try to adapt reality to your model.

Note! No figure available.

Axiom 3: Do not extrapolate your model's region of validity too far.

Contents

PART V: CONCLUSIONS AND PERSPECTIVES

PART I

General Introduction to Mathematical Modelling

Chapter 1

Introduction

All forms of life are dependent on water. The most prominent theories claim that life on this planet first developed in the surface layer of the prehistoric oceans approximately three billion years ago. Humans still spend the first nine months of their development completely surrounded by water and during an early phase of the growth of the fetus, it develops rudimentary gills. The human body consists of approximately 64 % water and requires 2–3 litres of drinking water daily to maintain a healthy balance within the body (during normal conditions). The immense quantity of water, covering 71 % of the earth's surface to a mean depth of 3.8 km, measures in volume about $1.4 \cdot 10^9$ km^3. But only a maximum of about 0.3 % of the world's total water resources is available for human usage.

In comparison with other life forms, man has a greater influence on the quality of the water around him. Man's influence is both negative and positive. Pollution, depletion of natural water resources and the deterioration of living conditions for other life forms are products of man's modern society. Nature's own self-cleansing mechanisms have been adopted by man to treat the wastes which are generated in his increasingly industrial society. As abundant supplies of clean, fresh water become more scarce, the need to protect these sources become more acute. The lack of clean drinking water is today probably the single most important factor for the spreading of various diseases in many parts of the world, leading to the deaths of thousands of people every day. On many occasions, wars have been initiated as a result of disagreements over the use of natural water resources.

The need for various aspects of sanitary engineering has been recognized for thousands of years. References are made in the Bible to sanitation laws and some ancient structures, such as the Cloaca Maxima built be the Romans more than 2500 years ago, are still in use today (Fuhrman, 1984). Epidemics like the Black Plague showed man's vulnerability to poor

hygienic conditions. Sir John Harington's invention of the flush toilet in 1596 was a significant technical development in the quest for improved hygienic conditions (although an arguable claim today) but its use was not common until much later. Prior to the mid 1800s there was little treatment of wastewaters as most utilities were constructed only for drainage purposes. The more general introduction of the flush toilet around the turn of the century led to a dramatic decrease of sanitary problems within the cities but instead created another problem – heavily polluted receiving waters. This initiated much of the development of wastewater treatment systems.

Today, modern and environmentally developed cities utilize vast sewage collection systems to collect and transport all types of wastewaters from homes, businesses and industries to wastewater treatment facilities. Once at the treatment plant, the wastewater is exposed to different processes which can remove most of the pollutants. The degree to which the wastewater must be purified depends on the ability of the recipient to accept, without harm, the effluent. In Figure 1.1 the development of wastewater treatment systems in Sweden in recent years is illustrated.

Note! No figure available.

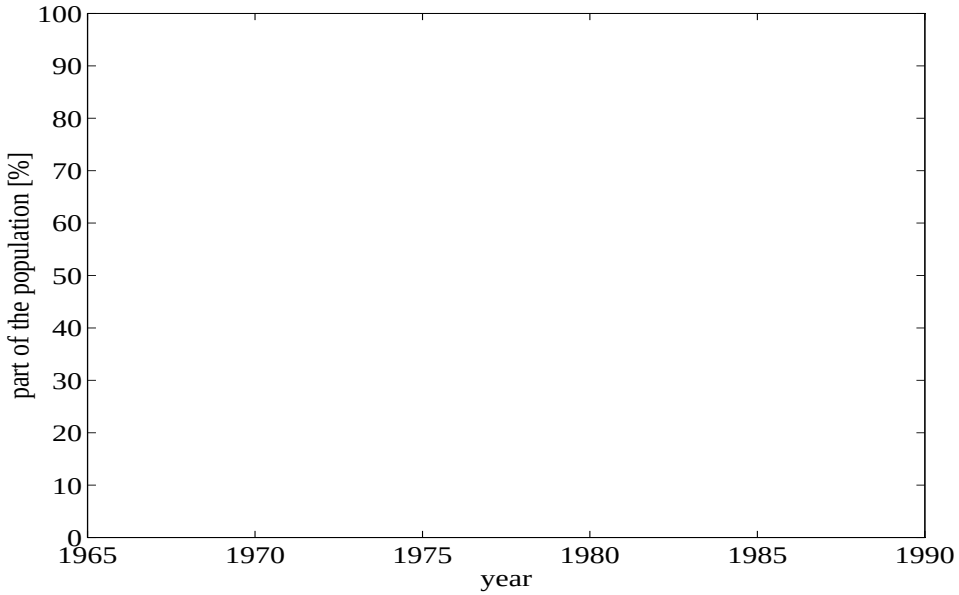


Figure 1.1 The development of the Swedish wastewater treatment systems in urban areas from 1965 to 1990 (Hultman, 1992).

Modern wastewater treatment (WWT) techniques have been in use for over a century. Many different processes have been developed and many variations tested. The activated sludge process and processes using biofilms (i.e., biological treatment) are two of the most common processes used today. Practically all wastewater treatment systems also use sedimentation

at some stage of the treatment process, to separate the solid matter from the liquid in a suspension. Some aspects of these three fundamental processes have been investigated from the viewpoint of mathematical modelling and will be presented in this work.

1.1 Motivation

Wastewater treatment processes can be considered as the largest industry in terms of treated mass of raw materials. In the European Community, for instance, a daily wastewater volume of approximately $40 \cdot 10^6$ m^3 has to be processed (Lens and Verstraete, 1992). However, studies have shown that even well attended WWT plants fail to meet the required effluent quality standards up to 9 % of the operation time (Berthouex and Fan, 1986), not including the short upsets lasting less than one day. The U.S. Environmental Protection Agency estimated that one out of three treatment works were not in compliance with discharge limitations (Ossenbruggen *et al.*, 1987), and in Germany and the Netherlands clarification problems were found to occur in almost half of the evaluated treatment plants (Chambers and Tomlinson, 1982). Besides poor design, overloading and inadequately trained operators, a lack of process control leading to excessive effluent quality variations, was reported as the main cause.

A closer look at the current operation of wastewater treatment plants shows that automation, while introduced in the late sixties, can still be considered minimal. Few plants are equipped with more than a few elementary sensing elements and control loops, mostly concerning flow metering and control, and for monitoring the basic plant performance over longer periods of time. Since the early seventies, when a major leap forward was made by the widespread introduction of dissolved oxygen control, little progress has been made. A number of reasons for this lack of *instrumentation, control and automation* (ICA) have been put forth (Beck, 1986; Olsson, 1993):

- *Understanding*: Insight in the treatment processes is still insufficient and there is a lack of suitable mathematical models.
- *Inadequate instrumentation*: Non-existent or insufficiently reliable technology.
- *Plant constraints*: Inapt and insufficient flexibility to manipulate the process.

- *Economic motivation*: There exists a lack of fundamental knowledge concerning benefits versus costs of automated treatment processes. In addition, WWT processes are not productive and automation can only contribute to a decrease of operating costs but does not directly lead to increased profit.
- *Education and training*: Operators are not always adequately trained to operate advanced sensor and control equipment and most environmental engineers would need more basic understanding of process dynamics and control in order to appreciate the potential of ICA.
- *Communication*: The interaction between operators, designers, equipment suppliers, researchers and government regulatory agents is often unsatisfactory and leads to poorly designed plants.

Over the last decade, the increased public awareness, as reflected in more stringent effluent regulations, has considerably increased the requirements imposed on treatment plants. Not only the organic carbon pollution of a wastewater must today be eliminated but also nutrients (i.e., nitrogen and phosphorus). With biological nutrient removal being the most economic way of treatment (in most cases), rather complex process configurations have resulted. The numerous interactions that occur among the different unit processes and the fact that the biological potential is taken to its limits, lead to nutrient removal plants being quite vulnerable to external disturbances or erroneous manipulations. Hence, the increased complexity is a major driving force for the introduction of advanced instrumentation that can provide the necessary information of the process condition. Moreover, as process complexity increases, more possibilities are required to act upon the process to guarantee satisfactory treatment performance. Finally, the increasing number of measured and manipulable variables gives rise to more complex control systems that take advantage of the new possibilities.

Many wastewater treatment plants are presently operated according to predetermined schemes with very little consideration to the variations of the material loads. Using on-line sensors for on-line control of the operation of the plants may enhance the ability to comply with assigned effluent standards. In general, a better understanding of the dynamic behaviour of the process, adequate mathematical models and an on-line identification of model parameters and influent loads in combination with the use of control systems have significant potential for solving operational problems as well as reducing operational costs. In addition, this knowledge may be used for reduction of volume holdings in the design of the plants to be constructed in the future.

Control of wastewater treatment plants relies on four building blocks (see Figure 1.2):

- insight in the plant operations and dynamics summarized in an appropriate mathematical process model;
- sensors that provide on-line data from some of the output variables of the process and disturbances acting upon it;
- adequate control strategies which try to minimize deviations from the control objectives;
- actuators which implement the controller outputs on the plant.

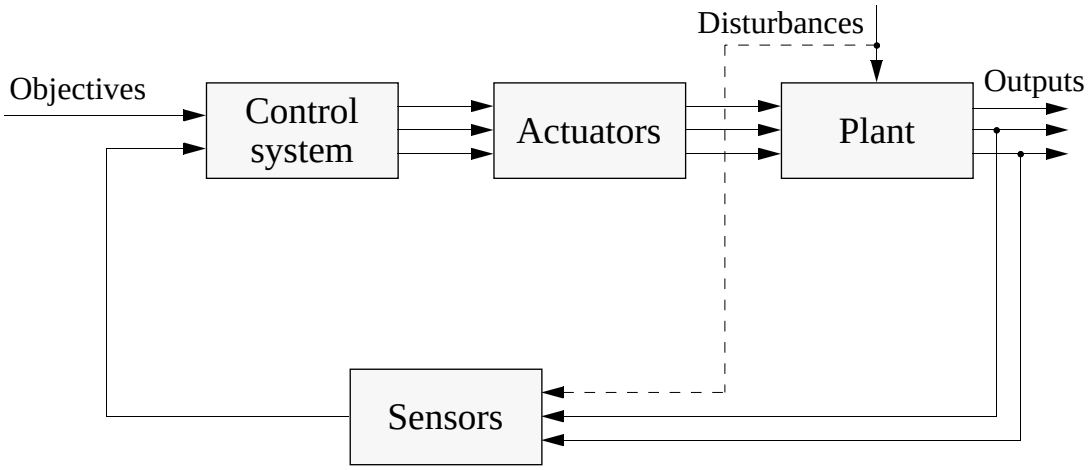


Figure 1.2 Schematic description of the control chain of a wastewater treatment plant (Vanrolleghem, 1994).

It must be realized that probably the most critical phase in the solution of any control problem is in the modelling stage. This is because nearly all control techniques require a knowledge of the dynamics of the system to be controlled before an attempt can be made to design a controller for it. This means that the primary building block of any modern control exercise is to construct and identify a model for the system to be controlled. Moreover, the quality of the control system obtained depends on the designer's understanding of the given system's dynamics and its limitations.

Many model construction and identification studies are conducted using general linear statistical models and corresponding identification algorithms. It is easy to assume more or less complete ignorance of a system and then let a general algorithm find an empirical model for it. This approach works fairly well on linear processes but a vast number of important systems (such as wastewater treatment processes) cannot be described adequately by linear equations. Another disadvantage of such an

approach is that it gives no or very little physical information about the system in question.

It is currently not possible to model and identify general non-linear systems because of their great variety and because of their structural complexity. There are numerous non-linear systems which must be controlled in the present industrial world. Thus, it is important to develop methods to model non-linear systems and to estimate the model parameters. Once the models are available, it is also important to investigate how they can be used to establish an understanding and an effective control of the relevant system.

The behaviour of biotechnological processes occurring in a bioreactor has a complexity unparalleled in the chemical industry. As a consequence, its prediction from information about the environmental conditions is extremely difficult. The number of reactions and organism species that are involved in the system may be very large. An accurate description of such complex systems can therefore result in quite involved models, which may not be useful from a control-engineering viewpoint. We can summarize some of the major problems when trying to model WWT processes in general as follows:

- lacking process knowledge (e.g., biofilm structure changes, hydrolysis, flocculation, settling characteristics);
- several different unit processes interconnected by various internal feedbacks;
- macroscopic modelling of microscopic reactions;
- highly non-linear processes;
- non-stationary processes;
- time varying process parameters (due to the adaptive behaviour of living organisms to changing environmental conditions);
- practically non-controllable and highly variable process inputs;
- lack of adequate measuring techniques.

1.2 Objectives and Contributions of the Thesis

Current research in the area of wastewater treatment process modelling is among other things concerned with the following items (Olsson, 1993; Vanrolleghem, 1994; Henze *et al.*, 1996):

- *Incorporation of the latest scientific insights in the different processes*: Significant efforts are made to model important processes, such as the
  1) phosphate removal;
  2) hydrolysis of substrates;
  3) fate of biopolymers;
  4) sedimentation process with special emphasis on the interaction between the biological phenomena such as filament growth and settling properties of the sludge.
- *Identifiability*: A discrepancy has developed between the amount of data needed to identify the increasingly complex models and the amount of information that can be obtained from the process. Especially if only on-line data can be used for model identification, serious problems may occur when trying to find unique parameter estimates. Even the combination of on-line and off-line data may be insufficient for accurate modelling. Current research is therefore directed towards the development of new monitoring equipment and off-line methodologies adapted to the information need of the complex models.
- *Verifiability*: The mathematical models that have been introduced recently are the result of fundamental studies aimed at elucidating the mechanisms of certain microbial processes. To explain the detailed experimental findings more precisely, state variables and parameters that are not directly measurable (e.g., active heterotrophs), have been introduced in the models. Hence, since verification of a model requires that all model predictions of the state variables can be compared with experimental data, current models have become intrinsically non-verifiable.
- *Model reduction for process control*: The identifiability and verifiability problems mentioned above means that considerable efforts must be devoted to the development of new sensor technology and experimental methods so that the new models may be used in model-based control systems. An alternative approach which attracts much attention is directed to the

reduction of the complexity of existing mechanistic models. The reduction is carried out to such a level that on-line identification with existing sensor technology is feasible while maintaining the necessary predictive capabilities of the major phenomena in the process.

In a broad perspective, the objective of this thesis is to cover different aspects of all the four items discussed above. As the various processes in wastewater treatment are closely related (activated sludge, sedimentation and different types of biofilm processes), it is the aim of this thesis to provide a good overview of the processes and in detail investigate some important modelling possibilities for each system. The introduced models should provide an adequate basis for future model development and refinements and also fulfil a potential need for different practical applications within the wastewater treatment industry.

The main objective of the work on activated sludge modelling is to combine knowledge of the process dynamics with mathematical methods for estimation and identification, in order to obtain the simplest possible models capable of describing the carbonaceous and nitrogenous activities in the process with reasonable accuracy. The objective of the models is that they should be identifiable from available on-line measurements and thereby provide a basis for future development of more sophisticated control strategies, such as feed-forward, adaptive and other types of model-based control. Due to the time varying characteristics of the process it is essential that the model parameters can be uniquely updated on-line. In order to make the reduced models more comprehensible and easy to use it is also important to maintain the basic mechanistic structure and model the reactions in a simplified but still physically reasonable way.

The behaviour of the secondary clarifier is often reported to be a bottleneck of the AS process. The need for adequate models is apparent. The highly complex two and three-dimensional models that have recently become available are still too complex to be used in practical simulations and are difficult to integrate with bioreactor models. The majority of used settler models are one-dimensional layer models. The objective of this thesis within the field of settler modelling is to enlighten and demonstrate the benefits of using a stringent mathematical analysis as the basis for model development. The model should not contain any *ad hoc* assumptions and predictions should be consistent with the analytical solution of the continuity equation, on which it is based. Moreover, it should be as simple as possible, computationally efficient, and easy to combine with models describing the bioreactor.

Effects due to higher order organisms in biofilm systems are normally neglected in current biofilm models, although it is known that the organisms play an important role. The objective of the biofilm modelling in this thesis is to introduce new knowledge (or a new hypothesis) into a biological model in an attempt to explain recent experimental findings concerning the influence of higher order organisms on the nitrification capacity of aerobic biofilms. However, as biofilm models are extremely complex in their current forms, a related objective is to describe the observed behaviour in a simplified manner. As the amount of information from the experimental system is limited, the primary aim is only to extend an existing model to explain the steady-state behaviour of the process.

The main results of this thesis are summarized in Chapter 9 together with some suggestions of topics for future research. The major contributions of the thesis are given below.

- The development of a set of reduced order models for the activated sludge process aimed at control applications. A detailed analysis of the identifiability properties and general dynamic behaviour of the models are given together with an investigation of both off-line and on-line methods for state and parameter estimation using currently available on-line measurements.
- A thorough evaluation of different one-dimensional settler models with respect to consistency and robustness. In particular, the coupling between a bioreactor model and a new robust settler model, derived from a stringent mathematical analysis based only on the constitutive assumption by Kynch, is presented in detail. This modelling of the entire AS process includes the prediction of the concentrations of the individual biological components as they propagate through the settler.
- A set of preliminary attempts to include the influence of higher order organisms on aerobic biofilm systems into existing biological models, without increasing the model complexity in any significant way. The modelling approach is based on the oxygen consumption of higher order organisms within the biofilm and will be further investigated in the future.

The thesis also includes an extensive bibliography, summarizing much of the innovative work performed within the field of modelling wastewater treatment processes, which the interested reader can use as a basis for further literature studies.

Finally, it is believed that the thesis (at least parts of it), with regard to its rather broad perspective on mathematical modelling of processes related to wastewater treatment, may serve as a good introduction for graduate students, when starting their work within this important field of research.

1.3 Outline of the Thesis

An interdisciplinary approach is often required to understand complex systems such as wastewater treatment processes. Unfortunately, within the field of wastewater treatment there appears to be a division between people working primarily with the process itself and people working with mathematical modelling and control aspects of the process. To enhance the understanding of this thesis for readers with different backgrounds, the basics of mathematical modelling are reviewed in Chapter 2 and descriptions of the investigated processes are included in Chapters 3, 5 and 7. Although this approach increases the scope of the thesis, it is believed to be beneficial in terms of allowing readers from different research fields to understand the purpose of the work better.

The thesis is organized in five different parts in order to make the content more comprehensible and also to make it easier for the reader to locate areas related to his/her special interests. These five parts are:

- Part I (Chapters 1 and 2): General introduction to problems related to wastewater treatment and various aspects of mathematical modelling;
- Part II (Chapters 3 and 4): Problems and possibilities related to modelling of the activated sludge process, focusing on reduced order models for future control applications;
- Part III (Chapters 5 and 6): Capabilities of one-dimensional layer models for an accurate description of the behaviour of the secondary clarifier, focusing on an analytically derived model using a consistent numerical algorithm;
- Part IV (Chapters 7 and 8): Possibilities of including the influence of higher order organisms on aerobic biofilms into a biological model in a simplified manner, validated against experimental data;
- Part V (Chapter 9): Conclusions and future perspectives.

In Chapter 2 we introduce several concepts related to mathematical modelling. A strategy for model building is presented together with a thorough discussion on model validation. Methods for model structure evaluation, model reduction, model identifiability, state and parameter estimation, etc., are reviewed. A small example is also used to demonstrate some of the problems related to mathematical modelling.

Chapter 3 provides a detailed background of the activated sludge (AS) process – the historical development, different types of processes available, and a review of the research within the field of modelling leading up to the models used today to describe the mechanisms of the AS process. The IAWQ AS Model No.1 is studied in detail. Finally, various problems related to the AS process, such as the lack of proper sensor technology and the limited flexibility of the process for control purposes, are discussed.

Chapter 4 deals with reduced order models for the AS process. Two reduced order models are developed (based on the IAWQ model) and investigated by means of numerical simulations. The identifiability problem is considered and the behaviour of the reduced models are compared with results of the IAWQ model for different operating conditions. State and parameter estimations are performed by both off-line and on-line methods, using the IAWQ model as a reference model.

The fundamentals of settler modelling are described in Chapter 5. The conservation law, which is the basis for all mechanistic settler models, is thoroughly presented from a mathematical viewpoint. The development of one-dimensional layer models is discussed together with problems related to the settling velocity functions used in these models, leading up to the models commonly used today. Finally, a new robust settler model that is derived using the knowledge of the analytical solution of the continuity equation and Kynch's constitutive assumption only, is described in detail.

In Chapter 6 problems related to the coupling of a settler model to a model of the bioreactor are investigated. A detailed analysis of the behaviour of the robust settler model compared with other one-dimensional models is carried out by means of numerical simulations. The models are investigated during steady-state and dynamic conditions, both for the settler model used as a stand-alone model and when coupled to the bioreactor (simulating an entire AS process). A special analysis is performed in relation to the problem of describing the propagation of the individual biological components through the settler.

Chapter 7 presents the fundamental mechanisms of biofilm processes together with a review of available process variants and some of the most important model developments during the last decades. A special description of the behaviour of higher order organisms influencing the biofilm process is also provided. Finally, a state-of-the-art biofilm model is discussed together with the numerical algorithm used to solve the resulting system of stiff partial differential equations.

An experimental system was used to investigate the influence of higher order organisms on the behaviour of a nitrifying aerobic biofilm. The results are presented in Chapter 8 together with a hypothesis on the role of this type of organism within biofilm systems. Based on the hypothesis, three different models are developed (extensions of the IAWQ model) and calibrated. Computer simulations are used to validate the behaviour of the models when compared with the experimental results (primarily the steady-state behaviour).

In the concluding chapter of this thesis the main results are summarized. Directions for future research and perspectives are also briefly discussed.

Chapter 2

Mathematical Modelling

In this chapter we give an introduction to the field of mathematical modelling and some relevant areas related to this subject. Model objectives, structures and construction are discussed as well as associated topics such as model identification, estimation, reduction and validation. Of necessity the introduction is limited, as mathematical modelling in general is an area that has provided enough material for hundreds (or possibly thousands) of books. Part of the material in this chapter is covered in [181].

2.1 Overview

Ever since Isaac Newton published his fundamental work *Mathematical Principles of Natural Philosophy* in 1687 where the fundamental laws of force and motion were formulated, the conclusion within the scientific community has been: *Nature has laws, and we can find them.* The importance of this statement cannot be overestimated. It implies that every system – mechanical, electrical, biological or whatever – can be accurately described by a mathematical model. Although proved wrong by the quantum theory or by the recently developed theory of chaos, the influence on the way scientists think has been enormous. In combination with the rapid development of computers during the last fifty years, the number of available models within every scientific area has exploded. The models can today also be applied in practice as the computers allow us to numerically solve process models of such complexity that could hardly be imagined a couple of decades ago.

In an ideal world, process modelling would be a trivial task. Models would be constructed in a simple manner yet in every way reproduce the true process behaviour. Not only would the models be accurate, but they would be concise, easy to use and reveal everything about the internal cause-effect relationships within the process. Each model would be built for a specific task to a prescribed accuracy. Unfortunately, our world is not ideal although the above modelling perspective may serve as an excellent long term goal for everyone dealing with modelling. In the real world it must be realised that a model is always a *simplification of reality.* This is especially true when trying to model natural systems containing living organisms. The common relationship between reality and a mathematical model is illustrated in Figure 2.1.

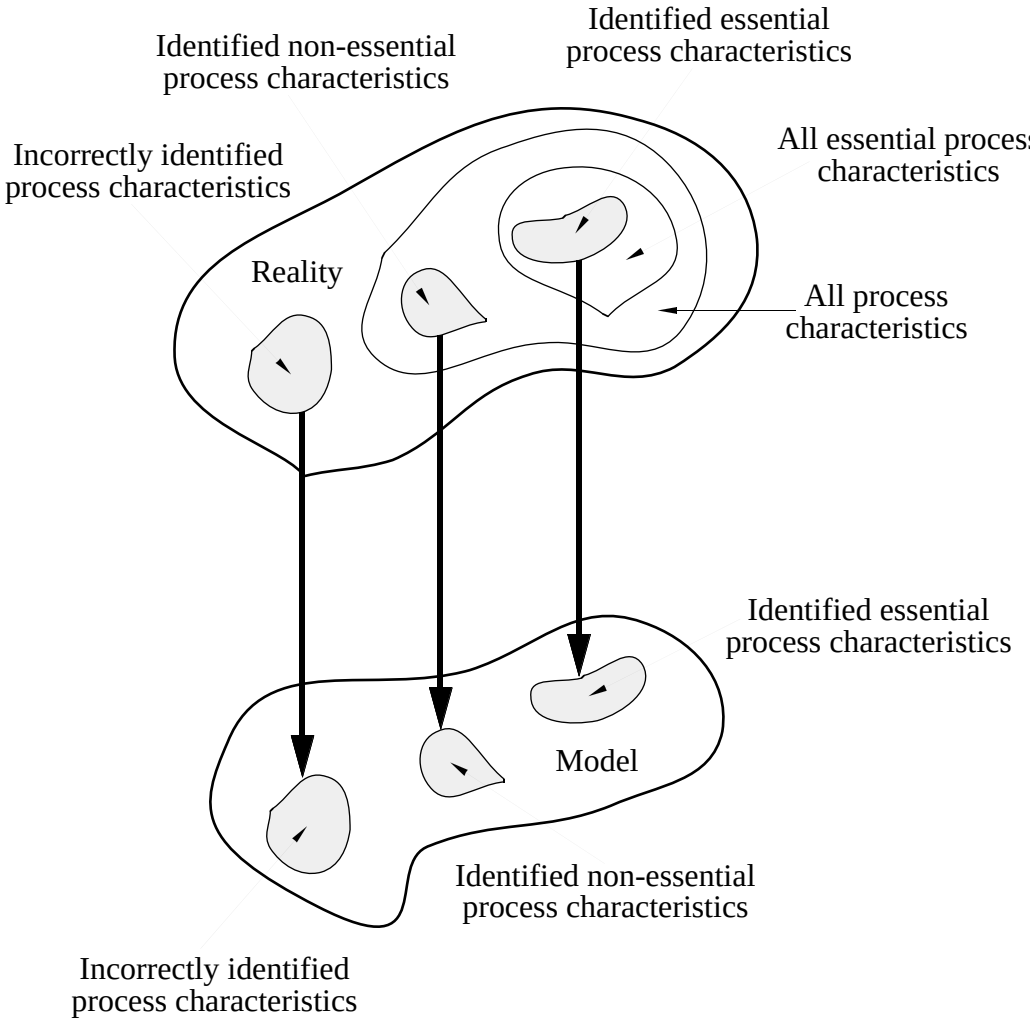


Figure 2.1 The relationship between reality and a model (Thensen, 1974).

Available Literature (Examples)

The number of books dealing with mathematical modelling in general is extensive. A few examples that describe the fundamentals of mathematical modelling, process dynamics and automatic control are Seborg *et al.* (1989), Åström and Wittenmark (1990), Murthy *et al.* (1990), Kuo (1991) and Olsson and Piani (1992).

There exist a large number of books dealing with system identification and related subjects. A few recent examples are Ljung (1987), Söderström and Stoica (1989), and Bohlin (1991). A thorough introduction to the important field of mathematical optimization is given in Fletcher (1987). As the main focus of the above books is aimed at linear and discrete-time models (for automatic control applications), the survey by Mehra (1980) may serve as a good complement, as it deals with system identification methods of non-linear systems. An excellent review of identifiability analysis and problems associated with this is given by Godfrey and DiStefano (1985; 1987). In the dissertation by Robertson (1992) several methods for model reduction are thoroughly explained and a number of case-studies are given.

Modelling of wastewater treatment (WWT) systems can be studied, for example, in Grady and Lim (1980), Patry and Chapman (1989), Andrews (1992), Henze *et al.* (1992) and Orhon and Artan (1994). A good introduction to the possibilities and difficulties of identification, estimation and control of WWT processes is given in Beck (1986; 1987). A large number of relevant references covering much of the work in this field until 1987 are provided. In Beck (1991) the concept of model calibration versus model uncertainty is further emphasized. Finally, the works by Bastin and Dochain (1990), Vanrolleghem (1994) and Reichert (1994b) cover many important aspects of on-line estimation and model structure characterisation for WWT applications.

The above examples of available literature will provide the enthusiastic reader with thousands of more references covering almost every possible aspect of mathematical modelling, system identification and automatic control.

Why Do We Need Mathematical Models?

The word ‘model’ has a wide spectrum of interpretations, e.g., mental model, linguistic model, visual model, physical model and mathematical

model. In this work we will restrict ourselves to mathematical models, that is, models within a mathematical framework where equations of various types are defined to relate inputs, outputs and characteristics of a system.

Primarily, mathematical models are an excellent method of conceptualising knowledge about a process and to convey it to other people. Models are also useful for formulating hypotheses and for incorporating new ideas that can later be verified (or discarded) in reality. An accurate model of a process allows us to predict the process behaviour for different conditions and thereby we can optimize and control a process for a specific purpose of our choice. Finally, models serve as an excellent tool for many educational purposes.

The mathematical model is a tool that allows us to investigate the static and dynamic behaviour of a system without doing – or at least reducing – the number of practical experiments. In practice, an experimental approach often has serious limitations that make it necessary to work with mathematical models instead. Some rather extreme examples of such limitations are given below (Finnson, 1994).

- Too expensive: It is somewhat expensive to launch rockets to the moon until one successfully hits the surface, then rebuild this type of rocket in order to use it for the intended purpose.
- Too dangerous: Starting to train nuclear power plant operators at full-scale running plants is not advisable.
- Too time-consuming: It would take far too much time to investigate all variations of combinations of mixtures, temperature and pressure in a complex chemical process to identify the optimum combination. With a few experiments, the rest of the experimental domain can be simulated by a model.
- Non-existing system: While designing a suspension bridge it is necessary to simulate how different designs will be affected by, for example, high winds.

2.2 General Modelling Strategy

The reasons why we need mathematical models suggested in the previous section are by no means exhaustive. However, once we have concluded that models are useful, we need a general strategy for model building. Such

a strategy will be discussed in this section. The formulation is inspired by the work of Robertson (1992).

In overview, the modelling of any system occurs in five rather distinct steps, as illustrated in Figure 2.2 (Murthy *et al.*, 1990). Step one is to delineate the system being modelled as a functional specification. A quantitative understanding of the structure and parameters describing the process is required. Typically for wastewater applications this functional specification may include such information as equipment type and size, flow-sheet layout, environment variables, nominal operating conditions.

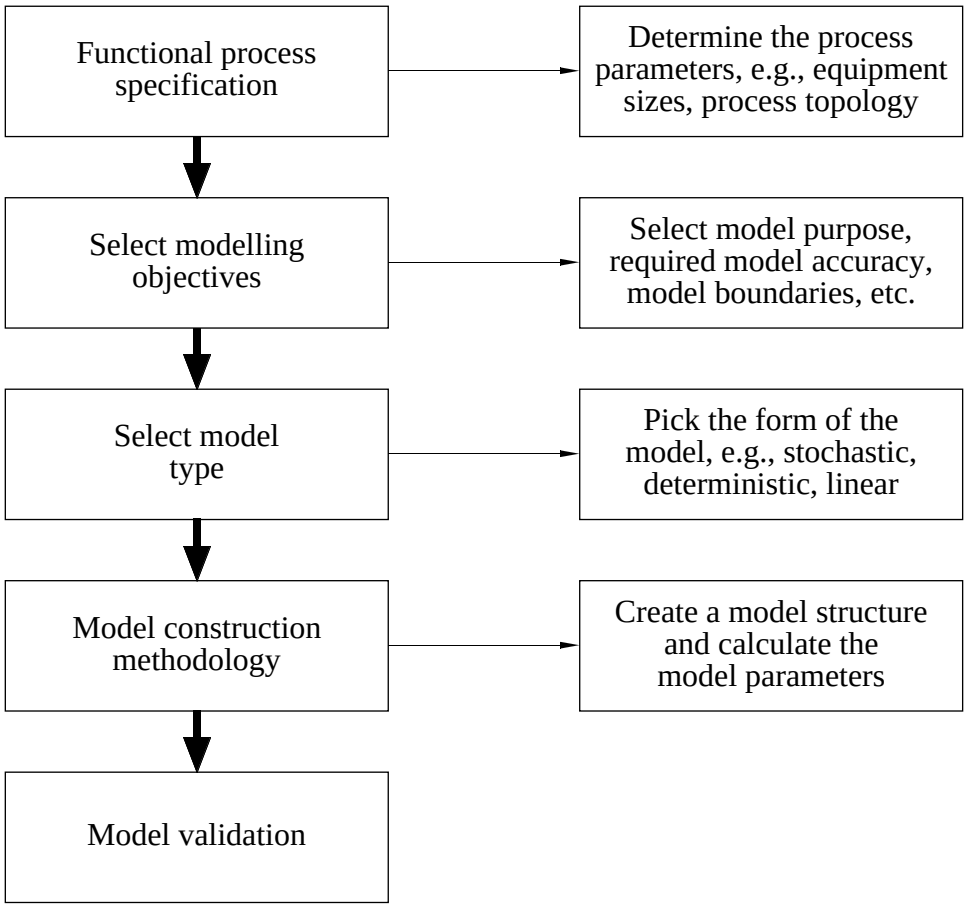


Figure 2.2 An overview of the modelling process.

The modelling objectives are then decided and then the desired model type selected. A model building strategy is then followed to arrive at the appropriate model for the desired application. In the following subsections we will assume that the first step, i.e., the functional process specification, has been successfully accomplished and look somewhat closer at the following four steps.

Modelling Objectives

Any given process may have different 'appropriate' models. The chosen appropriate model will depend on its objectives. These *a priori* decisions about the model must be made before the model construction can begin. Some of the more relevant objectives concern model *purpose*, *system boundaries*, *time constraints* and *accuracy*.

Model Purpose

A wide variety of models are possible, each of which may be suitable for a different application. For example, simple models which may be suitable for model-based control algorithms, may be totally inadequate for simulating and predicting the entire process behaviour for safety and operational analysis. A clear statement of the model intention is needed as a first step in setting the model objectives. This entails listing all the relevant process variables and the accuracy to which they must be modelled.

For example, within the field of wastewater treatment we can define a number of general purposes for mathematical models (also applicable to many other fields). These are listed below.

- Design – models allow the exploration of the impact of changing system parameters and development of plants designed to meet the desired process objectives at minimal cost.
- Research – models serve as a tool to develop and test hypotheses and thereby gaining new knowledge about the processes.
- Process control – models allow for the development of new control strategies by investigating the system response to a wide range of inputs without endangering the actual plant.
- Forecasting – models are used to predict future plant performance when exposed to foreseen input changes and provide a framework for testing appropriate counteractions.
- Performance analysis – models allow for analysis of total plant performance over time when compared with laws and regulations and what the impact of new effluent requirements on plant design and operational costs will be.
- Education – models provide students with a tool to actively explore new ideas and improve the learning process as well as allowing plant operators training facilities and thereby increasing their ability to handle unforeseen situations.

System Boundaries

The system boundaries define the scope of the model. A correct choice of the system boundaries is necessary so that all the important dynamics in the process are modelled. Choice of boundaries which include too many insignificant details will lead to an unduly large model. This may cloud an understanding of how and why the system dynamics are occurring as well as being computationally more expensive. Conversely the definition of boundaries which fail to include significant features of the real process could lead to inaccurate dynamic responses and a loss of confidence in the final model.

If uncertainty exists about the correct choice of boundary, a criterion for boundary selection is to check whether the streams crossing the proposed boundary are easy to characterize (e.g., constant, step impulses). If the streams are well characterized, then the correct boundary has been chosen.

Time Constraints

Time constraints are important model restrictions to be chosen before construction of dynamic models. Frequently the process under investigation will contain a wide range of dynamic activity with widely varying speeds of response. Characteristic time constants in the process may range over many orders of magnitude. Invariably the modeller is interested in a simulation over a defined period of time. For example, in an activated sludge process the dynamics of the dissolved oxygen concentration have a time constant in the range of minutes whereas the dynamics of the biomass population are more in the range of days-weeks.

To produce an appropriate model, the modeller should therefore identify a ‘time-scale-of-interest’ and not model any latent dynamic effects outside this time-scale. This identification should be in the form of maximum and minimum characteristic time constant. Selection of an appropriate time-scale will also have the added advantage of possibly avoiding ultra-stable or stiff problems in the model numerical solution. These numerical problems occur in systems with widely varying time constants or speed of response (Willoughby, 1974).

Accuracy

The appropriateness of the model depends on the ability to predict the system performance within a prescribed accuracy. The accuracy sought

will affect the degree of simplification which can be achieved in building the model. It is important that the desired accuracy of the model be specified before the model is constructed and that this accuracy reflect the purpose of the model. A measure of accuracy must be created to confirm this, or the accuracy must be confirmed during the model validation.

Model Types

Many different classifications have been produced for the different model types which are available (e.g., Murthy *et al.*, 1990; Jørgensen, 1992). It is possible to separate mathematical models based on the philosophy of the approach and with regard to the mathematical form of the model (sometimes also depending on the application area of the model). We start by looking at some different model philosophies.

Reductionist versus Holistic Models

Reductionist models are based on the attempt to include as many details as possible into the model and to describe the behaviour of a system as the net effect of all processes. In contrast to this approach, holistic models are based on a few important global parameters and on general principles.

Internal versus External Models

Internal (or mechanistic) models describe system response as a consequence of input using the mechanistic structure of the system, whereas external (or input/output, black-box, empirical) models are based on empirical relationships between the input and the output. Typical external models are time-series models (e.g., ARMAX models) and neural networks. A mechanistic model is a model based on fundamental engineering and scientific knowledge about the physical, chemical and biological mechanisms that affect a system. A model based on elementary principles tends to produce more reliable results when used for extrapolation (Andrews, 1992). In complex systems it can be very difficult to obtain the necessary fundamental relationships of the process and, consequently, a model must be based on empirical relationships. In practice, models are often a mixture of mechanistic and empirical models, using different concepts at different levels of resolution. As an example, microbial growth rates are in most cases parameterized empirically at the cell level, but macroscopic water flow and substance mass balances are treated in a mechanistic way. External models may even be used to obtain simplified

descriptions of situations in which the validity of an internal model is widely accepted. As an example, empirical parameterizations of turbulent correlations are used in equations describing mean values of turbulent flow, because the solution of the underlying Navier-Stokes equations is too difficult (e.g., Stull, 1988). A good example of a combined mechanistic-empirical approach (so called grey-box modelling) for wastewater applications are presented in Carstensen (1994).

Depending on the mathematical form of the models we can separate them in several ways. Some of the most common ones are presented below.

Dynamic versus Static Models

This classification arises between models that do or do not vary with time. Static models are often referred to as steady-state models. They model the equilibrium behaviour of the system. Conversely dynamic models account for the time varying responses of a system. Both these types are used extensively in engineering applications. This is evidenced by the large number of commercially available ‘simulators’ for both types. While it may appear that the dynamic simulators are dominating, they have received a more limited acceptance outside of an academic environment.

Deterministic versus Stochastic Models

Another classification arises between models that contain uncertainty or randomness in their final results and those that do not. Stochastic models are models in which the final outcome is not known with certainty but can be expressed as a distribution of all possible outcomes. In deterministic models all future outcomes are known with precision by the present state and the future values of external variables (inputs) of the model. Stochastic models also take into account the random influences of the temporal evolution of the system itself. Although the stochastic description of, for example, environmental systems may be more realistic, the large majority of environmental models formulated so far are deterministic. The main reasons for this fact may be the lack of data for the characterization of random variables, high requirements of computational resources for solving stochastic differential equations and the success of deterministic models in describing average future behaviour.

Continuous-Time versus Discrete-Time Models

Continuous-time models are based on formulations of the rates of change of state variables. The values of the state variables as functions of time are then obtained as the solution of a system of differential equations. In contrast to this approach, discrete-time models are based on a division of the time-scale into discrete intervals and specify the state variables in a given time interval as algebraic functions of the values in the immediately preceding time interval. When a model is simulated in a computer the model is always discretized as a digital computer is in itself discrete, although special algorithms and very short time steps may be used to mimic the behaviour of the original continuous system almost perfectly.

Distributed versus Lumped-Parameter Models

Many courses of events which are of interest to modelling are distributed not only in time but also in space. Mathematically, variables distributed in space can be described by partial differential equations and the resulting models are called distributed models. Application of such equations will, however, result in a complex simulation problem. A common way of overcoming this difficulty is to use the lumped-parameter approximation of these distributed equations. To use this approach, isotropic regions in the process are identified. These are regions in which composition, specific energy and momentum are approximately invariant with spatial dimension. The time-varying properties of this 'lump' are then calculated from the transfer of mass, energy and momentum over the boundary of the region.

Other possible ways of classifying mathematical models are, for example, linear versus non-linear models and continuous versus discrete-event models.

Model Construction Methodology

For the last thirty years the importance of building appropriate models has been recognized in many disciplines. Many attempts have been made to apply a systems approach to the development of a modelling methodology, (see Murthy et al., 1990). While a wide variety of different methodologies exist, all possess a number of common features. A general summary of these features is shown in Figure 2.3. After the definition of the modelling objectives and selection of model type, several key steps can be found in all current model building methodologies. These steps are outlined below.

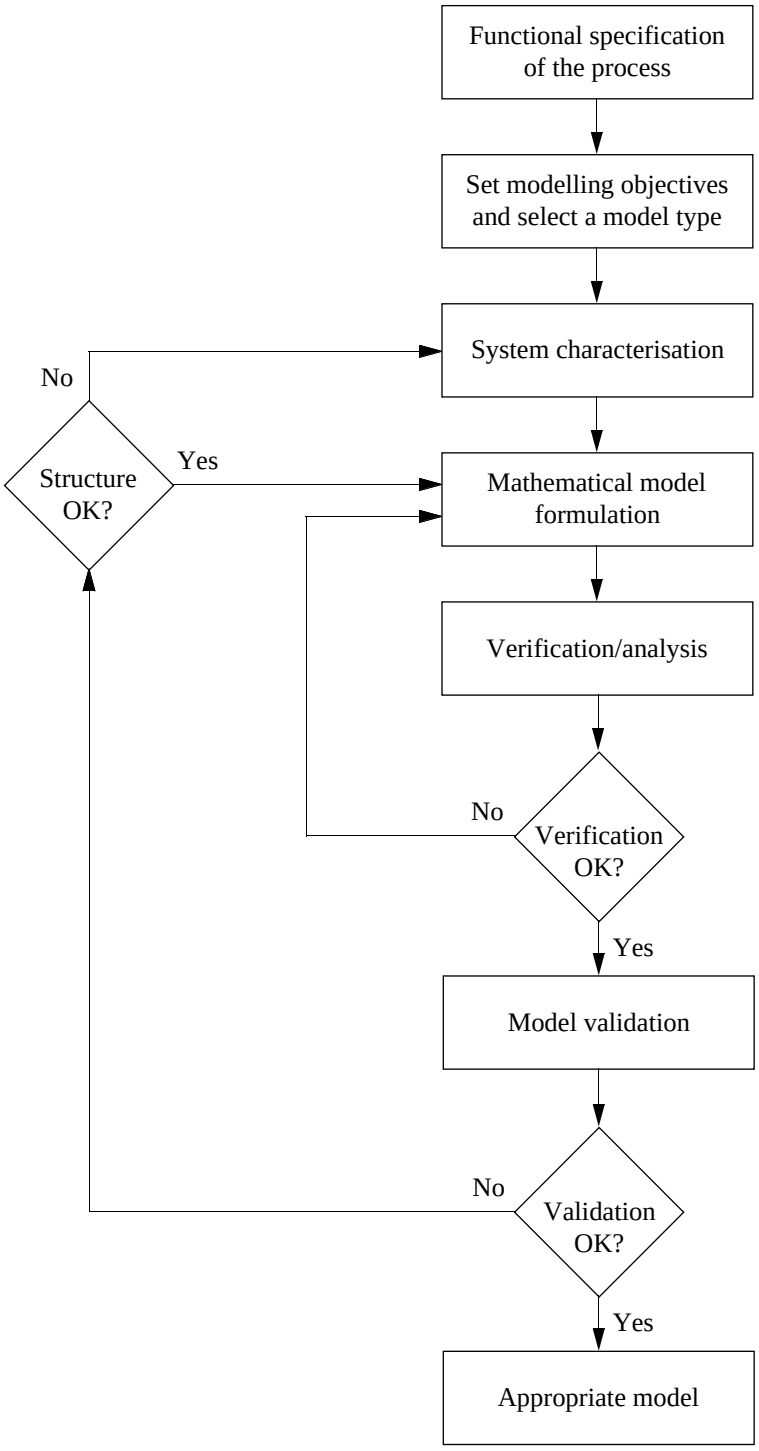


Figure 2.3 A generic modelling methodology.

System Characterization

The first step is to characterize the process. This is achieved by developing a set of axioms or descriptions of the process. These axioms can be formed by intuitively describing the process or in some manner idealizing or

approximating the process behaviour. They could be a result of some prior knowledge of the process or based on postulated mechanisms within the process. The end result of this step is a set of relationships that the modeller expects will adequately describe the process to the accuracy which is required.

The description invariably is incomplete so that a number of different possible models and processes could satisfy them. The axioms may also contain information which is incorrectly characterized from the process (cf. Figure 2.1). Identification of all important characteristics may, however, be impossible due to the complexity of the process and the limited expertise and knowledge of the modeller. Therefore, the task of the modeller is to strive for a good characterization, i.e., identify many of the most important characteristics whilst limiting incorrect and non-essential identifications.

Model Construction

In this step of the model building process, the axioms developed during the system characterisation are refined into mathematical relationships. This requires a quantitative assessment of the physical phenomena judged to be important during the characterisation step and involves determining a mathematical structure for the model and assigning parameters to the model.

The distinction between conservative and constitutive should be emphasized. Conservation relationships are fundamental physical laws whereas constitutive relationships are postulated mechanisms usually based on empirical evidence. In using constitutive relations, the modeller is incorporating more questionable but necessary information into the model.

In principle, the conservation laws can be applied at every point of a process. Mass, energy and momentum profiles are obtained throughout the process along all spatial directions. This is the earlier discussed distributed formulation and can be written as a general set of partial differential equations. The general formulation of these equations lies beyond the scope of this work. Instead the conservation law in combination with mass transfer by gravity settling is thoroughly discussed in Sections 5.2 and 5.4 for the sedimentation process and the conservation law is also discussed in Section 7.2 in connection to biofilm process modelling. The general set of equations describe the physical behaviour at any point of the system and are created by considering the mass, energy and momentum transfer from an infinitesimal control volume within the process. Transfers between

control volumes are governed by physical phenomena. Table 2.1 presents some of the more common constitutive mechanisms for mass, energy or momentum transfer experienced in process systems (Bird *et al.*, 1976).

Transfer mechanisms	Mass transfer	Energy transfer	Momentum transfer
Transport by molecular diffusion	Fick's law of diffusion	Fourier's law of conduction	Newton's law of viscosity
Transport in solids	Not available	Conduction	Not available
Laminar transport in fluids	Convection	Convection	Convection
Turbulent transport in fluids	Eddy diffusivity	Eddy thermal conductivity	Eddy viscosity
Interphase transport	Mass transfer coefficients	Conduction, convection, radiation	Friction factors
Generation/ accumulation	Chemical-biological reactions	Chemical-biological reactions	Not available

Table 2.1 Constitutive mechanisms for transport phenomena.

As discussed earlier, this type of partial differential equation system is almost impossible to solve and, therefore, it is more common to transform the model into a lumped parameter approximation. But even if the transport phenomena are modelled accurately, it still remains to include the physical, chemical and biological mechanisms that affect the process.

In a completely different direction, an arbitrary structure for the model can be chosen. Model parameters are then varied to achieve agreement between the process and the model. This type of model originates from the field of system identification and is only suitable for model realisation from observed process data. Different model structures and numerous methods for identification of such models are given in Ljung (1987) and Söderström and Stoica (1989), together with suitable experimental methods used to increase the amount of information in a certain data set. However, it should be noted that methods for parameter estimation are becoming increasingly important also for modellers working with deterministic models of complex systems. This topic will be discussed in Section 2.4.

Model Verification and Validation

Model verification and validation may be regarded as part of the model construction methodology or a concluding step (cf. Figure 2.2). However, it is clear that model construction and model validation are closely related and require an iterative procedure according to Figure 2.3.

The constructed model must be tested by simulations. In the first simulation the model behaviour is analysed/verified. This amounts to checking that the simulated responses are consistent with the axioms proposed during the process characterisation and the mathematical structure used in the model. It also means debugging the model code and ensuring that the simulated responses appear feasible.

The validation step involves checking that the model responses generated during the model analysis agree with that obtained from the true process. This is the ultimate check on the success of model building. A variety of methods are available for this purpose and some of them will be discussed in Sections 2.3-2.5.

If the process model fails the validation step, then it has to be reformulated and the verification/validation analysis repeated. If the mismatch between process and model is severe, then a new characterisation is required. A new model will result. When the mismatch is small, it may be possible to tune the model parameters to achieve a satisfactory agreement. In this case no new characterisation of the process will be required.

2.3 Model Structure Evaluation

In this and the following two sections we will discuss aspects on model structure, system identification and parameter estimation, and model validation. These areas can hardly be separated and should therefore be regarded as an entity when a model is evaluated. The discussion will primarily be focused on evaluation of internal, deterministic models (see Section 2.2) as this type of model is used throughout the work. We will also limit the discussion to models which are used in WWT applications, although many methods and procedures are naturally applicable to any type of model. Much of the inspiration for these three sections is due to Reichert (1994b).

Why Are Simple Models Needed?

As discussed before, a model is nothing more than a mathematical abstraction of a real process. The equation or set of equations that comprise the model are at best an approximation of the true process. Hence, the model cannot incorporate all of the features, both macroscopic and microscopic, of the real system. The engineer normally must seek a compromise involving the cost of obtaining the model, that is, the time and effort required to obtain and validate it and the expected benefits to be derived from its use. The ultimate application and purpose of the model finally determines how accurate it needs to be.

In general, modelling is still much of an art. The modeller must bring a significant level of creativity to the task, namely to make a set of simplifying assumptions that result in a realistic model. An 'optimal' model incorporates all of the important dynamic effects, is no more complicated in its structure than necessary, and keeps the number of equations and parameters at a reasonable level. The failure to choose an appropriate set of simplifying assumptions invariably leads to either a rigorous but overly complicated model or models that are overly simplistic. Both extremes should naturally be avoided.

For example, activated sludge models are often derived from simpler unit operations and then combined into larger plant models. Consequently, the model parameter values may not be the same. Moreover, several parameter combinations can often explain the same dynamical behaviour. This is further accentuated when the influent wastewater composition is taken into consideration; the consequence of a change in its characteristics can quite often be explained by kinetic parameter changes.

Even if a major problem concerning models for wastewater treatment has to do with the complex model structure and the large number of states and parameters to be identified, instrumentation problems amplify these difficulties (see Section 3.4). Available on-line sensors and laboratory procedures are usually not adequate to validate the details of a complex model. Furthermore, for a reliable identification result, the operation has to be perturbed (or purposefully disturbed) in such a way that all the interesting dynamical modes of the process are excited. This creates a demand not only in amplitude but also in the time frame of the disturbances.

It is practically impossible to develop a model for the activated sludge process, which is reliable on a microscopic level. Though available models are quite complex they are still greatly simplifying the representation of many species of organisms. As the microbial population changes this needs to be reflected in changing kinetic parameters and even by adding new state variables. For example, filamentous organisms ought to be represented during many operating conditions. On the other hand, quite simple models can be used effectively if the key model parameters are properly identified to the current operating conditions. This is especially important if the process is time variant.

A simplified model does not provide a fully explanatory model for every physical reaction. Several parts of the process are often lumped together in order to reduce the complexity. For a non-expert, the intuitive understanding of the process is, however, often enhanced by such models (if its basic structure is mechanistic). Furthermore, in many cases the model output needs only to be qualitatively significant, for example, show trends and whether a variable is increasing or decreasing, without providing exact quantitative results. The relative change of model parameters may also provide useful information. This may allow effective use of simplified models for highly complex processes.

The activated sludge process is suited for a hierarchical control structure based on several simple models. The process can in a natural way be divided into unit operations – aerobic reactor, anaerobic reactor, anoxic reactor, settler, sludge digester, etc. It can also be modularized based on different time constants of the processes (see Section 3.4). Each single model would be used to control and predict the behaviour of its specific area in some optimal way but would also be synchronized with a high level control system which optimizes the performance of the entire plant according to preset criteria, which are often contradictory. The inclusion of a knowledge-based system at the top level to allow for logical reasoning, diagnosis, and decision support would further enhance the capabilities. In Figure 2.4, a schematic view of such a hierarchical control system is demonstrated. A recent example of a full-scale implementation of an on-line hierarchical control system for the activated sludge process is presented in Nam *et al.*, (1996).

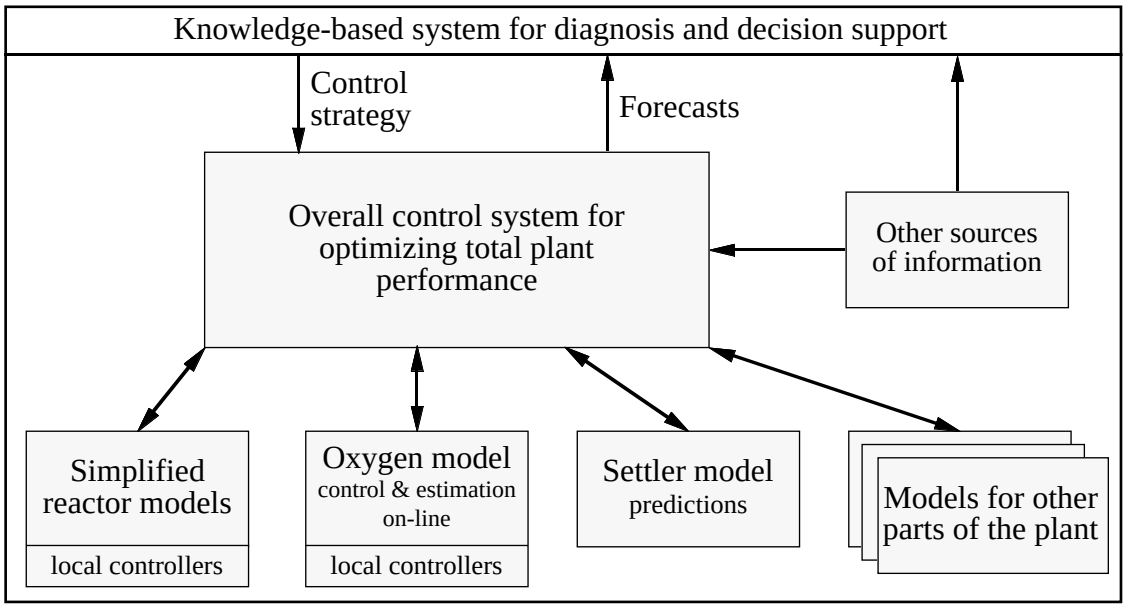


Figure 2.4 Hierarchical control structure of the WWT process.

Simplified models, like the ones aimed at in Chapter 4, should consequently not be evaluated and judged separately but be considered in a broader perspective and in the context of a full-scale hierarchical control structure. This type of distributed automation has been successfully applied to many complex industrial applications, for example, chemical, paper, and pulp processes. A similar approach could probably be applied to WWT plants as well.

If the main purpose of a model is control, the need for simplicity is evident. Due to the internal structure of a closed loop system, a reasonably small error will automatically be compensated for. As the practical control possibilities for the activated sludge process are quite limited, it is even more important to use the ones available to their full extent. Due to the large time constants and the difficulties in detecting problems early, traditional control strategies based on feedback are probably not sufficient. Methods based on feed-forward, predictive, or adaptive control appear to be better suited for this purpose. However, all these methods require a process model which is relatively simple in its structure, robust, uniquely identifiable, and possible to update on-line as the operational conditions change.

A model for operation and control has to be sufficiently complex to describe the major phenomena taking place but still so simple that its parameters can be updated while the plant is running normally, either by taking advantage of the natural disturbances of the process or by

introducing small deliberate perturbations. The need for highly complex models is recognized for design purposes so the operational model has to be considered as a special case, either for certain operational levels or for particular time scales.

The Basic Problem

For most physical and many chemical applications, the *a priori* knowledge is of such high quality that the system framework and most of the model structure can be deduced from it. The modelling methodology developed for these systems is adequate to estimate the parameters and solve the minor uncertainties in the model structure by using final validation experiments and eventually iterating a few times through the model building procedure (see Figure 2.3).

The behaviour of biotechnological processes occurring in a bioreactor has a complexity unparalleled in the chemical industry (Van Impe, 1993). The number of reactions and species that are involved in the system may be very large. An accurate description of such complex systems can therefore result in highly complex models, which may not be very useful from a control engineering viewpoint. The inherent characteristics of bio-processes, i.e., their non-linearity and non-stationarity, in combination with with the lack of adequate measuring techniques, make it necessary to modify the modelling methodology (Vansteenkiste and Spriet, 1982). More emphasis must be given to inductive reasoning to infer a larger part of the model structure from the scarce (or harder to obtain) experimental data. Consequently, structure characterisation methods (to infer the level of model complexity and the functional relationships between variables) become a more important tool, because the chance of obtaining an invalid model is much larger and, hence, the number of model building iterations may increase substantially (Vanrolleghem, 1994). The data scarcity also induce an important problem for parameter estimation. Identifiability of model parameters, i.e., the possibility to give a unique value to each parameter of a mathematical model, is a general concern in current WWT modelling efforts. The last problem is naturally more pronounced in on-line identification because one is relying much more on real-time information to perform the parameter estimation whereas off-line model calibration can take more advantage of the off-line data.

The three main sources of information that contribute to the model building process are:

- *a priori* knowledge;
- experimental data;
- modelling objectives.

It must also be realised that model structures with unidentifiable parameters are useless, and although *a priori* identifiability analysis is very useful, the estimation of accuracy and correlation of estimated parameters ultimately quantifies parameter identifiability.

Evaluation Criteria

Model structure evaluation consists of finding adequate model structures and of comparing their quality. In the text below we will assume that a number of different models are already available and discuss how to select the best one. The most important criterion for the comparison of models is that the deviations between measurements and model calculations should be as small as possible. This criterion cannot be used alone, because it favours the use of complex models with many parameters which are difficult to identify uniquely. For this reason, this criterion has to be complemented by a criterion of 'parsimony' leading to a preference for simple model structures.

The three most important techniques for deciding between competing model structures are:

- graphical or statistical searching for systematic deviations between calculations and measurements (i.e., the residuals);
- quantitative measures of model adequacy;
- recursive parameter estimation.

In most cases, graphical comparisons clearly show the existence or absence of systematic deviations between calculations and measurements. Such deviations can also be detected with the aid of statistical analyses such as residual plots, distribution or correlation tests of residuals, etc. The advantage of such statistical measures over graphical methods is mainly the fact that they facilitate the partial automation of model structure evaluation.

It is evident that a quantitative measure of the differences between calculated and measured values is an important criterion for the adequacy of a model. However, in order to avoid the above problem of favouring more complex models, which in turn will lead to problems in uniquely identi-

fying the model parameters, an extra criterion to promote simplicity is necessary. We can summarize the two criteria as follows (Spriet, 1985):

- quality of fit – the model structure should be able to represent the measured data in a proper manner;
- parsimony – the model structure should be as simple as possible compatible with the first criterion.

The first quantitative measure to include both these aspects was introduced in a series of papers by Akaike (see e.g., Akaike, 1981). It is a generalization of the maximum likelihood principle for parameter estimation of a given model (see Section 2.4). The Akaike information criterion (AIC) can be formulated as

$$\mathrm{AIC}(\boldsymbol{y}_{meas}) = -2\log\left(\max_{\theta}\left(\mathrm{L}(\boldsymbol{y}_{meas},\ \theta)\right)\right) + 2m \qquad (2.1)$$

where θ is the array of model parameters, $\boldsymbol{y}_{meas}$ the array of measured values, L the likelihood function of the model (probability density of the model at the measured values) and m the number of parameters. Equation (2.1) formulates the trade-off to be made between quality of fit and model complexity. The first term of the right-hand side increases with decreasing quality of fit, while the second term increases with increasing model complexity measured in terms of the number of parameters. The criterion is easy to use since the first term corresponds to the usual criterion for maximum likelihood parameter estimation, i.e., for all competing model structures, a maximum likelihood parameter estimation is performed. Then the model with the smallest value of AIC is chosen as the best model. The criterion has, however, been criticized because of the heuristic arguments for its justification and because it only accounts for model complexity in terms of the number of parameters. An alternative information criterion (B) was proposed by Schwarz (1978) as

$$\mathrm{BIC}(\boldsymbol{y}_{meas}) = -2\log\left(\max_{\theta}\left(\mathrm{L}(\boldsymbol{y}_{meas},\ \theta)\right)\right) + m\log(n) \qquad (2.2)$$

where n is the number of measured data points. Criterion (2.2) favours more simple models than does criterion (2.1) if 8 or more measured data points are used. For a large number of data points, the two criteria differ significantly from one another. Several other suggestions for such criteria exist, see e.g., Söderström and Stoica (1989) and Vanrolleghem (1994), and it is not clear which measure performs best, especially when con-

fronted with the general problem of discriminating between non-linear models such as those used for modelling WWT processes (Spriet, 1985).

In most cases, model parameters are assumed to be constant. In contrast to the usual parameter estimation, which compares the result of calculations performed using different – but fixed – parameters, recursive parameter estimation of time series data allow the parameters to change slowly in time (e.g., Ljung and Söderström, 1983 and Young, 1984). Recursive parameter estimation thus tests the model hypothesis of constant parameters and can therefore also be regarded as a *model* identification method (see Section 2.4).

Model Reduction

The discussion so far has dealt with the situation when a number of different candidate models are available and we want to select the best one. Another situation occurs when we have a reference model (a state-of-the-art model) available and we want to simplify this model for some specific purpose. For example, the reference model may be a complex model developed for providing deep insight in the behaviour of a process and we may want to reduce this model into a simpler one, which is more suited for control purposes. The reference model may also cover a wide range of time-scales whereas we may be interested in isolating the part of the model that describes the fast dynamics only. Some of latest modelling tools such as ASCEND (Piela *et al.*, 1991) and MODEL.LA (Stephanopolous *et al.*, 1990) have the capability to construct very detailed deterministic models, which are suited for reference model construction and later model reduction.

There are two principles for accomplishing this type of model reduction, namely

- intuitive model reduction;
- mathematical techniques for model reduction.

Intuitive model reduction implies that the modeller uses his knowledge and experience of the true process and its dynamics in combination with the defined purpose of the new model to infer a simplified model. Such a simplification may affect both the model structure and the functional relationships within the model. This is still the most common way used to simplify models (at least for complex non-linear systems) and is the

method that has been applied to a reference activated sludge model in Chapter 4. The simplified model is then validated in a traditional way in order to confirm or falsify the reduced model. A major disadvantage is that it is difficult to prove that the applied simplifications are the most relevant ones and another modeller may have a very different view on the relevance of the assumptions made.

A wide variety of mathematical methods for model reduction are available. The area has been the focus of much research particularly in the field of control. In Robertson (1992) approximately 20 different methods are described and evaluated according to four criteria:

- simplicity – the method should be easy to apply regardless of the model type or level of complexity;
- power of reduction – the method should be powerful so that the greatest reduction possible is achieved;
- error estimation – the method should possess a meaningful error estimate to allow for a quantitative assessment of the reduced model accuracy;
- structure preservation – the method should produce models that maintain the physical significance of the state variables.

The evaluation is restricted to dynamic-deterministic models classified into three different types:

- linear state-space models;
- non-linear state-space models;
- frequency domain models.

In this short review of Robertson's work we will not describe any details about the techniques but only discuss some major advantages and disadvantages of such methods. For details and references to the different techniques we refer to Robertson (1992).

The majority of model reduction techniques have been developed for linear models. Despite this, most linear model reduction methods are inappropriate because they are invariably structure destroying. Many reduction techniques such as *aggregation* or *principal component analysis* frequently make use of linear coordinate transformations to reduce the model dimensions. These transformations invariably lead to an alternative coordinate system in which state variables have no physical significance. A novel approach for reduction of linear models, is the use of *structural dominance*

concepts. Measures are developed which are indicative of the strength of the coupling between model components. By neglecting weak couplings, a reduced order model can be formed.

A large number of methods also exist for frequency domain models but most of them suffer from two major limitations. Firstly, they are all in practice limited in reducing only medium to small size models and do not work well for highly complex models. The other main problem is that when a deterministic state-space model is transformed into the Laplace domain, this transformation is in itself structure destroying. No physical significance can therefore be attached to the resulting frequency domain models.

The area of non-linear model reduction is still in its infancy with respect to the four selection criteria. Many of the methods resort to the use of non-linear transformations to turn the non-linear problem into a linear one (e.g., exact linearisation). This strategy, whilst destroying the original structure of the reference model will also lead to the limitations imposed on linear methods. Another serious inadequacy with the surveyed non-linear methods is that few methods allow for global reductions. Many techniques are limited to local reductions near fixed points in the solution manifolds (e.g., bifurcations, stationary points).

The conclusion of Robertson is that the combined non-linear methods of singular and regular perturbations (see e.g., Kokotovic *et al.*, 1976; Kevorkian and Cole, 1980; Jamshidi, 1983; Martinez and Drozdowicz, 1989) are the most appropriate for reference model reduction. These two methods are simple to apply, are both structure preserving and generate low-order reduced models. A reduction error estimate can also be calculated based on the linearized model before and after the reduction and the methods are applicable to non-linear process models (which is practically always the case). A small drawback when using these methods is that the model prior to reduction must be written in a particular format. Invariably the models encountered in the modelling of process systems do not always display this form. An identification step is therefore required to transform the reference model into the explicit form required for reduction.

The entire set of evaluated model reduction methods investigated by Robertson (1992) is provided in Table 2.2.

Model reduction methods		
Linear models	Frequence domain mod.	Non-linear models
Perturbations	Moments matching	Perturbations
Regular	Continued fraction	Regular
Singular	Stability equation method	Singular
Aggregation	Pade/Routh approximation	Lindstedt's method
Exact	Chebychev/Darlington func.	Centre manifolds
Model	Laguerre functions	Normal forms
Approximate	Shifted Legendre polynom.	Two variable expan.
Continued fraction		Averaging
Chained		Lie transforms
Error minimization		Liapunov-Schmidt red.
Principal comp. analysis		Exact linearisation
Balancing		Approx. linearisation
Quasi-Kalman decomp.		Integral manifolds
Impulse response match.		
Markov param. match.		

Table 2.2 Model reduction methods (Robertson, 1992).

2.4 System Identification

The task of system identification consists of making optimal use of the available information in order to find the most adequate model. The model structure evaluation discussed in the previous section is a part of this process as well as the final model confirmation or falsification (validation). In this section we will discuss two equally important subjects: model identifiability analysis and parameter estimation techniques. We also include the topic of state variable estimation (reconstruction) in this section.

Theoretical Identifiability

Theoretical identifiability analysis treats the problem of the uniqueness of the determination of model parameters resulting from a given input-output experiment with perfect data (acquired by simulation). Such an analysis can and should be performed prior to any real experimental investigations in order to investigate if the model structure is theoretically sound. For a given model and ideal measurements a parameter is called

- uniquely (globally) identifiable – if there exists a unique solution for the parameter;
- locally identifiable – if there exists a finite number of parameter values;
- unidentifiable – if there exists an infinite number of parameter values,

which make the model (exactly) reproduce the measurements. In the case of unidentifiable parameters, there exists different sets of parameter values which will lead to (exactly) the same model behaviour. In this case, it is important to investigate which combinations of parameters that are identifiable (e.g., if only the product of two parameters is used in a model, the parameters are not separately identifiable but the product may very well be identifiable).

Whereas there exist several methods for identifiability analysis of linear models, there is only one universal technique applicable to non-linear systems. This technique (Pohjanpalo, 1978) is based on a Taylor series expansion of the measured variables with respect to time. The coefficients of the power series contain the model parameters and the decision whether these parameters can be determined from the Taylor series coefficients is reduced to an algebraic problem. If the algebraic equations can be solved for the parameters then these are identifiable. An example of such an analysis is presented in Section 2.6. However, practical limitations make this method troublesome to apply for more complex models. The reason for this is that for non-linear systems there is no theoretical upper limit to the number of model differentiations which may provide new information. The use of computer algebra programs can partially improve the situation, but the fundamental problem still remains.

Another approach is to linearize the model around a suitable operating point (if such a point exists) and apply one of the many methods for identifiability analysis of linear systems. However, fewer identifiable parameter combinations than for the full non-linear model may result and

parameters of the non-linear model may not even appear in the linearized one. As a consequence, non-identifiability of a linearized system does not necessary indicate that the original non-linear model is unidentifiable.

An investigation of *local* identifiability is not very complicated. This can be done by examining the rank of the Jacobian of the model (Godfrey and DiStefano, 1985).

Practical Identifiability

Practical identifiability analysis treats the problem of parameter identification in the presence of noisy measurements. It is evident that practical identifiability of parameters requires their theoretical identifiability. Since practical identifiability is mainly a problem of the estimation of parameter uncertainty, it is not an objective characteristic of a model for a given experimental situation, but depends instead on the values of the measured data and on the desired accuracy of the parameters. Lacking practical identifiability means that unique sets of parameters can seldom be obtained, parameters estimated from data obtained during apparently similar conditions show considerable variations, and that the estimation methods show poor convergence properties. Parameter estimation results where the identified parameter values vary depending on the initial values is also an indication to proceed with care. Altogether, it is often easy to obtain sets of parameters which provide a good model fit but since these parameters may be far from the true ones, situations where they are given an exact physical/biological/chemical interpretation should be avoided.

Practical identifiability problems often arise as a result of the following factors:

- unsuitable model structure;
- poor sampling strategies, lack of reliable sensors and troublesome noise conditions;
- poor system ‘excitement’ during the identification experiment;
- unsuitable identification algorithms.

Improvement of the practical identifiability may be obtained by:

- changing the model structure (use reduced order models);
- improving the experimental design, available information and noise characteristics;
- model reparametrization (use combinations of model parameters).

There are two main techniques for practical identifiability analysis: sensitivity analysis (linear or non-linear) and parameter covariance estimation.

Linear sensitivity analysis consists of calculating a linear approximation to the change in a variable caused by a given change in a parameter. Depending on whether absolute or relative measures of the variable and of the parameter are used, the following four sensitivity function can be distinguished:

$$\delta_{a,a} = \frac{\partial x}{\partial \theta} \quad \text{(a)} \qquad \delta_{r,a} = \frac{1}{x}\frac{\partial x}{\partial \theta} \quad \text{(b)}$$
$$\delta_{a,r} = \theta\frac{\partial x}{\partial \theta} \quad \text{(c)} \qquad \delta_{r,r} = \frac{\theta}{x}\frac{\partial x}{\partial \theta} \quad \text{(d)} \tag{2.3}$$

Sensitivity function (2.3a) gives the absolute change in variable x per unit change in parameter θ, (2.3b) the relative change in x per unit change in θ, (2.3c) the absolute change in x for a 100 % change in θ, and (2.3d) the relative change in x for a 100 % change in θ. The two most often used sensitivity functions are (2.3c) and (2.3d) because the units of these functions do not depend on the units of the parameter. This makes the comparison of the sensitivity of a variable to different parameters possible.

The larger the values of the sensitivity functions, the more accurately a single parameter can be identified. In the case of several parameters, the sensitivity functions of the parameters as functions of the independent variable of the measurements (e.g., time) have to be linearly independent. Otherwise, the parameters are not individually identifiable because a change in one parameter can be compensated for by changes in the other parameters. The more different the patterns of the sensitivity functions are, the better the parameters can be identified.

Non-linear sensitivity analysis is based on the calculation of the probability distribution of calculated variables from the probability distributions of the parameters. This is done with the aid of Monte Carlo simulation. It is evident that this analysis, which takes the non-linearity of the model fully into account, gives much better information than does linear sensitivity analysis. The disadvantage of this method is that it requires a large computational effort and (at least approximate) knowledge of the probability distributions of the parameter.

Mathematical analysis based on the numerical properties of the covariance matrix of the estimated parameters, observability matrix or Fisher information matrix (the inverse of the parameter estimation error covariance matrix), etc. can be used to evaluate the quality of information with reference to the estimated parameters of the model. This kind of analysis is able to detect the unidentifiable cases and can also be useful for determining sampling strategies and experimental design, see for example, Vanrolleghem (1994).

Parameter Estimation

Parameter estimation consists of determining the 'optimal' values of the parameters of a given model with the aid of measured data. Although this procedure uses a given model structure, it is not completely independent of model structure evaluation, because the model may degenerate to a simpler structure for particular values of some parameters. Since the initial state of a simulation, the boundary conditions and the external variables can also be formulated with the aid of parameters, all these parameters, together with the model parameters, can be combined to yield a single array of parameters to be estimated simultaneously using the same estimation technique.

There are four important conventional techniques which can be used for parameter estimation, see for example, Beck (1987) and Ljung (1987):

- Bayesian estimation;
- maximum likelihood estimation;
- weighted least squares estimation;
- least squares estimation.

These four methods are listed in decreasing order of the amount of information that has to be provided by the user of the method, or, equivalently, in increasing order of the number of *a priori* assumptions already included in the method. For the most complicated case of Bayesian estimation, the probability distribution of the parameters and the conditional probability distribution of the measurements for given parameter values have to be parameterized, whereas the simplest case of least squares estimation can be performed without any extrinsic information. Weighted least squares and least squares estimation are special cases of the maximum likelihood method in which the measurements are assumed to be uncorrelated and normally distributed. Practical experience have shown

that the methods above do not always suffice, since distributions of real data are never known exactly. This has lead to the development of methods for robust estimation, see for example, Birkes and Dodge (1993).

Bayesian estimation treats both measurements and model parameters as random variables. If an *a priori* probability density $p(\theta)$ for the occurrence of the parameter vector θ and the conditional probability density $p(\boldsymbol{y}_{meas}|\theta)$ of the model for measuring the values $\boldsymbol{y}_{meas}$ for given parameter values θ are known, the probability density of the parameters for given values of the measurements can be written (according to Bayes' rule) as

$$p\left(\theta|\,\boldsymbol{y}_{meas}\right)=\frac{p\left(\boldsymbol{y}_{meas}|\,\theta\right)p(\theta)}{p\left(\boldsymbol{y}_{meas}\right)} \tag{2.4}$$

Equation (2.4) does not directly specify estimates of the parameters, but it yields a complete description of the distribution of parameter values for given measurements. Additional assumptions are necessary for the choice of parameter estimates. The central idea of Bayesian estimation is to update prior information on the distribution of parameters by taking measured data into account.

In contrast to Bayesian estimation, maximum likelihood estimation treats the parameters not as random variables but as constant parameters of the distributions of the measurements. Maximum likelihood estimation consists of maximizing the so-called likelihood function, L, which is the probability density of a model for the occurrence of the measurements for given parameters. The likelihood function is a complex function which depends on the probability distribution of the measurements. If we assume these to be uncorrelated normal distributions the likelihood function is given as

$$\mathrm{L}\left(\boldsymbol{y}_{meas}|\,\theta\right)=\frac{1}{\sqrt[n]{2\pi}}\prod_{i=1}^{n}\frac{1}{\sigma_{meas,i}}\exp\left(-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{y_{meas,i}-y_i(\theta)}{\sigma_{meas,i}}\right)^2\right) \tag{2.5}$$

where $y_i(\theta)$ is the calculated value of the model corresponding to $y_{meas,i}$ using the parameters θ and $\sigma_{meas,i}$ is the (estimated) standard deviation of $y_{meas,i}$. For given measurements $\boldsymbol{y}_{meas}$, the maximum likelihood estimates $\theta\left(\boldsymbol{y}_{meas}\right)$ of the parameters are those values of θ for which the likelihood function has its maximum. Maximizing (2.5) is equivalent to minimizing the function

$$\sum_{i=1}^{n}\left(\frac{y_{meas,i}-y_i(\theta)}{\sigma_{meas,i}}\right)^2 \tag{2.6}$$

The uncertainty of the estimates may in turn be estimated from the uncertainty of the measurements by studying the covariance matrix.

The weighted least squares method is a special case of the maximum likelihood estimation. In fact, the likelihood function (2.5) is actually the function used for weighted least squares estimation, which is then maximized as discussed above.

If not even the standard deviations $\sigma_{meas,i}$ of the measurements are known, all of them are assumed to be of equal size. In this case, the expression to be minimized for the simple least squares estimation is

$$\sum_{i=1}^{n}\left(y_{meas,i}-y_i(\theta)\right)^2 \tag{2.7}$$

The conventional parameter estimation techniques described above are derived using assumptions concerning the form of the distributions of the measured variables. If real data violate these assumptions slightly, the estimation techniques may give incorrect results. The application of robust estimation techniques makes the results much less dependent on slight violations of the parameterized probability distributions by the data (e.g., non-existing data points and outliers). One of the simplest robust estimation methods is the minimization of the median of the squares of the residuals instead of their sum, or, more efficiently, changing the sums in equations (2.6) and (2.7) to include only the smaller half of the squares of the residuals (Rousseeuw and Leroy, 1987). The only disadvantage of robust estimation techniques compared with the conventional ones is the increased computation time involved.

The statistical methods above can only be applied if reasonable amounts of measured data are available. An alternative approach, which can by applied in the case of bad resolution of data or even using semi-quantitative information on system behaviour, is the Monte Carlo filtering technique. The basic idea of the method is to fix the ranges of parameters which characterize reasonable system behaviour, to perform Monte Carlo simulations with the model, and to select sets of parameters which lead to the desired behaviour. This method does not lead to a unique set of parameter values but the sets of parameter values found to be compatible with the

data can be used for predictions, which then have to be statistically evaluated.

One broad distinction between different methods for estimation are:

- off-line estimation;
- on-line estimation.

When performing off-line estimation, a complete set of data is available in time-series form. It is then straightforward to apply the estimators discussed above to this data set. The procedure consists of performing a simulation with constant parameter values over the whole time interval containing measurements, and locating the minimum of (2.6) or (2.7). The estimation becomes an optimization problem. Several different algorithms are available for rapidly solving this optimization problem, e.g., simplex methods, Newton's method, quasi-Newton methods, conjugate direction methods, Levenberg-Marquardt methods (Fletcher, 1987). The Nelder-Mead simplex method (Nelder and Mead, 1965), which will be used in this work is described in Appendix D. The on-line methods (Kalman filter, recursive instrumental variable method, etc.) give estimates recursively as the measurements are obtained and are the only alternative if the identification is going to be used in an adaptive controller or if the process is time varying (Åström and Wittenmark, 1990). In many cases the off-line methods give estimates with higher precision and are more reliable, for instance in terms of convergence. However, it is often possible to reformulate an off-line method into a recursive equivalent.

State Estimation

It is often unrealistic to assume that all the internal state variables of a system and the disturbances can be measured. If a mathematical model of the system is available, the states can often be computed from measured inputs and outputs – *state estimation.*

The non-linear reduced order models that will be presented in Chapter 4 can be schematically described in the format

$$\begin{cases} \dfrac{d\boldsymbol{x}}{dt} = \boldsymbol{f}(\boldsymbol{x},t) + \boldsymbol{g}(\boldsymbol{u},t) \\ \boldsymbol{y} = \boldsymbol{h}(\boldsymbol{x},t) \end{cases} \tag{2.8}$$

where $\boldsymbol{x}$ is the state vector, $\boldsymbol{y}$ represents the measurable variables and $\boldsymbol{u}$ is the model input. $\boldsymbol{f}$, $\boldsymbol{g}$ and $\boldsymbol{h}$ are general functions that describe the relationships between the variables. The same system can also be described by the time discrete representation

$$\begin{cases} \boldsymbol{x}(t_{k+1}) = \boldsymbol{F}(\boldsymbol{x}(t_k), t_k) + \boldsymbol{G}(\boldsymbol{u}(t_k), t_k) \\ \boldsymbol{y}(t_k) = \boldsymbol{H}(\boldsymbol{x}(t_k), t_k) \end{cases} \tag{2.9}$$

where t_k and t_{k+1} are consecutive measurement times. Alternatively, it may be written on the form (2.10), when it is linearized for *every* time step. Note that the Φ, Γ and $\boldsymbol{C}$ matrices may depend on $\boldsymbol{x}$ and $\boldsymbol{u}$ as well as time.

$$\begin{cases} \boldsymbol{x}(t_{k+1}) = \Phi\boldsymbol{x}(t_k) + \Gamma\boldsymbol{u}(t_k) \\ \boldsymbol{y}(t_k) = \boldsymbol{C}\boldsymbol{x}(t_k) \end{cases} \tag{2.10}$$

If the model (2.10) is fully observable, the complete state vector can be directly calculated from the measured inputs and outputs. One disadvantage of such a method is that it may be sensitive to disturbances. But more important, the result depends on the model being sufficiently accurate. In models for wastewater treatment processes, several parameters (the Φ and Γ matrices) are time variant and it is essential to keep track of their values as conditions change as a function of time. Therefore, a direct method is not sufficient. However, it is possible to use the dynamic model to reconstruct the state variables as well as performing parameter estimation, simultaneously.

The method of reconstruction is based on the assumption that the true state $\boldsymbol{x}$ can be approximated by the state $\hat{\boldsymbol{x}}$ of the model

$$\hat{\boldsymbol{x}}(t_{k+1}) = \Phi\hat{\boldsymbol{x}}(t_k) + \Gamma\boldsymbol{u}(t_k) \tag{2.11}$$

which has the same input $\boldsymbol{u}$ as system (2.10). If the model (2.11) is perfect in the sense that the parameters are identical to those of system (2.10) and if the initial conditions of (2.10) and (2.11) are the same, then the state $\hat{\boldsymbol{x}}$ will be identical to the state $\boldsymbol{x}$ of the true system. If the initial conditions are different, then $\hat{\boldsymbol{x}}$ will converge to $\boldsymbol{x}$ only if system (2.10) is asymptotically stable (Åström and Wittenmark, 1990).

The reconstruction in (2.11) does, however, not make use of the measured output $\boldsymbol{y}$. Therefore, the method can be improved by introducing the

difference between the measured and estimated output as a feedback to obtain

$$\hat{\boldsymbol{x}}\left(t_{k+1}|t_k\right) = \boldsymbol{\Phi}\hat{\boldsymbol{x}}\left(t_k|t_{k-1}\right) + \boldsymbol{\Gamma}\boldsymbol{u}\left(t_k\right) + \boldsymbol{K}\left(\boldsymbol{y}\left(t_k\right) - \boldsymbol{C}\hat{\boldsymbol{x}}\left(t_k|t_{k-1}\right)\right) \quad (2.12)$$

The system (2.12) is called an observer and exists in many variations depending on how the $\boldsymbol{K}$ matrix is chosen. The notation $\hat{\boldsymbol{x}}\left(t_{k+1}|t_k\right)$ is used to indicate that it is an estimate of $\boldsymbol{x}(t_{k+1})$ based on measurements available at time t_k. In order to use this method for *simultaneous* state and parameter estimation it is has to be slightly modified and $\hat{\boldsymbol{x}}$ will be interpreted as a generalized state vector which contains not only the state variables but also the unknown parameters to be estimated. In this work we will use one such method, the extended Kalman filter, which is described in Appendix E.

2.5 Model Validation

The last step when evaluating a model consists of testing the model with independent data sets (data sets not used for model identification or calibration). It is, however, important to note that the absence of significant deviations between the model calculations and the measurements only proves that the model assumptions are compatible with the system behaviour. Thus in a strict sense, model validation is impossible, see Reichert (1994b). Because significant deviations between model calculations and measurements disprove a model, the goal of model confirmation should be to attempt to refute the model. It is important that the performed tests put the model in jeopardy and a goodness of fit is not a sufficient condition for model acceptance. Then, the confidence in the model assumptions increases as the model passes more and more severe tests. It is very desirable to quantify the result of such hypothesis tests with the aid of statistical criteria (Thomann, 1982). A related problem for model validation is that different postulated mechanisms may lead to the same mathematical description, thus making it impossible to verify certain mechanisms by traditional means. Situations may also arise where two or more models based on partly contradictory hypotheses explain experimental results equally well (Holmberg, 1981).

Just as it is important to investigate how well and under what specific conditions a model realistically mimics the true system behaviour, it is

equally important to consider the objective for which the model was developed. A model may, for example, be valid for describing

- steady-state behaviour (no transients);
- various types of dynamic behaviour and time horizons;
- certain operating conditions;
- certain input (amplitude, variability, frequency) conditions;
- specific noise distributions;
- qualitative differences.

This situation is especially true for highly complex processes such as wastewater treatment processes, where no such thing as *the true model* exists. Different models have different advantages, drawbacks and objectives. The final validation of a model can only be achieved by using it in practice for its intended purpose and critically evaluate the results over a longer period of time (Söderström and Stoica, 1989).

Since most models describing WWT processes are mechanistic, it is actually not sufficient that only the output of the model is validated against the true process. Many model parameters have a direct physical interpretation, which implies that the validation has to include an evaluation of those parameters when compared with the actual process parameters. However, mechanistic models nearly always include empirical qualities and it is unlikely that any biological or biochemical system has ever been described exactly by a theoretical model. Thus, a 'true' mechanistic model is one which describes the mechanisms well enough to assist understanding and to allow useful – but not exact – extrapolation. Clearly, the classification of biological models as empirical or mechanistic depends on what is expected of the ultimate model.

Although no true model validation may be possible it is necessary to investigate the model behaviour to the best of our abilities. We can distinguish between three general types of methods:

- use of plots and common sense;
- statistical methods based on the prediction errors;
- investigation of the underlying model assumptions.

Statistical validation is performed by calculating the residual difference between simulated and real process responses. For the model and the process to agree, the residuals should be small and devoid of any information. Often we can base the validation on four different criteria

(Söderström and Stoica, 1989). These criteria (not always applicable) state that for an adequate model, the residuals

- are zero mean white noise;
- have a symmetric distribution;
- are independent of past inputs;
- are independent of all inputs.

From these criteria several statistical tests can be constructed. In order to check if the residuals appear as white noise, the auto-correlation of the residuals should be zero or, equivalently, the power spectrum of the residuals should display no peaks. By testing the changes of sign of the residuals we can investigate the distribution of the residuals and cross-correlation tests can be performed to check the independence of the residuals on the inputs. Further details of these residual analysis methods can be found in Ljung (1987) and Söderström and Stoica (1989). Other methods for model validation are described in Sargent (1982).

From the discussion of the importance of model structure evaluation, identifiability, and validation in the last three sections it may be appropriate to extend the model validation part of the general modelling methodology shown in Figure 2.3. One possible approach is schematically outlined in Figure 2.5.

When an adequate model has been developed and calibrated for the behaviour of the system under consideration, an estimate of the uncertainty of the predictions is important. Especially if the model is to be used for management decisions it is important not to accept model results without questioning the model assumptions. Three main sources of uncertainty are distinguished (Beck, 1991):

- uncertainty in model structure;
- uncertainty in parameter values and initial state;
- uncertainty associated with external variables.

However, it is beyond the scope of this work to discuss any details of such an uncertainty analysis, instead we refer to Beck (1987; 1991) and Reichert (1994b).

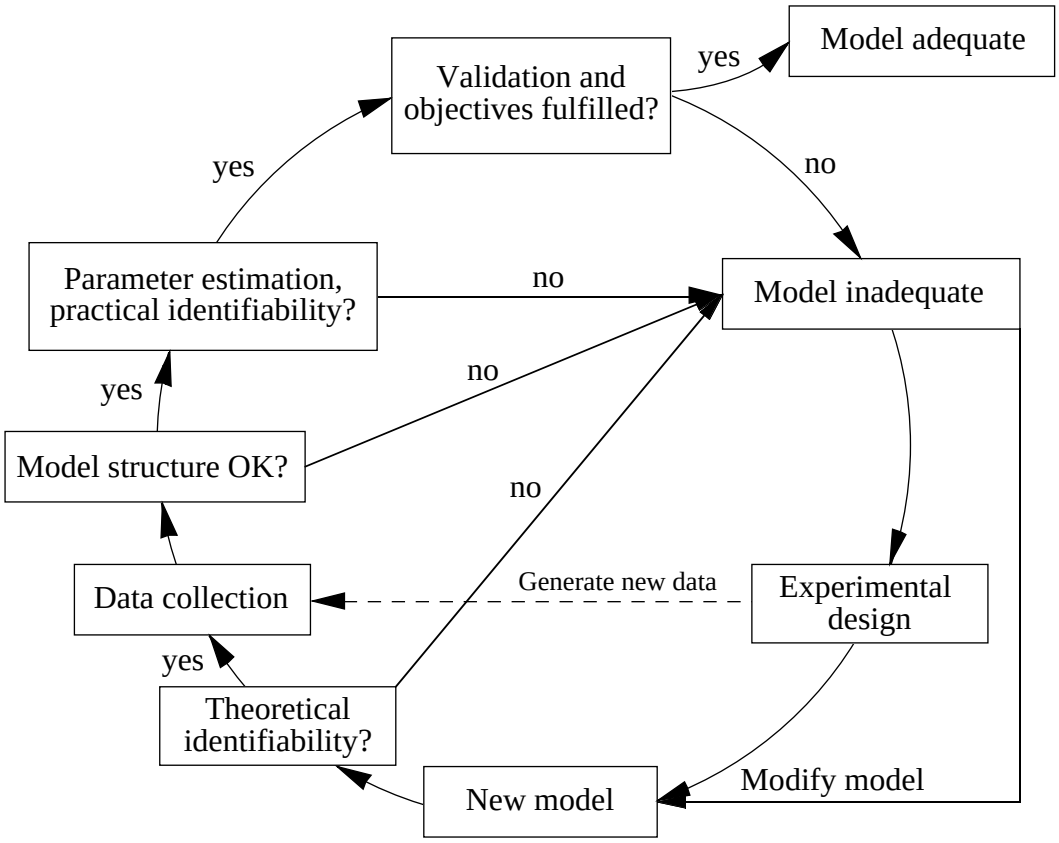


Figure 2.5 Model validation methodology.

2.6 A Small Example

In order to exemplify a few of the problems discussed in the previous sections, especially the ones dealing with identifiability, we present a small example from the field of bacterial growth and decay in this section.

One of the simplest possible models to describe growth of a single organism on a single substrate in a batch reactor with no other growth limitations, can be formulated as

$$\begin{cases} \dfrac{dX}{dt} = \mu(S)X - bX \\ \dfrac{dS}{dt} = -\dfrac{1}{Y}\mu(S)X \end{cases} \tag{2.13}$$

where: X = concentration of microorganisms [mg/l];
S = concentration of growth-limiting substrate [mg/l];
$\mu(S)$ = specific growth rate [day^{-1}];
b = decay rate [day^{-1}];
Y = yield factor [g cell COD formed (g COD oxidized)$^{-1}$].

We define $\mu(S)$ according to the famous Monod growth-rate expression (Monod, 1942), i.e.,

$$\mu(S) = \frac{\hat{\mu} S}{K_S + S} \tag{2.14}$$

where $\hat{\mu}$ is the maximum specific growth rate and K_S is the so called substrate half-saturation coefficient. We assume that both X and S are possible to measure directly and that only X and S are time variant, i.e., all model parameters are constant during the experiment. The measured data are further considered to be free from noise and continuously available. The model is obviously non-linear but not very complex.

Question: Can all four model parameters be uniquely determined from perfect data, i.e., is the system globally identifiable?

In Section 2.4 various methods for investigating identifiability were discussed. It was stated that there is only one universal technique applicable to non-linear systems, i.e., the *Taylor series expansion of observations* method. This analytical technique is applicable to small systems, such as (2.13), although it is often not practically feasible to use the method for complex systems.

We define the following nomenclature:

$$\begin{cases} X_0 = X(0) \\ S_0 = S(0) \end{cases} \tag{2.15}$$

$$\begin{cases} X'(0) = \dfrac{\hat{\mu} S_0}{K_S + S_0} X_0 - b X_0 = X_1 \\ S'(0) = -\dfrac{1}{Y} \dfrac{\hat{\mu} S_0}{K_S + S_0} X_0 = S_1 \end{cases} \tag{2.16}$$

$$\mu = \frac{\hat{\mu} S_0}{K_S + S_0} \tag{2.17}$$

Equation (2.16) can now be formulated in a less complex way as

$$\begin{cases} X_1 = (\mu - b) X_0 \\ S_1 = -\dfrac{\mu}{Y} X_0 \end{cases} \tag{2.18}$$

If equation (2.16) is differentiated one more time it yields, after some simplifications,

$$\begin{cases} X''(0) = (\mu - b) X_1 + \dfrac{\mu K_S X_0 S_1}{(K_S + S_0) S_0} = X_2 \\ S''(0) = -\dfrac{\mu}{Y} \left(X_1 + \dfrac{\mu K_S X_0 S_1}{(K_S + S_0) S_0} \right) = S_2 \end{cases} \tag{2.19}$$

From the first equation of (2.18) and the second equation of (2.19), an analytical expression for the substrate half-saturation coefficient can be determined as

$$K_S = \frac{S_2 S_0^2 X_0 - S_1 S_0^2 X_1}{S_1^2 X_0 - S_2 S_0 X_0 + S_1 S_0 X_1} \tag{2.20}$$

From the first equations of (2.18) and (2.19) together with (2.20), the expression for μ can be formulated as

$$\mu = \frac{X_2 X_0 S_1 - X_1^2 S_1}{S_2 X_0^2 - S_1 X_1 X_0} \tag{2.21}$$

from which an expression for the maximum specific growth rate can be determined by using equations (2.17) and (2.20), which yields

$$\hat{\mu} = \frac{\left(X_2 X_0 S_1 - X_1^2 S_1 \right) (K_S + S_0)}{S_2 S_0 X_0^2 - S_1 S_0 X_1 X_0} \tag{2.22}$$

The decay rate is easily calculated from the first equation of (2.18) and (2.21), which yields

$$b = \frac{S_1 X_2 - S_2 X_1}{S_2 X_0 - S_1 X_1} \tag{2.23}$$

Finally, the yield factor can be determined from the second equation of (2.18) and (2.21) as

$$Y = \frac{X_1^2 - X_2 X_0}{S_2 X_0 - S_1 X_1} \tag{2.24}$$

The above analysis, originally presented by Holmberg (1981; 1982), clearly shows that all four parameters of the model (2.13) and (2.14) are theoretically identifiable (i.e., when perfect data are available) from measurements of X and S if neither $X(0)$ nor $S(0)$ are equal to zero. Holmberg also performed a sensitivity analysis of the same model. This analysis revealed a possible difficulty of distinguishing between effects of $\hat{\mu}$ and K_S from non-perfect (i.e., true) measurements of X and S.

Question: How is the model behaviour and the identifiability affected by noise?

We assume the same batch reactor system as described above but now noise is included in the process. The added process noise (ν) is Gaussian with a mean value of zero and a standard deviation which is 10 % of the actual value of the state variable. The measurement noise (ε) is specified equivalently but an extra Gaussian noise component with zero mean value and a standard deviation of 2 mg/l is added to reflect the difficulties of measuring very low concentrations accurately. It should be noted that the chosen noise level is not very high when compared with real measurements (particularly for X, which is not possible to measure directly in a true process). Moreover, real measurements are normally affected by outliers, non-existent measurement values, noise with non-zero mean and changing variance, trends, etc. Such extra difficulties are neglected here and we define the following system:

$$\begin{cases} \dfrac{dX}{dt} = \mu(S)X - bX + \nu_1 \\ \dfrac{dS}{dt} = -\dfrac{1}{Y}\mu(S)X + \nu_2 \end{cases} \tag{2.25}$$

$$\begin{cases} X_{meas} = X + \varepsilon_1 \\ S_{meas} = S + \varepsilon_2 \end{cases} \tag{2.26}$$

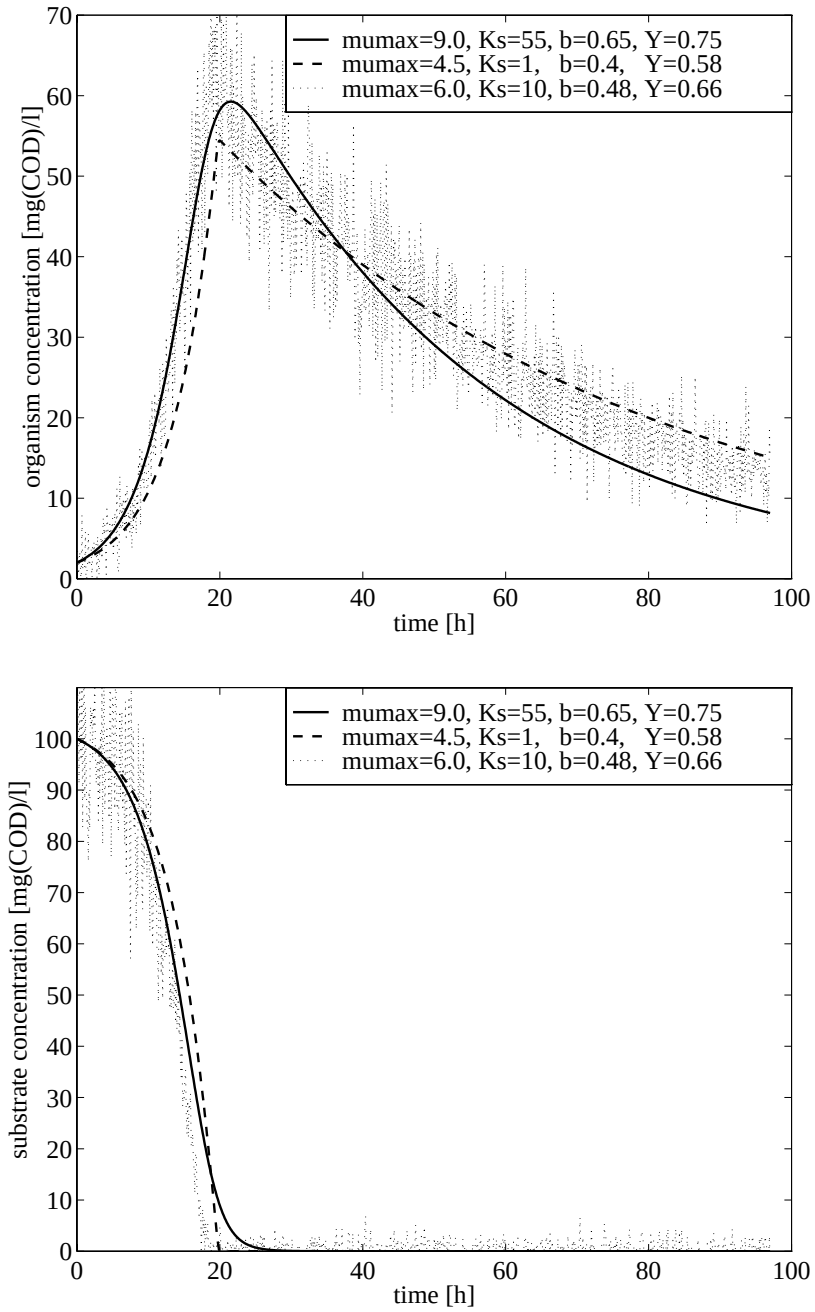


Figure 2.6 Simulation of system (2.13) using the Monod growth rate expression (2.14) with very different sets of model parameter values, compared with noise corrupted measurements of system (2.25) and (2.26). For all simulations $X(0) = 2$ mg/l and $S(0) = 100$ mg/l. For the noise affected system the chosen parameter values are: $\hat{\mu} = 6.0$ day^{-1}, $K_S = 10$ mg/l, $b = 0.48$ day^{-1}, $Y = 0.66$.

A simulation of the system (2.25) with noise conditions specified above is shown in Figure 2.6 (the measurable variables plotted). Although we earlier showed that the system was theoretically identifiable there is no

guarantee that this holds for true measurements, because the added noise makes it impossible to accurately calculate the derivatives that were used in the analytical analysis. By simulating model (2.25) for different sets of model parameters without noise, several sets can be found which produce a model output that are well within the noise deadband of the measurable variables. A few simulations with *very* different parameter values (compared to the ones used during the noisy conditions) are also presented in Figure 2.6, and it is clear that the simulation results are not very different. This indicates the importance of noise on the discussed system and why it has a significant effect on the identifiability analysis. The structure of the model leads to small variations of the assumed measurable quantities even when the internal model parameters are changing significantly and the small output differences are easily lost in the noise.

To improve the measurable outputs (2.26) various means of filtering may be used. An on-line filtered signal is always affected by an undesired time lag. If the outputs are manipulated off-line then this problem can be avoided. In this example a special low-pass filter with exactly zero-phase distortion (Little and Shure, 1988) is used to transform the measurable data into a more suitable form (see Figure 2.7). It should be noted that this is an ideal type of filter that cannot be physically implemented and any real-time filter will produce a poorer result. Because the applied noise in this example is chosen in a favourable way, filtering is not necessary for estimation purposes but still it is used to exemplify some difficulties with regard to filtering.

By applying an optimization algorithm to the filtered data, a set of model parameters, which provide the best possible fit according to a certain criterion (in this case minimizing the sum of squared residuals), can be determined. Two different algorithms have been tested (Fletcher, 1987):

- Nelder-Mead's algorithm (NM) – a simplex method which is very robust but exhibits slow convergence (see Appendix D);
- Gauss-Newton's algorithm (GN) – a generalized least squares method with linear search, which is less robust but converges faster than the NM algorithm.

Both methods above are suited for off-line parameter estimation but the GN method is quite sensitive to the chosen initial parameter values – large differences between the initial and true values cause divergence of the algorithm. The NM method is more reliable for optimizing the above type of system (i.e., (2.25)), and all optimization results presented in this section are based on this algorithm.

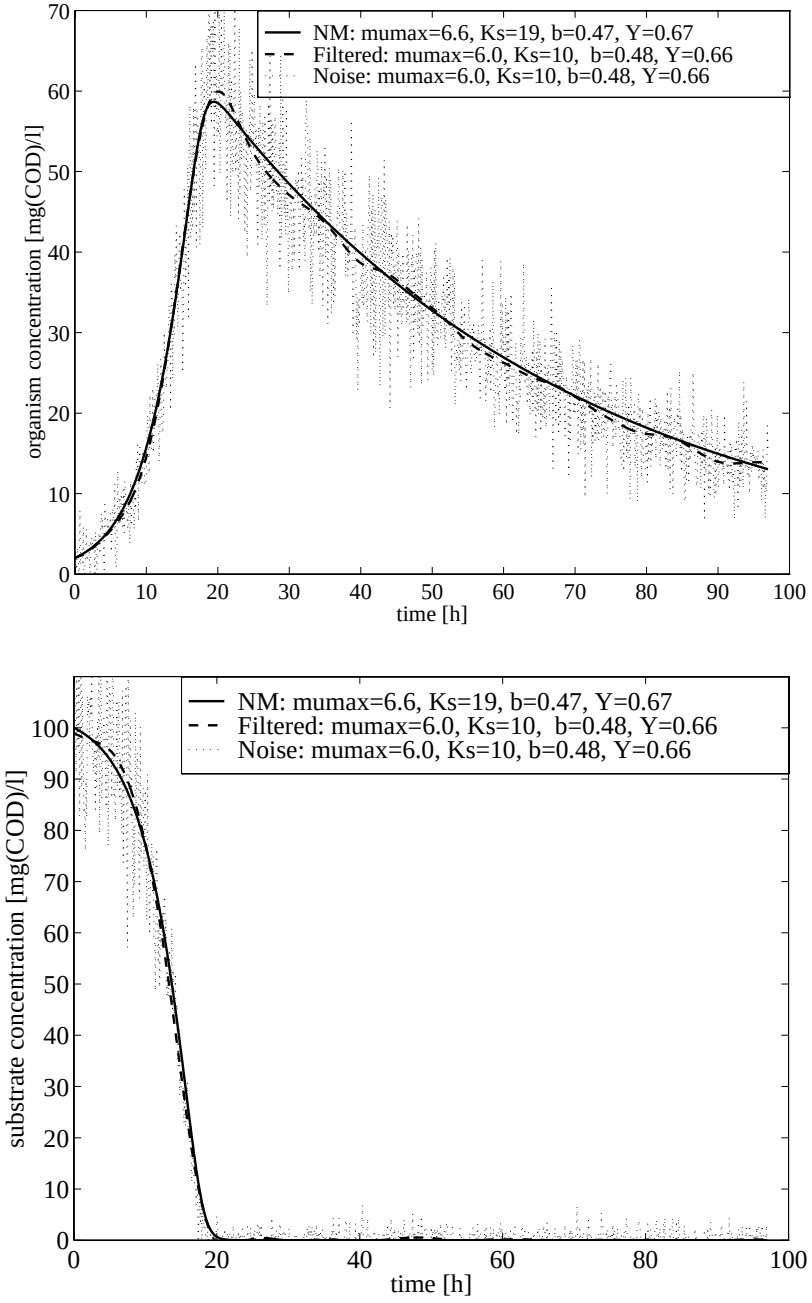


Figure 2.7 Simulation of system (2.13) using the Monod growth rate expression (2.14) with the optimum set of model parameter values determined by the NM algorithm, compared with noise corrupted measurements of system (2.25) and (2.26), filtered and non-filtered. For all simulations $X(0) = 2$ mg/l and $S(0) = 100$ mg/l. For the noise affected system the parameter values are: $\hat{\mu} = 6.0$ day^{-1}, $K_S = 10$ mg/l, $b = 0.48$ day^{-1}, $Y = 0.66$. The calculated optimum parameter set is: $\hat{\mu} = 6.6$ day^{-1}, $K_S = 19$ mg/l, $b = 0.47$ day^{-1}, $Y = 0.67$.

In Figure 2.7, a result of the NM algorithm is shown, when optimizing the parameter set from filtered noisy data. It should be noted that the optimized set of parameters provides an excellent fit to the filtered data, though the values are quite far from the true set.

It is difficult to determine whether the optimized parameter values are the best ones in a global sense or only a local optimum. By running the algorithm for a large number of different initial parameter values and comparing the final results, the global identifiability can be made plausible although it is not a proof. On the other hand, if the algorithm converges towards different values depending on the initial conditions we can conclude that the system is not globally identifiable from the available data. A very small change in the noise characteristics of the system or of the low-pass filter parameters will also greatly influence the optimum parameter set determined by the optimization algorithm. Especially the parameters $\hat{\mu}$ and K_S are difficult to determine accurately and should therefore be considered to be uncertain (if not practically unidentifiable), whereas the estimates of b and Y seem to be more reliable (from this type of idealized batch experiment).

If no noise is added (i.e., model (2.13)), both optimization algorithms locate the true parameter values practically independent of the chosen initial parameter values, i.e., the system is fully identifiable when perfect data are continuously available (which the analytical analysis also has shown). This is not the case even for ‘favourable’ noise conditions.

The discussed example points out some of the problems that may appear, even for simple models. Taking into account the fact that models used to describe WWT processes are much more complicated, the model parameters are usually time varying and functions of temperature, pH, etc., measurements are seldom continuously available and many of the state variables are not measurable at all, it is easily realized that the uncertainty of any estimated results from a true process are considerable.

The behaviour of the complex models will be discussed later in this work. Instead an attempt will be made to simplify model (2.13) in order to develop a practically identifiable model. As true measurements are usually scarce and uncertain, a low complexity model is easier to identify (may even be on-line identifiable), which may lead to more reliable model predictions even though the biological and physical interpretation of some model parameters may be lost.

One possible way to adjust the simple model in this example is to replace the Monod formulation (2.14) with the simpler expression

$$\mu(S) = \begin{cases} rS & \text{if } S < S_{sat} \\ rS_{sat} & \text{if } S \geq S_{sat} \end{cases} \tag{2.27}$$

where: r = reaction rate factor [l (mg day)$^{-1}$];
S_{sat} = growth saturation concentration [mg/l].

The growth function is reduced to a first-order rate expression for low substrate concentrations (the normal case) and a zero-order expression for high substrate concentrations, originally proposed by Blackman (1905). As the model in this example is to be used for a batch experiment, both parts of the growth expression must be included. This means that the number of parameters is not reduced (both r and S_{sat} have to be estimated), only the structure is simplified. However, if the model (2.27) is to be used for continuous-flow bioreactors treating municipal wastewater, the substrate concentrations are usually sufficiently low to motivate the use of a first order reaction only, which would imply a significant model simplification.

In Figure 2.8, a comparison of the behaviours of the traditional (2.14) and modified (2.27) growth rate expressions are presented. The model (2.13) is simulated using the two growth rate functions with the optimized parameter sets determined by the NM algorithm from noisy data and compared with the original disturbed system (2.25) and (2.26) using the standard Monod expression. The result for the organism concentration is presented in Figure 2.8. The result is equally good for the substrate concentration (although not shown). It is apparent that the results when using the simplified growth expression are very similar to that of the original Monod expression. If only a few measurements are available then the parameter r is much easier to identify than $\hat{\mu}$ and K_S.

For the parameter estimations exemplified in Figure 2.8, the measurements have been considered to be continuously available and distributed over the entire range of interesting substrate concentrations. However, this is often not the true case. In order to demonstrate how the sampling rate affects the results, two series of estimations are performed using different sampling rates to describe how frequently data are available for the estimation. The two investigated cases are (assuming both X and S to be measurable):

- sampling rate = 10 hour^{-1} (case A);
- sampling rate = 1 hour^{-1} (case B).

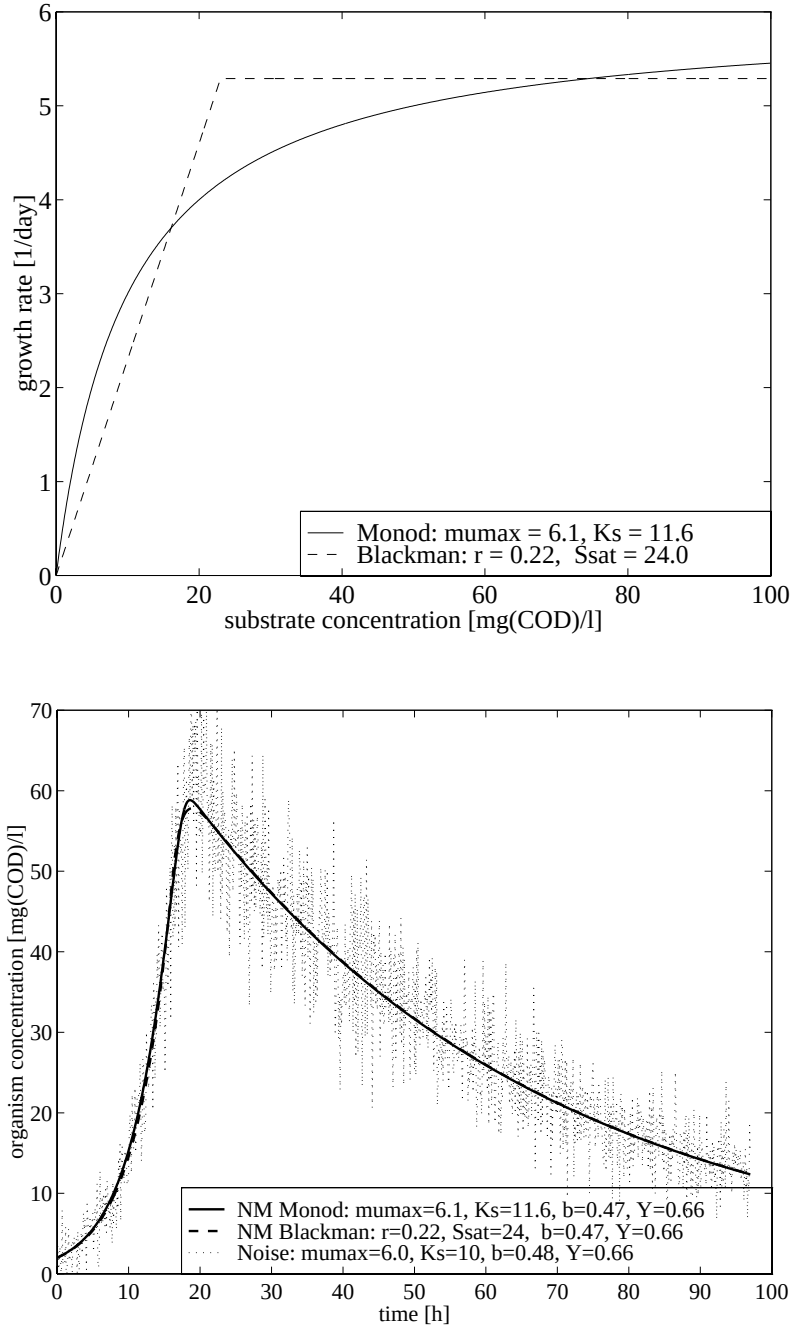


Figure 2.8 Illustration of the behaviour of the Monod (2.14) and Blackman (2.27) growth rate expressions (top). Simulation of system (2.13) using both growth rate expressions with the optimized sets of model parameter values determined by the NM algorithm based on non-filtered noise corrupted measurements of system (2.25) and (2.26) using the Monod function (bottom). For all simulations $X(0) = 2$ mg/l and $S(0) = 100$ mg/l. For the noise affected system the used parameter values are: $\hat{\mu} = 6.0$ day^{-1}, $K_S = 10$ mg/l, $b = 0.48$ day^{-1}, $Y = 0.66$.

For both cases the system (2.25) and (2.26) using the Monod growth rate expression is used to generate three different data sets on which the estimations will be based. The following three data sets are specified:

- no noise added;
- including noise (distributed as earlier described);
- including noise and low-pass filtering (as shown in Figure 2.7).

The NM algorithm is then used for the off-line parameter estimation and optimized sets of parameters are determined for model (2.13) using both types of growth rate functions. Various initial estimates are used in order to indicate possible global identifiability. The results are presented in Table 2.2. The importance of the number of samples is obvious when using noisy measurements as well as the significant effects of low-pass filtering. Note that the model fit to the available data is in all cases satisfactory (in the least squares sense) even though the estimated parameter values differ significantly. In all cases the convergence rates for Y and b are high whereas the parameters in the growth rate expressions are more difficult to determine. However, the convergence rates are significantly higher when using (2.27), i.e., when estimating r and S_{sat}, than when using (2.14) and estimating $\hat{\mu}$ and K_S.

In this example we have demonstrated the difficulty of globally identifying the parameters of a fairly simple model when the measurements are corrupted by noise. For the Monod growth parameters the difficulty is primarily caused by the internal correlations (the change of one parameter value can be compensated by changing the value of another parameter). The problem is especially prominent when the measurable data are not continuously available. Signal processing (e.g., filtering) also have a significant effect on the estimation results. Several parameter sets provide a good fit to the measurable data but it is difficult to determine whether the estimated parameters are the true ones. This is an especially important issue when the model parameters are given a direct physical or biological interpretation.

Conclusion: In order to improve the identifiability of a model it may be necessary to use simplified models and avoid the over-parameterized models that are so common within the field of WWT.

	Monod function	initial estimates				final estimates			
		$\hat{\mu}$	K_S	b	Y	$\hat{\mu}$	K_S	b	Y
Case A (1000 samples)	no noise	5.0 6.0 7.0	5 10 20	0.4 0.5 0.6	0.5 0.65 0.8	6.00 6.00 6.00	10.0 10.0 10.0	.480 .480 .480	.660 .660 .660
	including noise	5.0 6.0 7.0	5 10 20	0.4 0.5 0.6	0.5 0.65 0.8	6.13 6.13 6.13	11.6 11.6 11.6	.472 .472 .472	.664 .664 .664
	including noise and low-pass filtering	5.0 6.0 7.0	5 10 20	0.4 0.5 0.6	0.5 0.65 0.8	6.65 6.65 6.65	18.7 18.7 18.7	.473 .473 .473	.665 .665 .665
Case B (100 samples)	no noise	5.0 6.0 7.0	5 10 20	0.4 0.5 0.6	0.5 0.65 0.8	6.00 6.00 6.00	10.0 10.0 10.0	.480 .480 .480	.660 .660 .660
	including noise	5.0 6.0 7.0	5 10 20	0.4 0.5 0.6	0.5 0.65 0.8	7.07 7.07 7.07	23.2 23.2 23.2	.474 .474 .474	.677 .677 .677
	including noise and low-pass filtering	5.0 6.0 7.0	5 10 20	0.4 0.5 0.6	0.5 0.65 0.8	10.7 10.7 10.7	74.4 74.4 74.4	.488 .488 .488	.687 .687 .687
	Blackman function	r	S_{sat}	b	Y	r	S_{sat}	b	Y
Case A (1000 samples)	no noise	0.15 0.25 0.4	10 25 50	0.4 0.5 0.6	0.5 0.65 0.8	.224 .224 .224	23.7 23.7 23.7	.481 .481 .481	.660 .660 .660
	including noise	0.15 0.25 0.4	10 25 50	0.4 0.5 0.6	0.5 0.65 0.8	.222 .222 .222	24.0 24.0 24.0	.472 .472 .472	.663 .663 .663
	including noise and low-pass filtering	0.15 0.25 0.4	10 25 50	0.4 0.5 0.6	0.5 0.65 0.8	.161 .161 .161	33.4 33.4 33.4	.474 .474 .474	.666 .666 .666
Case B (100 samples)	no noise	0.15 0.25 0.4	10 25 50	0.4 0.5 0.6	0.5 0.65 0.8	.222 .222 .222	24.0 24.0 24.0	.481 .481 .481	.660 .660 .660
	including noise	0.15 0.25 0.4	10 25 50	0.4 0.5 0.6	0.5 0.65 0.8	.141 .141 .141	39.2 39.2 39.2	.476 .476 .476	.679 .679 .679
	including noise and low-pass filtering	0.15 0.25 0.4	10 25 50	0.4 0.5 0.6	0.5 0.65 0.8	.097 .097 .097	59.3 59.3 59.3	.491 .491 .491	.690 .690 .690

Table 2.2 Parameter estimates of model (2.13) using the Monod and Blackman growth rate expression from different data sets using the Nelder-Mead optimization algorithm.

PART II

Modelling the Activated Sludge Process

Chapter 3

Processes and Models – a Review

In this chapter we describe the principles for mathematical modelling of the activated sludge process. A short historical perspective of the development of the process and a review of the large number of existing process variants available today are provided. A literature review of different modelling approaches for the activated sludge process evolved during the last thirty years is also given. Special emphasis is put on the description of the IAWQ Activated Sludge Model No.1 (Henze *et al.*, 1987), which is the principal model used throughout this thesis. A simplified description of the most significant biological processes, essential to a nitrogen removal system, is provided alongside the model formulation. Finally, the possibilities of measuring and manipulating different process variables and parameters in an activated sludge system are commented upon. Parts of this chapter are covered in [181] and [257].

3.1 Historical Perspective

The basic idea of the activated sludge process is to maintain 'active sludge' suspended in wastewater by means of stirring or aeration. The suspended material contains not only living biomass, that is, bacteria and other microorganisms, but also organic and inorganic particles. Some of the organic particles may be broken down into simpler components by a process known as hydrolysis, while other organic particles are not affected (inert material). The biomass in the process will use the organic material as its energy source (usually in combination with oxygen or another oxidation agent), that is, the organic material will be removed from the wastewater while more biomass is produced. The amount of suspended material in the

process is normally controlled by means of adding a sedimentation tank at the end of the process, where the biomass is transported towards the bottom by gravity settling and is either recirculated back to the biological process or removed from the system as excess sludge, whereas the now purified wastewater is withdrawn from the top of the sedimentation tank and released either for further treatment or directly into a receiving water.

The concept of using supplemental aeration as a means of sewage purification dates back to the late 19th century. These early systems were based on the fill-and-draw approach, that is, wastewater was put into a reactor and aerated, and after a period of time the wastewater was released, the deposit of solids was removed and the process was repeated. In 1914, Arden and Lockett (1914), in England, pioneered one of the most popular processes in sewage treatment. Disregarding the current practice, they saved the flocculent solids and studied the effect of their repeated use in sewage treatment by aeration. These flocculent solids, which they called *activated sludge* proved to increase the purification capacity of simple aeration. The accelerating effect depended upon the proportion of activated sludge (AS) to the sewage treated. News of these findings spread rapidly to the United States, and during 1914, similar studies were undertaken at the University of Illinois, leading to the same conclusions. Efforts were then directed towards the adaptation of the process to operate under continuous-flow conditions. By 1917, two small-scale continuous-flow plants in England and a larger plant in Houston, Texas, were put into operation. Successful experience with these plants and the establishment of the diffused air process as a feasible means of air provision, encouraged the construction of other major plants, which were soon placed in operation. All were based on the continuous-flow principle, which had proven itself as the major practical method for activated sludge operation.

The early success of the activated sludge process did not persist for long. Rapid population expansion and industrial development greatly altered the magnitude and nature of sewage loads to existing wastewater treatment (WWT) plants, and the effect of flow and organic load variations became more pronounced. One of the most serious problems was caused by what was generally described as sludge bulking, a phenomenon that manifested itself as an appreciable reduction in settleability of activated sludge, often resulting in excessive suspended solids (SS) concentrations in the plant effluent. Extensive studies during the 1930s identified some environmental conditions causing this problem to be:

- improper balance of food caused by high carbohydrate levels;
- high carbon-to-nitrogen ratio, attributable to industrial discharges;
- low dissolved oxygen levels in the aeration tanks;
- increasing organic loads on the treatment plants.

Another serious problem that haunted the AS process was the shortage of oxygen, primarily at the head of the aeration tanks. It was only after the recognition of the importance of oxygen as a quantitative factor in the process and studies on the oxygen utilization during sewage treatment (Grant *et al.*, 1930; Bloodgood, 1938) that the relationship between oxygen and degradation of organic material became clear. It was suggested that appreciable changes might occur with regard to the character of the sludge as a function of the length of the oxygen deficiency period. Therefore, it was concluded that dissolved oxygen must be present at all points in the aeration tank.

Process Modifications

The difficulties encountered in the operation of AS plants, triggered the development of modified processes that would permit existing plants to treat larger flows and greater loads while maintaining a high effluent quality. The frequent shortage of oxygen in aeration tanks led to a modification of the process, known as *tapered aeration*. It involved sizing the aeration equipment as a function of anticipated oxygen requirements, that is, increasing the number of diffusers at the head of the aeration tank while decreasing the number of diffusers closer to the outlet (Kessler and Nichols, 1935). This modification was initiated around 1930 and today almost all AS plants include provision for tapered aeration.

Attention was also focused on the occasional load transients to which the process is exposed and which it must accommodate. It was noticed that the oxygen demand of a mixture of activated sludge and sewage could often exceed the ability to dissolve oxygen by means of conventional or modified aeration systems, if the entire incoming load together with the return sludge were applied to the head of an aeration basin. Around 1940, this led to the idea of adding sewage in regulated amounts, at multiple points along the tank, instead of applying tapered aeration (Gould, 1940). This process, known as *step aeration*, *distributed loading*, *step feed*, *multiple-point dosing* and *incremental dosing*, could produce an activated sludge with a good purifying capacity while maintaining the oxygen requirements at a

more uniform level throughout the aeration tank. The process could also reduce the effects of shock loads and produce savings in terms of less required tank volume.

Around the same time, two different process modifications were tested that never found any wide application. One was the *modified aeration* system, where the sewage was aerated for shorter periods and with a smaller quantity of biological solids in the aerators (Setter *et al.*, 1945). The sludge in this system settled rapidly and compacted to a concentration twice that of a conventional AS system. Less oxygen and recirculation was required but the effluent quality was generally reduced as well. The other process was the *activated aeration* system, where the excess sludge from a conventional AS process was to be utilized (Shapiro and Hogan, 1945). Therefore, this sludge was diverted into a second aeration unit to which part of the influent wastewater was directed. This added process operated without recycling of sludge.

Practical development of high-capacity aeration devices made it possible to develop the *high-rate* AS process. The basis of this process was that the conventional aeration period could be drastically reduced if adequate oxygen input to the system was ensured. In combination with a high degree of turbulence, it was possible to substantially increase the sewage load by an appreciable reduction in the floc size and improving the oxygen diffusion rate into smaller flocs. The short detention time and high substrate to biomass ratio could maintain the biomass in a very active phase, that is, the log-growth phase. The process was found to be as stable as conventionally operated plants, although it required a high sludge return rate. A major benefit was that the process was operated with comparably small aeration basins.

Large volumes of waste sludge that had to be treated and disposed of were a factor that led to the development of the *extended aeration* AS process. The process modification was established as a means of eliminating the problem of excess sludge handling, while producing a highly stabilized effluent and requiring a minimum of attention. On the other hand it requires large aeration basins allowing for a long hydraulic detention time. The process became popular during the 1950s. A feature of the process is its ability to contain a relatively large mass of sludge and thereby, for all practical purposes, totally remove the influent organic material in the wastewater. The effluent organic material is almost entirely due to the suspended solids that escape from the system. The process produces comparably small amounts of sludge, and the sludge is very stable as the

low substrate to biomass ratio maintains the biomass in the endogenous phase. An intentional sludge wastage is often provided to improve the effluent quality although the initial intent of the process was to be operated without any excess sludge production.

The concept of *sludge reaeration* was initially intended for improving the sludge settleability by aerating the recycled sludge in a separate tank prior to returning it to the aeration basin. The process could also be used as a way of providing a reservoir of sludge that was buffered against poisonous effects of short pulses of influent toxic material. The process later evolved into the *contact-stabilization* process (Ullrich and Smith, 1951). Basically the process employs two aerated reactors separated by a settling system. Either raw or presettled wastewater is mixed with activated sludge and aerated in a contact tank with a short detention time. Here the biomass adsorbs the influent organics and then settles out in the secondary clarifier. The concentrated sludge is then pumped into the stabilization tank and aerated for several hours prior to return to the contact basin. The process is still popular and it is often considered to increase the possible volumetric loading capacity and efficiency compared with a conventional-flow AS process, and to some extent improve the settleability of the sludge.

A way of improving the capacity of an AS process while maintaining the reactor volumes is to convert it to a *pure-oxygen* or *enriched oxygen* system. By applying oxygen (usually added under pressure) instead of air to the system it is possible to maintain high levels of oxygen in the aeration tank without excessive turbulence, which would break up the flocs and deteriorate the settleability. This allows for higher loadings, higher sludge concentrations and shorter detention times than in a conventional system. The process is especially favoured in areas with very limited space for plant expansions. Naturally, the cost for aeration is significantly increased by this approach.

The *sequencing batch reactor* (SBR) is an AS system operating on a fill-and-draw basis. As earlier discussed, this was the principle applied in the early AS plants. After the choice for continuous-flow processes, the interest for SBRs was revived in the early 1960s with the development of new technology and equipment. The SBR process is essentially composed of a single tank. The process modification basically consists of its semi-batch operation and the fact that biological conversion and settling take place in the same reactor in a cyclic operation. Most of the advantages of this process may be attributed to the very flexible nature of the operating parameters, as the process can be controlled by time rather than by space.

However, the SBR process normally requires a more sophisticated control strategy than a traditional continuous-flow system.

Processes for Nutrient Removal

The processes described above were developed primarily for the removal of organic material from the wastewater. However, during the last thirty years, nutrient removal has become a very important factor in WWT, that is, the removal of nitrogen and phosphorus components from the wastewater. In order to accomplish this in an AS process, a large number of different process configurations have been developed. One basic problem is that the microorganisms performing nitrification, denitrification and enhanced biological phosphorus removal (EBPR) require very different environments to function effectively, that is, a combination of aerobic, anoxic and anaerobic conditions. The term anoxic is frequently used to define a condition when oxygen is absent and nitrate or nitrite is present. Some of the most established processes for nutrient removal will be schematically described below.

Biological nitrogen removal AS systems are normally separated in two different categories – separate-sludge systems and single-sludge systems. The separate-sludge system is characterised by two sets of reactors with individual settling and sludge recycle, operated in sequence and sustaining two different types of microbial communities. The first reactor is aerated to achieve carbon removal and nitrification, whereas the second reactor provides an anoxic environment for denitrification. This process configuration is also known as a two-sludge system. Three-sludge systems with carbon removal, nitrification and denitrification in successive reactors are also a possible solution. Since the organic matter of the wastewater is almost completely consumed in the first part of the process, the anoxic reactor often requires the addition of an external carbon source, for example, methanol or ethanol. Another possibility is to bypass a portion of the influent wastewater to the anoxic reactor to provide the necessary carbon for the denitrification. While total separation of aerobic and anoxic processes enables optimum design and performance stability, economical considerations have been the major incentive in the development of combined or single-sludge systems. Basically, the combined process is applied in two different configurations – single-sludge predenitrification and single-sludge postdenitrification.

The postdenitrification process consists of two reactors in series, the first aerobic and the second anoxic, see Figure 3.1a. It was first suggested by Wuhrmann (1964) and is consequently known as the *Wuhrmann process*. The energy source for the denitrification process is provided by energy released by the sludge mass due to the death of organisms. However, since this rate is low, the denitrification rate is also low. Therefore, a very large anoxic volume is required for a high denitrification efficiency.

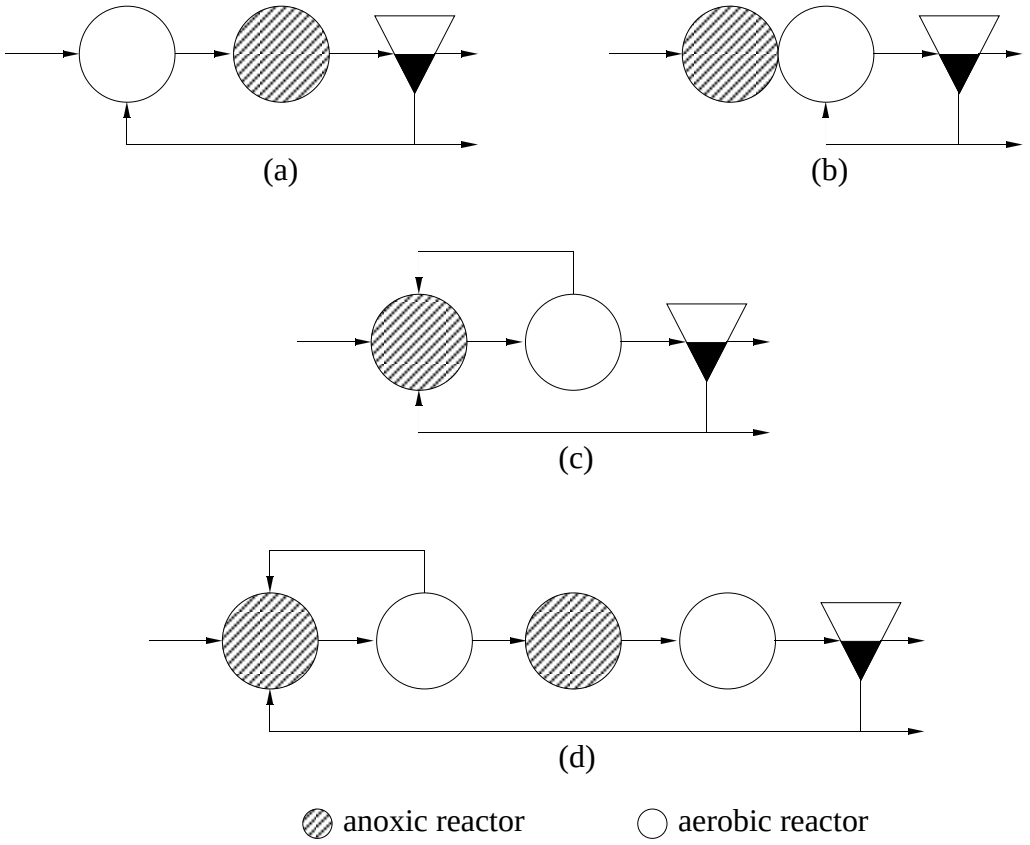


Figure 3.1 Single-sludge nitrogen removal systems: (a) the Wuhrmann process, (b) the Ludzack-Ettinger process, (c) the modified Ludzack-Ettinger process, and (d) the Bardenpho process.

The predenitrification process was first developed and proposed by Ludzack and Ettinger (1962) and, consequently, known as the *Ludzack-Ettinger process*. It consists of two reactors in series, partially separated, without intermediate settling, see Figure 3.1b. As there is only partial separation between the two reactors, a mixing of the nitrified and anoxic wastewater is induced, and the nitrate entering the anoxic reactor is reduced to nitrogen gas. This process was later modified by Barnard (1973), who completely separated the anoxic and aerobic reactors, recycling the settler underflow to the anoxic reactor, and providing an additional recycle from the aerobic to the anoxic reactor, see Figure 3.1c.

The process is known as the *modified Ludzack-Ettinger process* and the control of the process is significantly improved. However, with this flow scheme, complete denitrification is not possible, and the degree of denitrification depends upon the fraction of the total flow from the aerobic reactor not recycled but discharged directly with the effluent.

In order to overcome the incomplete denitrification, the *Bardenpho process* was proposed as a combination of the modified Ludzack-Ettinger and Wuhrmann processes, see Figure 3.1d. The low concentration of nitrate discharged from the aerobic reactor to the second anoxic reactor will be denitrified to produce an effluent free of nitrate. To strip the nitrogen bubbles generated in the secondary anoxic reactor attached to the sludge flocs, a flash aeration is introduced between the secondary anoxic reactor and the settler. This extra aeration is also considered necessary to nitrify the ammonia released through endogenous decay in the previous reactor.

In order to achieve EBPR in a conventional activated sludge process an anaerobic reactor can be added in front of the aerobic reactor, as shown in Figure 3.2a. This process is known as the *two-stage Phoredox process* or the *A/O process*. The real challenge was the development of an AS process capable of performing organic removal, nitrification, denitrification and EBPR within a single-sludge system. A number of such processes are schematically outlined in Figure 3.2, without further explaining any details about the flow schemes and reactor interactions of these complicated processes. Such a description is beyond the scope of this overview. An important event of the development of these processes was the observation of high and stable phosphorus removal in a pilot plant (Bardenpho process) designed by Barnard in 1975 for nitrogen removal (Barnard, 1975). Having observed that EBPR was possible when an anaerobic stage was followed by an aerobic stage, Barnard proposed to employ an anaerobic stage before the nitrogen removing Bardenpho system, thereby creating the *modified Bardenpho process* or the *five-stage Phoredox process*, see Figure 3.2b. When only partial nitrogen removal was required, it could be reduced to three stages – anaerobic, anoxic and aerobic (Barnard, 1983). This process is often referred to as the *three-stage Phoredox process* or the *A^2/O process*, see Figure 3.2c.

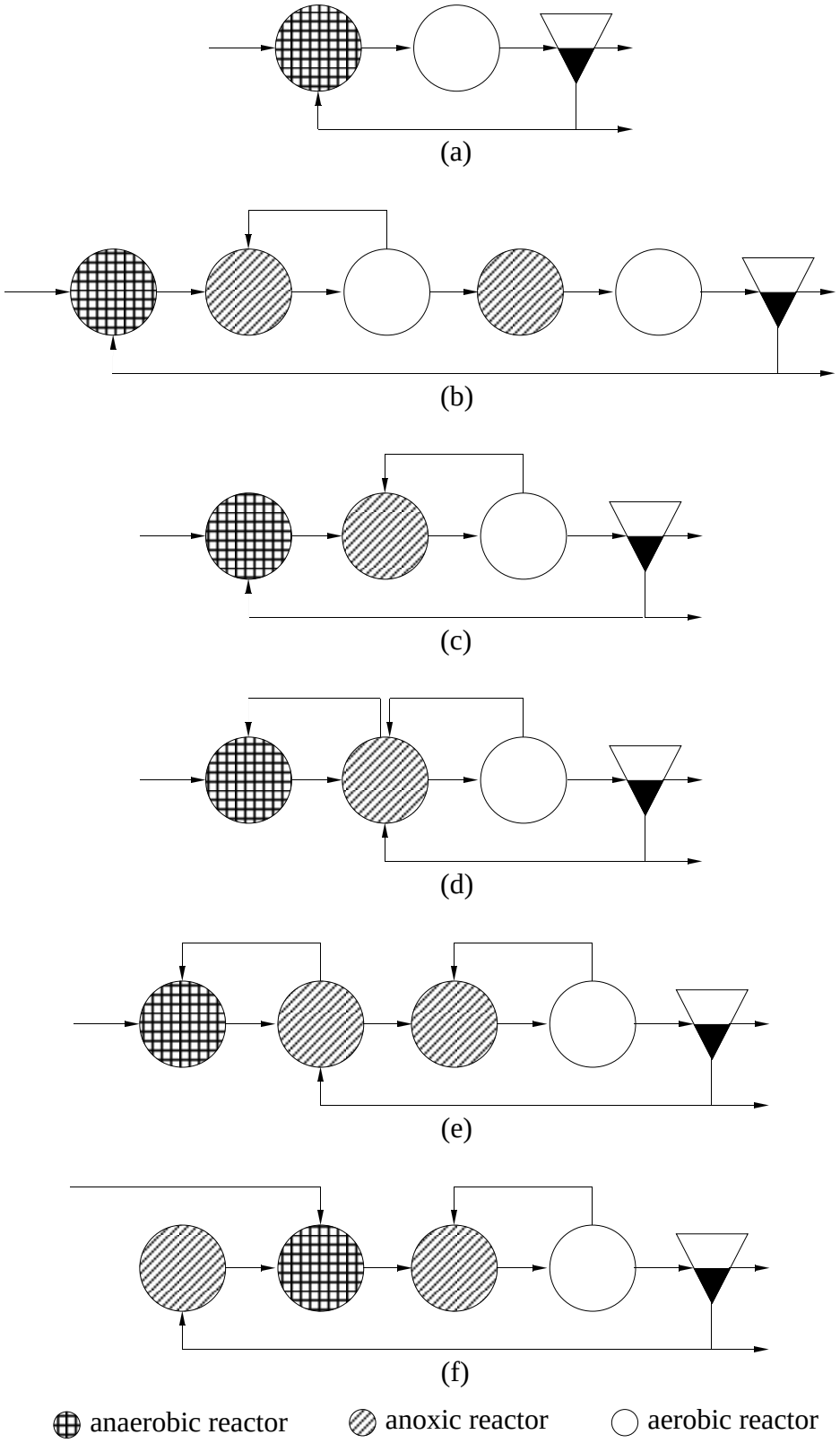


Figure 3.2 Single-sludge nutrient removal systems: (a) two-stage Phoredox (A/O) process, (b) five-stage Phoredox (modified Bardenpho) process, (c) three-stage Phoredox (A^2/O) process, (d) UCT process, (e) modified UCT (MUCT) process, and (f) Johannesburg process.

As researchers continued to develop biological nutrient removal AS systems, and discovered the significance of the sequence of anaerobic-aerobic stages, the inhibiting effect of nitrate recycle to the anaerobic stage on EBPR was also recognized (Nicholls, 1975; Barnard, 1976). These considerations later led to the introduction of the *University of Cape Town (UCT) process*, the *modified UCT (MUCT) process* and the *Johannesburg process* (Dold *et al.*, 1991), outlined in Figures 3.2d-3.2f.

It should be noted that modifications have also been made on other types of AS processes, such as the oxidation ditch and the SBR, converting them into nutrient removal systems. Two good examples of alternating processes (a development of the SBR process) including nutrient removal are the BIO-DENITRO (Christensen, 1975) and the BIO-DENIPHO processes (Einfeldt, 1992).

3.2 Model Development

There has been a long transition period between the promotion of the activated sludge method of wastewater treatment and the establishment of a theoretical framework that both quantitatively describes the process, and provides a rational basis for its design. The conflicting nature of the many hypotheses for the mechanistic explanation of the process, the difficulty of expressing them in precise mathematical models, and the contrived nature of the systems on which the models were developed were the main reasons for this slow transition. Due to the absence of basic rational guidelines, the early developments of plant design and operation have been more of an art than a science.

From the 1920s until the 1960s different hypotheses for explaining the mechanisms of organic matter removal by activated sludge, were proposed. In Arden and Lockett's original work, it was recognized that physical, chemical and biological mechanisms might be responsible in varying degrees for the purification of the wastewater, although no attempts were made to identify their existence or their relative importance. The hypotheses included theories recognized today as the *coagulation theory*, the *adsorption theory*, the *colloid theory*, the *biozeolite theory* and the *enzymatic theory.* A description of these theories is given in Orhon and Artan (1994). Although the adsorption theory was the dominating theory for many years it should be noted that already in 1923 a biological mechanism

was proposed as an alternative theory by Buswell and Long (1923). On the basis of their observations they reported that AS flocs were made of synthetic, gelatinous matrix enclosing filamentous and unicellular bacteria, as well as various protozoa and some metazoa. Their experimental evidence suggested that the purification was accomplished by ingestion and assimilation of the organic matter from sewage and their synthesis into the living material of the floc. This theory has later found universal acceptance and has remained virtually unchallenged.

By 1940, a number of experimental investigations had shown that adsorption could not be the only nor the predominant mechanism of organic matter removal. If adsorption did occur, it had to be accompanied by a biochemical reaction. McKee and Fair (1942) stated that the removal mechanism consisted of two distinct but interrelated steps:

- a physical process of adsorption and flocculation which proceeded rapidly and decreased as the organic substances were removed from sewage or as the contact surfaces became covered with these substances;
- a biological process of organic matter consumption for energy and cellular synthesis purposes, which was instituted simultaneously but proceeded more slowly.

The idea of the two-step adsorption/metabolism found immediate support, although there was no common agreement on the relative importance of each step in the overall removal mechanism. The concept was used as the basis for process modifications – especially for the development of the contact-stabilization process discussed in the previous section.

In all the theoretical speculations as well as the attempts to verify them experimentally, it must be noted that all organic material of the sewage was considered an an entity without any emphasis on its particulate and soluble components. It was first around 1955 that researchers began to consider the composition of the sewage and its impact on the reactions.

Operational difficulties encountered, together with new process extensions such as nutrient removal, have greatly increased the use of process modelling. This has led to an ever-increasing need for mathematical models incorporating the fundamental microbial mechanisms into a rational engineering description of the process. Consequently, a significant evolution in modelling practice has been experienced in the last three decades, from the single-component model advocated by McKinney

(1962), to the elaborate model including 19 components, 65 parameters and 19 different processes recently proposed as the IAWQ AS Model No. 2 for combined carbon, nitrogen and phosphorus removal (Henze *et al.*, 1995). The accumulated scientific information and ingenuity of the recent modelling efforts are noteworthy. However, the reliability of the proposed models depends on an increasing number of kinetic and stoichiometric parameters, which to a large degree depend on characteristics of the actual wastewater and must therefore be experimentally determined.

Bacterial growth and decay

The biological processes in a WWT plant are carried out by many different types of bacteria. The most important microorganisms in the AS process are bacteria, while fungi, algae and protozoa are of secondary importance. The different types of organisms that can be found in an AS process, are also found in the raw wastewater flowing into the plant or in the immediate surroundings of the plant. The predominant genera of bacteria in the activated sludge are mainly determined by the composition of the raw wastewater, the design of the plant, and to some extent the operation of the specific plant.

Bacteria constantly need energy in order to grow and to support essential life activities. Growing cells utilize substrate and nutrients located outside the cell membrane for growth and energy in a process, which can be described in a simplified form as

$$\text{Substrate} + \text{Nutrients} + \text{Oxygen} \rightarrow \text{Biomass} + \text{Energy}$$

The bacteria can also accumulate substrate and nutrients and store them internally in modified forms (typically polysaccharides, lipids and polyphosphates). The major part of bacteria in activated sludge (called heterotrophic bacteria) use organic carbon in the form of small organic molecules as substrate, and some bacteria (called autotrophic bacteria) which are essential to biological nutrient removal, use inorganic carbon as substrate. When the bacteria decay, the organic carbon of the bacteria is partly reused in the process.

For example, the formation of a typical biomass compound ($C_5H_7NO_2$) from a typical substrate ($C_{18}H_{19}O_9N$) in an aerobic environment with a typical yield coefficient is given by the following reaction:

$$C_{18}H_{19}O_9N + 0.74NH_4^+ + 8.80O_2 \rightarrow$$
$$1.74C_5H_7NO_2 + 9.30CO_2 + 4.52H_2O + H^+$$

The end-products on the right-hand side of the biochemical reaction is obviously harmless to the environment. It should be noted that in addition to the removal of organic matter, ammonia is removed by growth of bacteria. The above process is carried out by the heterotrophic bacteria in a WWT plant.

In order to mathematically describe the kinetics of the reactions taking place in a biological reactor, practically all models are based on the two fundamental processes discussed above, that is, microbial growth (3.1) and decay (3.2), usually described mathematically as

$$\frac{dX}{dt} = \mu X \tag{3.1}$$

$$\frac{dX}{dt} = -bX \tag{3.2}$$

where μ is the specific growth rate and b is the decay coefficient. Process stoichiometry is then used to relate substrate (S) utilization to microbial growth (3.3), as

$$\frac{dX}{dt} = -Y\frac{dS}{dt} \tag{3.3}$$

where Y is the yield coefficient. The decay process is generally defined by a first-order rate expression with respect to the biomass concentration (X). The above model is equivalent to the one investigated in the small example in Section 2.6. This description cannot differentiate between degradation of endogenous mass for the generation of maintenance energy, microbial death, cell lysis and interactions between predators and bacteria, but it reflects the overall combined effect.

A major issue has been how to mathematically describe the specific growth rate for a continuous culture of microorganisms growing in wastewater on a mixture of organic and inorganic substrates. The most commonly recognized rate expression is the hyperbolic expression proposed by Monod (1942; 1949), as an empirical deduction from pure culture studies, i.e.,

$$\mu = \hat{\mu} \frac{S}{K_S + S} \tag{3.4}$$

where $\hat{\mu}$ is the maximum specific growth rate, K_S is the half-saturation constant and S is the concentration of the growth-limiting substrate. This expression is in turn compatible with the Michaelis-Menten (1913) enzymatic reaction model. Monod deduced the above expression by investigating the different growth phases of pure bacteria cultures in a batch system, see Figure 3.3. The growth rate is very much depending on the living conditions for the organisms but the growth phases are generally categorized as:

I: lag phase II: acceleration phase III: exponential (log) phase
IV: retardation phase V: steady-state phase VI: declination phase

Note! No figure available.

Figure 3.3 The bacterial growth phases (Monod, 1942).

During the lag phase the bacteria adapt to the environment and no growth can be seen. The lag phase is followed by the acceleration phase, which is characterized by a fast increase of the growth rate. This later leads to the exponential phase. As a consequence of decreasing access to nutrients and increasing amounts of metabolic products and toxins the increase of the growth rate starts to decline and in the stationary phase the growth rate eventually stabilizes at a certain level. Finally, the bacteria die off during the declination phase.

The bacteria can be measured either as concentration of cells, that is, the number of individual cells per volume, or as bacteria density, i.e., the total dry weight of bacteria per volume (Finnson, 1994). A complicating fact when measuring the growth rate of bacteria is that in practice the size of

the microorganisms vary between the different growth phases (Monod, 1942). In general, the size of a cell reaches its maximum at the end of the lag phase, continue at constant size through the exponential phase and subsequently decrease in size. In the exponential phase the growth rate is proportional to the bacteria density (Monod, 1942), following that the size of the cell is constant. It must also be realised that a multiplicity of reaction mechanisms occur even in the simplest biological reaction. Adsorption, enzyme catalysis, inhibition and diffusion processes represent the major functional mechanisms that may control the uptake of a specific substrate. Furthermore, these mechanisms are dependent upon a number of physical, chemical and biological variables within a given system (cf. medical literature on glucose uptake by red blood cells).

The Monod expression is developed as an acceptable mathematical description of experiments conducted with pure bacterial cultures growing on single substrates. In WWT practice, the latter is substituted with non-specific parameters like BOD (biological oxygen demand) or COD (chemical oxygen demand). Although they are mathematically treated as single substrate components, these parameters include a great variety of organic compounds with different biodegradation characteristics. The influent wastewater also contains artificially manufactured chemical compounds and toxic materials, to which various organisms respond differently. Furthermore, conditions like the dissolved oxygen (DO) concentration and the pH may vary within the treatment plant. Consequently, in the biological reactors used for the removal of the mixture of organic compounds in wastewaters, there is no way to select a given microbial species, since a mixed microbial community develops as an enriched culture, resulting from natural selection. Taking the above into consideration, a slightly more realistic growth rate function for WWT plants, including growth on n multiple substrates where the different components exhibit a competitive inhibition effect on the utilization of the other components, may look like

$$\mu = \sum_{i=1}^{n} \frac{\hat{\mu}_i S_i}{K_{S,i} + \sum_{j=1}^{n} a_{i,j} S_j} \tag{3.5}$$

where $a_{i,i}=1$ and $a_{i,j}$ represents the inhibition effect of the j^{th} substrate on the utilization of the i^{th} substrate by the organism. However, this complex model structure only addresses a few of the problems discussed above. It would, on the other hand, be practically impossible to identify and verify such a model for values of i and j larger than 1, both because of the inherent model structure and due to the necessary detailed measurements.

In fact, a drawback of the original Monod expression is that the parameters as part of a biological model cannot be measured directly but must be estimated. This is difficult because even a model in the very simple form of (3.1)–(3.4) is not practically identifiable (Holmberg, 1981), although it is theoretically identifiable from perfect measurements of X and S in a batch experiment (Holmberg and Ranta, 1982), as discussed in Section 2.6. The lack of practical identifiability means that unique sets of parameters can rarely be obtained. Parameters estimated from data obtained during apparently similar conditions show considerable variations and parameter estimation methods show poor convergence properties (Holmberg, 1982). In Vialas *et al.* (1985) ways of improving the practical identification of the Monod equation by using different sample times depending on the current state of the process, is suggested. Ratkowsky (1986) proposes a different parametrization of the Monod equation to improve the identifiability of the model and enhance the convergence of estimation algorithms. Both linear and non-linear regression techniques are applied for estimating the growth model parameters as well as the yield and decay rate coefficients from true plant data in Vaccari and Christodoulatos (1990). The estimates are, however, not significant at the 95 % confidence level and, therefore, the use of a simple first-order rate equation is promoted instead of the Monod expression. A comparison of the non-linear Monod equation and a linear simplification is also performed in Derco *et al.* (1990a; 1990b). The investigations show that a linear rate model is not as good for predicting actual transient responses in biomass and substrate concentrations as the traditional formulation. On the other hand, the standard Monod expression does not provide perfect results either, when compared with true data. The advantage of the linear rate equation is that it improves the practical identifiability of the model. In a similar way, it has also been shown that by measuring the oxygen uptake rate (OUR) during very well-controlled conditions using a respirometer it is only possible to uniquely identify certain *combinations* of parameters and state variables in model (3.1)-(3.4), see Vanrolleghem (1994).

The relevance of the basic structure of the empirical Monod equation is also a matter of dispute. Depending on what mechanism is considered to be most important (biochemistry, adsorption, diffusion, etc.) the growth expressions are formulated differently. A large number of rival models that exhibit practically the same behaviour have been suggested and investigated from an identification point of view (Boyle and Berthouex, 1974; Dochain and Bastin, 1984). Some of the proposed variants that have been applied in WWT modelling are given below (all K and θ coefficients represent different model parameters).

The Tiessier model (Tiessier, 1936):

$$\mu = \theta_1\left(1 - \mathrm{e}^{\theta_2 S}\right) \tag{3.6}$$

The Blackman model (Blackman, 1905; Garrett and Sawyer, 1952):

$$\mu = \begin{cases} \dfrac{\hat{\mu} S}{K_\mathrm{B}} & \text{if } S < K_\mathrm{B} \\ \hat{\mu} & \text{if } S \geq K_\mathrm{B} \end{cases} \tag{3.7}$$

The Contois model:

$$\mu = \frac{\hat{\mu} S}{K_\mathrm{C} X + S} \tag{3.8}$$

The Powell model (Powell, 1967):

$$\mu = \theta_1\left(\frac{\theta_2 + \theta_3 + S}{2\theta_3}\right)\left(1 - \sqrt{1} - \frac{4\theta_3 S}{\left(\theta_2 + \theta_3 + S\right)^2}\right) \tag{3.9}$$

The Haldane model for inhibition kinetics:

$$\mu = \frac{\hat{\mu} S}{K_\mathrm{S} + S + \dfrac{S^2}{K_\mathrm{I}}} \tag{3.10}$$

The behaviour of some of the equations is exemplified in Figure 3.4. By choosing the model parameter values properly, the similarities of the equation behaviours are made obvious and indicate that the growth rate may be described by many different expressions.

There is evidence in the literature to show that Monod-type expressions provide reasonable models to describe the growth of the enriched culture sustained in WWT reactors, with the provision that the kinetic parameters be interpreted not as absolute values, but as average figures related to the predominant species in the particular growth conditions of the reactor. Bearing this in mind, it would seem appropriate to promote growth expressions that produce results similar to the original Monod expression but simultaneously enhance the identifiability of the model. Of the models discussed above, the Blackman model is the easiest one to identify,

especially if we consider the fact that in municipal WWT plants the substrate concentrations are generally so low that only the first part of the model (the first-order growth rate expression) is required.

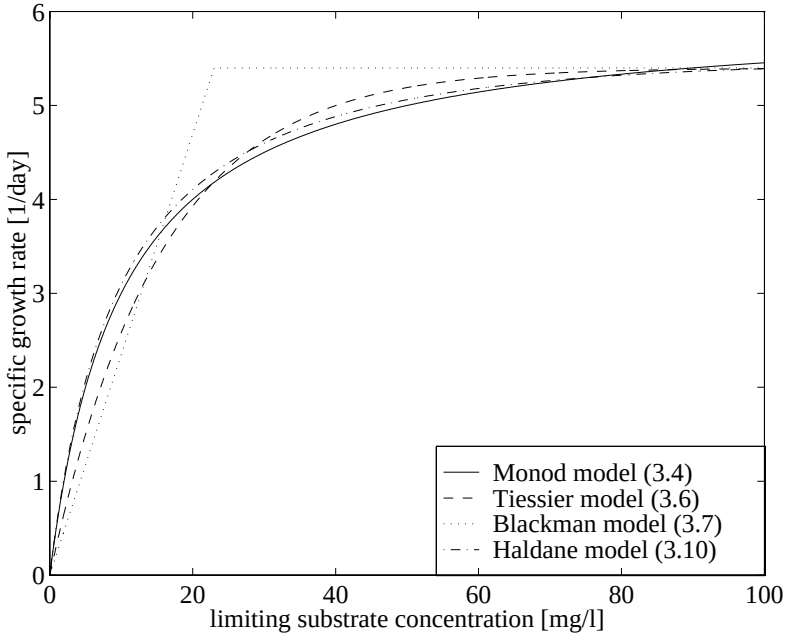


Figure 3.4 Behaviour of different bacterial growth rate equations.

Nitrification and Denitrification

Two more processes, which are of great importance for the AS process and, consequently, should be included in a mathematical model, are the nitrification and denitrification processes. Nitrification is a two-step microbiological process transforming ammonia into nitrite and subsequently into nitrate. The process is well-known from the biosphere, where it has a major influence on oxygen conditions in soil, streams and lakes. Soluble ammonia serves as the energy source and nutrient for growth of biomass of a special group of autotrophic bacteria (called nitrifiers). If ammonia is used only as a source of energy, the first step of oxidizing ammonia into nitrite is described as

$$NH_4^+ + 1.5O_2 \rightarrow NO_2^- + H_2O + 2H^+$$

and the second step of oxidizing nitrite into nitrate is

$$NO_2^- + 0.5O_2 \rightarrow NO_3^-$$

A typical representative for the first step is the bacteria of the genus Nitrosomonas and for the second step the bacteria of the genus Nitrobacter.

Because the reactions above only give a small energy yield, the nitrifying bacteria are characterized by a low biomass yield. This is an essential problem for the nitrification process in biological nutrient removal systems. Using a typical yield for autotrophic growth of biomass, the following reaction for the total nitrification process is obtained:

$$NH_4^+ + 1.86O_2 + 1.98HCO_3^- \rightarrow$$
$$0.02C_5H_7NO_2 + 0.98NO_3^- + 1.88H_2CO_3 + 1.04H_2O$$

where HCO_3^- is the form of soluble carbon-dioxide for pH-values in the range from 5 to 9. From the reaction above, it is seen that a large amount of alkalinity is consumed for every NH_4^+ being oxidized. Although the wastewater in many areas contains large alkalinity buffers, some WWT plants require the addition of lime or soda ash to maintain desirable pH-levels for nitrification. Normally, the nitrification is mathematically described as a one-step process in order to keep the models fairly simple.

Denitrification is a microbiological heterotrophic process transforming nitrate into nitrogen gas, using nitrate instead of oxygen as the oxidization agent. The conditions during which this process occurs, are called anoxic, because oxygen is not present and some heterotrophic bacteria are able to use nitrate for oxidation. Denitrification is also well-known from the biosphere, where it is common in soil and beneath the surface in stationary waters. Most of the heterotrophic bacteria are optional to the use of oxidation agent, but the energy yield of using nitrate is less than when using oxygen. Thus, if oxygen is present, the bacteria prefer to use oxygen. In practice, denitrification only takes place at low oxygen concentrations. The overall mechanism can be described by a typical microbial reaction of a saccharide with nitrate:

$$5C_6H_{12}O_6 + 24NO_3^- \rightarrow 12N_2 + 24HCO_3^- + 6CO_2 + 18H_2O$$

The lower energy yield for the heterotrophic bacteria during the anoxic conditions is also reflected in a somewhat lower biomass yield. Denitrifying bacteria using ammonia and the typical form of organic substrate ($C_{18}H_{19}O_9N$) in wastewater for bacterial growth with an observed yield coefficient of 0.47 g biomass/g substrate gives the following reaction:

$$0.61C_{18}H_{19}O_9N + 4.54NO_3^- + 0.39NH_4^+ + 4.15H^+ \rightarrow$$
$$C_5H_7NO_2 + 2.27N_2 + 6.70CO_2 + 5.12H_2O$$

Fortunately, some of the alkalinity lost by nitrification is gained by denitrification. A very important parameter for the denitrification process is the organic carbon/nitrogen-fraction (C/N ratio) of the raw wastewater, which also plays a significant role for the design and operation of the WWT plant. In practice, the C/N ratio of the raw wastewater should be at least 8-9 g COD/g N for a typical WWT plant, in order to assure a relatively high denitrification rate (Carstensen, 1994).

Dynamic Models

It was the 1950s before the first *dynamic* models of the AS process were proposed (Goodman and Englande, 1974). Prior to this time the models had dealt only with steady-state behaviour of the process. Initially, two state variables (substrate and biomass) were considered sufficient for a good dynamic description of the process, and degradation was modelled as a first order reaction as discussed above (Eckenfelder and O'Connor, 1955; McKinney, 1962; Eckenfelder, 1966). Later, saturation of the degradation capacity was included by introducing a Monod-type dependency of the removal rate on substrate concentration (Lawrence and McCarty, 1970). To describe new experimental findings, Andrews and coworkers introduced one of the first structured models. In this model the biomass was structured into three parts: active, stored and inert (Busby and Andrews, 1975). Another structured approach was suggested in McKinney and Ooten (1969) for conversion of carbonaceous material. These researchers proposed the following:

- the mixed liquor can be divided into three volatile solids fractions: active, endogenous-inert and inert (from the influent);
- a relationship between the mass of substrate utilized and the active mass of organisms was stated;
- an accumulation of endogenous-inert solids takes place because of endogenous respiration;
- a relationship between the oxygen demand and the organisms synthesized and the active mass loss due to endogenous respiration was stated;
- an accumulation of inert solids takes place due to the presence of this material in the influent wastewater.

An important factor that coincided with the development of dynamic models was the increasing computer power and falling prices for computers during the 1970s. This liberated mathematical modelling from many constraints. Systems of partial and ordinary differential equations could now be numerically solved and dynamic models could quickly be tested and validated.

Some of the most fundamental work concerning the development of dynamic models for the AS process has been performed at the University of Cape Town, Republic of South Africa. Based on the proposals of McKinney-Ooten above (except the suggestion for the rate of synthesis of active mass), Marais and Ekama (1976) developed a steady-state aerobic model from which a dynamical model evolved. Instead they accepted Lawrence-McCarty's proposal linking the specific organism growth rate to the concentration of substrate via the Monod relationship. They further suggested that influent carbonaceous material should be divided into three fractions:

- biodegradable;
- non-biodegradable particulate;
- non-biodegradable soluble.

The biochemical oxygen demand was rejected as a suitable parameter for defining the organic material and instead they accepted the electron donating capacity in its equivalent form, the chemical oxygen demand. The oxygen utilization rate was also recognized as the most sensitive parameter by which to verify the behaviour of proposed models to the activated sludge process. They further suggested that the influent nitrogen should be divided into the following four fractions:

- non-biodegradable soluble;
- non-biodegradable particulate;
- biodegradable organic;
- free and saline ammonia.

For the conversion of ammonia to nitrate they again followed the Monod approach, as set out by Downing *et al.* (1964).

Progressively the Marais-Ekama model evolved into a full dynamic model (Ekama *et al.*, 1979; Dold *et al.*, 1980; van Handel *et al.*, 1981) also including denitrification. Two key features of particular importance had then been included in the model, namely, the *bisubstrate* and *death-regeneration* hypotheses.

In accordance with practical experiments, it was proposed that the biodegradable COD in the influent wastewater consisted of two fractions: readily and slowly biodegradable COD (Ekama and Marais, 1979). This was the bisubstrate hypothesis included in the aerobic model (including nitrification) by Dold *et al.* (1980). The readily biodegradable COD was assumed to consist of simple molecules able to pass through the cell wall and immediately be used for synthesis by the organisms. The slowly biodegradable COD, which consisted of larger complex molecules, were enmeshed by the sludge mass, adsorbed and then required extracellular enzymatic breakdown (often referred to as hydrolysis) before being transferred through the cell wall and used for metabolism. The above approach was claimed to significantly improve the model predictions of the process under cyclic load and flow conditions.

The death-regeneration hypothesis was introduced in an attempt to single out the different reactions that take place when organisms die. The traditional endogenous respiration concept described how a fraction of the organism mass disappeared to provide energy for maintenance. However, practical experiments with varying anaerobic and aerobic conditions in a reactor showed that the endogenous respiration model was not satisfactory. It could not explain the rapid oxygen uptake rate that occurred when a reactor was made aerobic after an anaerobic period. In the death-regeneration model, the decayed cell material was released through lysis. One fraction was non-biodegradable and remained as an inert residue while the remaining fraction was considered to be slowly biodegradable. It could thus return to the process and be used by the remaining organisms as substrate through hydrolysis, consequently providing an explanation to the observation described above as a build up of biodegradable material during the anaerobic period.

Besides the carbonaceous conversion aspects described above, van Handel *et al.* (1981) showed that the bisubstrate and death-regeneration approach could be integrated in a consistent manner with the transformations of nitrogen.

The full UCT model (Dold *et al.*, 1991) consists of 14 different processes including 14 state variables and 21 parameters. It has provided the basis for most future mechanistic modelling approaches of the AS process, for example, for the IAWQ AS Model No.1 discussed in the next section. However, it is important to note that in order to successfully apply a model of such high complexity to a real process, an extensive measurement program in combination with estimation methods is required to determine

suitable values for the parameters and to characterize the influent wastewater. A description of methods for measuring these quantities is beyond the scope of this work. Some relevant references addressing these issues are (Ekama and Marais, 1984; Cech *et al.*, 1985; Dold and Marais, 1986; Ekama *et al.*, 1986; Henze *et al.*, 1987; Henze, 1988; Grady *et al.*, 1991; Ayesa *et al.*, 1991; Dold *et al.*, 1991; Sollfrank and Gujer, 1991; Henze, 1992; Kappeler and Gujer, 1992; Larrea *et al.*, 1992; Siegriest and Tschui, 1992; Vanrolleghem, 1994; Henze *et al.*, 1995).

The highly complex mechanistic models have initiated research to develop simpler, reduced order models for the AS process, more suited for on-line control and identification. One approach to develop a structured kinetic model for the activated sludge system is given in Padukone and Andrews (1989). The proposed model is stated to be the simplest one capable of giving a realistic description of the contact-stabilization process for carbonaceous removal although no validation of the model using experimental data is presented in the paper. Based on a traditional storage-metabolism hypothesis for the substrate, the rate equations are chosen in a way that reduce them to the Monod equation during 'balanced growth' (when the external conditions to which the cell is exposed change so slowly that its composition remains perfectly acclimated to them, for example in the completely mixed AS process). Because the rate equations are linear, the cell growth and substrate uptake in a stirred tank can be defined exactly in terms of the average composition of the biomass. The composition of the flocs is described by the ratio of stored substrate to active biomass. However, the type and number of parameters and state variables make this model difficult to verify and would require lengthy experiments in order to update the parameters for changing environmental conditions.

A simplified AS model is presented in Fujie *et al.* (1988). It predicts the concentration of organic material in the aeration basins and in the effluent from a wastewater treatment plant performing only carbonaceous removal. Only soluble organic substance is modelled since the particulate material is considered to be immediately adsorbed by the activated sludge and thereby remain within the system. The model is easily verified since practically all parameters and state variables are directly available through simple measurements. In the paper the predictions are validated against experimental data and they show a large degree of agreement. However, it has to be emphasized that the effluent concentration of organic substrate is not the most suitable variable for modelling a modern treatment plant receiving municipal wastewater. This concentration is in many cases so low that the

uncertainty of any measurement is considerable. Since many modern plants also perform nitrification/denitrification, the sludge age is usually so high that the effluent concentration of organic soluble material is more or less negligible.

A number of mechanistically simplified models for the organic substrate and the active biomass are presented and tested against each other in Sheffer *et al.* (1984). Ways of automatically selecting the best possible model for a certain purpose are also discussed as well as the need for on-line updating of model parameters. A similar comparison between different levels of mechanistic simplification of the IAWQ model to experimental data is given in Gujer and Henze (1991). Complete models for entire wastewater treatment processes, including primary settling, aeration, secondary settling, gravity thickening, anaerobic digestion, waste disposal etc., have also been proposed, for example by Tang *et al.* (1987). Such large models are usually only applicable for steady-state conditions and are mainly used to analyse the most cost-effective approach for operating an entire plant.

A somewhat different modelling approach is suggested in Benefield and Molz (1984). It is based on a modified Monod relationship and the transfer of nutrients into the flocs is modelled as spherical molecular diffusion depending on the floc radius. Biological phosphorus removal is also included in the model though in a very rudimentary form. The rate of the removal is simply stated to be directly proportional to the rate of microbial growth. The model is further investigated and validated in Benefield and Reed (1985).

In recent years, significant efforts have been made to mathematically model the processes involved in enhanced biological phosphorus removal. There are still many questions to be resolved concerning this highly complex process. Today, two main biochemical models can be recognized. They are referred to as the TCA model (Comeau *et al.*, 1985; Wentzel *et al.*, 1986) and the Glycogen model (Mino *et al.*, 1987). The basic principles of these models are discussed and compared in Wentzel *et al.* (1991). Both models recognize that stimulation of an EBPR process requires anaerobic/aerobic sequences and that VFA (volatile fatty acids), for example acetate, play a central role in the anaerobic phase. The models differ primarily with respect to the origin of the reducing equivalents (NADH) necessary for the production of poly-hydroxyalkanoate (PHA) from acetate. There is also questions with regard to which bacteria within the microbial community in a WWT plant that play the most important role

for EBPR. The mathematical models for describing EBPR are often added on as extensions to established models for carbonaceous and nitrogen removal in AS systems, applying many of the concepts incorporated in the original models. Some of the most recent proposals for modelling the EBPR process are Dupont and Henze (1989), Wentzel *et al.* (1989), Dold (1992), Wentzel *et al.* (1992), Johansson (1994), Henze *et al.* (1995) and Smolders *et al.* (1995). As EBPR is not considered in this work, the process will not be further discussed. A good reference for the interested reader is Aspegren (1995), where EBPR is investigated, both from a modelling and a practical perspective. A large number of relevant references concerning EBPR are also provided.

A very different approach when modelling the AS process is to use black-box models, e.g., stochastic processes or neural networks. A problem is the difficulty to incorporate any mechanistic knowledge about the processes into these types of models. Instead they rely heavily on identification and estimation algorithms in combination with a large and reliable database describing different dynamical aspects of the process. Although black-box models are reliable when it comes to interpolating results within a region which they have been trained for (calibrated), there is no guarantee that they will produce any relevant results when used to extrapolate data. Therefore, the use of black-box models to describe the full dynamics of the AS process is quite limited but they may prove useful for on-line control of certain well-defined parts of the process. A few relevant references dealing with black-box models (both stochastic processes and neural networks) in WWT applications are Bhat and McAvoy (1990), Hiraoka *et al.* (1990), Kabouris and Georgakakos (1991), Novotny *et al.* (1991), Boger (1992), Capodaglio *et al.* (1992), Yang and Linkens (1993), te Braake *et al.* (1994). A more fruitful use of black-box models may be to apply them *in combination* with established mechanistic models of the AS process for specific control purposes or for estimating and updating the mechanistic model parameters as the conditions of the process change.

Another, more promising approach is to use so called grey-box models for describing the AS process. In this type of model, some of the physical knowledge of the process is incorporated into a stochastic model, which means that many model parameters maintain their physical interpretation. Such a model has been successfully used for control and identification of an alternating process including nutrient removal (BIO-DENITRO and BIO-DENIPHO) and is thoroughly described in Carstensen (1994). The full potential of this modelling approach is still to be determined but significantly more research efforts are required within this new field.

3.3 A State-of-the-Art Model

In 1983, the International Association on Water Quality (IAWQ, formerly IAWPRC) formed a task group, which was to promote development, and facilitate the application of, practical models for design and operation of biological wastewater treatment systems. The first goal was to review existing models and the second goal was to reach a consensus concerning the simplest mathematical model having the capability of realistically predicting the performance of single-sludge systems carrying out carbon oxidation, nitrification and denitrification. The final result was presented in 1987 (Henze *et al.*, 1987) as the IAWQ Activated Sludge Model No. 1. Although the model has been extended since then, for example to incorporate more fractions of COD to accommodate new experimental observations (Sollfrank and Gujer, 1991), to describe growth and population dynamics of floc forming and filamentous bacteria (Gujer and Kappeler, 1992) and to include new processes for describing enhanced biological phosphorus removal (Henze *et al.*, 1995), the original model is probably still the most widely used for describing WWT processes all over the world. Due to its major impact on the WWT community it deserves some extra attention and it can still be considered as a 'state-of-the-art' model when biological phosphorus removal is not considered.

Many basic concepts were adapted from the UCT model discussed in the previous section, such as the bisubstrate hypothesis and the death-regeneration hypothesis. Again the Monod relationship was used to describe the growth rate of both heterotrophic and autotrophic organisms. COD was selected as the suitable parameter for defining the carbonaceous material as it provides a link between electron equivalents in the organic substrate, the biomass and the oxygen utilized. Furthermore, mass balances can be made in terms of COD.

Some substantial modifications were also proposed by the IAWQ task group with regard to the UCT model in terms of the enmeshment-adsorption (storage) and in the solubilization (hydrolysis) concepts. The task group rejected the view that the biodegradable particulate COD was adsorbed and stored on the organism mass. Instead they proposed that the enmeshed biodegradable material was hydrolysed to readily biodegradable COD, and released to the bulk liquid by the action of extracellular enzymes secreted by the organism mass. With regard to denitrification, the group separated the processes of hydrolysis and growth. Finally, the fate of the organic nitrogen and source of organic nitrogen for synthesis were treated

somewhat differently. The task group also introduced the concept of switching functions to gradually turn process rate equations on and off as the environmental conditions were changed (mainly between aerobic and anoxic conditions). The switching functions are 'Monod-like' expressions that are mathematically continuous and thereby reduce the problems of numerical instability during simulations. Furthermore, the work of the group promoted the structural presentation of biokinetic models via a matrix format, which was easy to read and understand, and consolidated much of the existing knowledge on the AS process. A complete description of all differences between the UCT and IAWQ models is given in Dold and Marais (1986). The full IAWQ AS Model No. 1, in the original matrix format, is provided in Appendix B.

As a comparison, the fourteen process equations of the UCT model were reduced to eight in the IAWQ model whereas the number of state variables were only reduced by one (from fourteen to thirteen). An evaluation of the two models (Dold and Marais, 1986; Dold *et al.*, 1991) revealed more or less identical predictions under most operating conditions when the models had been properly calibrated.

State Variables

The carbonaceous material in the IAWQ model is divided into biodegradable COD, non-biodegradable COD (inert material) and biomass, see Figure 3.5. The biodegradable COD is further divided into readily biodegradable substrate (S_S) and slowly biodegradable substrate (X_S). The readily biodegradable substrate is hypothesized to consist of simple soluble molecules that can be readily absorbed by the organisms and metabolized for energy and synthesis, whereas the slowly biodegradable substrate is assumed to be made up of particulate/colloidal/complex organic molecules that require enzymatic breakdown prior to absorption and utilization. Note that a fraction of the slowly biodegradable substrate may actually be soluble although it is treated as a particulate material in the model. The non-biodegradable COD is divided into soluble (S_I) and particulate (X_I) material. Both are considered to be unaffected by the biological action in the system. The inert soluble material leaves the system by the secondary clarifier effluent, whereas the inert particulate material is enmeshed in the sludge mass and accumulates as inert VSS (volatile suspended solids). The inert particulate material will be removed from the system by the removal of excess sludge and to some extent be present in the settler effluent as well. Moreover, the active biomass is divided into two types of organisms:

heterotrophic biomass ($X_{B,H}$) and autotrophic biomass ($X_{B,A}$). Finally, an extra state variable (X_P) for modelling the inert particulate products arising from biomass decay is included, similar to the endogenous mass of McKinney and Ooten (1969).

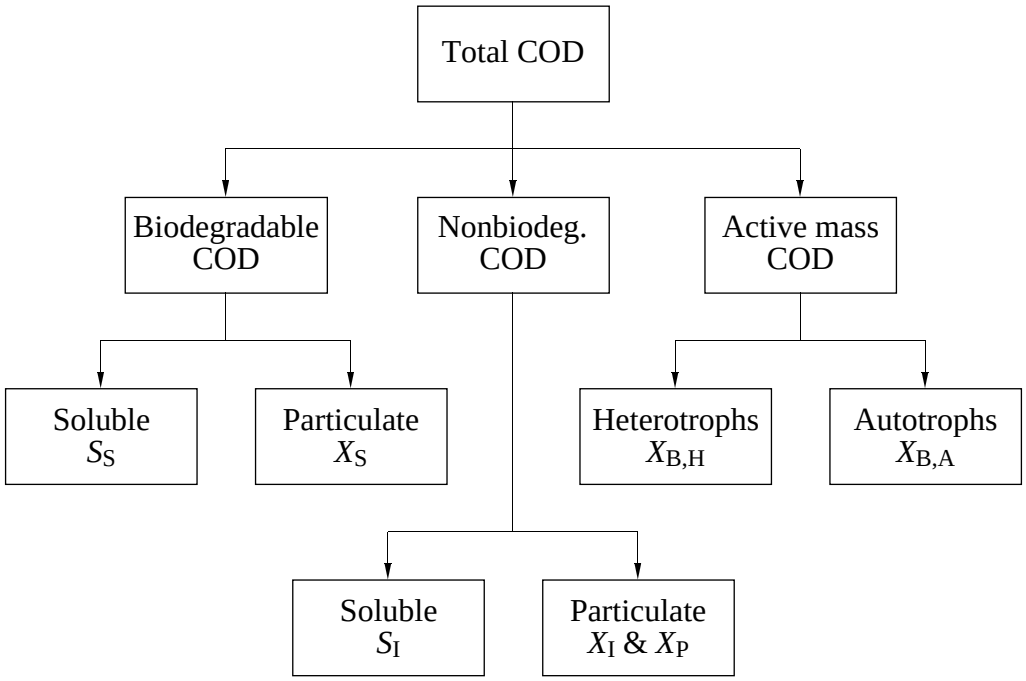


Figure 3.5 Wastewater characterization for carbonaceous components.

The nitrogenous material in the wastewater is divided according to Figure 3.6. Based on measurements of total Kjeldahl nitrogen (TKN), the nitrogen is divided into free and saline ammonia (S_{NH}), organically bound nitrogen and active mass nitrogen, that is, a fraction of the biomass which is assumed to be nitrogen. Similar to the division of the organic material, the organically bound nitrogen is divided into soluble and particulate fractions, which in turn may be biodegradable or non-biodegradable. It should be noted that only particulate biodegradable organic nitrogen (X_{ND}) and soluble biodegradable organic nitrogen (S_{ND}) are explicitly included in the model. The active mass nitrogen (X_{NB}) is included in the model only in the sense that decay of biomass will lead to a production of particulate biodegradable organic nitrogen. Organic nitrogen associated with the inert organic particulate products (X_{NP}) and the inert organic particulate matter (X_{NI}) can easily be calculated, although not described in the model matrix. No inert soluble nitrogen is modelled (S_{NI}). Finally, nitrate and nitrite nitrogen are combined into one variable (S_{NO}), as a way of simplifying the model.

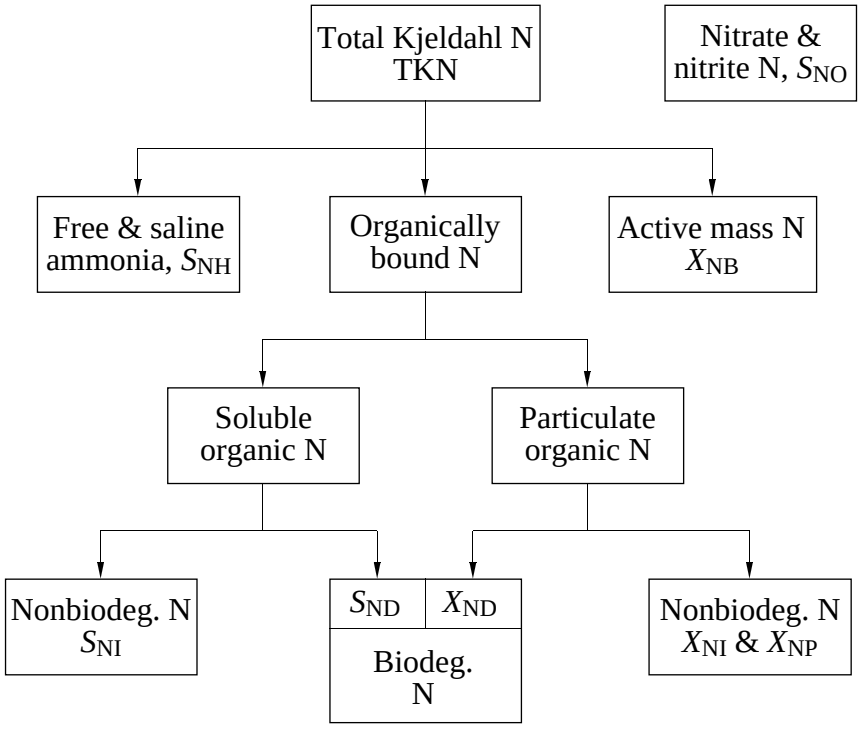


Figure 3.6 Wastewater characterization for nitrogenous components.

The last two components described in the IAWQ model are the dissolved oxygen concentration (S_O), expressed as negative COD, and the alkalinity (S_{ALK}).

Dynamic Processes

The different processes incorporated in the IAWQ model are briefly described below.

- *Aerobic growth of heterotrophs*: A fraction of the readily biodegradable substrate is used for growth of heterotrophic biomass and the balance is oxidized for energy giving rise to an associated oxygen demand. The growth is modelled using Monod kinetics. Ammonia is used as the nitrogen source for synthesis and incorporated into the cell mass. Both the concentration of S_S and S_O may be rate limiting for the growth process. This process is generally the main contributor to the production of new biomass and removal of COD. It is also associated with an alkalinity change.
- *Anoxic growth of heterotrophs*: In the absence of oxygen the heterotrophic organisms are capable of using nitrate as the terminal electron acceptor with S_S as substrate. The process will lead to a

production of heterotrophic biomass and nitrogen gas (denitrification). The nitrogen gas is a result of the reduction of nitrate with an associated alkalinity change. The same Monod kinetics as used for the aerobic growth is applied except that the kinetic rate expression is multiplied by a factor η_g (< 1). This reduced rate could either be caused by a lower maximum growth rate under anoxic conditions or because only a fraction of the heterotrophic biomass is able to function with nitrate as electron acceptor. Ammonia serves as the nitrogen source for cell synthesis, which in turn changes the alkalinity.

- *Aerobic growth of autotrophs*: Ammonia is oxidized to nitrate via a single-step process (nitrification) resulting in production of autotrophic biomass and giving rise to an associated oxygen demand. Ammonia is also used as the nitrogen source for synthesis and incorporated into the cell mass. The process has a marked effect on the alkalinity (both from the conversion of ammonia into biomass and by the oxidation of ammonia to nitrate) and the total oxygen demand. The effect on the amount of formed biomass is small as the yield of the autotrophic nitrifiers is low. Once again the growth rate is modelled using Monod kinetics.
- *Decay of heterotrophs*: The process is modelled according to the death-regeneration hypothesis (Dold *et al.*, 1980). The organisms die at a certain rate and a portion of the material is considered to be non-biodegradable and adds to the X_P fraction. The remainder adds to the pool of slowly biodegradable substrate. The organic nitrogen associated with the X_S becomes available as particulate organic nitrogen. No loss of COD is involved and no electron acceptor is utilized. The process is assumed to continue with the same rate under aerobic, anoxic and anaerobic conditions.
- *Decay of autotrophs*: The process is modelled in the same way as used to describe decay of heterotrophs.
- *Ammonification of soluble organic nitrogen*: Biodegradable soluble organic nitrogen is converted to free and saline ammonia in a first-order process mediated by the active heterotrophs. Hydrogen ions consumed in the conversion process results in an alkalinity change.
- *Hydrolysis of entrapped organics*: Slowly biodegradable substrate enmeshed in the sludge mass is broken down extracellularly, producing readily biodegradable substrate available to the organisms for growth. The process is modelled on the basis of surface reaction kinetics and occurs only under aerobic and anoxic conditions. The rate of hydrolysis is reduced under anoxic conditions compared with aerobic conditions by

a factor η_h (<1). The rate is also first-order with respect to the heterotrophic biomass present but saturates as the amount of entrapped substrate becomes large in proportion to the biomass.

- *Hydrolysis of entrapped organic nitrogen*: Biodegradable particulate organic nitrogen is broken down to soluble organic nitrogen at a rate defined by the hydrolysis reaction for entrapped organics described above.

Model Parameters

The selection of values for the kinetic and stoichiometric coefficients of a mathematical model is known as model calibration. In the case of activated sludge models, the calibration has traditionally been carried out through specific and well-controlled experiments at pilot and bench-scale plants assuming constant operating conditions. However, the values obtained in such a way may not be totally reliable for two prime reasons. The first reason being the difficulty of configuring and operating a small-scale plant in exactly the same way as a full-scale plant and thereby introducing a risk of changing the behaviour of the microorganism population and also the conditions that influence the values of the parameters which should be determined. The second reason is that the experiments and calculations are often based on the fact that the coefficients are constants. Since the experiments may take several days or even weeks to perform, they are not carried out very often. Many of the parameters are time variant and some of them may change considerably over a limited period of time. Factors such as plant configuration, operating conditions, microorganism population dynamics, degree of inhibition by toxic compounds, composition of the influent wastewater, temperature, pH, etc., all affect the values of the process parameters. The same type of problem is even more emphasized for characterizing the influent wastewater. While the parameters discussed above may change their values considerably over a period of a few days, the characteristics of the influent wastewater may change significantly within a few hours. The fact that the influence of the influent wastewater composition on the model behaviour is usually large, further amplifies these difficulties.

By examining the sensitivity, variability, and uncertainty of the model parameters, an indication is given as to which coefficients are most important to determine accurately. Such an investigation is performed in Henze (1988) for the IAWQ model. It is stated that for plants performing

nitrification and denitrification, the model show little sensitivity with regard to the COD due to the long mean cell residence time. The parameters that are considered to be the most important ones for this type of process are the

- decay rate of heterotrophs;
- growth rate for anoxic growth of heterotrophs;
- maximum specific hydrolysis rate;
- half-saturation coefficient for hydrolysis;
- correction factor for anoxic hydrolysis;
- maximum specific growth rate of autotrophs.

In Henze (1988) it is also demonstrated how different sets of parameter values may lead to approximately the same model behaviour. This is due to the fact that many model coefficients are correlated. It implies that parameters can often not be adjusted one by one, but rather a whole set must be tuned simultaneously. Some examples of such interrelations are given below.

- Growth rate and decay rate – increased growth and decay rate may produce an identical net growth rate but will increase the oxygen demand and speed up the substrate cycling.
- Yield and growth rate – increased yield and growth rate may outbalance each other with respect to substrate conversion rate but will increase the oxygen consumption.
- Yield and heterotrophs in the influent wastewater – high yield and a low concentration of heterotrophs in the influent wastewater is equal to a low yield and a high concentration of heterotrophs in the influent.

The situation outlined above is an indication that methods for identifying and estimating the non-measurable state variables and model parameters have to be employed. This should be done in order to extract all possible information from available on-line measurements as well as from laboratory investigations.

As an example, values for the model parameters suggested by the IAWQ task group are presented in Table 3.1. Note that many parameter values are strongly influenced by the environmental conditions and should be regarded more as average values indicating a reasonable order of magnitude. As a comparison, values commonly found in the literature are provided for some of the coefficients.

IAWQ model parameters	symbol	unit	20 ˚C	10 ˚C	literature
Stoichiometric parameters					
Heterotrophic yield	Y_H	g cell COD formed (g COD oxidized)$^{-1}$	0.67	0.67	0.38-0.75
Autotrophic yield	Y_A	g cell COD formed (g N oxidized)$^{-1}$	0.24	0.24	0.07-0.28
Fraction of biomass yielding particulate products	f_P	dimensionless	0.08	0.08	–
Mass N/mass COD in biomass	i_{XB}	g N (g COD)$^{-1}$ in biomass	0.086	0.086	–
Mass N/mass COD in products from biomass	i_{XP}	g N (gCOD)$^{-1}$ in endogenous mass	0.06	0.06	–
Kinetic parameters					
Heterotrophic max. specific growth rate	$\hat{\mu}_H$	day^{-1}	6.0	3.0	0.6-13.2
Heterotrophic decay rate	b_H	day^{-1}	0.62	0.20	0.05-1.6
Half-saturation coefficient (hsc) for heterotrophs	K_S	g COD m^{-3}	20	20	5-225
Oxygen hsc for heterotrophs	$K_{O,H}$	g O_2 m^{-3}	0.20	0.20	0.01-0.20
Nitrate hsc for denitrifying heterotrophs	K_{NO}	g NO_3-N m^{-3}	0.50	0.50	0.1-0.5
Autotrophic max. specific growth rate	$\hat{\mu}_A$	day^{-1}	0.80	0.30	0.2-1.0
Autotrophic decay rate	b_A	day^{-1}	0.20	0.10	0.05-0.2
Oxygen hsc for autotrophs	$K_{O,A}$	g O_2 m^{-3}	0.4	0.4	0.4-2.0
Ammonia hsc for autotrophs	K_{NH}	g NH_3-N m^{-3}	1.0	1.0	–
Correction factor for anoxic growth of heterotrophs	η_g	dimensionless	0.8	0.8	0.6-1.0
Ammonification rate	k_a	m^3 (g COD day)$^{-1}$	0.08	0.04	–
Max. specific hydrolysis rate	k_h	g slowly biodeg. COD (g cell COD day)$^{-1}$	3.0	1.0	–
Hsc for hydrolysis of slowly biodeg. substrate	K_X	g slowly biodeg. COD (g cell COD)$^{-1}$	0.03	0.01	–
Correction factor for anoxic hydrolysis	η_h	dimensionless	0.4	0.4	–

Table 3.1 Typical model parameter values at neutral pH (Henze *et al.*, 1987).

Model Formulation

Based on the above description, we can now formulate the full set of ordinary differential equations, making up the IAWQ AS Model No. 1 (not taking the flow terms into consideration). Each model equation is written explicitly, in order to demonstrate the full complexity which is somewhat hidden when using the matrix format.

The dynamic behaviour of the heterotrophic biomass concentration is affected by three different processes – aerobic growth, anoxic growth and decay – according to

$$\frac{dX_{B,H}}{dt} = \left[\hat{\mu}_H \left(\frac{S_S}{K_S + S_S} \right) \left\{ \left(\frac{S_O}{K_{O,H} + S_O} \right) + \eta_g \left(\frac{K_{O,H}}{K_{O,H} + S_O} \right) \left(\frac{S_{NO}}{K_{NO} + S_{NO}} \right) \right\} - b_H \right] X_{B,H} \tag{3.11}$$

The situation for the autotrophic biomass concentration is simpler since the autotrophs do not grow in an anoxic environment. Consequently,

$$\frac{dX_{B,A}}{dt} = \left[\hat{\mu}_A \left(\frac{S_{NH}}{K_{NH} + S_{NH}} \right) \left(\frac{S_O}{K_{O,A} + S_O} \right) - b_A \right] X_{B,A} \tag{3.12}$$

The concentration of readily biodegradable substrate is reduced by the growth of heterotrophic bacteria (in both aerobic and anoxic conditions) and is increased by hydrolysis of slowly biodegradable substrate and the differential equation describing this is

$$\frac{dS_S}{dt} = \left[-\frac{\hat{\mu}_H}{Y_H} \left(\frac{S_S}{K_S + S_S} \right) \left\{ \left(\frac{S_O}{K_{O,H} + S_O} \right) + \eta_g \left(\frac{K_{O,H}}{K_{O,H} + S_O} \right) \left(\frac{S_{NO}}{K_{NO} + S_{NO}} \right) \right\} + k_h \frac{X_S / X_{B,H}}{K_X + (X_S / X_{B,H})} \left\{ \left(\frac{S_O}{K_{O,H} + S_O} \right) + \eta_h \left(\frac{K_{O,H}}{K_{O,H} + S_O} \right) \left(\frac{S_{NO}}{K_{NO} + S_{NO}} \right) \right\} \right] X_{B,H} \tag{3.13}$$

The concentration of slowly biodegradable substrate is increased by the recycling of dead bacteria according to the death-regeneration hypothesis and decreased by the hydrolysis process according to

$$\frac{dX_S}{dt} = (1-f_P)\left(b_H X_{B,H} + b_A X_{B,A}\right) - k_h \frac{X_S/X_{B,H}}{K_X + \left(X_S/X_{B,H}\right)} \left\{ \left(\frac{S_O}{K_{O,H}+S_O} \right) + \eta_h \left(\frac{K_{O,H}}{K_{O,H}+S_O} \right) \left(\frac{S_{NO}}{K_{NO}+S_{NO}} \right) \right\} X_{B,H} \tag{3.14}$$

The shortest model equation is the one describing the concentration of inert particulate products arising from biomass decay, which is simply

$$\frac{dX_P}{dt} = f_P\left(b_H X_{B,H} + b_A X_{B,A}\right) \tag{3.15}$$

Similar to (3.14) the concentration of particulate organic nitrogen is increased by biomass decay and decreased by the hydrolysis process. The differential equation becomes

$$\frac{dX_{ND}}{dt} = \left(i_{XB} - f_P i_{XP}\right)\left(b_H X_{B,H} + b_A X_{B,A}\right) - k_h \frac{X_{ND}/X_{B,H}}{K_X + \left(X_S/X_{B,H}\right)} \left\{ \left(\frac{S_O}{K_{O,H}+S_O} \right) + \eta_h \left(\frac{K_{O,H}}{K_{O,H}+S_O} \right) \left(\frac{S_{NO}}{K_{NO}+S_{NO}} \right) \right\} X_{B,H} \tag{3.16}$$

The concentration of soluble organic nitrogen is affected by ammonification and hydrolysis, according to

$$\frac{dS_{ND}}{dt} = \left[-k_a S_{ND} + k_h \frac{X_{ND}/X_{B,H}}{K_X + \left(X_S/X_{B,H}\right)} \left\{ \left(\frac{S_O}{K_{O,H}+S_O} \right) + \eta_h \left(\frac{K_{O,H}}{K_{O,H}+S_O} \right) \left(\frac{S_{NO}}{K_{NO}+S_{NO}} \right) \right\} \right] X_{B,H} \tag{3.17}$$

The ammonia concentration is affected by growth of all microorganisms as ammonia is used as the nitrogen source for incorporation into the cell mass. The concentration is also decreased by the nitrification process and increased as a result of ammonification of soluble organic nitrogen. This leads to a complex differential equation formulated as

$$\frac{dS_{NH}}{dt} = \left[-i_{XB}\hat{\mu}_H \left(\frac{S_S}{K_S + S_S} \right) \left\{ \left(\frac{S_O}{K_{O,H} + S_O} \right) + \eta_g \left(\frac{K_{O,H}}{K_{O,H} + S_O} \right) \left(\frac{S_{NO}}{K_{NO} + S_{NO}} \right) \right\} + k_a S_{ND} \right] X_{B,H} - \hat{\mu}_A \left(i_{XB} + \frac{1}{Y_A} \right) \left(\frac{S_{NH}}{K_{NH} + S_{NH}} \right) \left(\frac{S_O}{K_{O,A} + S_O} \right) X_{B,A} \quad (3.18)$$

The concentration of nitrate is only involved in two processes – it is increased by nitrification and decreased by denitrification. The dynamic equation describing this is formulated below.

$$\frac{dS_{NO}}{dt} = -\hat{\mu}_H \eta_g \left(\frac{1 - Y_H}{2.86 Y_H} \right) \left(\frac{S_S}{K_S + S_S} \right) \left(\frac{K_{O,H}}{K_{O,H} + S_O} \right) \left(\frac{S_{NO}}{K_{NO} + S_{NO}} \right) X_{B,H} + \frac{\hat{\mu}_A}{Y_A} \left(\frac{S_{NH}}{K_{NH} + S_{NH}} \right) \left(\frac{S_O}{K_{O,A} + S_O} \right) X_{B,A} \quad (3.19)$$

Finally, the oxygen concentration in the wastewater is reduced by the aerobic growth of heterotrophic and autotrophic biomass, according to

$$\frac{dS_O}{dt} = -\hat{\mu}_H \left(\frac{1 - Y_H}{Y_H} \right) \left(\frac{S_S}{K_S + S_S} \right) \left(\frac{S_O}{K_{O,H} + S_O} \right) X_{B,H} - \hat{\mu}_A \left(\frac{4.57 - Y_A}{Y_A} \right) \left(\frac{S_{NH}}{K_{NH} + S_{NH}} \right) \left(\frac{S_O}{K_{O,A} + S_O} \right) X_{B,A} \quad (3.20)$$

We do not present the differential equation describing the dynamics of the alkalinity. Equations (3.11)–(3.20) clearly shows why the matrix format (see Appendix B) is preferred for describing this type of complex model. On the other hand, the matrix format creates an illusion for the non-experienced reader that, for example, the IAWQ model is not very complex. Equations (3.11)–(3.20) demonstrate the opposite.

A few final comments regarding the IAWQ model equations are required. The factor 2.86 in the stoichiometric expression for anoxic growth of heterotrophic biomass in (3.19) is the oxygen equivalence for conversion of nitrate nitrogen to nitrogen gas and it is included to maintain consistent units on a COD basis. The value is theoretical and means that if all the organic matter added to the denitrification reactor were only converted to CO_2 and H_2O, it would require $1/2.86 = 0.35$ g NO_3–N for each g COD removed. Similarly, the 4.57 term in the stoichiometric expression for aerobic growth of autotrophs in (3.20) is the theoretical oxygen demand associated with the oxidation of ammonia nitrogen to nitrate nitrogen, i.e., 4.57 g O_2/g NH_3–N is consumed. Due to the death-regeneration hypothesis used in the model, the heterotrophic decay rate is not the traditionally decay parameter used to describe endogenous decay, instead the value is significantly larger. If we denote the traditional decay rate by b'_H, the two decay rates are related according to

$$b_H = \frac{b'_H}{1 - Y_H(1 - f_P)} \tag{3.21}$$

Note that the specific decay rate coefficient for autotrophic bacteria, b_A, in the IAWQ model, is numerically equivalent to the traditional decay rate constant. This follows from the fact that the recycling of organic matter that results from decay occurs through the activity of the heterotrophic biomass and not by the autotrophic biomass. Also the coefficient f_P, representing the fraction of the biomass that ends up as inert particulate products following decay, is affected by the death-regeneration description. If the decay is modelled as endogenous decay, this value is usually assumed to be approximately 0.2 (i.e., 20 %), whereas the recycling of biomass by death-regeneration results in the use of a significantly lower value in order to end up with the same amount of particulate inert mass. If we denote the fraction of inert material following a traditional decay approach by f'_P, the two coefficients are related according to

$$f_P = \frac{(1 - Y_H)}{1 - Y_H f'_P} f'_P \tag{3.22}$$

It is naturally important to be aware of this type of special interpretation of various model parameters when attempting to calibrate the model to a real AS process.

Model Restrictions

A certain number of simplifications and assumptions must be made in order to make a model of a WWT system practically useful. Some of these are associated with the physical system itself, while others concern the mathematical model. A number of such restrictions concerning the IAWQ model are listed in Henze *et al.* (1987) and are summarized below.

- The system operates at constant temperature. In order to allow for temperature variations an Arrhenius equation may be used to adjust the model parameters within a certain region.
- The pH is constant and near neutrality. The inclusion of alkalinity in the model allows the user to detect potential problems with pH control.
- No consideration has been given to changes in the nature of the organic matter within any given fraction (e.g., the readily biodegradable substrate). Therefore, the coefficients in the rate expressions have been assumed to have constant values. This means that changes in the wastewater character cannot be properly handled by the model.
- The effects of limitations of nitrogen, phosphorus and other inorganic nutrients on the removal of organic substrate and on cell growth have not been considered. Thus, care must be taken to be sure that sufficient quantities of inorganic nutrients are present to allow for balanced growth.
- The correction factors for denitrification are fixed and constant for a given wastewater.
- The coefficients for nitrification are assumed to be constant and to incorporate any inhibitory effects that other waste constituents are likely to have on them.
- The heterotrophic biomass is homogeneous and does not undergo changes in species diversity with time. This means that effects of substrate concentration gradients, reactor configuration, etc. on sludge settleability is not considered.
- The entrapment of particulate organic matter in the biomass is assumed to be instantaneous.
- Hydrolysis of organic matter and organic nitrogen are coupled and occur simultaneously with equal rates.
- The type of electron acceptor present does not affect the loss of active biomass by decay.

3.4 Controllability of the Process

Traditionally, biological WWT processes have been regarded as more or less self-controlled and quite inflexible in their operation. The plants normally function under pseudo steady-state conditions for long periods of time, which are suddenly interrupted by abrupt failures. Some of these instabilities can be attributed to sudden, external disturbances of high amplitude but most are probably due to the propagation of slowly variable, internal perturbations in the largely inaccessible microbiological state of the system. The available control usually depends on the expertise of the human plant operators in combination with a few automatic, single-loop controllers. Since the early seventies, when a major leap forward was made by the widespread introduction of dissolved oxygen control, limited progress has been made.

Measurement Problems

The earliest models for describing the AS process were based on state variables which were quite readily measurable. Model calibration was performed using results obtained by operating continuous-flow plants at steady-state conditions for different sludge retention times. Today, mechanistic models have evolved considerably. In order to precisely explain the different phenomena occurring, many state variables and model parameters that are not directly measurable, have been introduced.

It is obvious that sensor technology for WWT applications has not evolved as fast as the complexity of the mathematical models. In order to use the highly complex models for the AS process available today for practical applications, a significant effort is required for model calibration in terms of exhaustive measuring campaigns, designing and performing special identification experiments, maintaining a sophisticated laboratory with highly trained technicians, etc. Process identifiability may be enhanced by means of exciting the system, perturbing the input and control signals in an optimal manner, properly choosing the sampling instants, using various methods of signal processing, etc. The proper design of identification experiments is a very troublesome task but of the utmost importance in order to produce reliable results, see (Ljung, 1987; Söderström and Stoica, 1989). A few major drawbacks are the large amount of resources required (both equipment and personnel), the high degree of uncertainty (two equally skilled persons may reach quite different results when performing

identical experiments due to the need for subjective interpretation of many results), the long time to perform certain experiments (some results may be obsolete by the time they are reached), and the lack of standardized methods (different methods for determining the same quantity may show considerable variations). Furthermore, many parameters used in complex models have to be considered as constants because of the practical difficulties of performing identification experiments as often as would be required to keep track of their variation.

Another possibility is to use simplified, reduced order models and take full advantage of the variables which are measurable on-line, that is, design models that are better adjusted to the current level of sensor technology. In combination with mathematical identification-estimation algorithms such models could be automatically calibrated on-line and always be tuned to the current situation of the plant (adaptive models). The drawback in this case is the lack of reliable on-line sensors (only a few variables can be accurately measured). This may lead to overly simplified models, which are not capable of producing any realistic predictions (the number of parameters which can be accurately estimated is related to the amount and quality of the available data). Moreover, the cost and need of maintenance for advanced instrumentation are quite high.

A combination of the two approaches is naturally an alternative. Depending on the purpose of the model and for what time scale the model is to be used, the best procedure can be selected. A model used for design of new plants simulates plant behaviour over long periods of time. The variations which are of real relevance are those with time constants of days, weeks or even longer. On the other hand, an oxygen regulator for an aerobic tank reacts within minutes and relevant parameter changes with time constants of seconds and minutes have to be detected. This can only be accomplished by on-line measurements.

It is interesting to note that the main measurement problem for WWT plants is usually not lack of data. On the contrary, large amounts of information are being measured and stored at a modern plant. The number of inputs from sensors usually varies between a few hundred and several thousand. The problem is often that the available measurements are not very relevant for modelling purposes. Moreover, the data must be available with adequate accuracy and frequency.

Traditionally, on-line measurements have been restricted to physical-chemical variables such as flow rates (of both water, sludge and air), power

to pumps, levels in reactors, temperature, pH, redox potential, water conductivity, etc. The introduction of on-line sensors for measuring the oxygen concentration of the wastewater led to intensive research and development of regulators, estimators and models for controlling the dissolved oxygen concentration as a key variable. This work has continued over the last two decades. The sensors are now very reliable and oxygen consumption is considered the key variable used for verifying mathematical models as well as for control purposes. On-line measurements of the oxygen concentration is today used in combination with model-based estimators to identify the oxygen dynamics of the activated sludge process, that is, on-line estimation of the oxygen transfer function ($K_L a$) and the respiration rate (oxygen utilization rate), see e.g., Koo *et al.* (1982), Holmberg and Olsson (1985), Holmberg *et al.* (1989), Holmberg (1991), Carlsson and Wigren (1993), Carlsson and Lindberg (1994).

During the last 5 to 10 years, it appears as if the problem of unreliable and unavailable on-line sensors has become more pronounced and a lot of research and development is currently directed towards this very important field. Not surprisingly, this new development coincides with the more widespread implementation of biological nutrient removal at many WWT plants. A comprehensive review of existing and new sensor technology was recently presented by Vanrolleghem and Verstraete (1993). Developments are many and increasingly sophisticated devices are proposed in an attempt to provide the necessary information on the complex processes needed to meet strict effluent standards. Table 3.2 summarizes the available sensor technology, the processes in which they can be implemented and the range of applicability, that is, the extent to which they are considered proven technology. Two significant trends in the developments of new on-line monitoring equipment are the application of ultrafiltration systems to bring automated wet chemistry to the WWT plant on the one hand and the combination of robust, proven sensor technology with extended data interpretation on the other hand.

However, not even sophisticated laboratory analyses and new sensor technology are enough to solve all problems related to the calibration of complex models. Many models are inherently unidentifiable. As an example, the IAWQ model contains five stoichiometric coefficients, fourteen kinetic parameters, and five non-measurable state variables (Larrea *et al.*, 1992).

Physical measurements			Physical-Chemical measurements			(Bio-)Chemical measurements		
Variable	Applicability Process ↓		Variable	Applicability Process ↓		Variable	Applicability Process ↓	
Temperature	G	⊕	pH	G	⊕	Respiration rate	2,3	⊕
Pressure	G	⊕	Conductivity	G	⊕	stBOD[4]	2,3	⊕
Liquid level	G	⊕	Oxygen			Toxicity	2,3	⊕
Flow rates			– Liquid	2,3	⊕	Sludge activity	2,3	⊕
– Liquid	G	⊕	– Gas	2,3	⊕	COD	1,2,3	∅
– Gas	1,2,3	⊕	Digester gas			TOC	1,2,3	∅
Suspended solids			– CH_4	1	⊕	NH_4	3	∃
– 0.0-0.1 g/l	4	∃	– H_2S	1	⊕	NO_3	3	∃
– 1.0-10 g/l	1,2,3	∃	– H_2	1	⊕	PO_4	3	∃
– 10-100 g/l	4	∃	CO_2	1,2,3	⊕	Bicarbonate	1,3	∅
Sludge blanket	4	∃	Flourescence			Volatile fatty acids	1,3	∅
Sludge volume	4	∃	– NAD(P)H	2,3	∃			
Settling velocity	4	∅	– F_{420}	1	∅			
Sludge morphology	G	∅	Redox	1,3	⊕			
Heat generation	1,2,3	∅	NH_4 (ISE[3])	3	∃			
UV absorption	G	∃	NO_3					
			– ISE	3	∅			
			– UV absorbance	3	∃			

[1]Applicability range
⊕: state of technology; ∃: applicable in certain cases; ∅: requires development work.
[2]Unit process in the wastewater treatment plant where the sensor can be implemented
1: anaerobic digestion; 2: activated sludge; 3: nutrient removal; 4: sedimentation; G: all.
[3]ISE: ion selective electrode.
[4]stBOD: short-term biological oxygen demand.

Table 3.2 On-line monitoring equipment for wastewater treatment processes (Vanrolleghem and Verstraete, 1993).

A general description of an extensive procedure to determine the unknown parameters of the model is given in Henze *et al.* (1987). It is a combination of practical experiments and curve-fitting procedures and it is clearly stated that an error introduced when determining certain coefficients will be compensated when determining another parameter. This might seem satisfactory but is actually an indication of lacking model identifiability (non-unique solutions), as different sets of model parameters will produce identical results.

Establishing Cause-Effect Relationships

During the last decade the complexity of the AS process has increased significantly with the introduction of biological nitrogen and phosphorus removal. This complexity in combination with the ever stricter legislative requirements on the effluent wastewater quality is today the major driving force for developing new control strategies, more sophisticated sensors and improving plant flexibility. It must also be recognized that suitable mathematical models of the processes are prerequisites for any successful implementation of sophisticated control strategies.

A WWT process is hardly ever in steady state, mostly due to load variations. As a consequence its dynamical properties and its response to changes has to be known if the plant is going to be consistently controlled towards a desired result. A traditional way to solve the problem of variable loads has been to increase the tank volumes. A better operation, however, can offer the possibility to calculate the trade-off between design and operational costs.

Since load variations have to be accepted, any operation has to make sure that the detrimental influence of any disturbance is minimized. The control and operation problem naturally has to focus on *disturbance rejection*. Still it is an open question how much operation (as opposed to design) can improve a WWT system.

One way of approaching the control problem is to realize that many of the control variables act in different time scales and, consequently, they may look as if they are quite independent of each other. Their influence can be observed via sensors or estimation procedures. In an advanced nutrient removal system it becomes increasingly important to monitor the system on the microbial level, and floc structures and organism compositions are of crucial importance.

A key question for nutrient removal systems is whether or not the process is sufficiently controllable with the existing control variables. This can qualitatively be expressed as: given an undesired operational condition (such as bulking sludge, excessive foam formation or poor sludge settleability), are there any operational procedures that can bring the plant from the present state to some desired state? At present there is insufficient knowledge available to answer this question completely. In particular, for undesired microbial conditions, there is no operational procedure that can be derived from known mathematical models, since the knowledge of these

conditions is far too limited (Albertson, 1991; Gabb *et al.*, 1991; Jenkins *et al.*, 1993). In order to further improve the controllability of plants, a much better understanding of microbial conditions and the plant parameters that may influence them, has to be established.

Often the goals of the plant operation are not clearly stated. One apparent goal is to satisfy the effluent requirements *consistently*. Furthermore the costs should be minimized while maintaining the water quality. Frequently, control criteria are mixed up with *constraints*. The goal of the operation is *not* to keep the dissolved oxygen concentration in the aerobic zone at 2 mg/l, maintaining the mixed-liquor suspended solids (MLSS) concentration at a fixed value or to ensure that the sludge retention time exceeds ten days. These values are chosen set-points that contribute to keep the plant running properly.

The main objectives of WWT plant operations may be categorized into the following groups:

- maintaining liquid and sludge inventories;
- maintaining required effluent quality;
- disturbance rejection;
- efficient operation and reduction of cost.

Liquid inventories are usually well taken care of for the total plant. However, as the settler unit is highly sensitive to hydraulic disturbances, any damping of hydraulic disturbances is important. Plants using parallel channels are designed for hydraulic symmetry. In practice, however, the flow symmetry has to be guaranteed by flow meters and control actions. For sludge inventories, more remains to be done at many plants and the mass balances have to be more closely maintained.

The product quality in terms of effluent carbon content is not a big issue anymore. Biological nitrogen and phosphorus removal are still challenges, in particular in cold climates. Most often the sludge settleability is not sufficiently controllable. Consequently, it may be difficult to guarantee the effluent suspended solids concentration.

For the disturbance rejection it is usually impossible to remove the source, even if this is an important option, especially for industrial effluents, where production control in the industry may be improved. Sometimes the magnitude of the disturbance can be reduced *before* it reaches the WWT plant. An integrated sewer-treatment plant control can attenuate hydraulic

disturbances, and some waters can be pretreated in order to avoid harmful effects on the plant. The control can also compensate for effects *within* the process. A common example is dissolved oxygen control.

An efficient operation can be obtained in many ways. A close monitoring of the plant by on-line measurements, estimation of non-measurable parameters and diagnosis of operational conditions are all important. Methods like statistical monitoring ought to be used to a greater extent. By maintaining certain concentration profiles of dissolved oxygen or sludge, more efficient operations can be obtained. Furthermore, alternative designs – such as sequential batch or alternating systems – can offer better operational flexibility.

The dynamics of the AS system spans several orders of magnitude, from seconds to months. The fact that the phenomena can be grouped into several classes will allow many control actions to be decoupled. As a result, slow phenomena can be controlled while considering the fast dynamics instantaneous. For the control of fast varying variables, the slow modes are considered constant. This approach is also useful for model development, in the sense that models can be simplified based on their range of applicability in time. From a complete set of models describing the different dynamics of the process, a fairly simple model suited for a specific purpose can be selected.

In a first attempt to systematically describe the many cause-effect relationships of the AS process, an incidence matrix is proposed, see Table 3.3. Manipulated variables and disturbances define the columns and measurable or estimated variables define the rows of the matrix. Each matrix element indicates the influence from one manipulated variable to a measurable variable of the AS process. The element is marked with a letter indicating the time scale of the dynamics of this particular interaction. However, the amplitude of the cause-effect is not indicated. The incidence matrix is by no means complete and there are still many unknown elements in the matrix, especially the relationships between manipulated variables and floc formation and microbial composition.

Table 3.3 (Next two pages) Incidence matrix of the AS process, based on Olsson and Jeppsson (1994a). F indicates fast (minutes), M means medium (hours) and S indicates slow (≥ days) dynamic influence. Empty boxes indicate too small or unknown effect, while ? means an unknown but probable influence.

CAUSE / EFFECT	Infl. flow	WAS flow	RAS flow	RAS distr.	NO_3 circ.	Chem. add.	C add.
REACTOR (measurements)							
DO total	M	S	M	M	F		M
DO profile	M	S	M	M	F		M
MLSS	M	S	M	M	M	F	S
COD	M	S	M	M	M		M
NH_4–N	M	S	M	M	M		
NO_2–N and NO_3–N	M	S	M	M	M		M
Redox			M		M		M
Water temp.	M						
REACTOR (estimations)							
Total mass	M	S	S	S	S	F-M	S
Organic load (F/M)	M	S	M	M	S	M	M
SRT, sludge age	M-S	S			?		
Oxygen transfer						F?	
Respiration rate	M	S	M	M	M		M
CO_2 production	M	S	M	M	M		M
Nitrification rate	M	S	M	M	(M)		
Denitrification rate	M	S	M	M	F		F
Sludge volume, SVI	S	?		?	?	M	
Biomass (VSS, enzymes, ATP)	M-S	S	S	S	S	F-M	S
Foam formation	?	?	?	?	?		
SETTLER (measurements)							
Effluent SS	F	?	F	M		F-M	
Sludge blanket level	F-M	M	F			M	
Return AS concentration	M		F			M	
Wastage AS concentration	M		F			M	
SETTLER (estimations)							
Settler influent flow rate	F	M	F	F			
Total mass	M	S	F			M	
VSS	M-S					M	
Initial settling velocity						M	
Clear depth	F	M	F			M	
Floc structure	?	?	?	?	?	?	
Bulking index	?	?	?	?	?	?	
Filamentous organisms	?	?	?	?	?	?	

Air total	Air distr.	CODs infl.	CODp infl.	NH_4 infl.	Toxic infl.	Temp. infl.	pH infl.	Super-natant	Back-wash
F	F	F	M	F	F	M-S	M	M	F
F	F	M	M	M	F	M-S	M	M	F
S	S	S	S	S	S	S	S	M	F
S	S	M	M	M	M	S	S	M	M
M	M	M	M	F	M	M		M	M
M	M			M	M	M		M	M
F	F			M	F				F
						M			F
S	S	S	S	S	S	S	S		M
S	S	M	M	M	M	S	M	M	M
					S	S			M
F	F					M			
F	F	F	M	F	F	M	M	M	M
F	F	F	M	F	F	M	M	M	M
F	F	M	M	F	F	M	M	M	M
F	F	M	M		F	M	M	M	F
?	?				M			?	?
S	S	S	S	S	S	S	S	M	M
?	?				?			?	?
?	?	?	?	?	M			M	F
					S				
					S				
					S				
?	?	?	?	?	S	?			?
					?				F
S?	S?	S?	S?		?				?
S?	S?	S?	S?		?				?
S?	S?	S?	S?		?				?

Table 3.3 indicates a major difficulty when attempting to control the AS process. When one variable in the process is manipulated it affects a large number of measurable variables in a very complex way. Over a longer period of time, the total system behaviour may change considerably. The ideal situation would naturally be that a change of one control variable would produce a well-defined response in one process variable, instead the situation is practically the opposite. This is an indication that it may be necessary to use control strategies based on multi-variable control to solve the problems in some cases.

From a modelling point of view a similar problem exists. A changing process behaviour will naturally have an impact on the microbial population in the system as they adapt to new environmental conditions (due to changing influent wastewater characteristics, operational modifications, physical-chemical variations, etc.). The model parameters can consequently not be regarded as constants but must be dynamically updated as the plant conditions change. More adequate and adaptive models are necessary in order to predict both the long and short term behaviour of the AS system and to determine and implement better control strategies.

Chapter 4

Reduced Order Models

In this chapter a set of reduced order models for the activated sludge process is developed based on a number of simplifying assumptions. A comparison of the dynamic behaviour of the reduced order models and the IAWQ model is performed, considering different types of process operations. The possibility to globally identify the parameters of the reduced order models is first investigated using an off-line optimization method. The identifiability of the parameters is tested based on different assumptions of what measurements are available (using the IAWQ model to simulate the 'true' process). The sensitivity of the reduced order models to parameter changes during normal operating conditions is also presented. Finally, an on-line estimation algorithm is tested for similar conditions as used for the off-line method, in order to evaluate the possibility of identifying the simplified models from on-line measurements. The material in this chapter is covered in [181], [182], [183] and [185].

4.1 Model Assumptions and Development

The basis for the development of reliable mathematical models is a thorough understanding of the involved processes. The understanding may to some extent be replaced by the use of stochastic models, a fast computer, proper software and a sufficient amount of experimental data. This is the field of system identification discussed in Chapter 2, which can in its ultimate form, be described as multi-dimensional curve-fitting procedures.

Physical modelling is, however, an analytical approach where basic laws from physics, chemistry, etc. are used to describe the behaviour of a process. Based on such process knowledge, a model suited for its predefined

purposes can be hypothesised. Its structure and behaviour may then be analysed using available tools (cf. Sections 2.3–2.5) and step by step further tested, modified and validated.

One of the main difficulties when developing a model is often to determine which reactions are the most significant ones and to describe these in a simple, yet comprehensive manner. A good physical model should realistically mimic the true dynamics of the process in question but still contain a minimum number of variables and parameters while maintaining the physical interpretation of those.

Overall Considerations

Activated sludge systems are usually described by mathematical models based on mass balance equations. These equations relate the changes of the state variables of the system (i.e., concentrations) due to transport and transformation mechanisms. Transport mechanisms are characteristic for the design and physical outline of a system (reactor configuration, distribution of the influent, mixing, excess sludge removal, etc.) but they leave the chemical structure of all material unchanged. A mass balance equation for a single component within a defined system boundary can be described according to

input – output + reaction = accumulation

The eight transformation processes used in the IAWQ AS Model No.1 (Henze *et al.*, 1987) were discussed in detail in Section 3.3. The model is considered to be particularly useful for the prediction of

- biological degradation of organic material and denitrification;
- nitrification;
- the distribution of oxygen consumption along a 'plug-flow' reactor in the course of diurnal variations;
- sludge production;
- variation in effluent quality during dynamic loading conditions.

In order for the above to be completely true there is normally a need to combine the biological model with other models describing for example the settling process and the oxygen transfer mechanisms. However, in this chapter we will focus on the biological mechanisms and problems related

to other process units will not be considered. Although the IAWQ model comprises much of the current knowledge of the biological reactions – in a fairly simple manner – a number of drawbacks exist. Apart from the model restrictions listed in Section 3.3 we have summarized some of those drawbacks below.

- Lacking identifiability – different sets of model parameters will produce inseparable results.
- Lacking verifiability – certain state variables and parameters are not directly measurable and, therefore, it is difficult to experimentally validate all aspects of the model.
- Limited understanding and knowledge about some of the described processes (for example, the hydrolysis mechanism).
- Troublesome practical characterisation of the influent wastewater, although essential for the model behaviour.
- Difficult to estimate and update the varying model parameters (functions of time, pH, load, temperature, etc.) on-line.
- Troublesome non-linearities (Monod functions, switching functions, etc.).
- Not useful for on-line control applications.
- Expert knowledge required to understand all internal model interactions (i.e., complex cause-effect relationships).
- Sophisticated instrumentation and laboratory facilities required for calibration and validation purposes in combination with expensive (time and money) measurement campaigns.

The aim of the work presented in this chapter is to approach some of the problems listed above and develop reduced order models which can adequately describe both carbonaceous and nitrogenous activities for the purpose of on-line control. A fundamental requirement is that the models contain a minimum number of state variables and parameters to allow for model identification based on available on-line measurements.

General mathematical methods for reduction of non-linear models are still difficult to apply (as discussed in Section 2.3). In order to maintain the basic mechanistic structure of the reference model and describe the significant reactions in a physically reasonable manner, the model reduction is instead based on traditional reasoning and means of analysis. Due to the widespread use, the general acceptance and the mechanistic structure of the IAWQ model, it was selected as the best reference model for the model reduction study.

Measurable Variables

Models for on-line control purposes must be related to quantities and variables which are possible to measure on-line. This is especially important as the model parameters are not constant but vary with time and operational conditions. Results from laboratory experiments and bench scale tests should be used for validation and further improvements of the model predictions whenever possible but not be of vital importance for the basic reliability and performance of the model.

Quantities and variables which are possible to measure and quantify on-line in the activated sludge process were discussed in Section 3.4. New types of sensors and measurement technologies are also continuously being developed. For the work presented in this chapter the following set of quality variables are assumed to be measurable on-line:

- biodegradable organic substrate concentration;
- ammonia nitrogen concentration;
- nitrate nitrogen concentration;
- volumetric flow rates.

In some cases, measurements of the oxygen uptake rate (OUR) or respiration rate are considered to be available as well (for example, by means of a respirometer).

Cost, accuracy, sensitivity and repeatability of the above measurements are not considered nor are the practical aspects of where to place the sensors, how data should be transferred to the computer systems, etc. Such questions are of great importance at a later stage of the work. But until the more basic and fundamental questions concerning the reduced order models have been investigated, the measurement quality issues are overlooked.

Simplifying Assumptions

The assumptions for simplifying the IAWQ model are from a physical and biological viewpoint mainly based on a discussion of how the following components are treated in the model:

- dissolved oxygen;
- organic matter;
- nitrogen;
- microorganisms.

For the reduced models, measurements of the dissolved oxygen (DO) concentration are not considered although the DO sensor is usually regarded as the most reliable on-line instrument for the activated sludge process. This is because the oxygen concentration is excluded as a state variable. It is assumed that the DO is controlled separately on a routine basis, so that corresponding growth expressions become independent of DO variations.

Still, the DO mass balance contains a lot of useful information. It is, for example, the basis for the estimation of the oxygen uptake rate, which is recognized as fundamental information for future control aspects. Consequently, the models describing the DO, oxygen transfer rate, blowers, etc., are considered to make up an important but separate module of an hierarchical control structure as outlined in Figure 2.4. The different submodels need to be synchronized with the overall control of the plant. Such an approach makes it possible to separate the biological model from the oxygen model on the first level of control. It also allows for a more strict boundary between the anoxic and aerobic environments from a modelling point of view. The existing DO control is assumed to provide a sufficient amount of oxygen in the aerobic reactor while minimizing the oxygen concentration into the anoxic reactor. The above model separation can also be motivated by considering the different time scale of the process dynamics. The dynamics with regard to the DO concentration have time constants in the range of seconds to a few minutes whereas the time constants for the biological reactions vary from hours all the way up to weeks.

The above model separation may appear as too rough a simplification. However, it must be remembered that we are discussing models for control and not models for design. Naturally the different models will interact but such interactions will be handled by a supervisory control system, whereas the various small models performs their specific tasks. Moreover, the approach does not imply that the biological models are insensitive to the DO concentration. The effects of DO changes are rather combined with other inhibitory circumstances and reflected as variations in the estimated growth rate factors whereas the IAWQ model uses a switching function to single out the effects of different DO concentrations.

The assumption of a constant DO concentration within a reactor at a specific time is valid only if the reactors are truly completely mixed. The situation is often quite different at real WWT plants. Experiments show that an aerobic reactor which is assumed to be completely mixed, with a DO probe in the centre of the tank and connected to a control system with a DO setpoint of 2 mg/l, may actually have a DO concentration of 0.5 mg/l

close to the influent and 5 mg/l close to the effluent. The expected value exists only close to the sensor. The reasons for such differences are mainly due to improper mixing, varying concentrations of available substrate and lacking control flexibility. There is no easy way of modelling these effects of non-ideal reactors. A method often applied to improve the situation is to model the aerator as a number of reactors (each completely mixed) in series to reduce the discrepancies. Another possibility is to use a plug-flow model and add on effects of turbulent diffusion or use partial differential equations to describe the spatial distribution (as well as the distribution in time) of the concentration variations for all the model variables including DO. Such approaches will result in more complex models. Compared with the problems discussed above, the assumption of having completely anoxic and aerobic reactors is only a small additional simplification. The major simplification in this respect is actually to assume completely mixed reactors.

The description of the organic matter represents the second considerable difference of the reduced models compared with the IAWQ model. In the IAWQ model four fractions of organic matter are considered:

- soluble inert organic matter (S_I);
- readily biodegradable substrate (S_S);
- particulate inert organic matter (X_I and X_P);
- slowly biodegradable substrate (X_S).

All the above fractions are replaced by a single variable in the reduced models (X_{COD}), which is considered to be made up of all biodegradable organic matter and is assumed to be directly measurable. The approach can be motivated in several ways.

The two inert fractions are not important from a biological point of view. The S_I fraction simply follows the wastewater flow and passes through a WWT plant without having any effect and the particulate inert fraction is used for predicting the total amount of sludge in the system to determine the wastage and recirculation rates. Variations of the amount of sludge in the system is a slow process and, consequently, it is not necessary to include this in a model for control with a predictive time horizon in the range of hours. On the other hand, the two biodegradable fractions are of the utmost importance for describing the biological reactions. S_S is considered to be directly available for the microorganisms while X_S first has to be enzymatically broken down into S_S (the hydrolysis mechanism) before the organisms can use it for metabolism.

The hydrolysis process is, however, not very well understood. The IAWQ description of it is quite complex but still a simplification of the true reaction. Due to the uncertainty and complex description of the mechanism it is not included in the reduced order models.

Another reason for lumping the biodegradable organic matter together is the difficulty of measuring the S_S and X_S fractions separately. In a laboratory scale experiment it is possible to monitor the oxygen uptake rate of a small batch reactor and thereby determining an average value of the two fractions. To do this on-line at a full-scale plant is more difficult although a respirometer can be used to determine the so called short-term COD content of a wastewater. In practice, COD measurements on filtered samples of the wastewater is often considered to be equal to the amount of S_S. However, there is no evidence that all soluble biodegradable matter is readily biodegradable and that all particulate biodegradable matter is slowly biodegradable. Moreover, the time constant for the hydrolysis process may be very different for various organic components. Recent extensions of the IAWQ model suggests that X_S should be divided into two variables, rapidly and slowly hydrolysable COD, where the rapidly hydrolysable organics are primarily soluble in nature (Sollfrank and Gujer, 1991; Henze, 1992).

The organic matter that is received by a WWT plant includes all kinds of different molecular structures. Different organisms deal with different substrates in different time scales, which makes it probable that an entire set of biodegradation processes with time constants ranging from fast to slow biodegradability is at work here. Since there is no apparent upper limit to the number of substrates which would really need to be included, the opposite solution is suggested in this work, that is, we model only one type of organic biodegradable substrate. This means that the reduced models do not take rapid uptake phenomena into consideration. Instead it makes some averaging of biosorption and growth by combining soluble and stored organic substrate. Consequently, fast dynamics (in the order of less than an hour) are neglected. Together with the earlier discussed way of modelling the DO concentration, these simplifications make the models less stiff, that is, the ratio between the smallest and the largest time constants is reduced. The complexity and the number of model parameters are naturally also significantly reduced and the possibility to end up with an identifiable model structure is increased.

The third major difference between the IAWQ model and the reduced models concerns the nitrogen components. In the IAWQ model four fractions of nitrogen are considered:

- nitrate and nitrite nitrogen (S_{NO});
- ammonia nitrogen (S_{NH});
- soluble biodegradable organic nitrogen (S_{ND});
- particulate biodegradable organic nitrogen (X_{ND}).

The only two nitrogen fractions included in the reduced order models are the nitrate nitrogen and the ammonia nitrogen, which are both assumed to be measurable on-line. The reason for this is firstly to reduce the complexity of the model structure and the number of parameters. Secondly, the two organic nitrogen fractions S_{ND} and X_{ND} primarily describe the internal formation of S_{NH} by hydrolysis and ammonification (see Section 3.3). In the reduced models, ammonia nitrogen is assumed to be measured and, therefore, its formation mechanism is not considered to be crucial for control purposes. Moreover, as the hydrolysis mechanism was excluded to describe the transformation of organic matter it should consequently not be used to describe the transformation of nitrogen (as it is basically the same process according to the IAWQ model).

The two types of microorganisms described in the IAWQ model (and many other AS models) are maintained in the reduced models, i.e.,

- active heterotrophic biomass ($X_{B,H}$);
- active autotrophic biomass ($X_{B,A}$).

In the reduced models heterotrophs are considered to grow in both anoxic and aerobic environments whereas autotrophs only grow in an aerobic environment. A death-regeneration principle (see Section 3.2) is also used to describe the decay of the organisms but in a modified way. The decayed biomass is considered to transform into biodegradable COD and ammonia nitrogen directly. In the IAWQ model the decay material is suggested to be partly inert (a small fraction) and partly transformed into X_S and X_{ND} which after hydrolysis and ammonification become available as S_S and S_{NH}, respectively.

The Reduced Order Model

The simplifications discussed in the previous subsection has reduced the number of state variables to five compared with the twelve state variables of the original IAWQ model (alkalinity is not considered). The reaction mechanisms for hydrolysis of entrapped organics, hydrolysis of entrapped organic nitrogen, and ammonification of soluble organic nitrogen have also

been removed, mainly due to measurement problems, uncertainties of the actual processes, and the need to reduce the overall model complexity.

It is possible to continue the simplification procedure one step further. As illustrated by the example in Section 2.6 there are good reasons to reconsider the parametrization of the Monod growth rate expression and the similar switching functions. The idea is to approximate the Monod functions by linear functions (cf. Figure 2.8), that is, use a first-order reaction followed by a zero-order reaction (Blackman, 1905). If we further note that for most operating conditions of WWT plants receiving municipal wastewater the substrate concentrations are generally quite low, then there is only a need to model a first-order reaction.

As a result of the assumed existing DO control, the DO influence on the switching functions is constant. The switching functions are regarded solely as functions describing growth limitation due to low concentrations of DO or different types of substrates. Therefore, the estimated parameters of the first-order rate equations include both maximum specific growth rates and possible limitations by DO, COD, nitrate, etc. The switching functions are consequently removed.

The differential equations for the first reduced order model (model A) of the activated sludge process can now be formulated. It describes carbonaceous oxidation as well as nitrification and denitrification according to the simplifications discussed above. Altogether three summary processes are proposed to describe the anoxic environment – growth of heterotrophs, decay of heterotrophs and decay of autotrophs – and four parameters should be estimated – r_H, Y_H, b_H and b_A – preferably on-line. For anoxic conditions the following simplified model is suggested:

$$\frac{dX_{COD}}{dt} = -\frac{1}{Y_H} r_H X_{COD} X_{B,H} + b_H X_{B,H} + b_A X_{B,A} \tag{4.1}$$

$$\frac{dS_{NH}}{dt} = -i_{XB}\left(r_H X_{COD} X_{B,H} - b_H X_{B,H} - b_A X_{B,A}\right) \tag{4.2}$$

$$\frac{dS_{NO}}{dt} = -\frac{1-Y_H}{2.86 Y_H} r_H X_{COD} X_{B,H} \tag{4.3}$$

$$\frac{dX_{B,H}}{dt} = \left(r_H X_{COD} - b_H\right) X_{B,H} \tag{4.4}$$

$$\frac{dX_{B,A}}{dt} = -b_A X_{B,A} \tag{4.5}$$

In an aerobic environment four main mechanisms are defined – growth of heterotrophs, growth of autotrophs, decay of heterotrophs and decay of autotrophs – and six parameters need to be updated – r_H, r_A, Y_H, Y_A, b_H and b_A. The following set of equations is suggested:

$$\frac{dX_{COD}}{dt} = -\frac{1}{Y_H} r_H X_{COD} X_{B,H} + b_H X_{B,H} + b_A X_{B,A} \tag{4.6}$$

$$\frac{dS_{NH}}{dt} = -i_{XB}\left(r_H X_{COD} X_{B,H} - b_H X_{B,H} - b_A X_{B,A}\right) - \left(i_{XB} + \frac{1}{Y_A}\right) r_A S_{NH} X_{B,A} \tag{4.7}$$

$$\frac{dS_{NO}}{dt} = \frac{1}{Y_A} r_A S_{NH} X_{B,A} \tag{4.8}$$

$$\frac{dX_{B,H}}{dt} = \left(r_H X_{COD} - b_H\right) X_{B,H} \tag{4.9}$$

$$\frac{dX_{B,A}}{dt} = \left(r_A S_{NH} - b_A\right) X_{B,A} \tag{4.10}$$

where: r_H = reaction rate factor for heterotrophs [l (mg day)$^{-1}$];
r_A = reaction rate factor for autotrophs [l (mg day)$^{-1}$];
Y_H = yield factor for heterotrophs;
Y_A = yield factor for autotrophs;
b_H = decay rate coefficient for heterotrophs [day^{-1}];
b_A = decay rate coefficient for autotrophs [day^{-1}];
i_{XB} = mass N/mass COD in biomass.

The model is also described using the traditional matrix format in Appendix C. By comparing equations (4.1)–(4.10) with the full set of equations describing the IAWQ model, that is, equations (3.11)–(3.20), the simplicity of the reduced order model is obvious. The factor 2.86 (see equation 4.3) in the stoichiometric coefficient for anoxic growth of heterotrophic biomass is the oxygen equivalence for conversion of nitrate nitrogen to nitrogen gas included to maintain consistent units on a COD basis. The parameter i_{XB} is considered to be a constant with a value of 0.086, as suggested by the IAWQ task group (Henze *et al.*, 1987). The other parameters are considered to be time varying and need to be identified for adequate model performance. Note that the model parameter values are not assumed to be

identical for anoxic and aerobic conditions and should be estimated within their specified environment.

For purposes of comparison and estimation, the oxygen uptake rate is in some cases considered to be measurable. It is modelled in the same way as in the IAWQ model apart from the use of first-order reaction rates and the absence of switching functions according to

$$\mathrm{OUR} = \frac{1-Y_{\mathrm{H}}}{Y_{\mathrm{H}}} r_{\mathrm{H}} X_{\mathrm{COD}} X_{\mathrm{B,H}} + \frac{4.57-Y_{\mathrm{A}}}{Y_{\mathrm{A}}} r_{\mathrm{A}} S_{\mathrm{NH}} X_{\mathrm{B,A}} \qquad (4.11)$$

The factor 4.57 in the stoichiometric coefficient for aerobic growth of autotrophs is the theoretical oxygen demand associated with the oxidation of ammonia nitrogen to nitrate nitrogen.

Theoretical Identifiability

Prior to any further investigations it is important to determine if the suggested model is theoretically identifiable. If the model does not pass this test then it must be reformulated. Methods for investigating the theoretical identifiability of a model were discussed in Section 2.4. For nonlinear models the only universal technique is the Taylor series expansion of observations method. The details of this method were demonstrated in Section 2.6. By applying the same method to the reduced order model it is possible to examine its theoretical identifiability.

We start by investigating the anoxic part of the model, that is, equations (4.1)–(4.5). Assume that all five state variables are possible to measure continuously with no noise and that the model parameters are constant during the experiment. If we differentiate the model equations one more time, the following system results:

$$\frac{d^2 X_{\mathrm{COD}}}{dt^2} = -\frac{r_{\mathrm{H}}}{Y_{\mathrm{H}}}\left(X'_{\mathrm{COD}} X_{\mathrm{B,H}} + X_{\mathrm{COD}} X'_{\mathrm{B,H}}\right) + b_{\mathrm{H}} X'_{\mathrm{B,H}} + b_{\mathrm{A}} X'_{\mathrm{B,A}} \qquad (4.12)$$

$$\frac{d^2 S_{\mathrm{NH}}}{dt^2} = -i_{\mathrm{XB}}\left(r_{\mathrm{H}}\left(X'_{\mathrm{COD}} X_{\mathrm{B,H}} + X_{\mathrm{COD}} X'_{\mathrm{B,H}}\right) - b_{\mathrm{H}} X'_{\mathrm{B,H}} - b_{\mathrm{A}} X'_{\mathrm{B,A}}\right) \qquad (4.13)$$

$$\frac{d^2 S_{\mathrm{NO}}}{dt^2} = -\frac{1-Y_{\mathrm{H}}}{2.86 Y_{\mathrm{H}}} r_{\mathrm{H}}\left(X'_{\mathrm{COD}} X_{\mathrm{B,H}} + X_{\mathrm{COD}} X'_{\mathrm{B,H}}\right) \qquad (4.14)$$

$$\frac{d^2 X_{\mathrm{B,H}}}{dt^2} = r_{\mathrm{H}}\left(X'_{\mathrm{COD}} X_{\mathrm{B,H}} + X_{\mathrm{COD}} X'_{\mathrm{B,H}}\right) - b_{\mathrm{H}} X'_{\mathrm{B,H}} \tag{4.15}$$

$$\frac{d^2 X_{\mathrm{B,A}}}{dt^2} = -b_{\mathrm{A}} X'_{\mathrm{B,A}} \tag{4.16}$$

From equation (4.5) we can immediately deduce that

$$b_{\mathrm{A}} = -\frac{X'_{\mathrm{B,A}}}{X_{\mathrm{B,A}}} \tag{4.17}$$

By combining equations (4.2), (4.4) and (4.5), an analytical expression for the i_{XB} parameter can be determined as

$$i_{\mathrm{XB}} = \frac{S'_{\mathrm{NH}}}{X'_{\mathrm{B,H}} + X'_{\mathrm{B,A}}} \tag{4.18}$$

From the equations (4.4) and (4.15) the expression for r_{H} can be formulated as

$$r_{\mathrm{H}} = \frac{X'_{\mathrm{B,H}} X''_{\mathrm{B,H}} - \left(X'_{\mathrm{B,H}}\right)^2}{X'_{\mathrm{COD}}\left(X'_{\mathrm{B,H}}\right)^2} \tag{4.19}$$

By continuing the analysis we can derive analytical expressions for the remaining parameters (Y_{H} and b_{H}) as well.

In a similar way it can be shown that the aerobic part of the reduced order model, that is, equations (4.6)–(4.10) are theoretically identifiable if all five state variables are assumed to be known with perfect accuracy.

If only measurements of X_{COD}, S_{NH} and S_{NO} are assumed to be available, the identifiability analysis is more complicated. For the anoxic part of the model we can formulate an expression for $X_{\mathrm{B,H}}$ using equation (4.3) that only depends on the available state variables according to

$$X_{\mathrm{B,H}} = \frac{2.86 S'_{\mathrm{NO}}}{\left(1 - 1/Y_{\mathrm{H}}\right) r_{\mathrm{H}} X_{\mathrm{COD}}} \tag{4.20}$$

and from equations (4.1) or (4.2) we can define an expression for $X_{\mathrm{B,A}}$ (in combination with (4.20)), which depends only on the three measurable

variables. Based on these expressions defining $X_{B,H}$ and $X_{B,A}$, together with equations (4.1)–(4.5) and their derivatives, the same type of identifiability analysis can be performed. However, in this case the set of equations becomes more complex and will require lengthy calculations (or a suitable computer program for symbolic mathematical analysis, e.g., Maple™ or Mathematica™). A detailed analysis of this case (or the equivalent case describing the aerobic part of the model) has not been performed in this study although preliminary results show that b_A cannot be identified within the anoxic reactor if measurements of $X_{B,A}$ are not available. This fact may actually be directly concluded by observing equations (4.1)–(4.5), since b_A and $X_{B,A}$ always occur as a combined variable ($b_A X_{B,A}$).

Further Simplifications

While maintaining the basic structure of model A (equations (4.1)–(4.10)), it is possible to impose some further simplifications to improve the identifiability. It is not unrealistic to assume the decay rates for heterotrophs and autotrophs, respectively, to be equal for both anoxic and aerobic conditions. This will reduce the total number of parameters to be estimated from ten to eight.

Taking this approach one step further, b_H and b_A can be lumped together into one single decay rate, b, equal for all conditions. This reduces the number of unknown parameters to seven. The simplification is not totally realistic but on the other hand not very drastic – the assumption of having only two kinds of microorganisms representing the entire microbial population in a WWT plant is definitely a much more severe simplification. Moreover, it is really the net growth rate of the organisms that determine their behaviour and since the growth rates are estimated separately for the different types of organisms this simplification should be valid.

A more practical reason for such a simplification is due to the difficulty of estimating decay rates during normal plant operations. The small example in Section 2.6 indicated that the decay rate parameter was easily estimated. This was primarily because during the final stage of such a batch experiment, the decay rate is the sole factor to influence the behaviour of the process (when all available substrate has been consumed). Its identifiability is therefore enhanced. From continuous-flow reactors, however, the effects of microbial growth and decay are difficult to separate, especially when the oxygen uptake rate is not monitored. In the previous subsection we also concluded that b_A is not theoretically identifiable in the anoxic reactor if

the autotrophic organism concentration is not monitored. These problems will be more closely investigated later in this chapter.

The second proposed reduced order model which will be investigated, referred to as model B (see Appendix C), contains the same principal equations as model A. The only difference is that all parameters b_H and b_A have been replaced by a single decay rate, b, which is assumed to be identical for both anoxic and aerobic conditions.

In much the same way as the growth and decay rates may be difficult to separate, identifiability difficulties may arise when trying to estimate the yield and reaction rate factor simultaneously. A higher value for both the yield and reaction rate factor will outbalance each other with respect to the substrate conversion rate. In the example (Section 2.6), the situation was improved because both the substrate and organism concentrations were assumed to be measurable, in combination with the fact that a batch experiment excites all modes of the system as it covers a wide range of different concentrations. This is not true during traditional continuous-flow plant operation. For this reason model B may be even further reduced by assuming the same value for the heterotrophic yield, Y_H, for both anoxic and aerobic conditions (model C). It is clear from equations (4.7) and (4.8) that a similar problem exists for the growth rate and yield coefficients describing the autotrophic organisms. It may prove necessary to estimate the ratio of these parameters rather than their individual values and use some additional measurement to separate them (for example, the OUR). These final simplifications leave a total of six model parameters to be estimated. This model will, however, not be further investigated in this study.

As for the reaction rate factor of the heterotrophic organisms, r_H, it does not seem realistic to assume this parameter to be identical for both anoxic and aerobic conditions. Experiments have shown that either only a fraction of the heterotrophic biomass is able to function with nitrate as the terminal electron acceptor or the maximum specific growth rate is lower for anoxic conditions (Batchelor, 1982). The reaction rate factor must therefore be separately identified to take such effects into account. The absolute minimum realistic number of parameters to be updated would therefore be the six parameters suggested for model C – r_H (both anoxic and aerobic), r_A, Y_H, Y_A and b – where some additional measurements may be required to separate r_A and Y_A.

The above types of highly simplified models will naturally not predict all the intricate details of the more complex models, since many reactions and variables have been lumped together. Instead the models should be evaluated in view of their specific purpose, that is, on-line model-based control. However, the basic mechanistic structure of the models have been maintained when possible and, consequently, the possibility to directly interpret the results. The presented models serve a major purpose as an experimental platform to investigate how far the model reduction principles can be pursued without losing the possibility of adequately predicting the main phenomena of the true processes while simultaneously gaining the possibility to determine unique estimates of the model parameters and perform on-line model calibration.

4.2 Methods for Model Analysis

A large number of methods for analysing the properties of a model exist. Some of those were discussed in Sections 2.3–2.5. In this work we focus on investigating the practical identifiability of the reduced order models depending on which variables are assumed to be measurable. Therefore, we will study the possibility to estimate the model parameters by using one off-line and one on-line estimation method and also perform a sensitivity analysis of the reduced model for continuous-flow plant operation.

Estimation as an Optimization Problem

Parameter estimation problems can be formulated as an off-line optimization problem where the best model is the one that best fits the data according to a given criterion. Such a criterion (J) is often based on the difference between the real measurements $\boldsymbol{y}$ and the model outputs $\hat{\boldsymbol{y}}$ written in the form

$$J = \sum_{k=1}^{n} \left[\boldsymbol{y}(t_k) - \hat{\boldsymbol{y}}(t_k)\right]^{\mathrm{T}} \boldsymbol{W} \left[\boldsymbol{y}(t_k) - \hat{\boldsymbol{y}}(t_k)\right] \tag{4.21}$$

where n is the number of discrete time measurements and $\boldsymbol{W}$ is a weight matrix. J is a function of all the unknown model parameters. These are adjusted until J has obtained a minimum. If there is only one unique minimum for J (independent of the initial estimates) then the system is defined

as globally identifiable. The loss function (4.21) is an example of a weighted least squares criterion (cf. Section 2.4).

In this work the weight factors (the diagonal elements of $\boldsymbol{W}$) have been chosen so that a 10 % difference between the measured values and the model estimates around a specified steady state gives approximately the same contribution to the loss function (J) for all measured variables. This means that the measurements are considered to be of the same quality and that the model is capable of estimating all variables with the same accuracy. If this was not considered to be the case, less weight could have been given to some of the residuals. It is also possible to let the weight factors vary with time and the actual value of the measurements.

The Nelder-Mead simplex method (Nelder and Mead, 1965; Fletcher, 1987) used in this work is an *ad hoc* optimization method that applies a type of random search by calculating and examining the function value (i.e., the loss function) at several points in the state space of the model – together forming a so called simplex – and moving towards lower values until convergence. The method has already been used in the example in Section 2.6 with good results. The main advantage of the algorithm is its robustness and its insensitivity to noise, whereas the convergence rate is slow and the computational effort goes up rapidly (typically as 2^n) with the dimension (n) of the model. A more detailed description of the simplex optimization algorithm is given in Appendix D. Results of the method applied to the reduced order models are presented in Section 4.4.

The Kalman Filter

One of the most commonly used methods for on-line state estimation is the Kalman filter. It is based on the reconstruction algorithm (2.12). By updating the gain matrix $\boldsymbol{K}$ in a special way, the estimation of the states is optimal in the sense that the variance of the reconstruction error is minimized. The problem is that the disturbances and properties of the noise have to be fairly well known. For the extended Kalman filter not only the modelled states $\hat{\boldsymbol{x}}$ are updated from the available measurements but also the unknown model parameters. The filter algorithm can be divided into two phases: *prediction* and *correction.* The principal structure of the on-line identification procedure is illustrated in Figure 4.6.

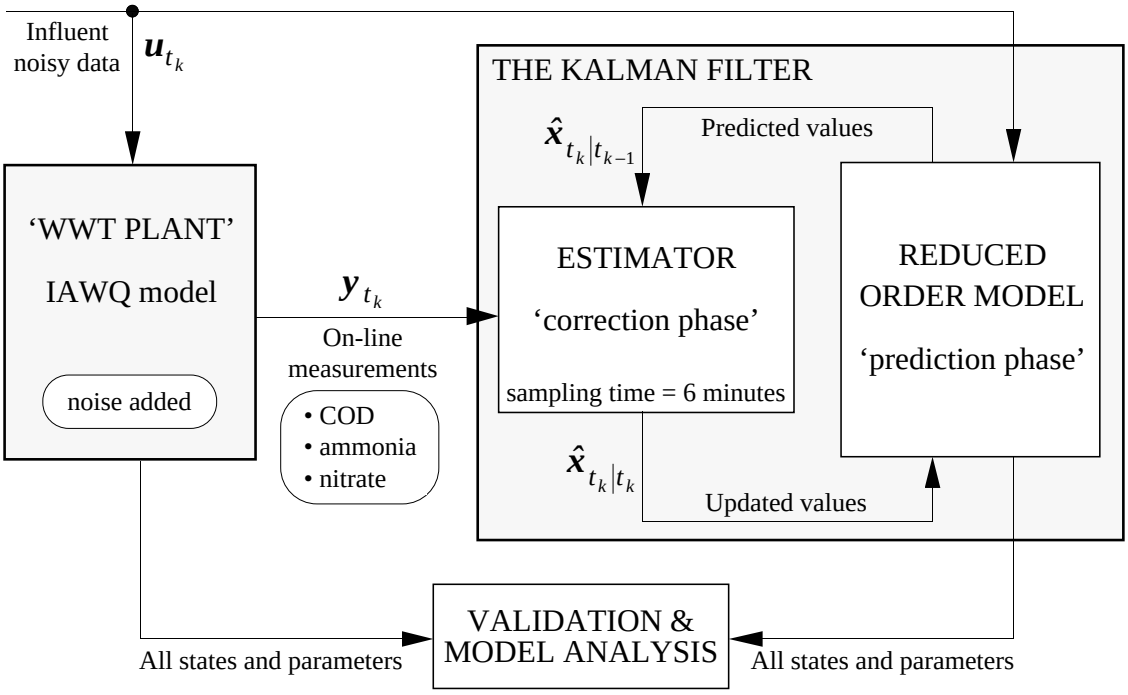


Figure 4.1 Structured identification using an extended Kalman filter.

During the prediction phase, the dynamic equations of the model are integrated between two measurements, from time t_{k-1} to time t_k (using a much smaller time step) as shown below.

$$\hat{\boldsymbol{x}}\left(t_k|t_{k-1}\right) = \boldsymbol{\Phi}\hat{\boldsymbol{x}}\left(t_{k-1}|t_{k-1}\right) + \boldsymbol{\Gamma}\boldsymbol{u}\left(t_{k-1}\right) \tag{4.22}$$

The states (in this case the concentrations) and the model parameters are now based on measurements up until time t_{k-1}. As new measurements are acquired at time t_k, they are used to update the generalized state vector. This latter part is called the correction phase and is based on the calculation

$$\hat{\boldsymbol{x}}\left(t_k|t_k\right) = \hat{\boldsymbol{x}}\left(t_k|t_{k-1}\right) + \boldsymbol{K}\left(\boldsymbol{y}\left(t_k\right) - \boldsymbol{C}\hat{\boldsymbol{x}}\left(t_k|t_{k-1}\right)\right) \tag{4.23}$$

A Kalman filter is, however, based on the assumption that the dynamics are linear, which is not the case for this application. In order to calculate $\boldsymbol{K}$, the dynamic equations are linearized around the existing operating point for each measurement instance. At the time for correction, $\boldsymbol{K}$ is calculated from the linearized equations at time t_k and depends not only on the linearized state equations but also on the properties of the noise that affects both the process states and the available measurements.

In this work the gain matrix $\boldsymbol{K}$ has been kept constant in order to simplify the computations (a constant gain extended Kalman filter (Hendricks, 1992)). The values of $\boldsymbol{K}$ will influence the convergence speed of the parameters towards their final values. The chosen values of $\boldsymbol{K}$ have been calculated according to equations (E.19) and (E.20) using the calculated steady state values of the IAWQ model (which is used to simulate the true process) as the operating point (the conditions are defined in Section 4.3) together with the selected noise characteristics. In Appendix E, a more detailed description of the extended Kalman filter is given and some of the computational results are presented in Section 4.5.

The state variables in the models have significantly different values, expressed as mg/l. In order to obtain reasonably accurate identification results, it is mandatory to *scale* the equations or *normalize* them to a reference point so that all the concentration values are expressed in the same order of magnitude. This is even more important when the parameter values are considered, that is, when performing simultaneous state and parameter estimation. If no scaling is done, the $\boldsymbol{K}$ matrix will contain such different elements that the estimation becomes numerically infeasible.

Sensitivity Analysis

Sensitivity analysis is another important tool when analysing model characteristics. It expresses the influence of a small parameter change on the state variables and can therefore provide strong indications as to which parameters are most difficult to identify either because of their limited influence on the total system behaviour or due to the fact that several parameters compensate for the effects of others. A good illustration of the latter was the combined effect of $\hat{\mu}$ and K_S in the example of Section 2.6. Several parameter sets produced approximately identical results because the effect of one parameter was compensated by another.

The method is also useful for experimental planning and design as well as for model reduction. Since the aim of this study is to identify the parameters of the reduced order models for normal operating conditions the analysis is performed for similar conditions, i.e., varying influent flow rates and concentrations, both anoxic and aerobic zone active, changing internal and sludge recirculation rates (the exact conditions are described in Section 4.3). By simulating such a process using model A, changing one parameter slightly and rerun the simulation repeatedly (Monte Carlo simulations) and storing the value of the loss function (4.21) for each

simulation, a rough 'map' describing the influence of the parameters on the model behaviour is acquired. The relative change of the parameters is identical for every simulation to make a fair comparison possible. The results of such an analysis are presented and discussed in Section 4.4.

It is often more practical to do a sensitivity analysis on decoupled systems (only anoxic or aerobic reactor) for batch conditions and without any feedback (recirculation). Results from this type of unit operation may show quite different results, which are not applicable for real continuous-flow plant operation because special modes that do not appear during normal plant operation are often emphasized. For example, the decay rate coefficient in the example was easily identified from a batch experiment because its effect was enhanced when no substrate was available (as a sensitivity analysis will show). However, a situation like that will not occur during normal operating conditions and in a continuous-flow reactor the identification will be much more difficult.

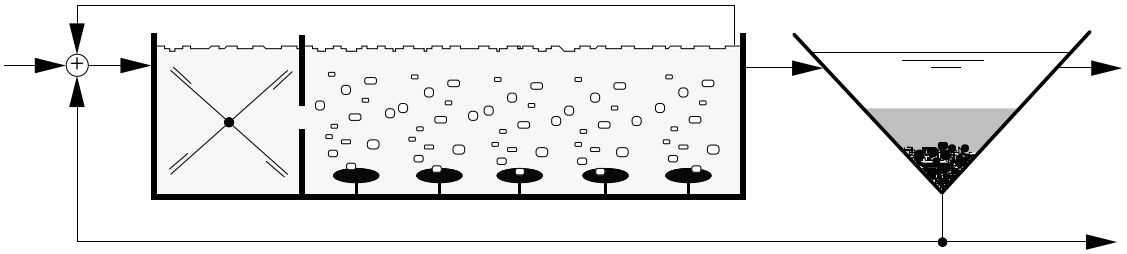
Therefore, since the purpose of this work is aimed at on-line parameter identification during normal operating conditions, it is advantageous to perform the sensitivity analysis for the same type of condition. On the other hand, if the work is aimed at experimental design, the possibility of investigating how different conditions affect the identifiability may prove very useful (e.g., Vanrolleghem, 1994).

4.3 Simulated Plant Configuration

So far no validation has been performed using real data but only with simulated data mainly based on the IAWQ model. Comparisons are made of model outputs from the IAWQ and the reduced order models to determine if they incorporate the same dynamical phenomena – both qualitatively and quantitatively – of importance when subjected to the same type of model input (i.e., influent wastewater). There are several reasons for such an approach. Since the analysis is mainly theoretical, fundamental model weaknesses can be more thoroughly investigated using simulated data because it is possible to change the noise characteristics, repeat an experiment for identical conditions but for a change in one specific variable, control the inputs to the system, etc. Furthermore, the time and effort required to collect this type of detailed data from a full-scale WWT plant have been considered to be outside the scope of this

work. When the structural modelling problems have been satisfactory solved then investigations based on real data are called for.

In order to investigate the behaviour of the reduced order models, the IAWQ model has been used to simulate the 'real' AS process. The physical outline of the simulated WWT plant includes a completely mixed anoxic reactor for pre-denitrification followed by a completely mixed aerobic reactor and a secondary clarifier. The process includes an internal recirculation stream from the aerobic to the anoxic reactor as well as sludge recycling from the thickener to the anoxic reactor. All influent wastewater is fed into the anoxic reactor. The system is operated with a sludge age of ten days and a hydraulic retention time of ten hours (in steady state). The default set of parameters for the IAWQ model at 20 °C is used for the simulations. A more detailed description of the volumes, flow rates, influent wastewater characteristics, etc., is presented in Figure 4.2.



Operational variables			Model parameters (Henze *et al.*,1987)		
Influent flow rate (Q_{in})	=	3000 m³/day	$\hat{\mu}_H$	=	6.0 day^{-1}
Recycle flow rate	=	$0.5*Q_{in}$	K_S	=	20 mg COD/l
Internal recycle flow rate	=	$3*Q_{in}$	$K_{O,H}$	=	0.2 mg O_2/l
Anoxic tank volume	=	250 m³	K_{NO}	=	0.5 mg NO_3–N/l
Aerobic tank volume	=	1000 m³	b_H	=	0.62 day^{-1}
Settler volume	=	1250 m³	$\hat{\mu}_A$	=	0.8 day^{-1}
Sludge age	=	10 days	K_{NH}	=	1.0 mg NH_3–N/l
Hydraulic retention time	=	10 hours	$K_{O,A}$	=	0.4 mg O_2/l
			b_A	=	0.2 day^{-1}
Influent wastewater characteristics			k_a	=	0.08 mg/(mg COD day)
			k_h	=	3.0 mg COD/(mg COD day)
S_S	=	30 mg COD/l	K_X	=	0.03 mg COD/(mg COD)
X_S	=	70 mg COD/l	h_g	=	0.8
S_I	=	0 mg COD/l	h_h	=	0.4
X_I	=	10 mg COD/l	Y_H	=	0.67 mg COD/(mg COD)
S_{NO}	=	2 mg N/l	Y_A	=	0.24 mg COD/(mg COD)
S_{NH}	=	10 mg N/l	f_P	=	0.08
S_{ND}	=	1 mg N/l	i_{XB}	=	0.086 mg N/(mg COD)
X_{ND}	=	1 mg N/l	i_{XP}	=	0.06 mg N/(mg COD)
S_O	=	0 mg (–COD)/l			
$X_{B,H}$	=	0 mg COD/l			
$X_{B,A}$	=	0 mg COD/l			

Setpoint for oxygen concentration in the aerobic reactor = 2.0 mg O_2/l

Figure 4.2 Configuration of the simulated WWT plant.

It is necessary to include a model of the secondary clarifier to describe the behaviour of the entire plant. Models for the settling process will be studied in detail in Chapters 5 and 6. For this study, a simple model was used because we are not interested in predicting the true effluent and underflow concentrations. The thickener is modelled as a constant compaction ratio (γ) of the underflow sludge concentration and the average sludge concentration in the reactors (Olsson and Andrews, 1978). Based on a steady state relationship over the settler, the compaction ratio can be expressed in terms of flow rates and the sludge retention time as

$$\gamma = \frac{Q_{in} + Q_r - V/\theta_X}{Q_r} \tag{4.24}$$

where: Q_{in} = influent flow rate to the WWT plant;
Q_r = sludge recycle flow rate from the settler;
θ_X = sludge retention time;
V = total bioreactor volume.

This model is highly idealized in the sense that the compaction ratio will be adjusted so that for any given flow rates the required sludge age will be maintained. The hydraulic retention time of the settler is taken into account by a subsequent time lag. It should be noted that all the biodegradable organic matter of the reduced models (X_{COD}) is considered to be part of the flocs (which is why the IAWQ nomenclature for particulate matter, X, is used) and, consequently, settles. In the IAWQ model only the X_S fraction is particulate whereas the S_S fraction is considered soluble. This difference will not have any major impact on the process behaviour since the S_S concentration is normally very low when the wastewater reaches the settler, due the long sludge age of the simulated plant. The different definitions of the biodegradable substrates in the models will have a more significant effect when the total biodegradable substrate concentration (i.e., $S_S + X_S$ in the IAWQ model) of the influent wastewater changes. In the IAWQ model only the S_S fraction is directly available for microbial growth whereas the entire X_{COD} fraction in the reduced models can be immediately used for microbial growth.

The importance of process excitation was discussed in Section 2.4. In this study the system was perturbed using pulse disturbances of the following input variables:

- the influent flow rate;
- the influent biodegradable organic substrate concentration;
- the influent ammonia concentration.

The reduced order models display a somewhat different behaviour than the IAWQ model, especially for the transient behaviour. Thus, pulse disturbances appear to be a decisive model test. The model input is not optimized from an identifiability point of view although step changes of the input signals have the advantage that they excite several modes of the process whereas sinusoidal variations excite only one specific frequency.

At a real continuous-flow WWT plant it is often possible to excite the influent flow rate in a step wise manner by changing the input pumping capacity of the plant. It is more difficult to produce a similar change of the biodegradable substrate concentration and the ammonia concentration. At certain times, abrupt changes of the above concentrations may occur due to external events and may then be used for identification purposes. It is important to take advantage of such natural variations of the influent wastewater characteristics to improve the possibility of producing good estimation results.

For the off-line estimation results presented in Section 4.4, the input variables are in most cases perturbed according to Figure 4.3. The identification is based on measurements during a ten day period which begins with a steady state period lasting for three days (based on the values given in Figure 4.2) followed by a 50 % increase of the three input variables discussed above during a two day period (each disturbance lasts for one day, see Figure 4.3). Measurements for five more days are included during which the process slowly approaches steady state. Such a time series includes steady state behaviour, fast dynamics during the disturbance period, and slow dynamics as the system settles down towards the initial steady state. For the on-line identification results in Section 4.5, other types of process perturbations have been used.

The possibility to estimate the parameters of the reduced order models will be investigated for three different cases. The basic case assumes the three following concentrations to be directly measurable on-line:

- biodegradable organic matter concentration;
- ammonia concentration;
- nitrate concentration.

The sampling time for the measurements is chosen to be six minutes. It is assumed that the above variables are measured not only in the influent wastewater stream but also in both the anoxic and aerobic reactor. The influent flow rate is also assumed to be continuously available.

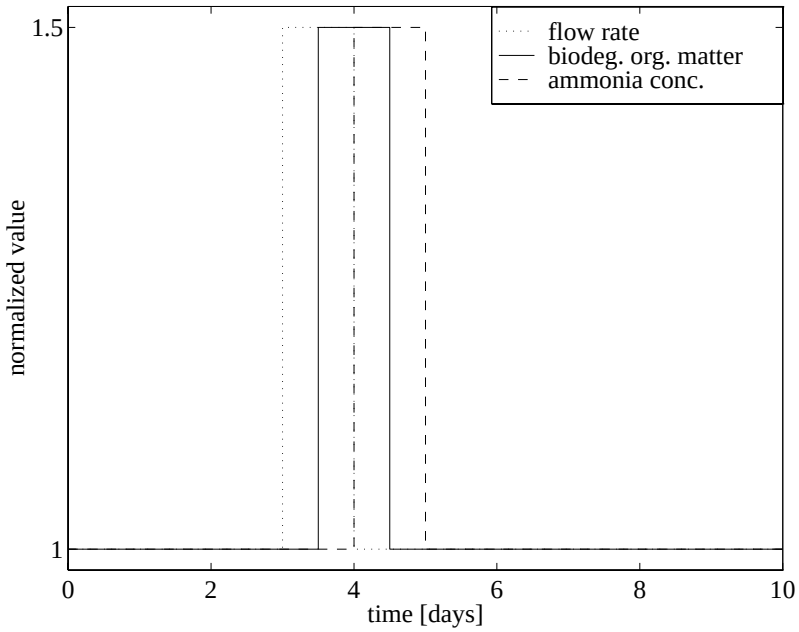


Figure 4.3 Perturbations of the influent variables (normalized values).

The second investigated case also considers the oxygen uptake rate to be continuously available as a measurable variable in the aerobic reactor. For the third case the concentrations of microorganisms (both heterotrophic and autotrophic organisms) are assumed to be measurable. Although this is not a realistic assumption, this case is included to examine the model behaviour if all state variables of the reduced order models are possible to measure directly.

The noise conditions used for the different simulations vary significantly. For the off-line estimation problem, noise is usually not added to the process because the main purpose is here to investigate the basic identification properties. If the amount of data is sufficiently large then similar results will usually be achieved whether noise is added or not (for the chosen noise distributions in this study) as was illustrated in the example in Section 2.6.

In the simulations where noise is added to the system (mainly for the on-line identification in Section 4.5), this is done in two ways. *Process* noise is simulated by adding Gaussian white noise to the variables of the IAWQ model input (influent flow rate, biodegradable organic matter and ammonia concentration, etc.). The white noise has a mean value of zero and a standard deviation which is 10 % of the steady-state value of each individual variable, that is, all input variables are exposed to the same relative noise level. *Measurement* noise (with the same properties) is added to the measurable variables in the same way as to the input variables, which

implies that all measurements are assumed to be equally uncertain. The chosen noise level is quite realistic although related problems, such as 'outliers', trends, uncalibrated sensors and sensor failures, have not been considered.

All computer simulations in this chapter have been carried out using the simulation platforms Simnon™ (SSPA Systems, 1991) and Simulink™ (MathWorks, 1995). More detailed descriptions of these two simulation programs are provided in Appendix F.

4.4 Off-Line Estimation

In order to perform an investigation of the behaviour of the reduced order models and the possibility to identify the parameter sets from the type of data available from full-scale WWT plant, a number of simulations are carried out. These analyses do not provide proof whether a model is fully identifiable or not but they give strong indications of the major characteristics of the models and point out some of their potential weaknesses.

The assumed physical outline of the plant, the variations and character of the influent wastewater and the measurable variables, were all defined in the previous section together with the chosen values of the IAWQ model parameters used to simulate the true WWT plant. The off-line optimization algorithm (the Nelder-Mead simplex method) and the type of loss function applied, were discussed in Section 4.2 and the algorithm is described in more detail in Appendix D. In Section 4.1 the reduced order models were developed and discussed, based on the biological and physical processes involved. In this section they will be further investigated mainly from an identifiability point of view. Note that all the off-line estimations presented in this section are based on simulated data without any noise added.

The results and conclusions are based on a large number of massive computations. Only a limited number of these can be presented here. The case studies to be discussed below are selected to illustrate how the results depend on the following situations:

- availability of different measurable variables;
- unit or coupled optimization of the two reactor types;
- identification based on models A or B.

An examination of the sensitivity of the loss function (4.21) to parameter changes will also be presented using model A for both unit and coupled optimization. Such an analysis may explain some results of the parameter estimations.

All results and graphs in this chapter are presented using the units milligramme [mg] (mg COD or mg N depending on the variable), litre [l] and day [day]. During the actual computations the variables are scaled to avoid numerical problems.

Case 1 – Anoxic Reactor Using Model A

The first presented case is an investigation of the part of model A describing an anoxic environment, that is, equations (4.1)–(4.5). Data are generated by simulating the IAWQ model of an entire plant according to the description given in Section 4.3. The variables describing the total input to the anoxic reactor (i.e., the combination of influent wastewater, internal recirculation and sludge recirculation) as well as the internal variables of the reactor are stored and then used for the off-line optimization. The anoxic part of model A is then simulated using the stored influent data.

The loss function on which the optimization is based, is calculated as the sum of weighted squares of the residuals, see (4.21). These residuals are the difference between the measurable variables of the IAWQ and the reduced model in the anoxic reactor. The simplex method finally suggests a new set of parameters for model A and the procedure is then repeated until an optimum solution is reached. The optimization procedure is illustrated in Figure 4.4.

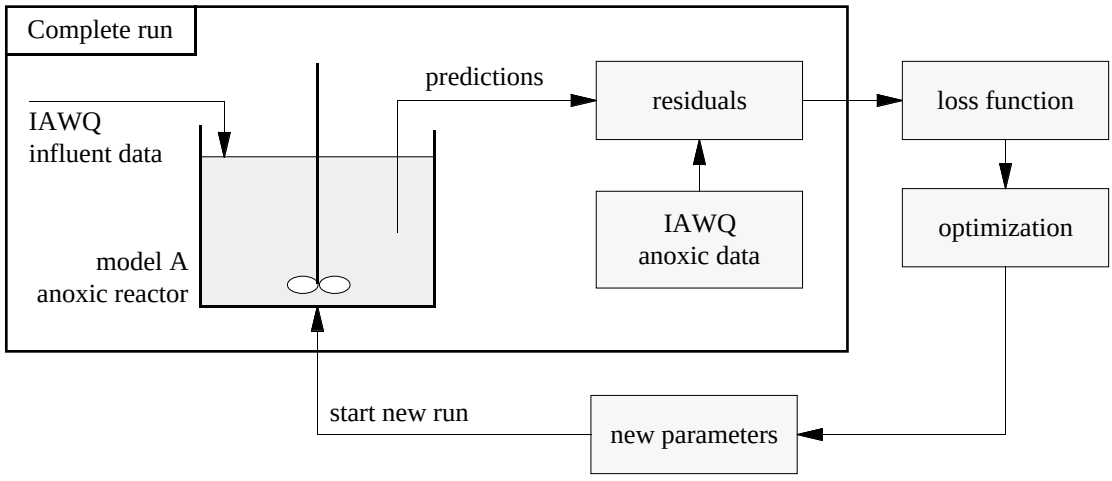


Figure 4.4 Optimization procedure for an anoxic reactor.

Two special cases are examined depending on which variables are assumed measurable. In case 1A the X_{COD} (i.e., $S_S + X_S$ of the IAWQ model), S_{NH} and S_{NO} concentrations in the anoxic reactor are assumed to be measurable and in case 1B the above variables plus the $X_{B,H}$ and $X_{B,A}$ concentrations are assumed possible to measure. This difference affects the value of the loss function and thereby the optimization. In both cases all five quantities are assumed available from the influent data (necessary if an optimization is to be performed for a single anoxic reactor without modelling the aerobic reactor and the settler of the true process).

It should be noted that the generated data for the true plant are based on a simulation with both anoxic and aerobic reactors, settler, sludge recirculation, etc., and not a special identification experiment using an isolated anoxic reactor. This is because the aim of the study is to identify the models during normal plant operation. However, the situation is simplified since the anoxic reactor of model A is simulated as a single unit although the input data are generated from a simulation of an entire plant.

This is a first test to determine whether the anoxic part of the reduced model is at all capable of mimicking the behaviour of the IAWQ model in an anoxic environment. It is also a first rough test of the identifiability of model A since several sets of initial parameter values are used for the optimization and different variables are assumed to be measurable for the different test cases. Results of the optimizations are presented in Table 4.1.

Optimization	initial estimates				final estimates				value of loss func.
model A, anoxic part	r_H	Y_H	b_H	b_A	r_H	Y_H	b_H	b_A	
Case 1A	.024	.35	.46	.06	.013	.498	.000	.014	19.3
(measured:	.046	.69	.94	.12	.013	.502	.000	.207	19.5
X_{COD}, S_{NH}, S_{NO})	.057	.86	1.18	.15	.025	.663	.326	.000	31.9
	.068	1.03	1.42	.18	.013	.498	.000	.000	19.3
Case 1B	.024	.35	.46	.06	.013	.498	.000	.014	19.5
(measured:	.046	.69	.94	.12	.013	.506	.000	.383	20.4
X_{COD}, S_{NH}, S_{NO},	.057	.86	1.18	.15	.013	.500	.000	.079	19.5
$X_{B,H}$, $X_{B,A}$)	.068	1.03	1.42	.18	.013	.502	.000	.159	19.6

Table 4.1 Results of the parameter optimization for case 1.

An analysis of the final estimates suggests some possible conclusions. Firstly, the autotrophic decay rate is extremely difficult to estimate for the applied conditions. The concentration of autotrophs does not change much due to reaction mechanisms in the anoxic reactor (see equation (4.5)) but

more due to the variations of the input data (since an aerobic reactor was included to generate the data). Moreover, the volume of the anoxic reactor is only 20 % of the total reactor volume. Therefore, b_A may assume practically any small value and its effect will be negligible. The situation is emphasized by the fact that the estimations are not improved for case 1B when $X_{B,A}$ is assumed to be measurable. The effect of b_A on the X_{COD} and S_{NH} concentrations by the transformation of dead microorganisms is also small due to the low $X_{B,A}$ concentration. Part of the above also holds for b_H and its effect may be compensated for by the value of r_H for the applied conditions.

Secondly, the parameters r_H and Y_H determine the main behaviour of the investigated system. In most situations these two parameters converge more or less globally (if the initial estimates are 'reasonable') for both cases 1A and 1B. The effect of including measurements of $X_{B,H}$ and $X_{B,A}$ in the loss function is small (the same optimum is reached for both cases). However, for the initial estimates of case 1A(row 3), a completely different optimum is reached which shows that the model is not globally identifiable from the available data. When $X_{B,H}$ and $X_{B,A}$ are included in the loss function, their influence is significant enough to draw the optimization algorithm away from this local optimum and towards the truly best parameter set, even when the values for the local optimum are used as initial seeds for the optimization (not shown).

The reason why both optima produce similar model outputs (see Figures 4.5–4.9) is due to the fact that the numerical values of the net reaction rate expressions $(r_H X_{COD} - b_H)$ and $(r_H X_{COD}/Y_H + b_H)$ are practically identical for both parameter sets (cf. equations (4.1) and (4.4)). By optimizing the system for several initial parameter sets and examining the value of the loss function it is possible to detect this type of problem.

However, it is not realistic to assume $X_{B,H}$ and $X_{B,A}$ to be measurable. The situation may be improved if instead measurements of the denitrification rate are included in the optimization. Such measurements can be performed with reasonable accuracy (though not on-line) but the possibility to include this information has not been tested in this study.

In order to verify the behaviour of the anoxic part of model A after the optimization is concluded, the system is simulated with the obtained parameter sets and the values of the internal state variables are compared with those of the IAWQ model when using the same set of input data. Such a comparison is illustrated in Figures 4.5–4.9.

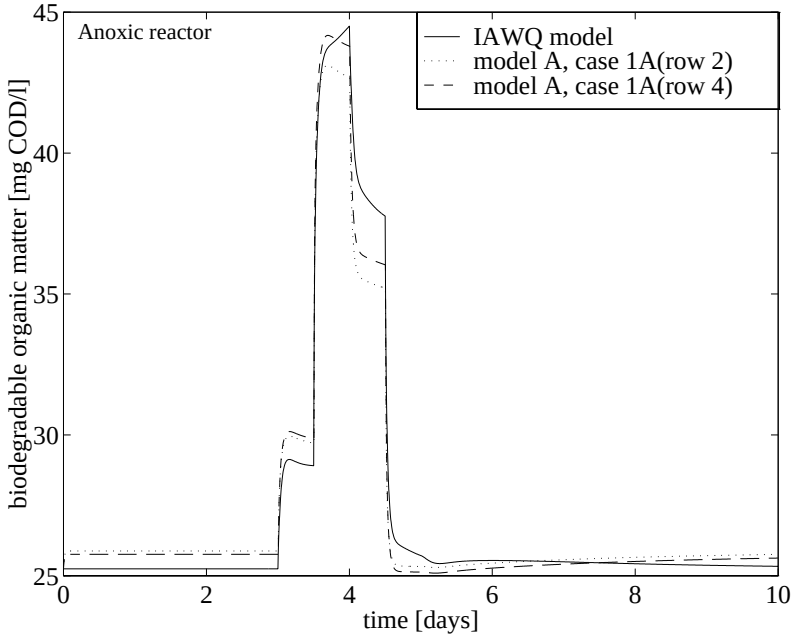


Figure 4.5 Biodegradable organic matter concentration in the anoxic reactor using the IAWQ and A models (model A with two different sets of parameters).

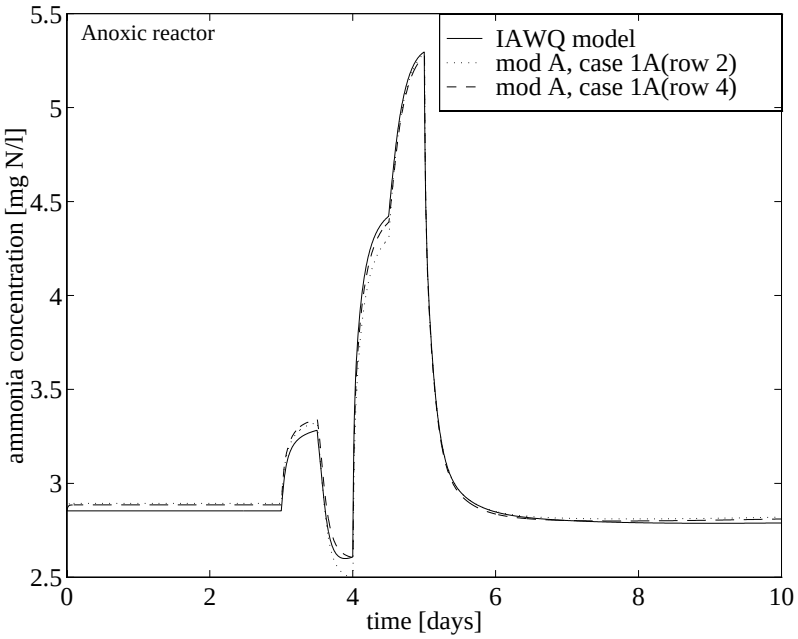


Figure 4.6 Ammonia concentration in the anoxic reactor using the IAWQ and A models (model A with two different sets of parameters).

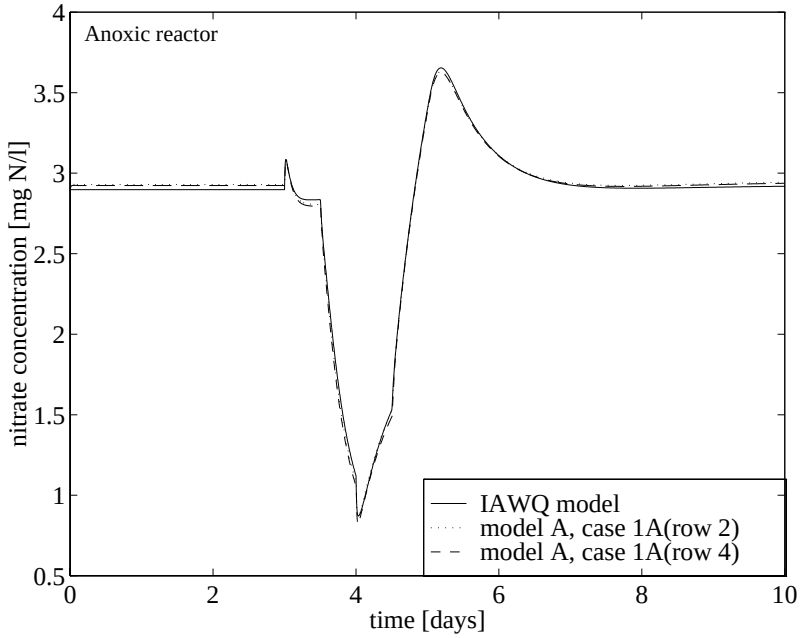


Figure 4.7 Nitrate concentration in the anoxic reactor using the IAWQ and A models (model A with two different sets of parameters).

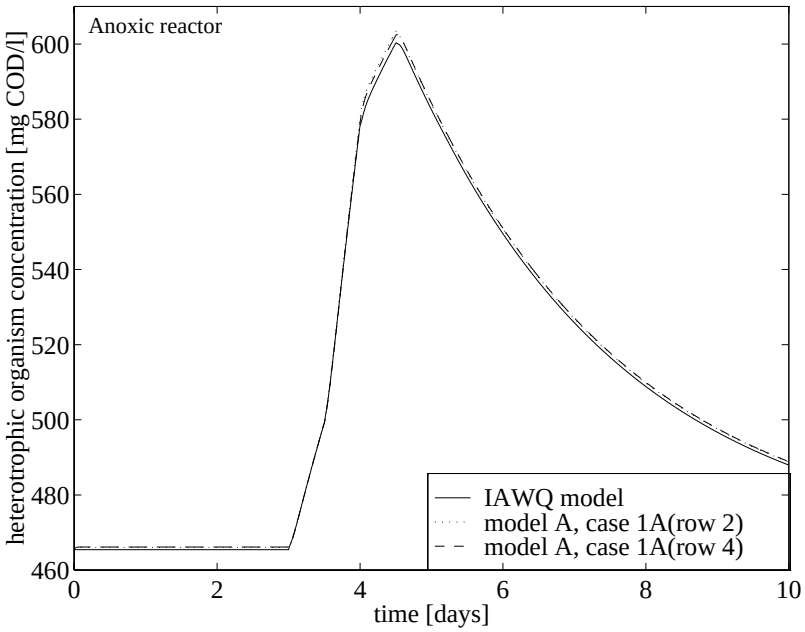


Figure 4.8 Heterotrophic organism concentration in the anoxic reactor using the IAWQ and A models (model A with two different sets of parameters).

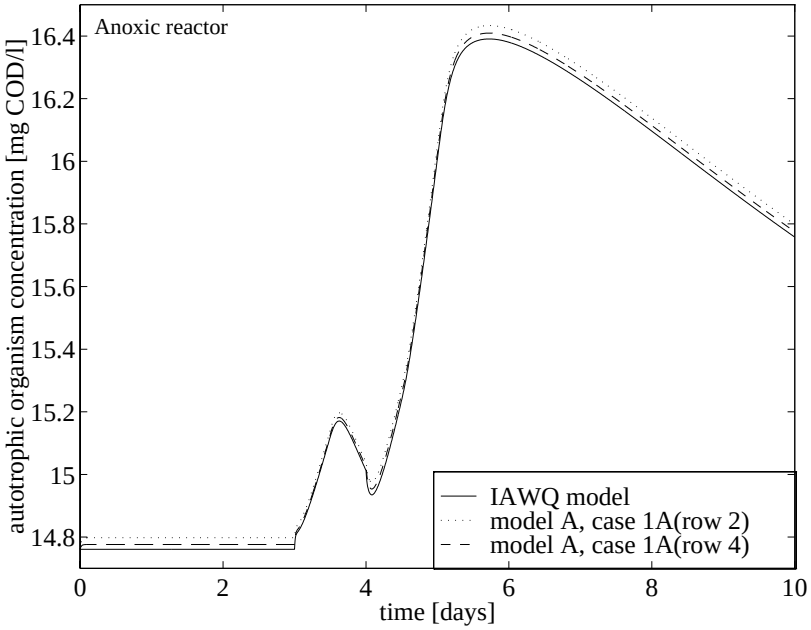


Figure 4.9 Autotrophic organism concentration in the anoxic reactor using the IAWQ and A models (model A with two different sets of parameters).

It is obvious from the graphs that the behaviour of the anoxic part of the reduced model is similar to that of the IAWQ model, both during transient and steady-state situations. The main difference can be observed for the biodegradable organic substrate since this variable is described quite differently by the two models. The discrepancy, however, is not significant.

Case 2 – Aerobic Reactor Using Model A

The second case is an investigation of the *aerobic* part of model A, that is, equations (4.6)–(4.10). It is carried out using the same principle as was described for case 1. Data are generated by simulating the IAWQ model for an entire plant according to Section 4.3. The variables of the wastewater flowing into the aerobic reactor as well as the internal variables of the reactor are stored and used for the optimization. The aerobic part of model A is then simulated using the stored data as model input. The loss function on which the optimization is based, is calculated as the sum of weighted squares of the residuals. The residuals are the differences of the assumed measurable variables of the IAWQ and the reduced model output. A new set of model parameters is proposed by the algorithm and the procedure is repeated until an optimum parameter set is achieved.

Four special cases are examined depending on the set of variables that are assumed to be measurable in the aerobic reactor. They are defined below.

- Case 2A: measurements of X_{COD}, S_{NH} and S_{NO}.
- Case 2B: measurements of X_{COD}, S_{NH}, S_{NO} and OUR.
- Case 2C: measurements of X_{COD}, S_{NH}, S_{NO}, $X_{B,H}$ and $X_{B,A}$.
- Case 2D: measurements of X_{COD}, S_{NH}, S_{NO}, $X_{B,H}$, $X_{B,A}$ and OUR.

The investigation shows whether the aerobic part of the reduced model is capable of mimicking the basic behaviour of the IAWQ model or not. It is also a preliminary test of the identifiability of the aerobic part of model A since several sets of initial parameter values are used for the optimization and different variables are assumed to be measurable for the various cases. Some results of the optimizations are presented in Table 4.2.

Two different situations can immediately be observed from the results – when the OUR is included in the optimization and when it is not. The significance of the OUR measurements leads to similar results for cases 2B and 2D, whereas the results for cases 2A and 2C are more scattered. As was also seen for case 1, it is extremely difficult to estimate the autotrophic decay rate factor b_A with any relevance for the operational conditions assumed in this simulation. The impact of b_A on the model behaviour is practically negligible and is compensated for by other model parameters.

It is interesting to observe that several parameters appear to converge almost globally – r_H, Y_H and b_H. For example, it was not possible to determine b_H realistically in case 1 (converged towards zero in most cases). The optimum parameter set is quite different when the OUR is included in the calculations due to the extra information.

The parameters r_A and Y_A do not appear to converge globally. However, the ratio r_A/Y_A always converges towards practically the same values (0.89 for cases 2A and 2C, 0.98 for cases 2B and 2D). This implies that the small difference of the $X_{B,A}$ concentration is not sufficient to separate the effects of the two parameters even when this variable is assumed measurable – especially as b_A varies as well. Only the combined effect of the parameters is possible to determine for the conditions used in this simulation. A modification of the weight factors defined for the loss function may improve matters slightly. It should also be noted that for case 2A there exist a number of local optima close to each other depending on the strong correlation between r_H and b_H. This is quite natural since the optimization in case 2A is based on the smallest amount of information.

Optimization model A, aerobic part												value of loss func.
initial estimates						final estimates						
r_H	r_A	Y_H	Y_A	b_H	b_A	r_H	r_A	Y_H	Y_A	b_H	b_A	
Case 2A (measured: X_{COD}, S_{NH}, S_{NO})												
.046	.111	.33	.10	.42	.064	.029	.004	.494	.004	.172	.000	376
.093	.222	.66	.20	.86	.126	.035	.066	.558	.075	.227	.000	376
.116	.278	.83	.25	1.08	.157	.034	.000	.552	.000	.213	.072	369
.140	.333	.99	.30	1.30	.188	.032	.000	.528	.000	.219	.096	375
Case 2B (measured: X_{COD}, S_{NH}, S_{NO}, OUR)												
.046	.111	.33	.10	.42	.064	.044	.043	.582	.045	.358	.000	472
.093	.222	.66	.20	.86	.126	.044	.000	.588	.000	.364	.153	468
.116	.278	.83	.25	1.08	.157	.044	.001	.583	.000	.357	.018	466
.140	.333	.99	.30	1.30	.188	.043	.000	.581	.000	.342	.309	469
Case 2C (measured: X_{COD}, S_{NH}, S_{NO}, $X_{B,H}$, $X_{B,A}$)												
.046	.111	.33	.10	.42	.064	.033	.012	.554	.013	.207	.116	380
.093	.222	.66	.20	.86	.126	.034	.000	.554	.000	.208	.042	375
.116	.278	.83	.25	1.08	.157	.034	.000	.552	.000	.210	.080	377
.140	.333	.99	.30	1.30	.188	.033	.025	.544	.028	.209	.000	374
Case 2D (measured: X_{COD}, S_{NH}, S_{NO}, $X_{B,H}$, $X_{B,A}$, OUR)												
.046	.111	.33	.10	.42	.064	.044	.042	.583	.043	.362	.000	474
.093	.222	.66	.20	.86	.126	.044	.000	.584	.000	.359	.009	470
.116	.278	.83	.25	1.08	.157	.044	.000	.589	.000	.368	.000	472
.140	.333	.99	.30	1.30	.188	.043	.024	.579	.025	.352	.000	472

Table 4.2 Results of the parameter optimization for case 2.

The behaviour of the aerobic part of the reduced model is quite similar to that of the IAWQ model for both steady state and transient situations (see Figures 4.10 and 4.11). The differences are, however, more significant than for the anoxic part of the model (partly due to the fact that the aerobic reactor volume is 80 % of the total reactor volume and the internal process mechanisms therefore play a more important role). This is obvious when the values of the loss functions for cases 1 and 2 are compared. For case 2 the values are more than ten times higher, although the weight factors in both cases are chosen to show the same impact for the same relative value of the squared residuals. The behaviour of the model is investigated by simulating the system with the optimized parameter sets and comparing it with data generated with the IAWQ model. In Figures 4.10 and 4.11 such a

comparison is illustrated for two variables, namely S_{NH} and OUR. For the other variables the differences are not significant.

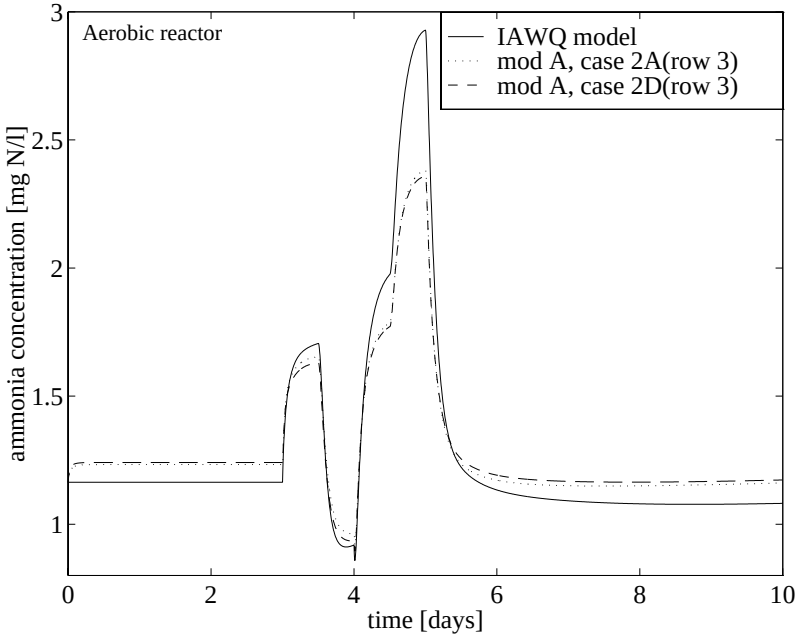


Figure 4.10 Ammonia concentration in the aerobic reactor using the IAWQ and A models (model A with two different sets of parameters).

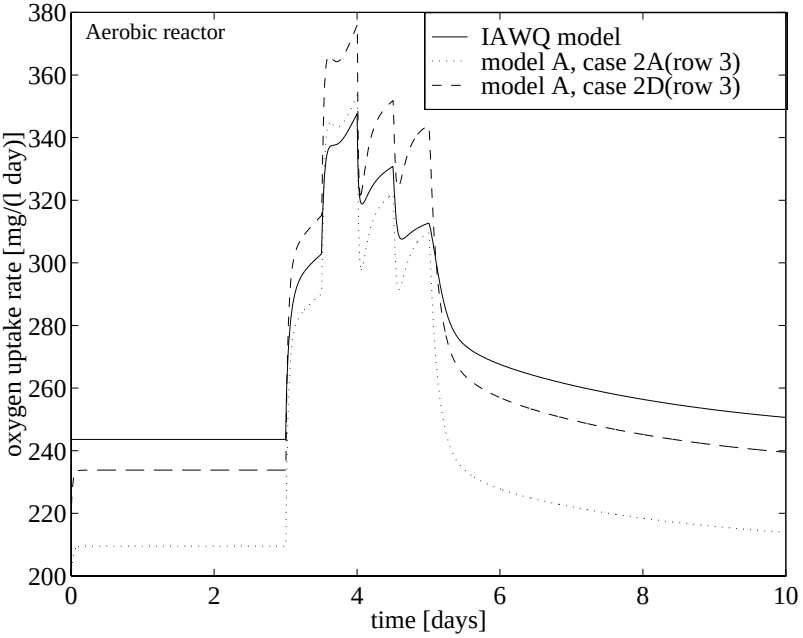


Figure 4.11 Oxygen uptake rate in the aerobic reactor using the IAWQ and A models (model A with two different sets of parameters).

It is clear from the graphs in Figure 4.10 that the difference for the S_{NH} concentration when compared with the IAWQ model is larger in the aerobic reactor than in the anoxic one (cf. Figure 4.6). A main reason for the difference is that the growth rate of autotrophs in the reduced model is described by a first-order reaction, while the IAWQ model uses the Monod growth rate expression with a very low value for the ammonia half-saturation coefficient ($K_{NH} = 1$ mg N/l). Therefore, the autotrophic growth rate in the IAWQ model (in an aerobic environment) is almost constant for ammonia concentrations higher than a few mg N/l while the growth rate expression used in the reduced model will increase proportionally to the ammonia concentration and, consequently, transform more ammonia into nitrate. Consequently, it may be better to assume a zero-order autotrophic growth rate for the reduced order model. The graphs of the OUR in Figure 4.11 show that the reduced model produces reasonable results even when the optimization is not based on information of this variable (dashed line). The result is further improved when OUR is included in the calculation of the loss function, especially in the steady state region (dotted line).

The amount of readily biodegradable substrate (S_S) is dependent on the hydrolysis of slowly biodegradable substrate (X_S) in the IAWQ model, especially in the aerobic reactor since the influent S_S is primarily consumed in the first anoxic reactor. Variation of the S_S/X_S ratio is consequently an important factor for the behaviour of the IAWQ model whereas the reduced model does not respond to such changes as long as the $S_S + X_S$ concentration is fairly constant. Changes in the ratio of organic substrate fractions are therefore troublesome to mimic with the reduced model (especially in an off-line estimation approach) and differences will occur because the reduced order model does not include a hydrolysis process.

If the optimization of cases 1 and 2 is carried out using only steady-state data or data generated using small disturbances, the parameters will not converge towards the values found in this study. In such a case the results will to a large extent depend on the initial values of the parameters because the amount of information in the data is insufficient to locate a true optimum set.

Sensitivity Analysis of Cases 1 and 2

In cases 1 and 2, the anoxic and aerobic parts of model A were investigated for their identifiability properties. Parameter optimization was performed using different sets of initial parameter values and assuming various variables to be measurable. The reactor models were identified separately and the effect of the recirculation due to the behaviour of the reduced model was therefore not taken into account (although the model inputs were generated from a complete WWT plant simulated with the IAWQ model including recirculation etc.). It was demonstrated that the reduced model was capable of mimicking the behaviour of the IAWQ model during these conditions although it proved troublesome to identify all the model parameters – especially b_H and b_A but also r_A and Y_A – in a global sense.

In order to investigate the sensitivity of the model to parameter variations, the value of the loss function is analysed. This is not to be considered as a complete sensitivity analysis of the model but since the result of the optimization algorithm is based on how the value of the loss function changes, it may serve as a good indicator of convergence problems.

The case which will be examined in this subsection is the one where the loss function is calculated from measurements of X_{COD}, S_{NH} and S_{NO} (i.e., the earlier described cases 1A and 2A). By simulating each reactor type repeatedly with identical input data, introducing a small change in one parameter for each run and storing the value of the loss function, a 'map' describing how the loss function varies can be created (i.e., a type of Monte Carlo simulation). The initial sets of parameters are the optima found in case 1A(row 1) and case 2A(row 2) and each parameter (four for the anoxic and six for the aerobic part of the model) varies ±20 % around its initial value in steps of 2 % for each simulation.

The situation is first illustrated in Figure 4.12 for the anoxic part of model A. The model is simulated for the same conditions as described in case 1. Contour plots show how the value of the loss function is affected as the parameters change. In the upper plot of Figure 4.12, the parameters b_H and b_A are kept constant at the optimized values and in the lower plot r_H and Y_H are kept constant. Consequently, not all possible parameter interactions are shown (only two parameters vary within each plot). To show the effect of all the parameter variations simultaneously would require a four-dimensional plot which would be difficult to interpret.

A small complication occurs when the sensitivities of b_H and b_A are to be examined. Since the optimized values for these parameters are close to zero, the effect of a 20 % change would be negligible. Instead the values for b_H and b_A are set to vary from 0.0 to 0.4 day^{-1} in the lower plot of Figure 4.12 while the other two parameters are kept constant. To be able to compare the two plots, two contour lines next to each other indicate that the value of the loss function has changed with a value of 10 (the same scale is used in Figure 4.13).

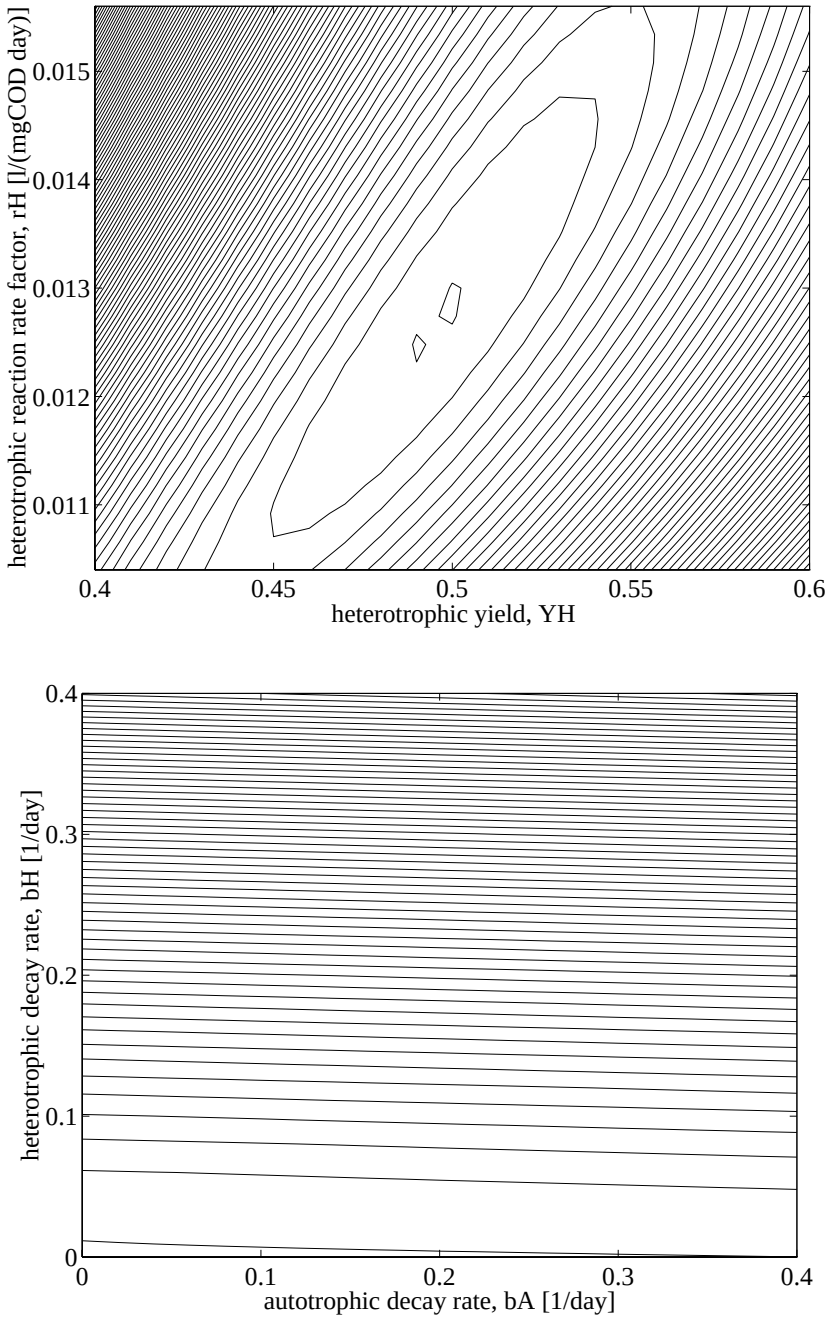


Figure 4.12 Contour plots showing the sensitivity to parameter changes for the anoxic part of model A (case 1A).

The first plot in Figure 4.12 shows that the loss function is sensitive to small changes of both r_H and Y_H and the direction of the gradient is well defined. This implies a reliable convergence of the optimization algorithm although a problem occurs as the optimum is approached. Two different optima appear, very close to each other. The values of the parameters are almost the same for both optima but it is an indication that the information on which the loss function is based, is not sufficient or that there is a structural problem in the model. If the region of analysis was extended, a third optimum would appear for the parameter set found in case 1A(row 3). It is obvious that the ratio of the two parameters are of importance for the model behaviour. Within the narrow valley seen in the plot – indicating a low value of the loss function – this ratio is almost constant.

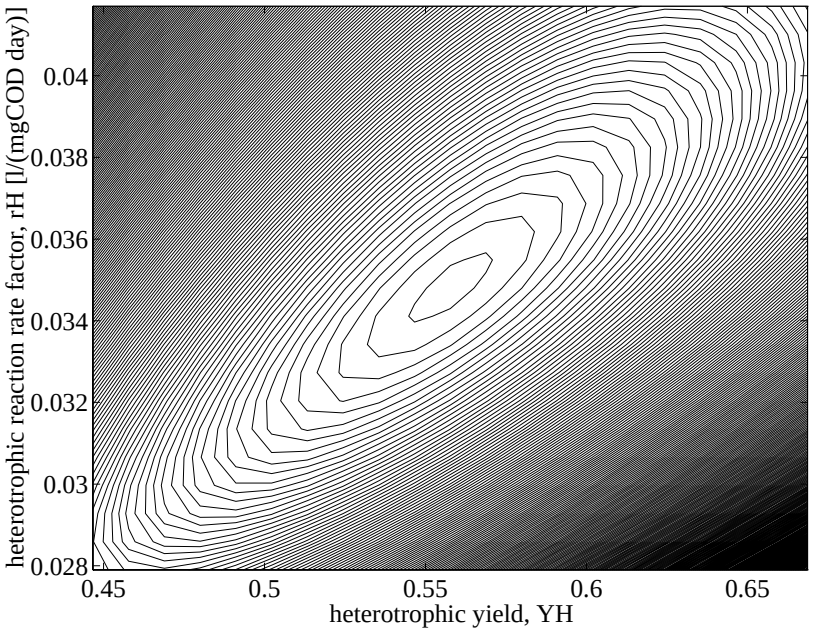
The second plot illustrates the previously discussed problem of identifying b_A. The gradient of the loss function indicates how the optimization algorithm would change the value of b_H to improve the result but a significant change of b_A does practically not affect the loss function. The model is not sensitive to this parameter for the applied conditions which explains the results from case 1. In Table 4.1 it was shown that the sensitivity is not significantly improved when $X_{B,H}$ and $X_{B,A}$ are also included in the calculation of the loss function (cf. case 1B).

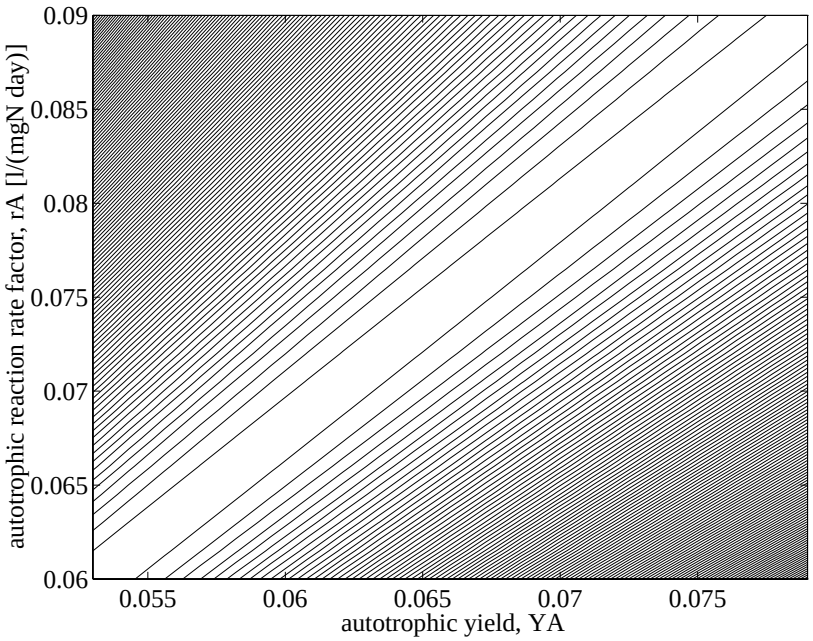
In Figure 4.13 results from the same type of analysis as described above are shown for the aerobic part of model A. The initial parameter values are the optimized set found in case 2A(row 2) and the simulation conditions are identical to the ones described there. Due to the fact that the initial value of b_A is almost zero, its value is set to vary between 0.0 and 0.4 day^{-1} in the last plot.

The first plot of Figure 4.13 illustrates the high sensitivity of the model to the two parameters r_H and Y_H. A well-defined optimum appears although from the results shown in Table 4.2 it is clear that the best parameter set for r_H and Y_H is also influenced by the convergence of the other model parameters.

The second plot shows that the model is also sensitive to variations of the parameters r_A and Y_A. However, their individual values are obviously of practically no significance for the applied conditions; it is only the ratio of the two that matters. The optimum valley is extremely long and narrow which indicates that the optimization algorithm has no problems determining the best ratio but cannot find the correct individual values of the parameters. This explains some of the results earlier discussed for case 2.

The final plot demonstrates the much smaller influence of b_H and b_A on the model behaviour. Variations of the value of b_A has a very limited effect. It should also be noted that the true optimum for b_A in this case appears to be a negative value. However, a deliberate restriction built into the optimization algorithm hinders any of the model parameters to assume negative values even if such a change would further lower the value of the loss function. Negative parameter values do not have any relevance when interpreting the parameters in a physical/biological sense.





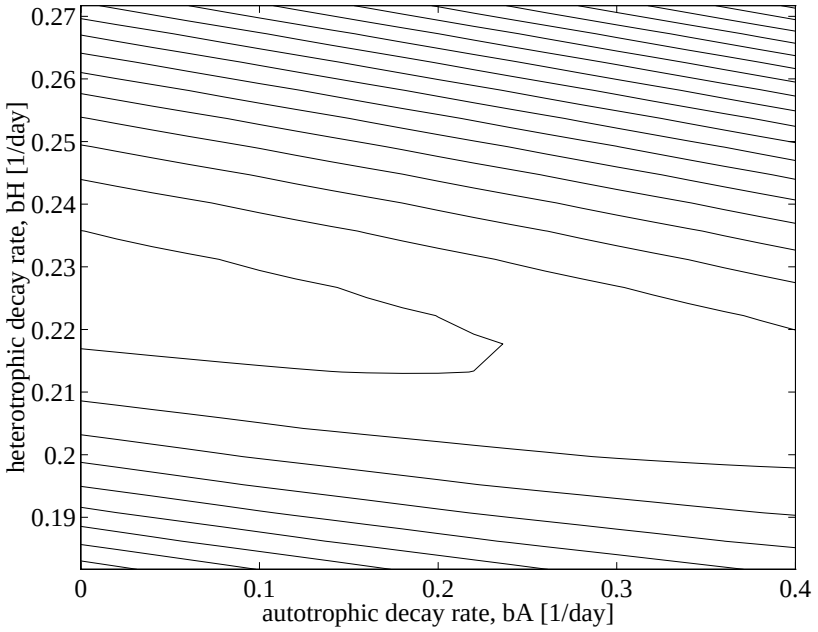


Figure 4.13 Contour plots showing the sensitivity to parameter changes for the aerobic part of model A (case 2A).

Case 3 – Anoxic/Aerobic Reactor Combination Using Model A

In this subsection the optimization is generalized. A natural extension would be to investigate the anoxic and aerobic part of model A separately but for conditions where the effect of the recirculation is included in a direct way, that is, simulating a plant with either an anoxic or an aerobic reactor in combination with a settler. However, for practical reasons the number of different cases has to be limited and, therefore, the optimization algorithm is now used for the complete model (i.e., coupled operation of the two reactor types) and all parameters are simultaneously estimated. The principle of the optimization is shown in Figure 4.14.

The necessary data are generated by simulating the IAWQ model (describing the true WWT plant) as discussed earlier. The character of the influent wastewater to the plant is identical to what was used for the earlier cases (see Figures 4.2 and 4.3). The reactor volumes, recirculation rates, settler model, etc., are naturally the same for both the IAWQ and the reduced model, according to Figure 4.2. Measurements are assumed to be available from the influent wastewater, the anoxic reactor and the aerobic reactor.

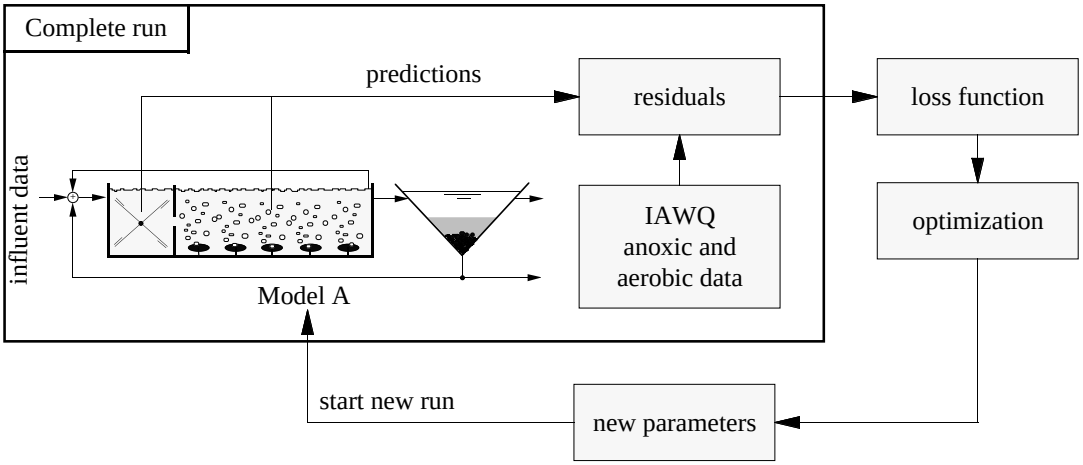


Figure 4.14 Optimization procedure for the complete model A.

In the influent wastewater four variables are always considered to be measurable – X_{COD}, S_{NH}, S_{NO} and Q_{in} (the concentration of micro-organisms in the influent wastewater is assumed to be zero). Depending on which variables are considered to be measurable in the two reactors, five different cases will be examined.

- Case 3A: measurements of X_{COD}, S_{NH}, S_{NO} in both the anoxic and aerobic reactor;
- Case 3B: measurements of X_{COD}, S_{NH}, S_{NO} in both reactors plus OUR in the aerobic reactor.
- Case 3C: measurements of X_{COD}, S_{NH}, S_{NO} in both reactors plus OUR, $X_{B,H}$, $X_{B,A}$ in the aerobic reactor.
- Case 3D: measurements of X_{COD}, S_{NH}, S_{NO} in both reactors plus $X_{B,H}$, $X_{B,A}$ in the anoxic reactor.
- Case 3E: measurements of X_{COD}, S_{NH}, S_{NO}, $X_{B,H}$, $X_{B,A}$ in both reactors plus OUR in the aerobic reactor.

The results of the optimizations are presented in Tables 4.3 and 4.4 (next opening) for different sets of initial estimates. The two tables show the results for the anoxic and aerobic part, respectively, but the results should be interpreted simultaneously. They are separated into two tables only to make the results easier to read. Consequently, the values of the loss function given in Table 4.4 represent the values of the total loss function for both the anoxic and aerobic part of the model.

It is obvious that the values of the loss function are much larger for case 3 than would be expected if the results from cases 1 and 2 could simply be added together to describe coupled operation. This is because the input

characteristics to the unit reactors in cases 1 and 2 were identical to the input calculated by the IAWQ model. In case 3 only the input to the entire plant is the same. A small error in the prediction of the behaviour in the anoxic reactor is in case 3 propagated into the aerobic reactor, further amplified, and propagated back to the anoxic reactor by the recirculation (the internal feedback of the model). This emphasizes the differences between the reduced and the IAWQ model.

The different way the organic matter is described in the IAWQ and reduced order models also has an influence on the behaviour, especially in this case where the settler model is included for both models. On an average more biodegradable organic matter is recirculated when using the reduced model (γX_{COD}) than when using the IAWQ model ($S_S + \gamma X_S$). As the recirculated sludge is fed back to the anoxic reactor, this difference will also have an impact on the behaviour of the models, which was not included in cases 1 and 2.

In order to avoid the influence of the initial transients when the parameters are updated during the optimization, the new steady state of the reduced order model is first calculated for each new parameter set before the actual optimization is initiated.

The results for case 3A show large variations. Ten parameters are optimized based on six measurable variables and several local optima are detected. Even the most sensitive parameters, that is, r_H and Y_H, converge towards different values depending on the initial estimates. However, a strong correlation between r_H and b_H is apparent for both reactors. A high value of r_H is always accompanied by a high value of b_H and vice versa. This is because the algorithm attempts to minimize the residuals of X_{COD}. A high r_H indicates a high growth rate and, consequently, a large consumption of organic matter. This effect is compensated by a high b_H, which leads to a high conversion rate of decayed material into X_{COD}. The mathematical relationship between the two parameters is not clear due to the different values of Y_H (among other things). However, the model sensitivity to the heterotrophic decay rate is increased by the recirculation when compared with cases 1 and 2.

Optimization	initial estimates				final estimates			
model A, anoxic part	r_H	Y_H	b_H	b_A	r_H	Y_H	b_H	b_A
Case 3A	.024	.35	.46	.06	.023	.723	.372	.111
(measured:	.046	.69	.94	.12	.014	.726	.131	.071
X_{COD}, S_{NH}, S_{NO})	.057	.86	1.18	.15	.008	.711	.015	.073
	.068	1.03	1.42	.18	.030	.626	.579	.040
Case 3B	.024	.35	.46	.06	.050	.872	1.10	.090
(measured:	.046	.69	.94	.12	.013	.660	.117	.057
X_{COD}, S_{NH}, S_{NO})	.057	.86	1.18	.15	.021	.743	.302	.086
	.068	1.03	1.42	.18	.047	.862	1.02	.158
Case 3C	.024	.35	.46	.06	.043	.744	.830	.120
(measured:	.046	.69	.94	.12	.032	.691	.521	.139
X_{COD}, S_{NH}, S_{NO})	.057	.86	1.18	.15	.019	.573	.172	.110
	.068	1.03	1.42	.18	.029	.671	.452	.101
Case 3D	.024	.35	.46	.06	.031	.675	.510	.116
(measured:	.046	.69	.94	.12	.037	.711	.644	.128
X_{COD}, S_{NH}, S_{NO},	.057	.86	1.18	.15	.020	.566	.185	.070
$X_{B,H}$, $X_{B,A}$)	.068	1.03	1.42	.18	.041	.733	.782	.136
Case 3E	.024	.35	.46	.06	.027	.647	.360	.209
(measured:	.046	.69	.94	.12	.028	.652	.340	.075
X_{COD}, S_{NH}, S_{NO},	.057	.86	1.18	.15	.025	.630	.335	.163
$X_{B,H}$, $X_{B,A}$)	.068	1.03	1.42	.18	.027	.646	.380	.286

Table 4.3 Results of the parameter optimization for case 3 (anoxic part).

The same type of correlation appears to exist for r_A and b_A in the aerobic reactor due to the minimization of the residuals of S_{NH}. This conclusion is, however, more uncertain as a part of the formed S_{NH} originates from the decay of heterotrophs (i.e., depends on b_H). The values for the heterotrophic yield coefficient in the anoxic reactor appears quite stable whereas the other yield coefficients and the anoxic b_A display large variations.

Another important factor which must also be considered, does not show in the tables. As the concentrations of microorganisms are not assumed to be measurable, the different optimized parameter sets lead to very different values of the $X_{B,H}$ and $X_{B,A}$ concentrations. Based on the measurements in case 3A, the optimization algorithm cannot determine whether the true system contains a high concentration of organisms with a low reaction rate or vice versa. This affects all parameter values and is the major reason why the optimized parameter sets are so different.

Optimization model A, aerobic part												value of loss func.
initial estimates						final estimates						
r_H	r_A	Y_H	Y_A	b_H	b_A	r_H	r_A	Y_H	Y_A	b_H	b_A	
Case 3A (measured: X_{COD}, S_{NH}, S_{NO})												
.046	.111	.33	.10	.42	.064	.078	.156	.766	.202	.780	.042	1430
.093	.222	.66	.20	.86	.126	.041	.180	.728	.059	.385	.097	1238
.116	.278	.83	.25	1.08	.157	.019	.250	.643	.716	.123	.174	979
.140	.333	.99	.30	1.30	.188	.042	.162	.478	.052	.385	.069	1499
Case 3B (measured: X_{COD}, S_{NH}, S_{NO}, OUR)												
.046	.111	.33	.10	.42	.064	.025	.142	.555	.129	.203	.030	1361
.093	.222	.66	.20	.86	.126	.028	.192	.602	.141	.234	.107	1112
.116	.278	.83	.25	1.08	.157	.029	.196	.576	.707	.248	.105	1200
.140	.333	.99	.30	1.30	.188	.028	.214	.568	.500	.239	.108	1370
Case 3C (measured: X_{COD}, S_{NH}, S_{NO}, $X_{B,H}$, $X_{B,A}$, OUR)												
.046	.111	.33	.10	.42	.064	.091	.232	.714	.208	.944	.138	2359
.093	.222	.66	.20	.86	.126	.041	.223	.539	.204	.406	.124	2146
.116	.278	.83	.25	1.08	.157	.050	.238	.586	.219	.503	.152	2121
.140	.333	.99	.30	1.30	.188	.044	.244	.556	.222	.439	.160	2150
Case 3D (measured: X_{COD}, S_{NH}, S_{NO})												
.046	.111	.33	.10	.42	.064	.042	.141	.528	.128	.409	.023	1369
.093	.222	.66	.20	.86	.126	.038	.208	.501	.188	.374	.107	1413
.116	.278	.83	.25	1.08	.157	.043	.221	.530	.199	.421	.138	1376
.140	.333	.99	.30	1.30	.188	.033	.201	.466	.180	.315	.095	1429
Case 3E (measured: X_{COD}, S_{NH}, S_{NO}, $X_{B,H}$, $X_{B,A}$, OUR)												
.046	.111	.33	.10	.42	.064	.044	.148	.548	.136	.443	.021	2149
.093	.222	.66	.20	.86	.126	.041	.189	.526	.173	.401	.097	2169
.116	.278	.83	.25	1.08	.157	.044	.189	.549	.172	.442	.075	2156
.140	.333	.99	.30	1.30	.188	.042	.166	.536	.152	.418	.015	2153

Table 4.4 Results of the parameter optimization for case 3 (aerobic part).

In an attempt to improve the behaviour of model A and enhance the optimization, OUR is included as a measurable variable in case 3B. The effect is apparent on the parameters r_H, Y_H and b_H in the aerobic reactor, which now converge towards approximately the same values independently of the initial estimates. Consequently, the predicted value of $X_{B,H}$ is much more stable (although not identical to the concentration predicted by

the IAWQ model). Since the concentration of autotrophic organisms is much smaller than $X_{B,H}$, its effect on the OUR is small. Therefore, $X_{B,A}$ converges towards different values and the estimates of the autotrophic parameters r_A, Y_A and b_A are not significantly improved by the added information. The same result could be observed for case 2B. In the anoxic reactor the earlier discussed correlation between r_H and b_H still holds although the results are not improved when compared with case 3A.

In case 3C the concentrations of heterotrophs and autotrophs in the aerobic reactor are also assumed to be measurable. This will draw the $X_{B,H}$ and $X_{B,A}$ concentrations of the reduced model towards the values calculated by the IAWQ model and thereby eliminating one of the problems discussed for cases 3A and 3B. Note that the concentration of microorganisms will be approximately the same in both reactors due to the recirculation, although it is only assumed to be measurable in the aerobic reactor. The result of row 1 is clearly a special case where the algorithm has converged towards a local optimum quite far from the best one. The other three optimizations produce almost identical values of the loss function *but* with quite different parameter sets although the convergence of the autotrophic parameters is remarkably consistent. For case 3B the situation was practically the opposite. There are a number of reasons for this.

Apart from the earlier discussed correlation between r_H and b_H in the two reactors a new correlation can be observed. A low r_H in the aerobic reactor leads to a high r_H in the anoxic reactor and vice versa (the same correlation holds for the ratio r_H/Y_H). In case 3A the opposite relationship could be observed. This implies that the reduced model deals with dynamic disturbances differently in the two reactors which can be observed by comparing dynamic simulations but not by the value of the loss function where all errors are lumped together. The value of $(1-Y_H)r_H/Y_H$ in the aerobic reactor is practically the same for all examples of case 3C due to the available OUR and $X_{B,H}$ measurements.

A problem is that the measurements of the organism concentrations are in conflict with the OUR measurements. To reach the best estimated result of the OUR behaviour, the reduced model requires a set of parameters which leads to a significantly higher $X_{B,H}$ concentration whereas the measurements of this variable force the optimization algorithm away from that set of parameters in order to predict the low organism concentration calculated by the IAWQ model. Therefore, the values of the loss function are much higher for cases 3C and 3E than for cases 3B and 3D.

In case 3D the concentrations of microorganisms in the anoxic reactor are considered to be measurable. Only the three basic variables are assumed to be measurable in the aerobic reactor. The results of the optimization are similar to the ones for case 3C. The values of the loss function are almost identical for all examples although much lower than for 3C. The parameter values found in cases 3C and 3D are in the same region, which is an indication that the weight factor for the OUR residuals should perhaps be increased to enhance the influence of the OUR measurements. However, it is obvious that the information of $X_{B,A}$ is essential for estimating many of the autotrophic parameters.

In the final case, all state variables of model A are assumed to be measurable in both reactors. Furthermore, the OUR is also available. It is not realistic to have so much information available. Parameters that do not converge globally for this case will probably not be possible to estimate directly in a continuous-flow plant. The conclusion would be that the influence of such parameters is practically negligible.

The results for case 3E clearly indicate that most model parameters converge globally when all this information is available. The slightly differing parameter values are mainly due to the slow convergence of the optimization algorithm and in some cases the optimization has been stopped prematurely. However, the autotrophic decay rate coefficient does not converge globally for the chosen conditions in neither the anoxic nor the aerobic reactor (which to a small extent also influence the values of r_A and Y_A though the ratio r_A/Y_A is perfectly constant). A low b_A value in the anoxic reactor is compensated by a higher value in the aerobic reactor and vice versa. The recirculation then equalises the differences of the $X_{B,A}$ concentrations in the two reactors. A small correlation is also apparent between the values of r_H and b_H.

The low concentration of autotrophic organisms (2–10 %) compared with the heterotrophs, implies that b_A – from the model point of view – is only of importance for keeping the $X_{B,A}$ concentration at the correct level (if it is assumed to be measurable). The parameter has no real significance on the process of transforming decayed material into X_{COD} and S_{NH}. This is a common problem and motivates the modification of model A into model B, which is further discussed in the next subsection.

An investigation of the sensitivity for the complete model A during coupled operation has been performed. By varying the model parameters (the same principle as was shown for cases 1 and 2), the model sensitivity

to parameter changes has been examined. Although not presented here, the results of this analysis further motivate the conclusions discussed above.

The behaviour of model A is validated by simulating the process and comparing the results to the predictions of the IAWQ model using identical influent wastewater characteristics, flow rates, recirculation rates, etc. In Figures 4.15–4.19 such a comparison is illustrated for some of the state variables using a few of the parameter sets determined by the optimization.

All graphs show that the qualitative behaviour of the IAWQ model and model A is similar. Figures 4.15–4.17 also indicate that the behaviour of model A with regard to X_{COD}, S_{NH} and S_{NO} is similar for both cases 3A and 3E although the optimized parameter sets are based on different amounts of information. The reasons for the difference in X_{COD} concentration between the IAWQ and the reduced models were commented in a previous subsection (e.g., the hydrolysis process, the changing ratio of S_S and X_S, and the behaviour of the settler).

Figure 4.17 shows a significant difference in the behaviour of the two models with regard to the predicted ammonia concentration, especially around day five. This is primarily because the ammonia half-saturation coefficient of the IAWQ model is set to 1 mg N/l. As the S_{NH} concentration is increased from 1 to 3 mg/l (i.e., the most non-linear region of the Monod growth rate expression used for the autotrophs) in the aerobic basin it is impossible to achieve the same behaviour for model A when using a first-order approximation to describe the autotrophic growth rate. Therefore, the nitrification rate will increase more rapidly in the reduced model when the ammonia concentration increases and the ammonia peak predicted by the IAWQ model is flattened. As a consequence, a higher concentration of nitrate in the anoxic reactor is also predicted by the reduced model during the same period due to the recirculation from the aerobic reactor (see Figure 4.16).

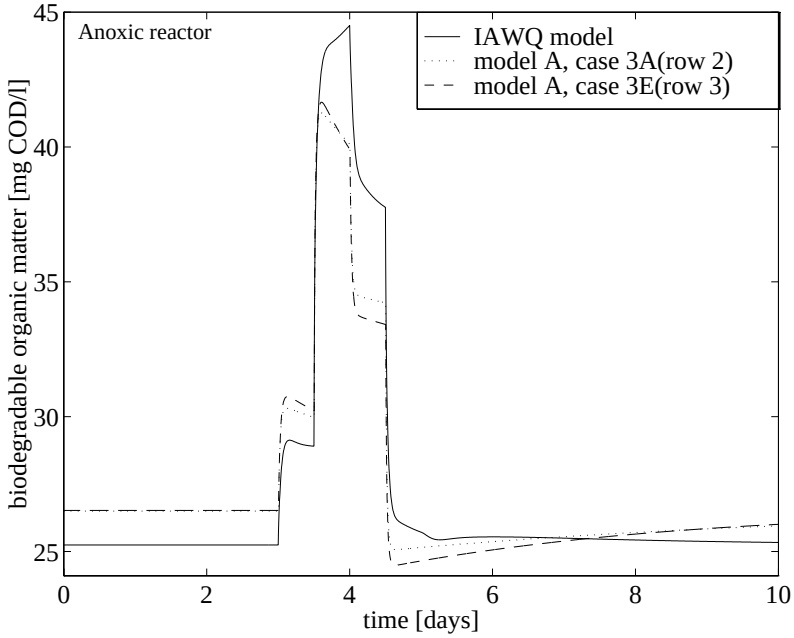


Figure 4.15 Biodegradable organic matter concentration in the anoxic reactor using the IAWQ and A models (model A with two different sets of parameters) to simulate an entire AS process.

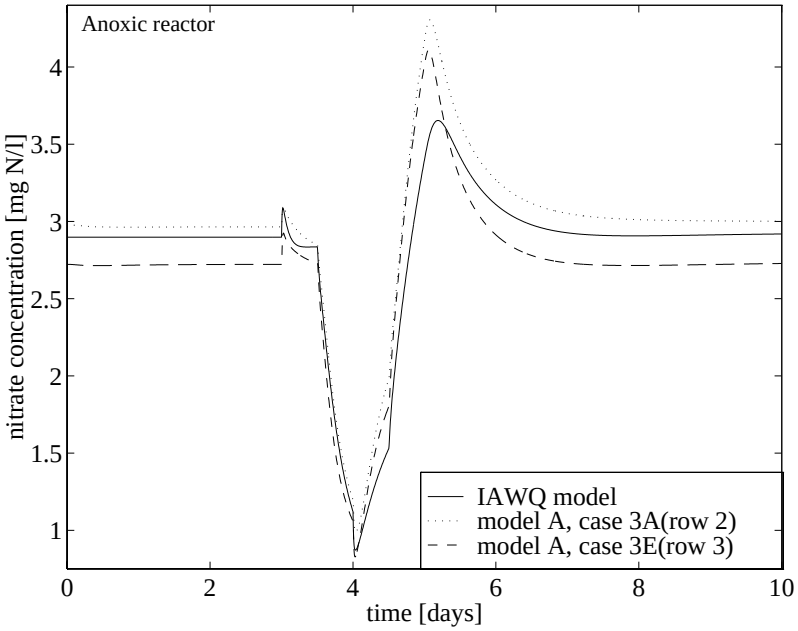


Figure 4.16 Nitrate concentration in the anoxic reactor using the IAWQ and A models (model A with two different sets of parameters) to simulate an entire AS process.

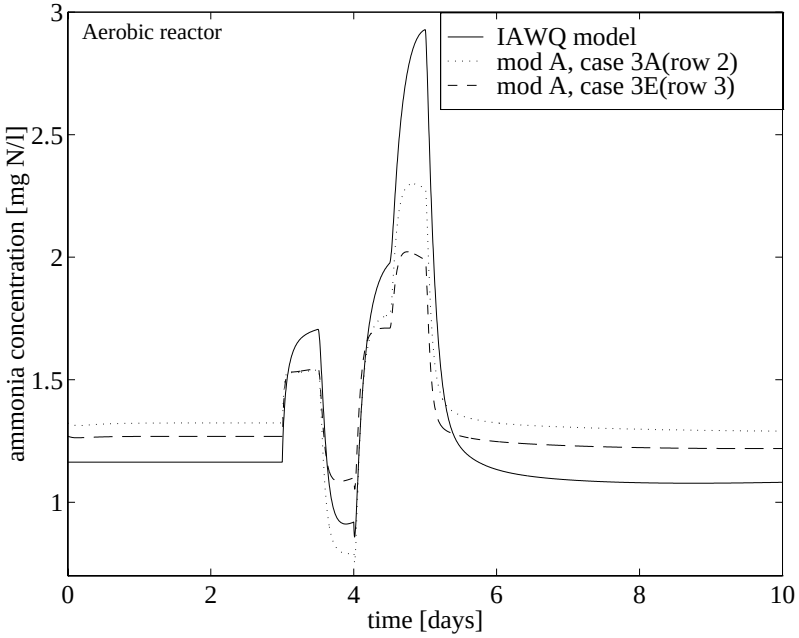


Figure 4.17 Ammonia concentration in the aerobic reactor using the IAWQ and A models (model A with two different sets of parameters) to simulate an entire AS process.

Figure 4.18 illustrates how differently the quantitative concentrations of microorganisms are predicted by the reduced model for cases 3A and 3E. This is the major reason for the different sets of parameters found by the optimization algorithm. If instead the organism concentrations are normalized around their steady state values the results are almost inseparable, that is, the qualitative behaviour of the model for the different cases is identical.

Finally, the OUR is shown in Figure 4.19. It appears remarkable that the predictions from case 3A are better than from case 3E, although OUR measurements were not included for the optimization in case 3A. The large offset for case 3E compared with the result of the IAWQ model is due to the simultaneous optimization of the organism concentrations. While the predictions of $X_{B,H}$ and $X_{B,A}$ are improved for this case, the prediction of the OUR is not satisfactory. If the OUR for case 3B would have been plotted, it would be in perfect agreement with the results of the IAWQ model. On the other hand, the steady-state heterotrophic concentration for case 3B is close to 800 mg COD/l, which should be compared with the steady-state value of 460 mg COD/l predicted by the IAWQ model.

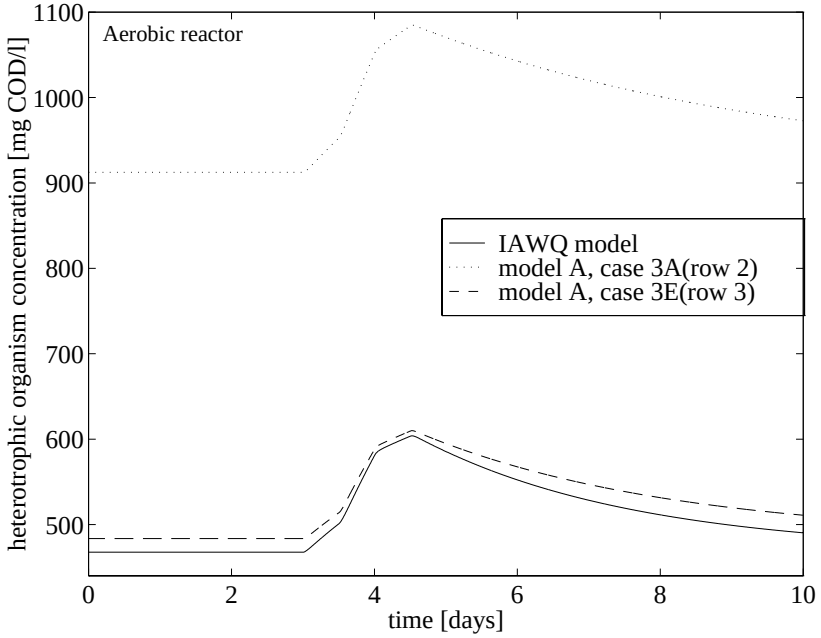


Figure 4.18 Heterotrophic organism concentration in the aerobic reactor using the IAWQ and A models (model A with two different sets of parameters) to simulate an entire AS process.

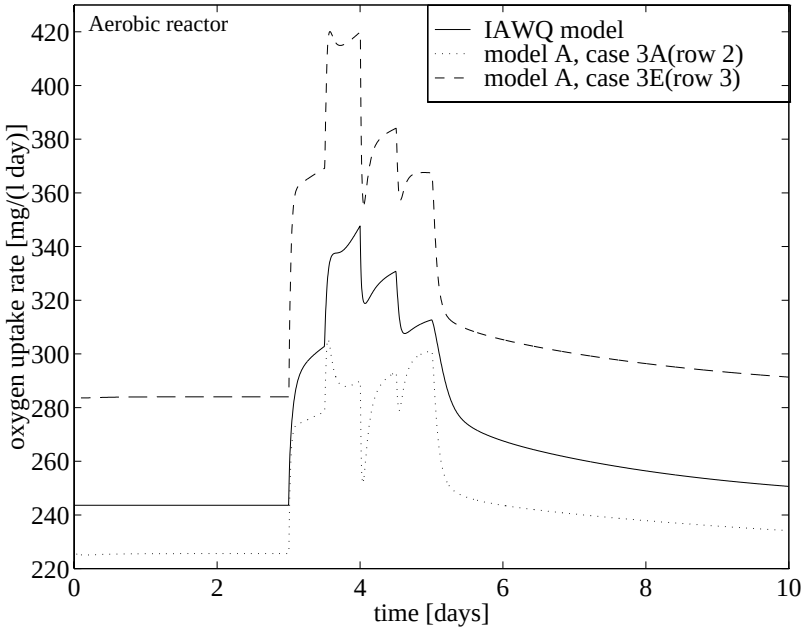


Figure 4.19 Oxygen uptake rate in the aerobic reactor using the IAWQ and A models (model A with two different sets of parameters) to simulate an entire AS process.

Case 4 – Anoxic/Aerobic Reactor Combination Using Model B

The results from case 3 show that the problems concerning global identifiability are apparent. Even when a large number of variables are assumed to be measurable it is difficult to obtain an optimum global set of parameters. Part of the problem is due to the model complexity. Although model A is a simple model for the activated sludge process when compared to the IAWQ model, there are still ten parameters to be identified simultaneously (if one anoxic and one aerobic reactor are modelled) during normal operating conditions (no special identification experiments) for a continuous-flow WWT plant.

In an attempt to reduce the degrees of freedom for the model and improve its global identifiability, model A is further reduced into model B. As discussed in Section 4.1, this is done by assuming the decay rate factors b_H and b_A to be identical, both for anoxic and aerobic conditions. The number of model parameters are hereby reduced from ten to seven. Moreover, the difficulty of separately identifying b_H and especially b_A, which has been discussed for the previous cases, may also be reduced since the effect of each of the four parameters on the model behaviour are now combined into one. Therefore, the possibility for the optimization algorithm to detect significant changes in the value of the loss function is improved, that is, the model sensitivity to the decay parameter is enhanced.

From a biological point of view there is no motive to differentiate the decay rate factors depending on the applied condition (anoxic or aerobic) since the microorganisms are circulated through the plant and exposed to both environments during their life cycles. As for the assumption that the decay rates for both heterotrophic and autotrophic bacteria are the same, this is more debatable. However, both b_H and b_A are rough average values because heterotrophs and autotrophs are large groups that consist of many different organism species with individual variations.

It should be noted that the different values for b_H and b_A suggested in the IAWQ model (see Table 3.1) depends on the fact that b_A represents a traditional decay rate coefficient due to endogenous decay, whereas the value of b_H is affected by the death-regeneration hypothesis, as described in Section 3.3, equation (3.21). Using the default parameter values of Table 3.1 and inserting them into equation (3.21), it is obvious that the decay rate for heterotrophs in the IAWQ model (0.62 day^{-1}) is equivalent to a traditional decay rate of 0.24 day^{-1} (which is practically the same value as used to describe the autotrophic decay rate). However, in the reduced order

models both types of organisms contribute to the recycling of decayed organic matter in the process, as hydrolysis is not included. This implies that for the reduced order models it is correct to assume practically identical decay rates for the different organisms because the recycling of decayed material is modelled differently than in the IAWQ model.

In most AS models the decay coefficient is important in order to predict the sludge production, sludge age and the oxygen consumption. This is not the purpose of the simplified models presented in this chapter. As the reduced models only describe active biomass and the fact that all material which results from decay is directly transformed into X_{COD} and S_{NH}, the specific decay rate factor is in this case actually a rate coefficient, describing this transformation. To compare identified values of b_H and b_A from the reduced models with values normally used in traditional AS models or determined by laboratory experiments, may therefore not be relevant. This is an important drawback of the reduced models.

The reasons discussed above motivate the simplification of model A into model B. It is also considered important to apply a model structure which enhances parameter identification and automatic model calibration from full-scale continuous-flow plant operation. The possibility of determining the traditional heterotrophic decay rate from laboratory experiments by monitoring the OUR exists, although such experiments usually require a significant amount of time and is based on a very small sample of the sludge. To experimentally determine the special b_H of the IAWQ model is much more complicated (unless Y_H and f_P are assumed to be exactly known, in which case equation (3.21) may be used), since it includes the transformation of decayed material into organic substrate. Also b_A is difficult to determine with any true meaning (Henze *et al.*, 1987).

In order to investigate the behaviour of model B, the off-line optimization algorithm is used to estimate the model parameters for exactly the same conditions as were described in case 3. The selection of cases to investigate (4A to 4E), depending on which variables are assumed to be measurable, are also identical to case 3. In Table 4.5 the results of the optimizations are presented for different sets of initial estimates.

When the results are compared to case 3, there are three important observations that can be made immediately. The first is the fact that model B is capable of mimicking the behaviour of the IAWQ model practically as well as model A although the number of model parameters have been reduced from ten to seven (compare the values of the loss functions in

Tables 4.4 and 4.5). A small increase of the loss function values (10–15 %) is noticeable for the cases where the microorganism concentrations are considered to be measurable (cases 4C, 4D and 4E), whereas the results for cases 4A and 4B are actually improved.

The second observation is that the variations of the values for the optimized sets of parameters are significantly smaller for model B as compared to the results when using model A. These smaller variations also mean that the loss function values are much more consistent for each investigated case when using model B. This fact more than adequately compensates for the somewhat higher values for cases 4C, 4D and 4E. The possibility of finding the truly optimum set of parameters (or values very close to it) is thereby significantly enhanced.

Finally, the algorithm appears to converge towards a global optimum parameter set not only for case 4E but also for case 4B. This is a dramatic improvement since the assumption of being able to monitor the different organism concentrations on-line (case 4E) is not practically possible. In contrast to this, the assumption regarding what measurements are available in case 4B is much more realistic.

When the five investigated cases are more closely examined, most parts of the conclusions from case 3 still hold. In case 4A a strong correlation between b, and r_H and r_A in the aerobic reactor is indicated. As the concentrations of microorganisms converge towards different values depending on the initial estimates, the yield coefficients vary significantly.

The inclusion of OUR in the calculation of the loss function in case 4B enables the optimization algorithm to determine an almost unique set of parameters. In case 3B only the convergence of the heterotrophic parameters in the aerobic reactor were improved to this extent. The sole major discrepancy is seen in row 3 for the autotrophic parameters. With this set of parameters the concentration of $X_{B,A}$ is about one fourth of what is predicted by the parameters in the other rows and, consequently, Y_A is four times smaller. The principal reason for this is that the autotrophic parameters are less sensitive to the OUR because the concentration of $X_{B,A}$ is very low when compared to $X_{B,H}$. Apart from this problem the results are remarkably consistent and this is clearly the main advantage of model B.

Optimization model B														value of loss func.
initial estimates							final estimates							
anoxic							anoxic							
r_H	Y_H	r_H	r_A	Y_H	Y_A	b	r_H	Y_H	r_H	r_A	Y_H	Y_A	b	
Case 4A (measured: X_{COD}, S_{NH}, S_{NO})														
.024	.35	.046	.111	.33	.10	.20	.011	.772	.018	.209	.635	.577	.114	992
.046	.69	.093	.222	.66	.20	.40	.011	.754	.017	.197	.615	.160	.101	958
.057	.86	.116	.278	.83	.25	.50	.014	.700	.012	.167	.434	.226	.073	1171
.068	1.03	.140	.333	.99	.30	.60	.015	.780	.027	.302	.679	.188	.209	1075
Case 4B (measured: X_{COD}, S_{NH}, S_{NO}, OUR)														
.024	.35	.046	.111	.33	.10	.20	.016	.697	.024	.282	.564	.236	.189	1192
.046	.69	.093	.222	.66	.20	.40	.016	.697	.025	.295	.568	.209	.202	1193
.057	.86	.116	.278	.83	.25	.50	.016	.697	.021	.257	.539	.055	.163	1185
.068	1.03	.140	.333	.99	.30	.60	.016	.700	.025	.287	.568	.206	.195	1189
Case 4C (measured: X_{COD}, S_{NH}, S_{NO} plus $X_{B,H}$, $X_{B,A}$, OUR in aerobic reactor)														
.024	.35	.046	.111	.33	.10	.20	.018	.631	.033	.393	.476	.363	.309	2363
.046	.69	.093	.222	.66	.20	.40	.015	.567	.023	.290	.403	.258	.196	2507
.057	.86	.116	.278	.83	.25	.50	.018	.564	.028	.330	.443	.299	.238	2431
.068	1.03	.140	.333	.99	.30	.60	.034	.699	.052	.609	.586	.550	.533	2509
Case 4D (measured: X_{COD}, S_{NH}, S_{NO} plus $X_{B,H}$, $X_{B,A}$ in anoxic reactor)														
.024	.35	.046	.111	.33	.10	.20	.022	.600	.027	.343	.418	.310	.253	1585
.046	.69	.093	.222	.66	.20	.40	.021	.588	.031	.374	.447	.330	.281	1604
.057	.86	.116	.278	.83	.25	.50	.017	.531	.019	.244	.330	.216	.147	1659
.068	1.03	.140	.333	.99	.30	.60	.020	.572	.021	.274	.356	.248	.181	1614
Case 4E (measured: X_{COD}, S_{NH}, S_{NO}, $X_{B,H}$, $X_{B,A}$, OUR)														
.024	.35	.046	.111	.33	.10	.20	.026	.633	.032	.398	.465	.364	.314	2450
.046	.69	.093	.222	.66	.20	.40	.024	.622	.034	.408	.479	.372	.323	2448
.057	.86	.116	.278	.83	.25	.50	.025	.628	.034	.408	.474	.371	.323	2447
.068	1.03	.140	.333	.99	.30	.60	.025	.631	.035	.425	.487	.385	.340	2449

Table 4.5 Results of the parameter estimation for case 4.

Case 4C shows a number of local optima yielding approximately the same values of the loss function but with very different parameter sets. Row 2 and 4 illustrate this problem well. For the parameters in row 2, the OUR is somewhat better predicted, whereas the $X_{B,H}$ concentration is more accurately predicted by the parameter set in row 4. The differences are, however, quite small and the total value of the loss function is the same. The parameter set of row 4 is clearly a local optimum, while the parameter values of rows 1, 2 and 3 are more similar but still vary significantly. The increased values of the loss function when compared to cases 4A, 4B and

4D indicate the difficulties to simultaneously predict both the organism concentrations and the OUR when using the reduced model B. The same fact was also observed for model A. The main reason for this is once again the different description of the organic substrate in the reduced models and the IAWQ model. Note that the ratios of r_A and Y_A are identical for all examples in cases 4C, 4D and 4E (when $X_{B,A}$ is considered to be measurable) and the correlation between b, and r_H and r_A in the aerobic reactor for case 4C is apparent.

In case 4D the OUR is excluded from the optimization. This leads to that also the ratios of r_H and Y_H converge towards the same values for all examples in both the anoxic and the aerobic reactor.

Finally, a global optimum is reached for case 4E. Due to the assumed measurements of $X_{B,H}$ and $X_{B,A}$ in both reactors the effects of the OUR measurements on the optimization are reduced and the convergence is improved when compared to case 4C.

The behaviour of model B is validated against data generated by the IAWQ model by simulating the two systems separately using identical influent wastewater characteristics, flow rates, recirculation rates, etc. In Figures 4.20–4.22 such a comparison is illustrated for some of the state variables. The behaviour of the other variables (organic substrate, ammonia, and nitrate) are similar to the ones earlier presented in Figures 4.15–4.19.

Figures 4.20 and 4.21 show how the concentration of heterotrophs is overestimated and how the concentration of autotrophs is underestimated when the OUR is used for the optimization and measurements of the organism concentrations are not available (i.e., case 4B). On the other hand, the prediction of the OUR is more accurate for this case, see Figure 4.22.

In order to more accurately predict the concentration of heterotrophs for case 4E compared to case 4B, the heterotrophic yield factor (Y_H) is reduced (among other things). Consequently, to increase the predicted concentration of autotrophs, the autotrophic yield factor (Y_A) is almost doubled (cf. cases 4B and 4E in Table 4.5). These parameter changes also have an immediate effect on the OUR. The factor $(1-Y_H)/Y_H$ in equation (4.11) is doubled, which, however, is compensated for by the fact that the $X_{B,H}$ concentration is reduced by almost 50 %. But since r_H is simultaneously increased to maintain the good predictions of X_{COD} and S_{NO}, the predicted oxygen uptake rate will be too high. The OUR for the autotrophs are on the whole fairly constant for cases 4B and 4E since the factor $(4.57-Y_A)/Y_A$ is

decreased by approximately 50 % while the factor $r_A X_{B,A}$ is increased by the same amount and, consequently, compensates for the different autotrophic yield factors determined for cases 4B and 4E.

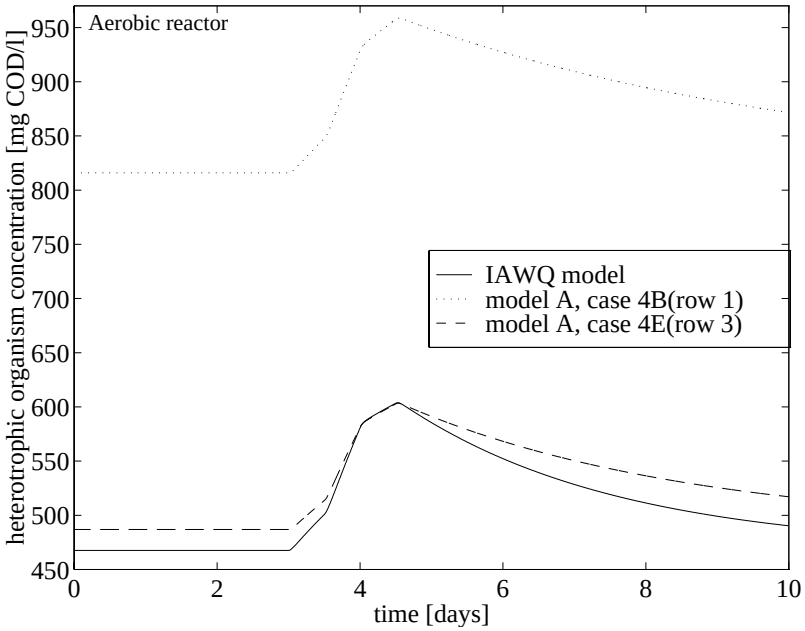


Figure 4.20 Heterotrophic organism concentration in the aerobic reactor using the IAWQ and B models (model B with two different sets of parameters) to simulate an entire AS process.

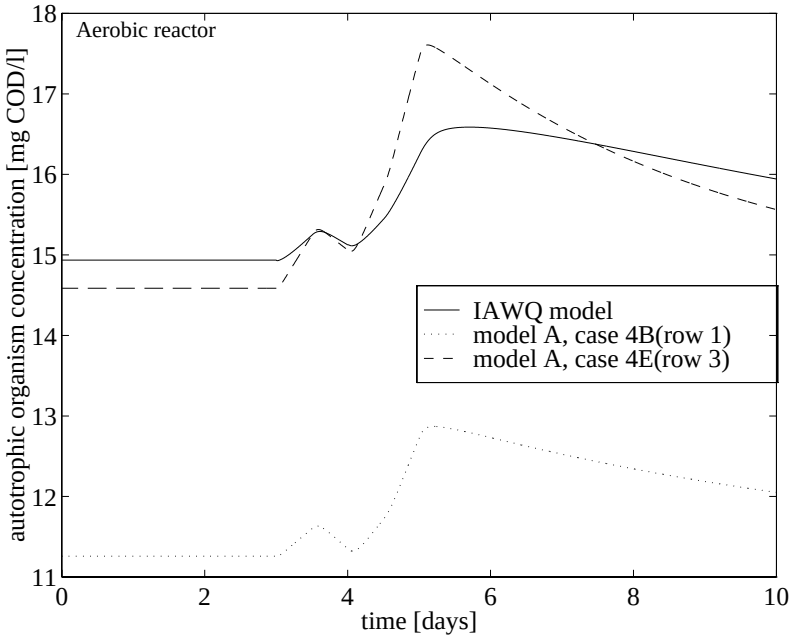


Figure 4.21 Autotrophic organism concentration in the aerobic reactor using the IAWQ and B models (model B with two different sets of parameters) to simulate an entire AS process.

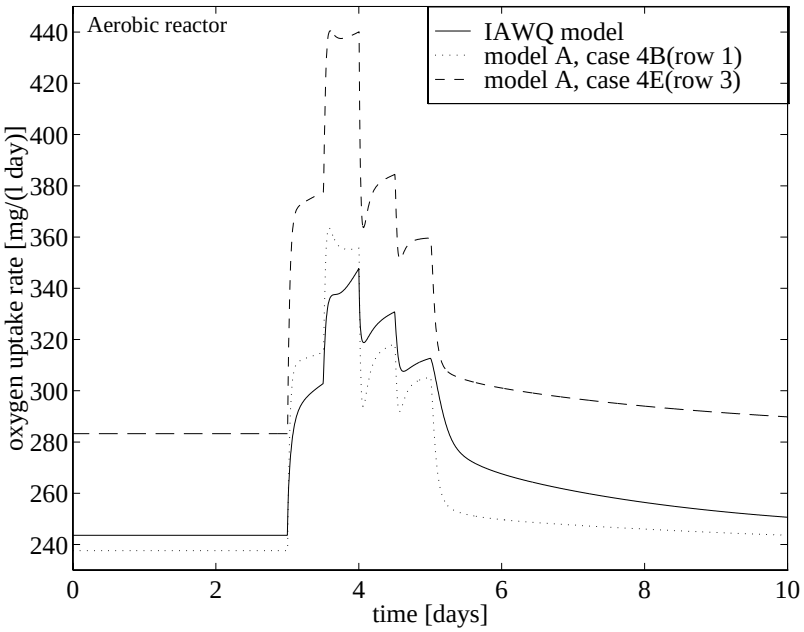


Figure 4.22 Oxygen uptake rate in the aerobic reactor using the IAWQ and B models (model B with two different sets of parameters) to simulate an entire AS process.

Summary Discussion

A number of case studies have been presented in this section to illustrate the behaviour and identifiability properties of the reduced order models with regard to operational conditions, assumed measurable variables, initial parameter values, etc. The investigations do not cover all important aspects of the models but provide a basis for further analysis.

In cases 1 and 2 it was shown that the reduced model A was capable of mimicking the behaviour of the IAWQ model with reasonable accuracy in unit operation (only one type of reactor). The main differences were due to the fact that the reduced models do not include a hydrolysis process for the organic matter which affects the time constants of the models. Difficulties to identify global parameter sets for the simplified model were also clearly indicated.

This difficulty was further investigated by examining the sensitivity of the model to parameter changes. The model proved to be especially insensitive to variations of the decay rate coefficients but also r_A and Y_A were almost impossible to determine separately for the applied conditions and assumed available measurements.

In case 3 the complete model A was used to simulate a WWT plant with both anoxic and aerobic reactors and the results were compared with simulations of an identical plant simulated with the IAWQ model. The recirculation introduced new problems when trying to determine a global optimum parameter set by means of optimization. Not even when highly favourable conditions were assumed, could the autotrophic decay rate coefficient be uniquely determined.

This motivated the use of a further simplified model – model B. In this model all the decay rate coefficients were identical. This approach significantly improved the optimization results without deteriorating the general behaviour of the model, as was shown in case 4. What appeared to be an almost global optimum parameter set could be determined without assuming the concentrations of the different microorganisms to be measurable.

However, it appears troublesome to uniquely identify all the parameters of the models based on measurements of X_{COD}, S_{NH} and S_{NO} alone (at least for normal continuous-flow operating conditions). In many cases when measurements of the OUR were considered to be available, the estimation results were significantly improved, especially for parameters describing the heterotrophic biomass in case 3 and for all parameters in case 4. New sensor technologies, such as on-line respirometers which are currently being developed, may therefore provide a very useful tool for identification and validation of models describing the AS process. An important reason for this is that the OUR provides direct information concerning the activities of the microbial processes while traditional measurements of concentrations only provide indirect knowledge.

In some cases it was assumed possible to measure the concentration of active heterotrophic and autotrophic biomass. This is unrealistic but useful when fundamental properties of a model are investigated. When $X_{B,H}$ and $X_{B,A}$ are not measurable, these model variables may converge towards very different values when the model is optimized, depending on the initial parameter estimates. This leads to non-unique parameter sets because high concentrations of microorganisms with a low activity in the system result in approximately the same model behaviour as low concentrations of microorganisms with a high activity.

Since it is practically impossible to measure the concentration of *active* biomass directly, it may prove useful to base an AS model on estimations of actual reaction rates instead of rate coefficients. Knowledge of the reaction rates are important for process control, whereas the rate factors

may be more important for biological interpretations concerning the state of the process. For example, estimation of $r_H X_{B,H}$ as a combined parameter and not as two separate variables may improve the identifiability of a model, especially as the $X_{B,H}$ concentration is not measurable. All the parameters r_H, r_A, b_H and b_A of the reduced models always appear in multiplicative combinations with $X_{B,H}$ or $X_{B,A}$ and, consequently, the identifiability may be global in the sense that the parameter combinations converge towards identical values but may assume practically any value if all parameters are analysed separately. Fairly simple measurements of the suspended solids concentration or the total COD content may also prove useful since they contain a certain amount of information about the microorganism concentration (though in a complex combination with other matter). The development and analysis of an AS model which uses the concept of direct estimation of reaction rates is an interesting topic for future work.

For all case studies presented, the measurements were assumed to be completely free from noise. An analysis of the impact of noise (Gaussian noise with reasonable variance) shows that the effect is negligible for the type of optimization presented here since it is based on a large number of measurements. The reduced models with optimized sets of model parameters have also been compared with the IAWQ model using different types of input dynamics than have been presented in this section. The different behaviours of the reduced models and the IAWQ model were, however, most prominent for step variations, which motivated the choice.

No thorough analysis has been performed to determine how the process should be perturbed to further enhance the model identifiability. Influent variations synchronized with the time constants of the model, repeated rapid input variations of high amplitude, etc., may improve the identifiability. However, since the possibilities to control and manipulate the influent wastewater are quite limited for most WWT plants, such an investigation would probably produce results which are not practically feasible, though theoretically interesting.

As a final remark the computational effort is commented upon. In order to determine *one* set of optimized parameters the system of differential equations describing the dynamics of the plant has to be simulated between 500 and 1000 times (the simplex algorithm converges slowly). This means that the required CPU time for one optimization on a ‘standard’ workstation is measured in hours and even days rather than seconds and minutes. The work presented in this section is based on hundreds of such optimizations.

For practical use the optimization algorithm will have to be modified (for example, combined with faster but less robust algorithms once a reasonable set of parameters have been determined) or the error tolerance increased.

4.5 On-Line Estimation

The main disadvantage of the off-line optimization approach used in the previous section is that the resulting model is fitted to certain operational and influent conditions. The model is capable of predicting the behaviour of the real process as long as these conditions are not significantly changed. However, over longer periods of time, small variations will also accumulate and affect the plant performance and the behaviour of the microorganisms. Therefore, it may prove necessary to update the model parameters on a regular basis. A possible solution is to perform a new optimization when needed but better still is to automatically update the model and track the parameters on-line as new measurements become available. This will guarantee that the model predictions are as reliable as possible since the model will always be calibrated for the current conditions. In this section results from such an approach are illustrated and discussed.

Off-line and on-line methods should produce approximately the same results when used to estimate process parameters during identical operating conditions. Off-line methods are, however, often easier to use and also more robust. Therefore, the simplex method has been used for the principal model investigations reported in Section 4.4 although on-line parameter estimation is of great importance for time varying systems, such as the AS process. The Kalman filter (described in Appendix E) can be shown to be an optimum on-line estimation algorithm under certain conditions. However, as a result of the recirculation and the long time delays of the activated sludge system difficulties may occur. When a parameter is adjusted the effect on the model behaviour is only partly ‘instantaneously’ noticeable. After a period of time (depending on the design of the plant but normally several hours) the model input is also changed by the impact of the recirculation, which is interpreted as a new disturbance of the system and, consequently, may induce new parameter adjustments. In certain cases this can lead to stability problems when the parameters are over-compensated back and forth. Especially the effect on the microorganism concentrations is slow and it may take days before significant variations

can be observed. To some extent it is possible to modify the estimation algorithms to take the correlation of a changing parameter and a later change in the model feedback into account. Such modifications, however, have not been tested in this study. Instead the maximum rate of change for the model parameters is set to a low value (by adjusting the gain matrix of the Kalman filter) to avoid oscillations in the parameter estimates.

The reasons above motivate the use of a combination of on-line and off-line methods. By first applying an off-line optimization algorithm, a set of reasonable parameter values can be determined. These values may then be used as initial seeds for the on-line estimation algorithm, thereby reducing the risk of divergence and unwanted oscillations.

When performing on-line estimation using real data it is more important to determine the noise characteristics of the process and the measurements. Whereas the off-line methods are based on a large number of samples and the effect of noise is reduced by the averaging calculations, the on-line algorithms respond immediately to sudden changes in the measured data. By performing a thorough investigation of the noise distribution of the process, this information can be included in the identification algorithm to prevent the estimated parameters from changing rapidly in an unrealistic manner due to the influence of noise. Furthermore, the measurements should be 'logically' analysed before the data are used for on-line parameter identification. The reason for this is to detect trends, outliers, drastic sudden changes, etc., which may indicate that a certain sensor is not properly calibrated or is out of operation, before the identification procedures are applied and produce an erroneous result. This can be performed on-line and would be one important function of the top level knowledge-based system outlined in Figure 2.4.

A number of cases will be presented in this section to determine the basic behaviour of the reduced order models when used for on-line parameter identification. The assumed conditions for the simulated WWT plant will be those already described in Section 4.3 and previously used for the off-line optimizations. The following situations will be investigated:

- a case study of a *modified* model A for various assumed available measurements and initial estimates (the same principle as used in Section 4.4 for cases 3A, 3B and 3E (without the OUR));
- the effects of both measurement and process noise on the identification results;
- the effect of a change in one of the IAWQ model parameters on the reduced model behaviour during on-line estimation.

Results from Tables 4.3 and 4.4 indicate that the decay rates (b_H and b_A) may assume almost any values especially when only X_{COD}, S_{NH} and S_{NO} measurements are available. If all ten parameters of model A are identified simultaneously with the extended Kalman filter this is manifested as divergence. The correlation between the decay rate factors and the other parameters often lead to an oscillatory behaviour of the estimates, that is, one parameter changes at a certain rate and another variable changes with the same relative rate and the total effect on the measurable state variables is negligible. Finally, the identification process breaks down.

To avoid this problem the elements of the Kalman gain matrix related to the decay parameters are set to zero, that is, the decay rate coefficients are assumed to be constants. The more variables that are measurable the more parameters may be estimated successfully. If the OUR or the organism concentrations are assumed to be measurable as well, one or two of the decay parameters may be estimated too. All ten parameters, however, cannot be estimated simultaneously for those cases either. In order to allow the results of the different on-line estimations to be compared, all decay parameters are constant in this section. The model used for the on-line estimations is consequently a simplification of models A and B since only six parameters are assumed to vary (ten and seven parameters are considered to vary in models A and B, respectively).

One basic problem was already discussed in Section 4.4. It is 'easy' to estimate the net reaction rates for the organisms but much more difficult to determine the growth rates and the decay rates separately. Future models should be developed taking this fact into account.

In Figure 4.23 the results from an on-line estimation are shown for some key variables. The measurements assumed to be available for the identification in this case are the ones described for case 3A (i.e., X_{COD}, S_{NH} and S_{NO} in both reactors) but the process is now in steady state. In order to initiate the on-line estimation with reasonable parameter values, the off-line results from case 3A(row 2) are used (solid line). This parameter set is then increased (dashed line) and decreased (dotted line) by 15 % (not the decay rate coefficients) and the identification procedure is repeated for these new initial parameter sets to investigate the possibility for global convergence.

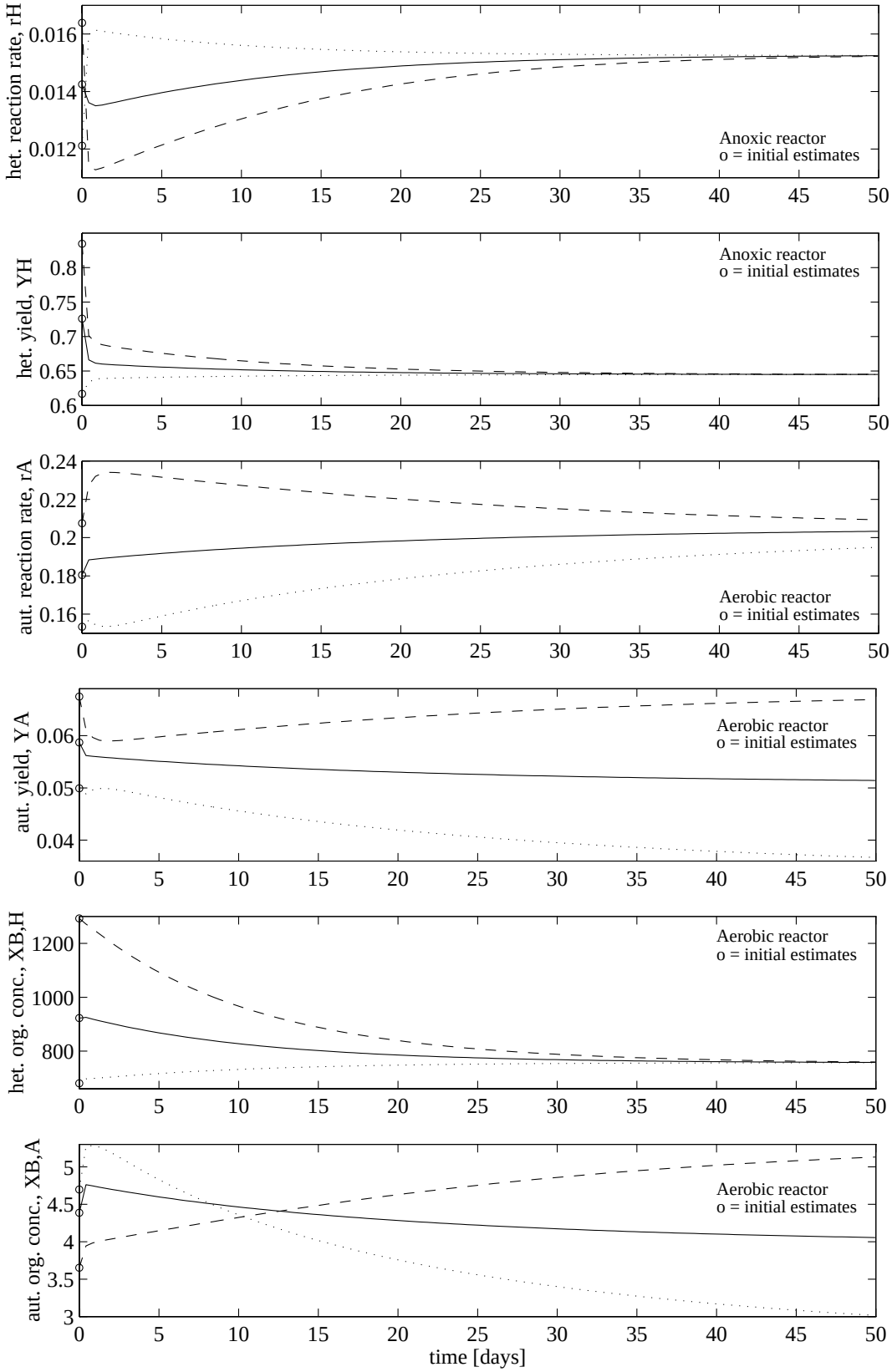


Figure 4.23 On-line state and parameter estimation of a modified model A based on measurements of organic substrate, ammonia and nitrate concentrations from a stationary AS process.

The graphs in Figure 4.23 illustrate that the heterotrophic parameters (shown for the anoxic reactor) and, consequently, the heterotrophic organism concentration converge towards identical values independently of the initial estimates. This is possible since the decay rate coefficients are identical for all simulations (if not, the identification algorithm would find different estimates). For the autotrophic parameters, however, the available measurements are not sufficient. The autotrophic reaction rate converges towards the same value, whereas the yield coefficient converges towards different values. Since $X_{B,A}$ is not assumed to be measurable, this variable may compensate for the effects of the different yield values. A high value of the yield factor leads to a high concentration of autotrophs and vice versa. The ratios of $X_{B,A}$ to Y_A converge towards identical values but the Kalman filter cannot determine a global estimate for the two variables separately, especially not based on data from a process in steady state. The entire model is also less sensitive to variations of the autotrophic parameters due to the low concentration of autotrophs when compared with the concentration of heterotrophs.

The low sensitivity is a major reason why the convergence rate is considerably slower for the autotrophic parameters, which is clearly illustrated in Figure 4.23. This is an expected result as the time constants for the dynamics of the autotrophic biomass are larger than for other mechanisms in the activated sludge process. Furthermore, the convergence rates for all parameters shown in Figure 4.23 are very slow because of the used steady-state data. Such data produce small residuals and, consequently, a slow convergence rate. The simplified Kalman filter used in this study (constant gain matrix) further emphasises this fact. The rate of convergence is also correlated to the number of measured variables (see Figure 4.25) and the number of parameters to be estimated. Comparisons of the convergence rates of the different parameters for different cases may consequently provide some valuable information.

The estimation procedure described above is repeated to obtain the results shown in Figure 4.24. In this case the identification is based on the assumption that the OUR is also measurable. The first initial parameter set is taken from case 3B(row 2) (solid line) and then changed ±15 % (the same principles as were used in the previous example, +15 %: dashed line; –15 %: dotted line). The same data, generated by simulating the IAWQ model for a complete WWT plant during steady-state conditions, are used.

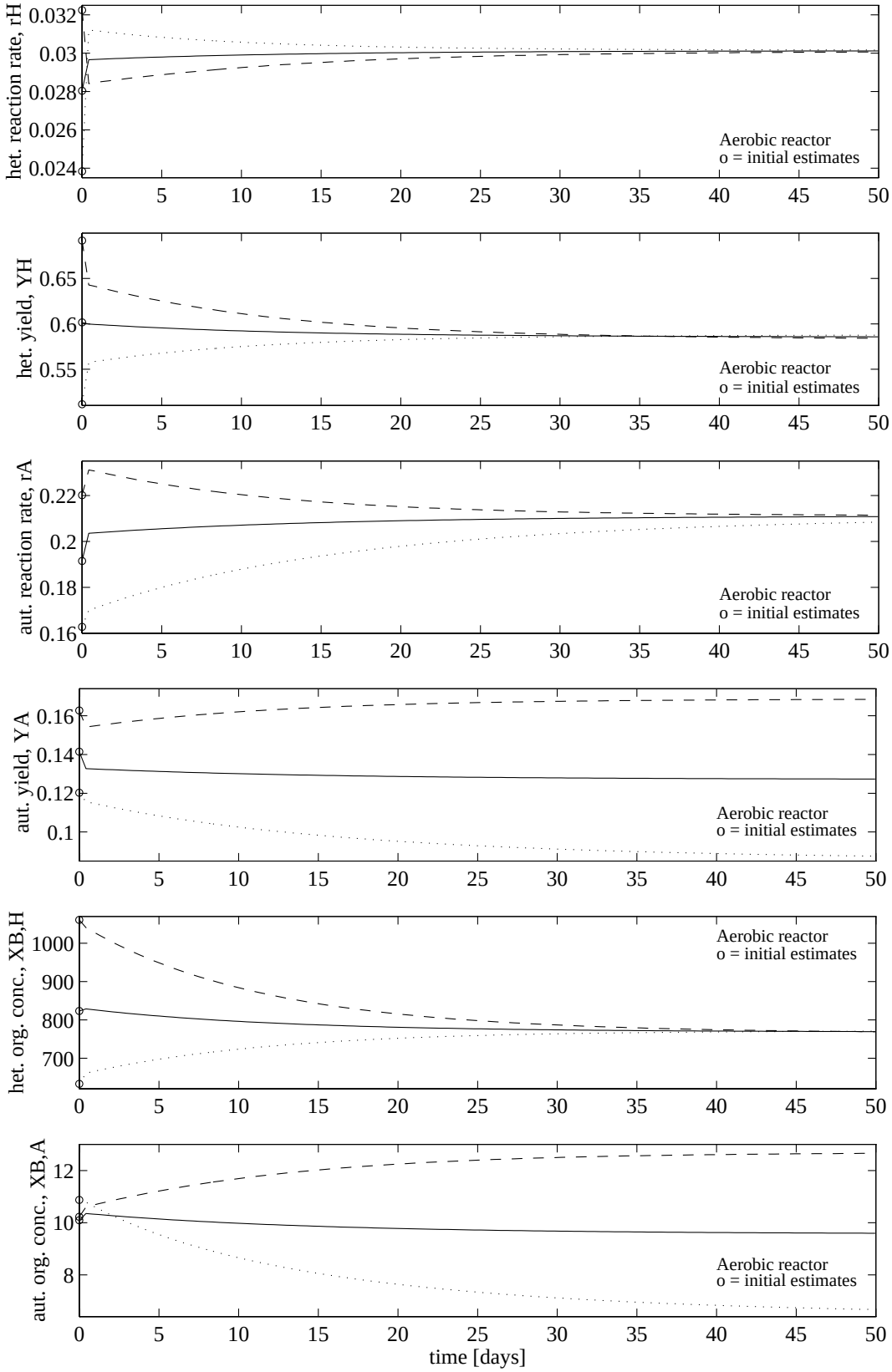


Figure 4.24 On-line state and parameter estimation of a modified model A based on measurements of organic substrate, ammonia and nitrate concentrations plus the OUR from a stationary AS process.

The graphs in Figure 4.24 show that the additional information from the OUR measurements is not sufficient for the identification algorithm to estimate a global set of parameters. It was already discussed in Section 4.4 that the effect of the OUR is most prominent for the heterotrophic parameters because of the low concentration of autotrophs. The autotrophic yield factor and the concentration of autotrophic biomass still converge towards different values depending on the initial parameter set. The situation is, however, somewhat different when compared with the previously investigated case. The ratios of $X_{B,A}$ to Y_A do not converge globally. Rather the weighted mean value of this ratio and the expression $(4.57 - Y_A) X_{B,A}/Y_A$, which is part of the calculation of the oxygen uptake rate (4.11), converge in a global sense. Thereby the reduced model may provide a good fit to data from the IAWQ model for several sets of parameter estimates also when the OUR is assumed to be available.

In Figure 4.24 no estimations of the heterotrophic parameters in the anoxic reactor are shown. In this reactor the OUR does not provide any new information and the estimation results are therefore similar to what was illustrated in Figure 4.23. An examination of the autotrophic variables in Figures 4.23 and 4.24 show a higher rate of convergence when the OUR is measurable. The effect of the OUR measurements is not significant enough to determine a global set of parameter estimates but it increases the convergence rate. Although not shown, the convergence rate for the heterotrophic parameters also increase when compared with the previous example. The more information available, the faster and more accurate is the convergence of the identification algorithm.

Finally, the identification procedure is tested on a new case. Now the OUR is not considered measurable, instead the organism concentrations ($X_{B,H}$ and $X_{B,A}$) in both reactors are assumed to be available. Some results are presented in Figure 4.25. The initial parameter set is taken from case 3E(row 3) (solid line) and then changed ±15% (+15%: dashed line; –15%: dotted line). Note that two different time scales are used in the graphs. Neither of the time scales are the same as the one previously used in Figures 4.23 and 4.24.

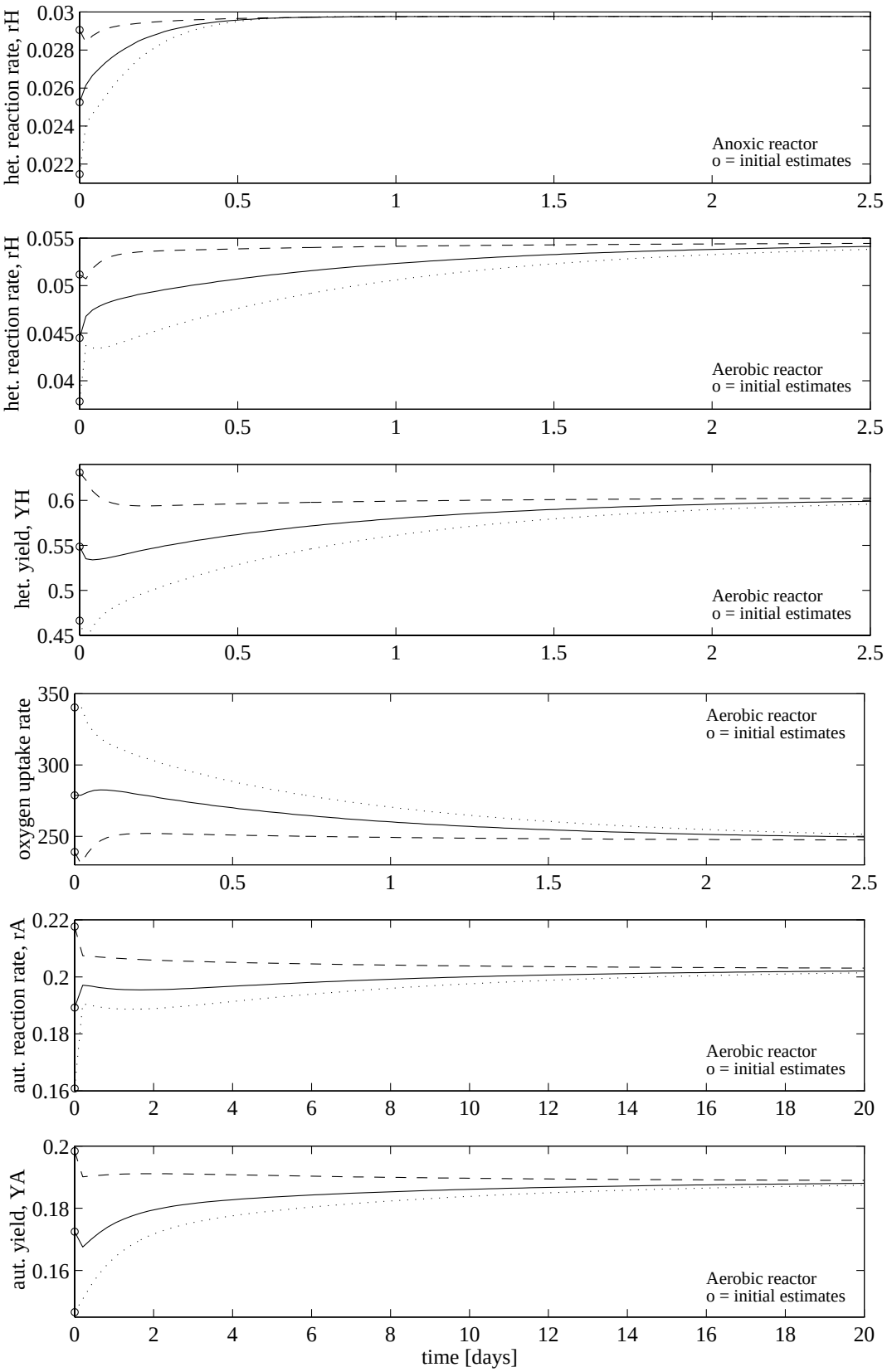


Figure 4.25 On-line parameter estimation of a modified model A based on measurements of organic substrate, ammonia, nitrate and two organism concentrations from a stationary AS process.

The graphs in Figure 4.25 illustrate how all parameters converge globally for different initial parameter sets. The rate of convergence is also dramatically higher than for the earlier shown cases. In order to determine a global set of estimates it would actually be sufficient to use measurements of $X_{B,A}$ from the aerobic reactor. Possibly, measurements of the total organism concentration $(X_{B,H} + X_{B,A})$ would also be sufficient instead of requiring individual data of both types of organisms. Such approaches would, however, slow down the convergence rates. The prediction of the oxygen uptake rate based on the estimated parameters also converges rapidly. Moreover, it converges towards the identical value as determined by the IAWQ model (244 mg O_2/(l day)), although it is only estimated by using the reduced order model.

The actual values of the parameters estimated in this section are not identical when compared with the optimized sets determined in Section 4.4. The reason for this difference is that the weighting of the residuals performed by the gain matrix of the Kalman filter is not exactly the same as for the loss function used by the simplex method. The off-line optimizations were also based on data generated during dynamic conditions, whereas the on-line estimations are performed using steady-state data.

The large initial transients for some of the estimates shown in Figures 4.23–4.25 are not an error produced by the identification algorithm. Instead, they are the result of the very large residuals that occur at the early stage of the estimation. The reason for these large residuals is that the reduced model is simulated towards steady state prior to the identification procedure is initiated and the steady-state values are then used as initial predictions by the simplified model. Consequently, the values may be quite different when compared with the assumed measured variables of the IAWQ model in steady state. The use of a constant gain matrix for the Kalman filter further emphasizes the large initial variations.

In order to investigate how noise affects the results of the on-line estimation algorithm the available measurements are corrupted by Gaussian white noise with a mean value of zero and a standard deviation which is 10 % of the steady-state value of each variable. To complicate matters further, noise with the same distribution is also added to the flow rate, organic substrate concentration and ammonia concentration of the influent wastewater, which are inputs to the IAWQ model used to simulate the true WWT plant. Otherwise the conditions are identical to those that have already been described in the previous examples. Note that the noise components added to the different variables are uncorrelated.

To demonstrate the magnitude of the chosen noise distribution added to the model variables and to illustrate the difficulty the identification algorithm has to deal with, two noise corrupted variables are shown in Figure 4.26.

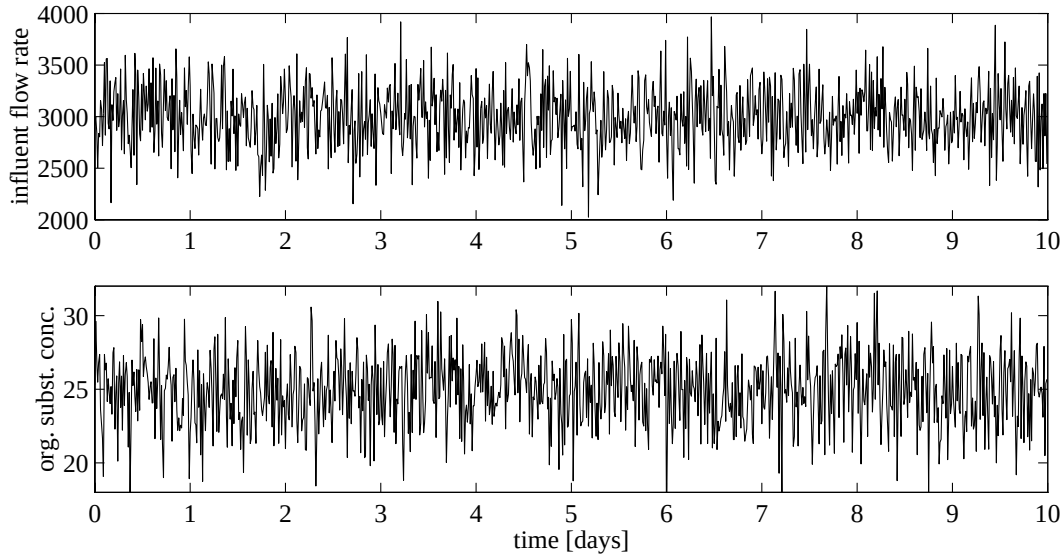


Figure 4.26 The magnitude of the used noise distribution shown for the influent wastewater flow rate and measurements of the organic substrate concentration in the anoxic reactor.

Noise causes problems for the identification algorithm. The residuals by which the Kalman filter determines how the parameters are to be updated are dramatically affected by the noise and it may very well conceal the true differences between the measurements and the model predictions, thereby causing the identification to fail. This problem is more apparent when the process is close to steady state (i.e., small residuals) and the situation can be improved by perturbing the system in a suitable manner. Since the influent wastewater characteristics to a real WWT plant is only to a small extent controllable (in practice only the flow rate is to a certain extent controllable), such perturbations may be difficult to effectuate.

For this reason the identification procedure is performed using noisy data for steady-state conditions, thereby giving the algorithm every possibility to fail. Investigations have been performed for all three cases shown in Figures 4.23, 4.24 and 4.25. For all cases the behaviour of the estimates was approximately the same as already shown in those figures. The identification algorithm converge towards the same values as when the data was free from noise, independently of the initial parameter values. The estimates are naturally corrupted by the added noise although the Kalman filter reduces its effects. By using a time-variable extended Kalman filter the results could be improved further (on-line updates of the gain matrix).

A few results from the estimations including noise are illustrated in Figure 4.27. The measurable variables and the initial estimates are the same as those used for one of the estimations shown in Figure 4.25. In order to allow for a closer comparison with the noise-free results, the estimates are filtered using a traditional low-pass filter and then plotted together with the estimates produced when no noise was present. Such a comparison is shown in Figure 4.28 for some key variables. It is clear that the correct information with regard to the parameter values is available also from the noisy estimates.

The last example in this section demonstrates how the parameters of the reduced model can be tracked and updated on-line to maintain a good fit of the model to the measurements as conditions change. In order to simulate this, the maximum specific hydrolysis rate (k_h) of the IAWQ model is increased by 50 % over a period of one day starting at $t = 1$ day (a ramp disturbance) and then maintained at this higher level throughout the simulation. The hydrolysis rate was selected as a suitable parameter because of its significant influence on the overall behaviour of the IAWQ model. In reality this change would reflect that the character of the influent slowly biodegradable organic substrate had changed and could now more easily be transformed into readily biodegradable substrate. A step increase of the influent flow rate by 50 % (starting at $t = 6$ and lasting for one day) is also simulated to investigate the behaviour of the reduced model to such a disturbance. The process is initially in steady state. Note that no noise has been added to the measurements in this example.

The effects on the behaviour of the IAWQ model, when the hydrolysis rate is increased, are mainly a considerable reduction of the amount of slowly biodegradable substrate in both reactors (in this case about 50 % in steady state) and a lower concentration of nitrate in both reactors as the produced readily biodegradable substrate improves the denitrification. The concentration of heterotrophs is also somewhat increased. The effects on the ammonia concentration and the autotrophic biomass are negligible.

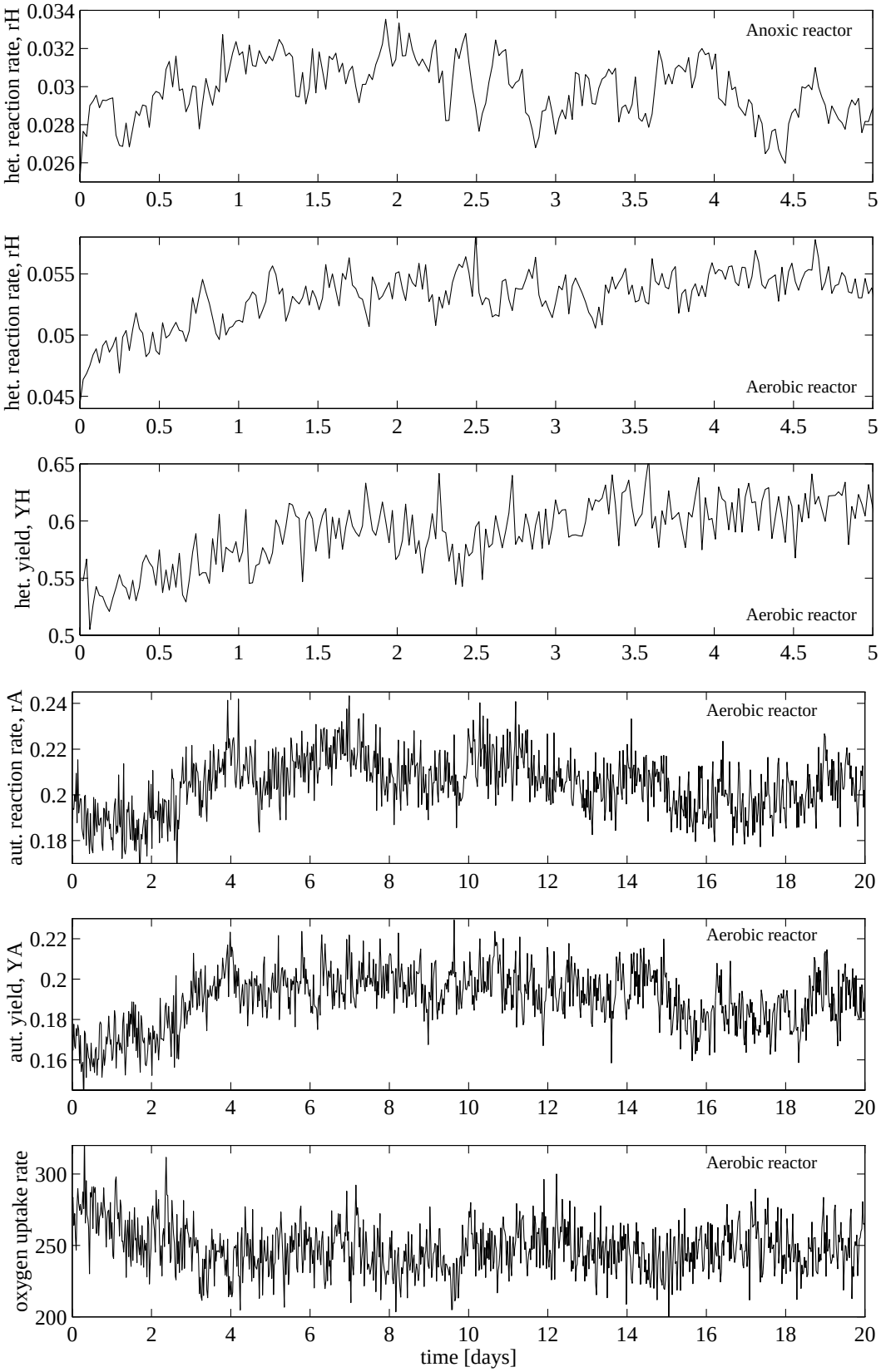


Figure 4.27 On-line parameter estimation of a modified model A based on noisy measurements of organic substrate, ammonia, nitrate and organism concentrations from a stationary AS process.

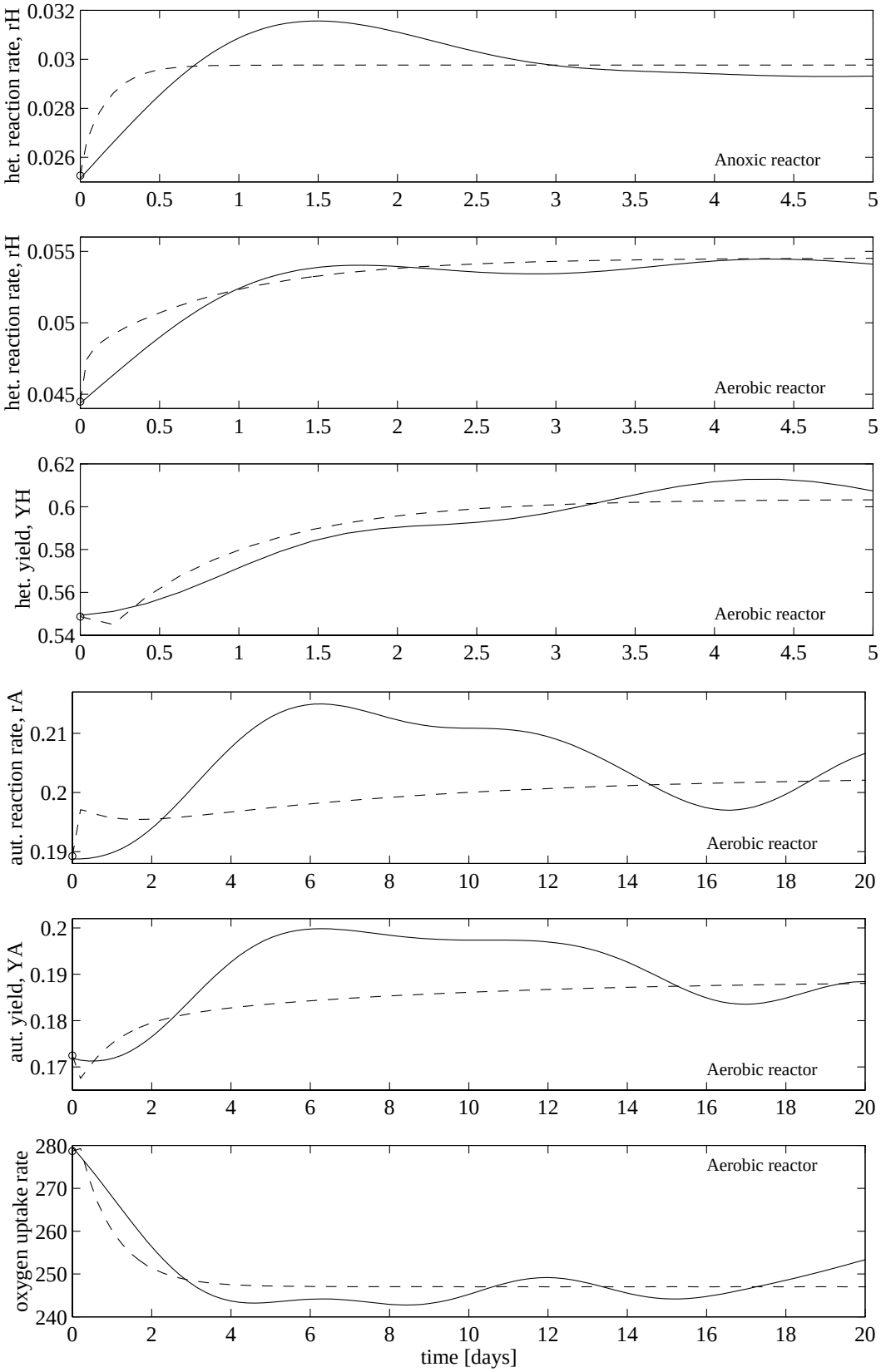


Figure 4.28 The estimates from Figure 4.27 after low-pass filtering (solid) compared with the same estimates (previously shown in Figure 4.25) determined from noise-free measurements (dashed).

All three cases earlier described in this section (based on which measurements are assumed to be available) have been investigated. It is interesting to note that the model parameters are qualitatively updated in the same way for all three cases, as shown for a few examples in Figure 4.29. Note that the convergence rates are not a problem for any of the cases because the process is now in a transient state due to the changes of the hydrolysis rate and the influent flow rate. Consequently, the rate of convergence is much higher than for the previous examples where steady-state data were used for the estimations. Prior to $t = 0$, the Kalman filter has been used to estimate the model parameters with the initial steady-state data so that the parameters of the reduced model have converged, in order to avoid transients in the graphs shown in Figures 4.29 and 4.30.

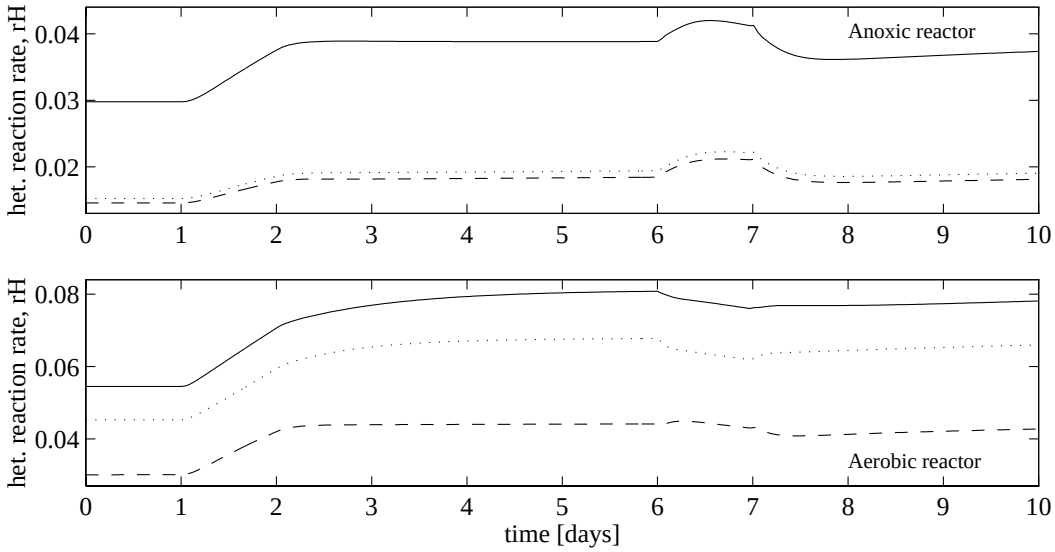


Figure 4.29 On-line parameter tracking based on different available measurements – X_{COD}, S_{NH} and S_{NO} (dotted line), plus OUR (dashed), plus $X_{B,H}$ and $X_{B,A}$ (solid). The hydrolysis rate is increased by 50% from $t = 1$ to $t = 2$ (ramp) and then maintained at this level while the influent flow rate is increased by 50% at $t = 6$ (step) and set back to its original value at $t = 7$.

Figure 4.30 illustrates some more detailed results of the on-line parameter estimation when X_{COD}, S_{NH}, S_{NO}, $X_{B,H}$ and $X_{B,A}$ are assumed to be measurable. Predictions of the oxygen uptake rate is also shown. As the hydrolysis rate of the IAWQ model increases, the total amount of biodegradable organic substrate ($S_S + X_S$) is decreased and, consequently, the measured values of X_{COD} are reduced. As an effect r_H increases rapidly (in both reactors) but during the most transient stage (from $t = 1$ until 2) the heterotrophic yield is also affected. This is the way for the reduced model to rapidly reduce the X_{COD} concentration while maintaining the correct

concentration of heterotrophs during the transient phase – r_H and Y_H are correlated. As the process settles down at the higher hydrolysis rate, r_H reaches an optimum value and Y_H returns to approximately the original value, which is realistic as the concentration of active biomass has not changed in any significant manner. Measurements of the OUR would reduce the variations of Y_H. The effects on the autotrophic parameters caused by the changing hydrolysis rate are quite small, as would be expected.

The influent flow rate disturbance is introduced at $t=6$. If the reduced model was a perfect replica of the true process (in this case the IAWQ model), this perturbation would not require any model parameters to be updated since both model and process would then react in exactly the same way to such an external disturbance. This almost holds for the heterotrophic parameters where only small adjustments (≈ 2–$3\,\%$) are required to maintain a good fit of the model. For the autotrophic parameters larger adjustments are required ($\approx 10\,\%$). The reaction rate factor r_A is decreased primarily because the autotrophic growth rate of the IAWQ model is in its most non-linear region for the concentrations considered in this example. The sudden increase of the ammonia concentration caused by the increased flow rate leads to an overly large predicted growth rate (cause by the linear relationship used in the reduced model) for the autotrophs and, consequently, the value of r_A must be reduced. For the same reason Y_A is increased. The predicted values of the OUR are during the whole experiment fairly accurate when compared with the calculated values of the IAWQ model (not shown in the graph).

Note that if the initial and final parameter values shown in Figure 4.30 are compared, only the heterotrophic reaction rate factor has increased (for both reactors) with any real significance. All other parameters are approximately the same. This is a realistic consequence of the imposed disturbance of the process. A higher rate of the hydrolysis mechanism produces more readily biodegradable substrate and, consequently, this substrate is consumed at a higher rate by the active heterotrophic biomass, that is, in total an increased reaction rate of the heterotrophic organisms.

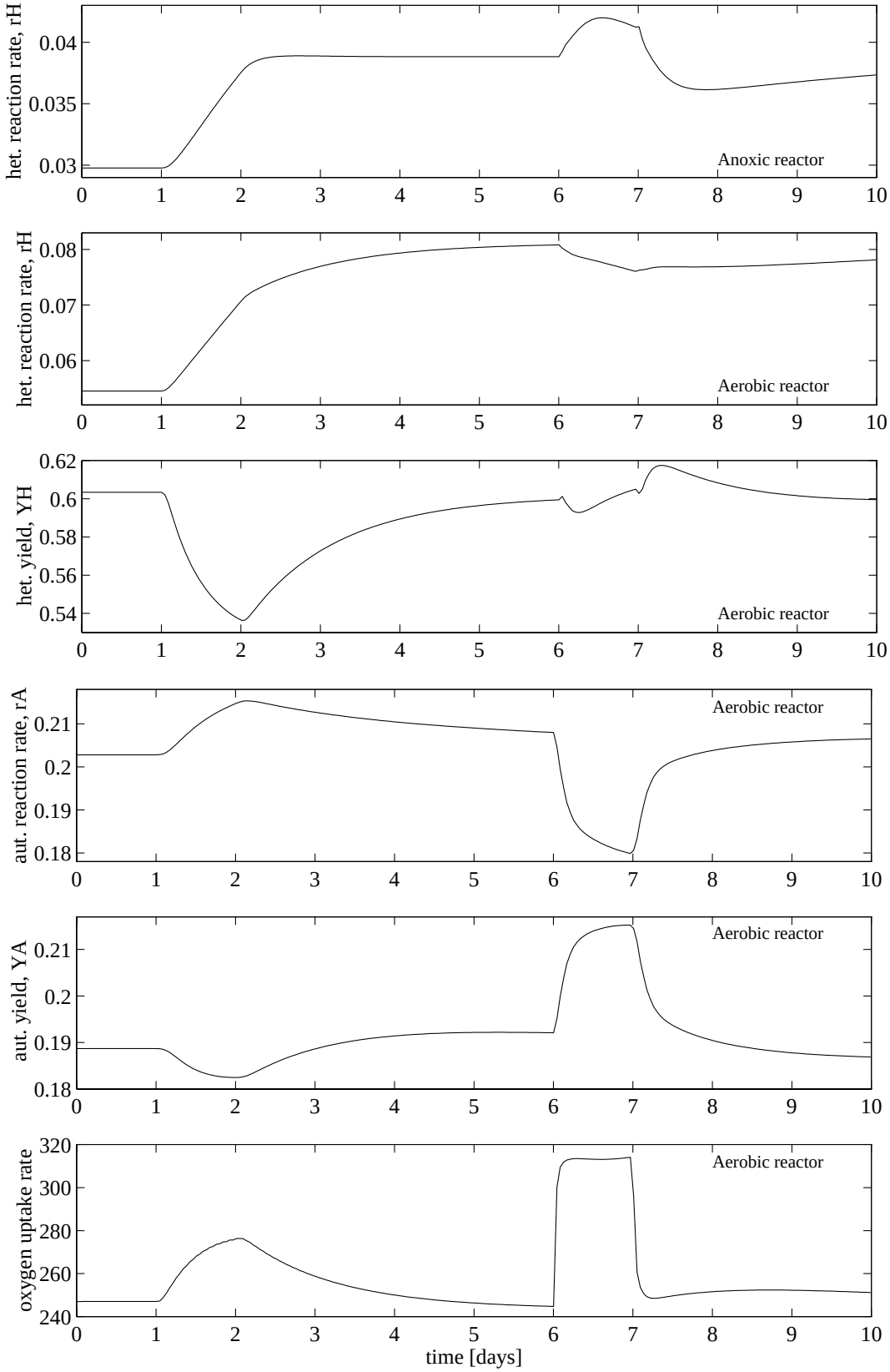


Figure 4.30 On-line parameter identification based on measurements of X_{COD}, S_{NH}, S_{NO}, $X_{B,H}$ and $X_{B,A}$ during a change of the maximum hydrolysis rate (starting at $t=1$) and a disturbance of the influent flow rate (starting at $t=6$).

Summary Discussion

The examples presented in this section illustrate the behaviour of the reduced model during on-line state and parameter estimation using an extended Kalman filter. The model was investigated assuming different variables to be measurable, for different sets of initial parameter sets, regarding the sensitivity to noise, etc. As a result of the off-line optimizations in Section 4.4 it was clear that all ten parameters of model A could not be estimated simultaneously for the conditions considered in this study. Early on-line estimation studies confirmed this conclusion. Trying to estimate all ten model parameters led to divergence and unstable estimates. Therefore, a further simplified model with constant decay rate factors was used. A total of six model parameters and a number of state variables were estimated on-line based on assumed measurements from a continuous-flow, single-sludge AS process with pre-denitrification during normal operating conditions.

The estimates of the model converge for all tested cases even when steady-state data are used. However, in order for all parameters to converge globally some information with regard to the concentration of micro-organisms is required. Otherwise, $X_{B,A}$ and Y_A cannot be uniquely determined – instead the ratio of the two variables converges globally. The rate of convergence of the extended Kalman filter is also significantly higher when measurements of the organism concentrations are available.

Even when a significant amount of noise is added to the measurements and the simulated process, the estimated parameters for all cases converge towards approximately the same values as when no noise was added. Also when the real process is exposed to both internal and external disturbances, the on-line parameter tracking system still produces reasonable and fairly accurate results.

PART III

Modelling the Settling Process

Chapter 5

Modelling Approaches – a Review

In this chapter we focus on the mathematical modelling of the secondary clarifier. A historical perspective is provided and the traditional modelling concepts based on one-dimensional layer models are described. The solids flux theory by Kynch (1952) and the continuity equation, both fundamental for settler modelling, are discussed. Finally, a new robust settler model introduced by Diehl (1995a) is presented. This chapter covers material from [90], [95], [186], [188] and [189].

5.1 Fundamentals

The separation and concentration of the active biomass in an activated sludge (AS) process is performed in a settling basin, often referred to as the secondary clarifier, secondary settler or secondary thickener. The force that makes the sedimentation of the particles in the liquid possible, originates from gravity and the density differences between the particles and the liquid. From the bioreactor the mixed liquor enters the secondary clarifier where it should be sufficiently clarified in order to produce an effluent of acceptable quality. The sludge should also be adequately thickened so that the desired solids level in the bioreactors can be maintained through the sludge recirculation and enhance an effective treatment of the wasted activated sludge. This means that the settler combines the functions of clarification and thickening into one unit. Should the settling tank fail with respect to either of these functions, the result would be a rapid increase of suspended solids in the effluent or a deterioration of the AS process. Practical experience has shown that the secondary clarifier is often the main bottleneck of the entire AS process.

The mixed liquor is a flocculent suspension in which larger particles can be formed by the coalescing of particles which have collided. These larger particles generally enhance settling characteristics. The particle distribution is bimodal with primary particles (microflocs) in the 0.5 to 5 μm range and flocs (macroflocs) in the 10 to 5000 μm range (Billmeier, 1978; Parker 1983), as illustrated in Figure 5.1. The settling properties of a particular sludge depends both on the distribution of primary and floc particles and on how easily the primary particles are entrapped into larger flocs. Other factors that influence the settling behaviour are the hydraulic regime, temperature, basin design, flow and feed variations, sludge characteristics, predators consuming dispersed bacteria, etc.

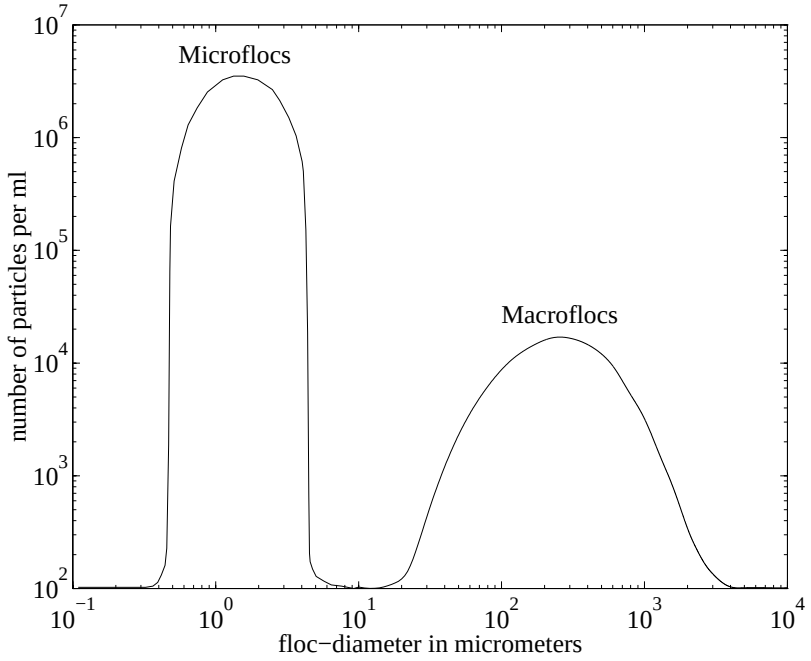


Figure 5.1 Particle diameter distribution in activated sludge according to Billmeier (1978).

The complex behaviour of the secondary clarifier and its great importance for the successful operation of the activated sludge process have made the settling process a major issue for researchers working within the field of mathematical modelling. An excellent review of the historical development of models for the secondary clarifier is given in Lumley (1985), of which a short summary is presented here. The foundation of sedimentation theory can be traced back to the work of Hazen (1904). Hazen developed a theory for the continuous sedimentation of discrete particles having an identical settling velocity. The models were developed for both quiescent (non-turbulent) and turbulent conditions. For quiescent settling, Hazen found the fraction removal to be a discontinuous function of the relative overflow

rate (settling velocity/hydraulic loading rate) with an inflection point where the settling velocity equals the hydraulic loading rate. To model the removal under turbulent conditions he used a series of equally sized completely mixed cells.

Camp (1936) modified Hazen's quiescent settling theory to include discrete particles that have a distributed settling velocity. Camp assumed an ideal basin with homogeneous horizontal flow, even inlet distribution, free settling and particle removal when they reached the bottom of the basin. The effluent concentration then depends only on the overflow rate and the particle settling velocity distribution, and is independent of depth and detention time (Camp, 1946). Within this framework the earlier work of Hazen for quiescent flow becomes a special case of Camp's generalization.

Dobbins (1944) developed a model for predicting concentrations in a settler for single velocity particles under isotropic turbulence with no bottom scour and starting from a constant inlet concentration. He found good agreement with lab scale tests using Lucite moulding powder as suspended material. In a later work (Cordoba-Molina *et al.*, 1978) it was shown that the model equations by Dobbins could for the limiting case be reduced to Hazen's model for both quiescent and turbulent flow. Cordoba-Molina *et al.* also extended Dobbins' model to include a settling velocity distribution that is a property only of the sediment concentration and independent of the turbulence.

The models of Hazen, Camp and Dobbins have the disadvantages of many assumptions such as ideal flow conditions, ideal basin design, no turbulence or infinite turbulence, no bottom scour, no tank depth effects, only discrete sedimentation, no cohesion between particles and no hysteresis effects. Note that many of these assumptions still apply in models currently used. A major criticism was that the early models only focused on the removal of solids from the liquid (and not vice versa), that is, they did not consider any phenomena occurring within the part of the settler with high solids concentrations (Dick, 1970).

The above continuous sedimentation theory as a basis for settler design and operation ignores the thickening phenomenon prevalent in activated sludge systems. A settler used to separate flocculent, compressible particles, as those found in activated sludge systems, is usually divided into four zones, referred to as the discrete particle, flocculent, hindered settling and compression zones, see Figure 5.2.

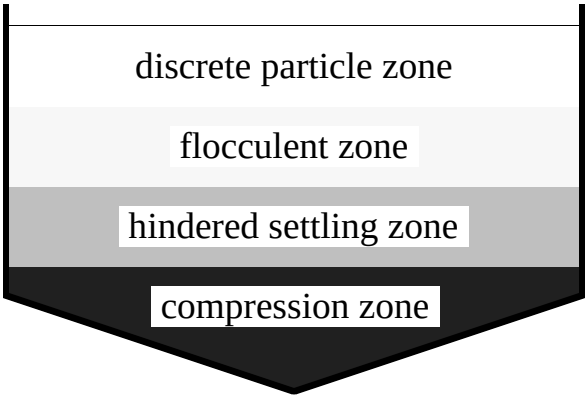


Figure 5.2 Settling zones for activated sludge.

The compression phase begins when the critical concentration, a characteristic of the suspension, is reached (Eckenfelder and Melbinger, 1957). In this region, the settling velocity is drastically reduced due to the high concentration of solids.

The thickening of the sludge is in turn influenced by a number of factors (Ingersoll *et al.*, 1955), such as:

- nature of the mixed liquor particles (density, shape, floc structure, type of microorganisms, electrostatic charges, etc.);
- dissolved substances in the substrate;
- temperature;
- depth of the sludge blanket;
- surface area of the sludge blanket;
- effects due to mechanical actions, vibrations, pressure, etc.;
- concentration of settleable solids in the mixed liquor.

The concentration at the bottom of the settler is also affected by the time allowed for compaction.

In a fundamental work by Kynch (1952) a theoretical analysis of sedimentation was made, based on the theory of Coe and Clevenger (1916). Kynch concluded that the concentration of settleable solids in the mixed liquor was of the utmost importance when describing the settling process, that is, he focused on only one of the many factors listed above. The settling in batch reactors was analysed as a process where levels of constant concentrations moved upwards due to the downward movement of particles. Kynch's theory for batch reactors was later extended for continuous reactors by Yoshioka *et al.* (1957). The main four assumptions of Kynch's theory are that

- the settling velocity of a particle depends only on the local concentration of particles;
- all the particles have the same shape, size and density;
- the particle concentration is constant within each horizontal cross-section of the settler;
- in continuous sedimentation the total settling velocity is a function of both the settling rate of particles relative to the liquid and of the downward flow of the suspension due to the underflow withdrawn from the bottom of the thickener.

The first assumption is the fundamental one. This means that all other forces acting on a particle are in equilibrium. Dixon *et al.* (1976) found that inertial effects cannot be ignored by comparing simulation models of Kynch's continuum theory and discrete settling theory and thus questioned the validity of Kynch's assumptions. Another study (Hultman and Hultgren, 1980) showed how flocculent suspensions did not strictly follow Kynch's assumptions. In a work by Dick and Ewing (1967) it is stated that Kynch's theory is highly idealized and requires an 'ideal slurry' to be directly applicable – which activated sludge cannot be said to be. However, in Dick (1970) and Dick and Young (1972) it was concluded that the mass-flow concept could be applied to a flocculent suspension, such as activated sludge, as a reasonable approximation. Some more recent references dealing with the analysis and possible extension of the solids flux theory (primarily for batch sedimentation) are, e.g., Fitch (1983), Concha and Bustos (1987), Font (1988) and Fitch (1993).

Based on the solids flux theory a number of methods have been developed to determine the steady-state behaviour of the secondary clarifier which could be used for design purposes. Yoshioka *et al.* (1957) presented a simple geometric technique to find the limiting values from solids flux curves and Keinath *et al.* (1976) introduced the concept of a state point applied to the solids flux theory to define a safe operational zone for the settler. Another design method is the Coe and Clevenger method described by Dick (1970). From a mass balance over the settler the limiting flux could be determined and a required cross-sectional area of the settler calculated. It should be noted that many of the methods above were developed prior to the widespread use of computers.

Vesilind (1968a) reviewed the different solids flux methods and found that all were sensitive to the accuracy of the solids flux curve which must be determined empirically. He also assumed that the initial settling velocities

measured in batch settling tests are representative of the settling characteristics of a large settler and that the initial settling velocity is dependent on concentration, test-cylinder diameter, stirring and flocculation characteristics of the sludge.

Even though the solids flux theory by Kynch contains idealized assumptions and generalizations that are not fully applicable to the type of solids present in an activated sludge process, its simplicity and deterministic background have attracted a lot of researchers to continue working with it. As computers have become available, a number of simulation models based on the continuity equation and Kynch's theory have been presented. This type of model is today the most commonly used. Due to its great importance and the fact that the new settler model by Diehl (1995a), presented and investigated later in this work, is based on the same concept, a more complete description of the theory is given in Section 5.2.

Novel Modelling Approaches

The very simple empirical models describing the behaviour of the secondary clarifier (e.g., Pflanz, 1966) based on extensive experimental work are today replaced by more sophisticated models. With the development of high speed computers, other types of mechanistic and numerical models have been developed to describe the behaviour of the secondary clarifier. One approach is to use regression models based on empirical data (e.g., Olsson and Chapman, 1985). Olsson and Chapman conducted research to examine the transient performance of the settler and developed a dynamic model of minimal order based on effluent data. A second-order structured model could satisfactorily explain a large number of different types of behaviour found in the tested sedimentation tank. The basic problem with this type of 'black-box' model is that it seldom explains or increase the understanding of the underlying phenomena. However, such models may be useful for investigating correlations between different process variables and for practical implementation at specific WWT plants. Due to the non-linear behaviour of the settler, the data used for identification of the models must include much process dynamics. Otherwise, extrapolation of model results for situations not included in the data used for calibration may produce highly erroneous results. This type of model must also be recalibrated on a regular basis in order to account for changing conditions, such as the time varying properties of the sludge.

A fairly recent model extension is to include biological processes within the settler model in order to explain effects such as additional denitrification and rising sludge, which appear in the settler (Henze *et al.*, 1993). For example, a settler model included in the WWT plant simulator program GPS-X (Hydromantis, 1992), uses a modified IAWQ model within every single concentration layer of the settler. This makes it possible to predict the dissolved oxygen and nitrate concentration profiles in the settler as well as the amount of nitrogen gas generated (which in turn may lead to problems with rising sludge).

Another promising approach seems to be the inclusion of hydrodynamic phenomena into traditional models and the extension to two or even three dimensions. Effects such as turbulent dispersion and mixing, bottom density currents, buoyant density currents, short circuiting, density waterfall and recirculation within the settler can then be described more accurately.

The suspended solids transport through an area in a settling tank is governed by the processes of advection, diffusion and settling (Vitasovic *et al.*, 1994). Since the former two effects are much determined by the turbulence of the flow, it is obvious that the hydrodynamics play an important role for the behaviour of a settler, especially during transient conditions. Hydrodynamic models make it possible to investigate effects of baffle sizes, skirt radius, inlet zone design and other details in the design of the settler.

The study of flow patterns in settling tanks was initiated by Anderson (1945), and extensive field and laboratory investigations on the hydrodynamics and sedimentation in clarifiers were presented by Larsen and Gotthardsson (1976) and Larsen (1977). Larsen (1977) and Imam *et al.* (1983) separately developed similar numerical models to describe the settling process in rectangular clarifiers. Recent studies of hydrodynamic effects within the settler are numerous, for example, Krebs (1991a; 1991b), Bretscher *et al.* (1992), Lyn *et al.* (1992), Samstag *et al.* (1992a), Samstag *et al.* (1992b), Zhou and McCorquodale (1992a), Zhou and McCorquodale (1992b), Zhou *et al.* (1992) and Szalai *et al.* (1994). The models often require sophisticated finite element or finite difference methods in order to solve the numerics, require very powerful computers and are still restricted to simulating the behaviour of the clarifier decoupled from the rest of the wastewater treatment plant. Although promising results have been presented, the hydrodynamic models are still too complex to be implemented in commonly used simulation programs (time-consuming identifi-

cation and simulation, difficult to verify, etc.) and a further description of these models lies beyond the scope of this work. For practical reasons, the most widely used models of the secondary clarifier are *one-dimensional layer models*.

5.2 Introduction to the Conservation Law

Practically all one-dimensional layer models are based on the solids flux theory developed by Kynch (1952) and later advanced by Dick (1970) and Shin and Dick (1980). It states that the solid flux of particles due to gravity sedimentation J_s depends on the sludge concentration X and its settling velocity v_s

$$J_s(X) = v_s(X)X \tag{5.1}$$

In order to apply this assumption we need to put (5.1) into a mathematical framework. Many physical phenomena obey the *conservation law.* This basic law states that the change of the total amount of some physical entity (mass, momentum, etc.) in a region of space is equal to the inward net flux across the boundary of that region (provided no sources or sinks are present). Such conservation laws are used to model phenomena in, for example, gas and fluid dynamics and traffic-flow analysis. They also apply to sedimentation of solid particles in a liquid.

The discussion below will be restricted to one-dimensional problems. Consider a settler performing batch sedimentation and define the z-axis along its vertical side and let $X(z,t)$ denote the concentration of particles at depth z at time t. Let f denote the flux of particles, that is, the mass of particles per unit time passing a point z at time t. We define (z_1, z_2) as an arbitrary interval of the z-axis. The conservation law written in mathematical terms in integral form is

$$\frac{d}{dt}\int_{z_1}^{z_2} X(z,t)dz = f|_{z=z_1} - f|_{z=z_2} \tag{5.2}$$

which means

increase of mass per unit time = flux in per unit time – flux out per unit time

If we assume that the concentration $X(z,t)$ is differentiable, then the left-hand side of (5.2) can be written

$$\int_{z_1}^{z_2} \frac{\partial X}{\partial t} dz \tag{5.3}$$

If the right-hand side of (5.2) is written

$$-\int_{z_1}^{z_2} \frac{\partial f}{\partial z} dz \tag{5.4}$$

then we get

$$\int_{z_1}^{z_2} \left(\frac{\partial X}{\partial t} + \frac{\partial f}{\partial z} \right) dz = 0 \tag{5.5}$$

Since this holds for every interval (z_1, z_2) (and assuming that the integrand is continuous) it follows that

$$\frac{\partial X}{\partial t} + \frac{\partial f}{\partial z} = 0 \quad \Leftrightarrow \quad X_t + f_z = 0 \tag{5.6}$$

This partial differential equation is called the *continuity equation* or the *conservation law*.

A common dependence on z and t of the flux function f is of the form

$$f = f\big(X(z,t)\big) \tag{5.7}$$

where $f(X)$ is assumed to be a smooth function. For the settling problem it is according to Kynch's theory reasonable to assume that the settling velocity is dependent only on the local concentration of particles at the point z at time t, that is, $v_s = v(X)$, and we can use any continuous function to describe the relationship between v_s and X. The most common function used to define the settling velocity is the *exponential* settling velocity function (Thomas, 1963; Vesilind, 1968b)

$$v_s = k\mathrm{e}^{-nX} \tag{5.8}$$

where k and n are parameters used for calibrating the function to experimental data. Other types of settling velocity functions are discussed in

Section 5.3. The total particle-flux function is thus

$$f(X) = v(X)X = k\mathrm{e}^{-nX}X \tag{5.9}$$

The graphs of (5.8) and (5.9) are shown in Figure 5.3.

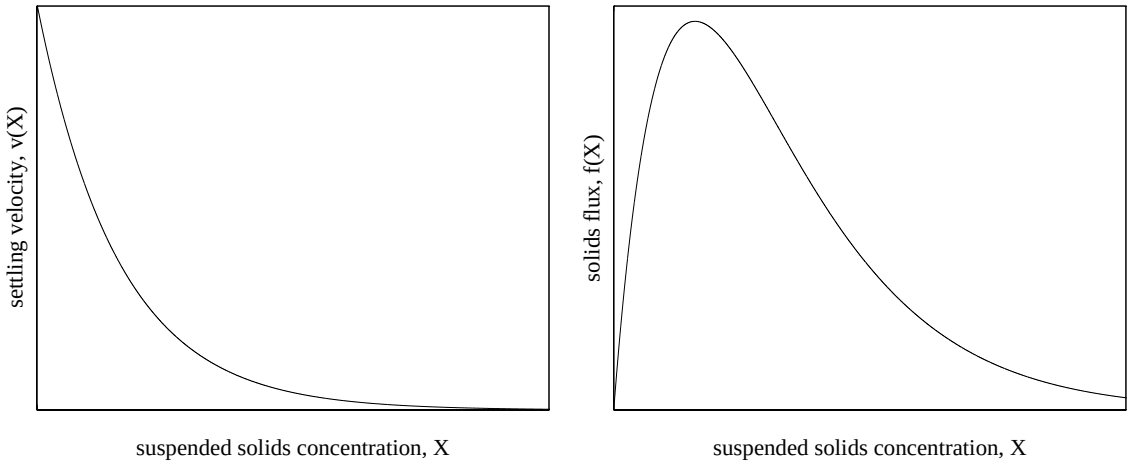


Figure 5.3 Graphs of the settling velocity and the corresponding settling flux function.

Using (5.7) we can write equation (5.6) as

$$X_t + f(X)_z = 0 \quad \Leftrightarrow \quad X_t + f'(X)X_z = 0 \tag{5.10}$$

This equation is called *quasi-linear*, since it is linear in the derivatives but the coefficient of X_z depends on X.

Characteristics

To be able to solve the partial differential equation (5.10) some initial concentration distribution must be given at $t = 0$. Then we get the initial value problem

$$\begin{aligned} &X_t + f(X)_z = 0 \qquad z \in \Re, t > 0 \\ &X(z,0) = X_0(z) \quad z \in \Re \end{aligned} \tag{5.11}$$

Let $z = z(t)$ be a level curve in the z–t plane, i.e.,

$$X\big(z(t),t\big) = \text{constant} = X_0 \tag{5.12}$$

Differentiating with respect to t and using $X_t = -f'(X)X_z$ gives

$$0 = X_z z'(t) + X_t = X_z\left(z'(t) - f'\left(X_0\right)\right) \tag{5.13}$$

Generally it must hold that $z'(t) = f'\left(X_0\right)$, which means that the level curve is a straight line with slope $= 1/f'\left(X_0\right)$ in the z–t plane. Such a line is called a *characteristic*. (If $X_z \equiv 0$ in some region, then the conservation law implies that $X_t \equiv 0$ too, hence the solution is constant, say $\equiv X_0$, and we can still define the level curves to be straight lines by the equation $z'(t) = f'\left(X_0\right)$). The value $f'\left(X_0\right)$ is called the *signal speed*, since a wavefront or a disturbance will propagate with this speed. Note that the non-linearity consists in that the signal speed is dependent on the solution X. A geometrical construction of a solution for given initial data $X(z,0) = X_0(z)$ can be done as follows: Through each point X_0 on the z-axis, draw a straight line with speed $f'\left(X_0\left(z_0\right)\right)$ in the z–t plane. Along this line the solution has the value $X_0(z_0)$. Analytically we can write the solution in implicit form: The connection between the value X, the point z and time t is

$$\begin{cases} z = f'\left(X_0\left(z_0\right)\right)t + z_0 \\ X = X_0\left(z_0\right) \end{cases} \tag{5.14}$$

Shock Waves and the Jump Condition

If the characteristics intercept at time t it is not possible to define a continuous solution after this time. Even for differentiable initial data, discontinuous solutions may appear after a finite time. This can be seen in the following way. A differentiable solution is obtained if we can solve the first equation of (5.14) for z_0, thus formally $z_0 = z_0(z,t)$, and then substitute this expression into the second equation of (5.14). According to the implicit function theorem this can be done if

$$\frac{dz}{dz_0} = f''\left(X\left(z_0\right)\right)X_0'\left(z_0\right)t + 1 \neq 0 \tag{5.15}$$

This is true for small $t > 0$, which proves that if the function $X_0(z)$ is smooth, then there exists a smooth solution $X(z,t)$ for small $t > 0$. The smallest time for which $dz/dz_0 = 0$ holds is called the *critical time*.

To be able to continue the solution after a discontinuity appears, one has to generalize the concept of solution. The conservation law in (5.10) is multiplied by a special test function φ and after partial integration one arrives at the condition

$$\int_0^\infty \int_{-\infty}^{\infty} \left(X\varphi_t + f(X)\varphi_z \right) dz dt + \int_{-\infty}^{\infty} X(z,0)\varphi(z,0) dz = 0 \qquad (5.16)$$

A function X that satisfies (5.16) is called a *weak* solution of the conservation law (5.10).

The conservation law also states how a discontinuity moves. Let X be a piecewise differentiable solution of the conservation law, with a discontinuity curve $z = z(t)$ in the z-t plane. Let $X^{\pm} = X\left(z(t) \pm 0, t\right)$ denote the values of the solution on the left and right side of the discontinuity curve. It can then be shown that the speed of the discontinuity satisfies

$$z'(t) = \frac{f\left(X^+\right) - f\left(X^-\right)}{X^+ - X^-} \qquad (5.17)$$

(5.17) is the so called *jump condition* or *Rankine-Hugoniot condition*. Note that the speed is actually the slope of the straight line through the points $\left(X^-, f\left(X^-\right)\right)$ and $\left(X^+, f\left(X^+\right)\right)$ on the graph of f. It can also be shown that if $X(z,t)$ is a piecewise smooth function, which satisfies the initial data $X(z,0) = X_0(z)$, then $X(z,t)$ is a weak solution of (5.16) if and only if

- the conservation law is satisfied at points where X is differentiable;
- the jump condition is satisfied at discontinuities.

Viscous Waves and the Entropy Condition

The problem by introducing general solutions (weak solutions) is that we may obtain different solutions for the same initial data. In order to select a unique, physically relevant solution, an additional condition must be imposed, a so called *entropy condition*. This condition will pick out the physically correct shocks and discard others and can be motivated by studying what happens when diffusion or viscosity is also taken into account. Use $f := f(X) - \varepsilon X_z$ in the derivation instead of (5.7). The term $-\varepsilon X_z$, with $\varepsilon > 0$, comes from Fick's law of diffusion. Then we obtain the viscous equation

$$X_t + f(X)_z = \varepsilon X_{zz} \tag{5.18}$$

Generally, solutions of (5.18) are smooth for $\varepsilon > 0$. If ε is small we get approximately the same solutions as of the conservation law (5.10), but the shocks are now slightly smoothed. Consider a solution of (5.10) with a single discontinuity moving with the speed defined by (5.17). Such a shock is allowed if the solution of (5.10) can be obtained as a limit solution of (5.18) as $\varepsilon \to 0$. It can be shown (Diehl, 1995a) that such an admissible shock from X^- to X^+ satisfies the entropy condition

$$\frac{f\left(X^+\right) - f\left(X^-\right)}{X^+ - X^-} \leq \frac{f(x) - f\left(X^-\right)}{x - X^-} \tag{5.19}$$

for all x between X^- and X^+

Hereby we can use the generalized solutions of the conservation law, select a unique, physically relevant solution and allowing shock waves to propagate through the settler. The speed of a shock wave is given by the jump condition (5.17), which follows directly from the conservation law. A remaining problem is how to define the numerical flux terms when the model is discretized as well as how to introduce source and sink terms necessary to describe continuous sedimentation. This is discussed in Section 5.4.

Graphical Representation of Sedimentation

To establish the gravity settling flux curve, a number of batch sedimentation tests are usually performed at different initial concentrations. It is then possible to plot the hindered settling velocity as a function of the initial concentration. From such a graph, the solids flux curve due to gravitational forces is directly obtained (see Figure 5.3). An example of a solids flux curve is illustrated in Figure 5.4.

Assume a cylinder of depth D with an initial concentration of particles, X_m, in the entire volume. A point of inflection, X_{infl}, can often be found on the batch sedimentation curve. As a consequence of the shape of the curve in Figure 5.4, a batch settling test results in two shock waves, one moving downwards from the top and one moving upwards from the bottom of the test cylinder.

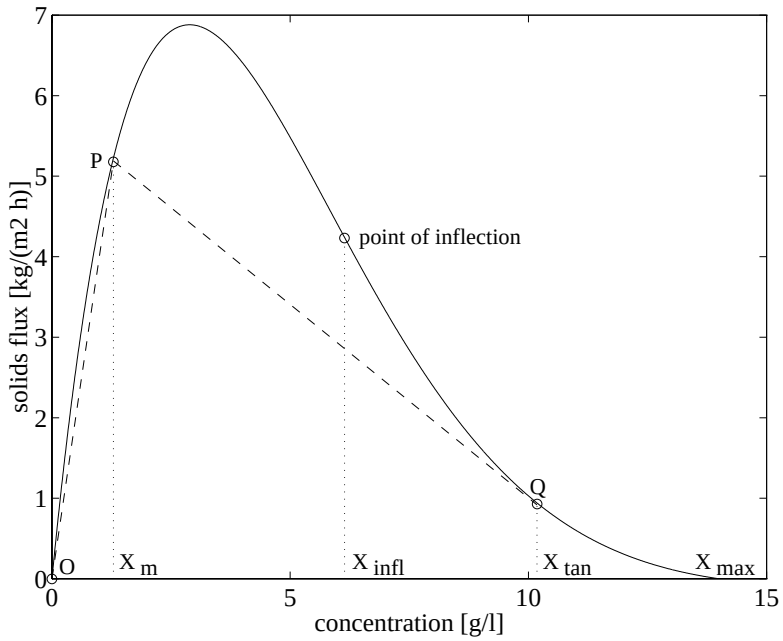


Figure 5.4 Batch sedimentation curve, *f(X),* and graphical determination of limiting (X_{tan}) and final ($X_{\max}$) sludge concentration.

Note that if the settling velocity curve is of the form (5.8), that is, there is no defined value $X_{\max}$ where the settling velocity is zero, the conservation law will yield an infinite concentration at the bottom of the settler in steady state, which is not realistic. Therefore, the curve in Figure 5.4 has been modified slightly when compared with (5.8).

By constructing the characteristics of a batch settling test in the *z*–*t* plane we can interpret Figure 5.4 better. This has been done in Figure 5.5. The speed of the shock wave moving downwards from $z = 0$ is determined by equation (5.17), that is, the slope of the line OP in Figure 5.4. The speed of the characteristics above this discontinuity is $f'(0)$ since the concentration is zero, and below, the speed of the characteristics is $f'(X_m)$ and the concentration equals X_m. The second shock wave is moving upwards from $z = D$ with the speed determined by the expression $f'(X_{tan})$, i.e., the slope of the line PQ (see Figure 5.4). The concentration immediately below this discontinuity is X_{tan} and is continuously increasing towards $X_{\max}$ at the bottom of the cylinder. At a certain time the two shock waves will meet and the resulting discontinuity will continue to move downwards with a speed that is $f(x)/x$ where x is continuously increasing from X_{tan} to $X_{\max}$ until the concentration in the entire region below the shock wave is equal to $X_{\max}$ and the concentration is zero above the discontinuity. In Figure 5.5, the boundary below which the concentration is constant ($= X_{\max}$) is plotted with a thick solid line. Note that this is not a discontinuity, since the concentration is increasing continuously from X_{tan} to $X_{\max}$.

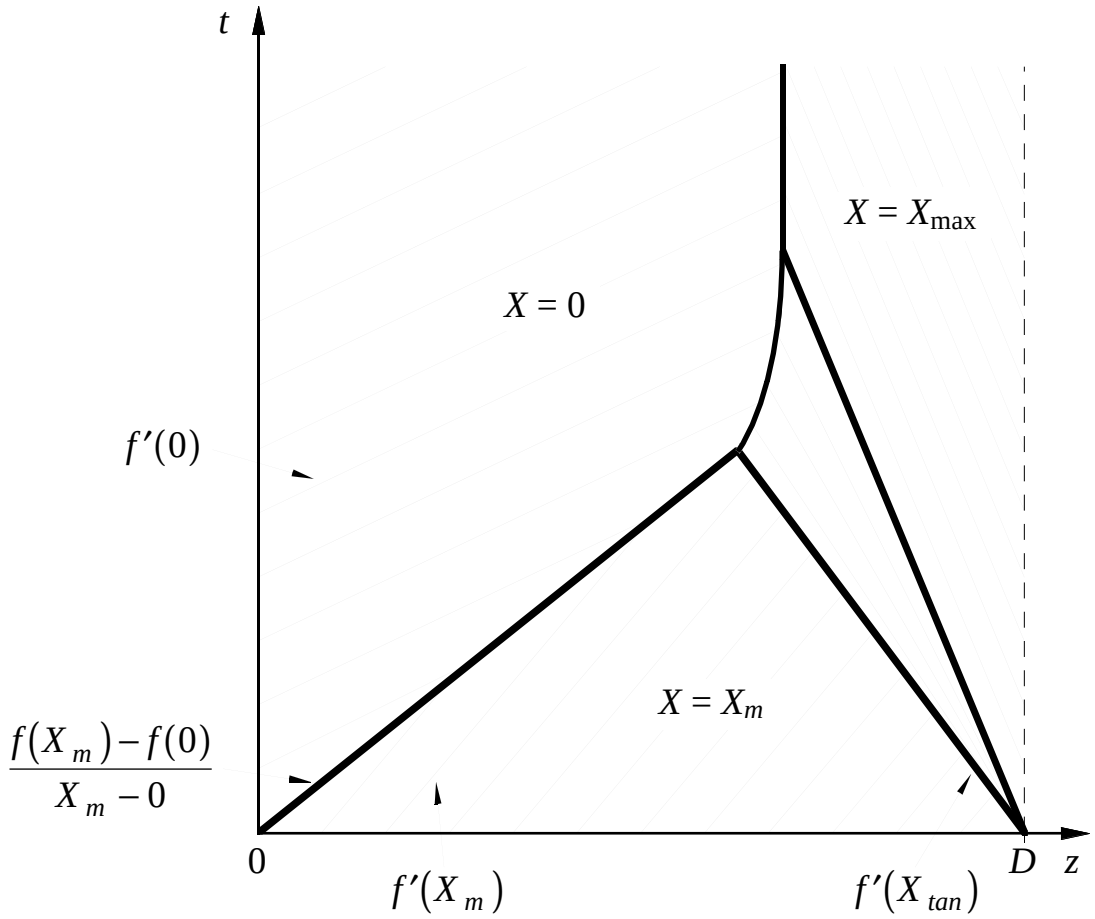


Figure 5.5 Characteristics of a batch sedimentation test showing the propagation of the shock waves.

If the batch sedimentation curve is concave, that is, there is no inflection point, then the characteristics will be similar to the ones shown in Figure 5.5. The only difference is that when the two initial shock waves intercept, a final stationary discontinuity is formed with the concentration $X = 0$ above and $X = X_{\max}$ below it.

For continuous sedimentation, the downward bulk flux is added to the batch settling curve yielding $f(X) + Q_u X/A$. The limiting flux, $J_{\lim}$, that is, the maximum mass-flux capacity of the thickening zone at steady state, can in this case be determined graphically as the point on the curve tangent to the horizontal line. This is shown in Figure 5.6 for a critically loaded settler, that is, when the incoming flux equals the limiting flux and a stable discontinuity (the sludge blanket) is maintained. In this case X_{tan} is equal to the solids concentration between the sludge blanket and the bottom of the settler, and the concentration where the limiting flux and the bulk flux line intercept is equal to the settler underflow concentration X_u.

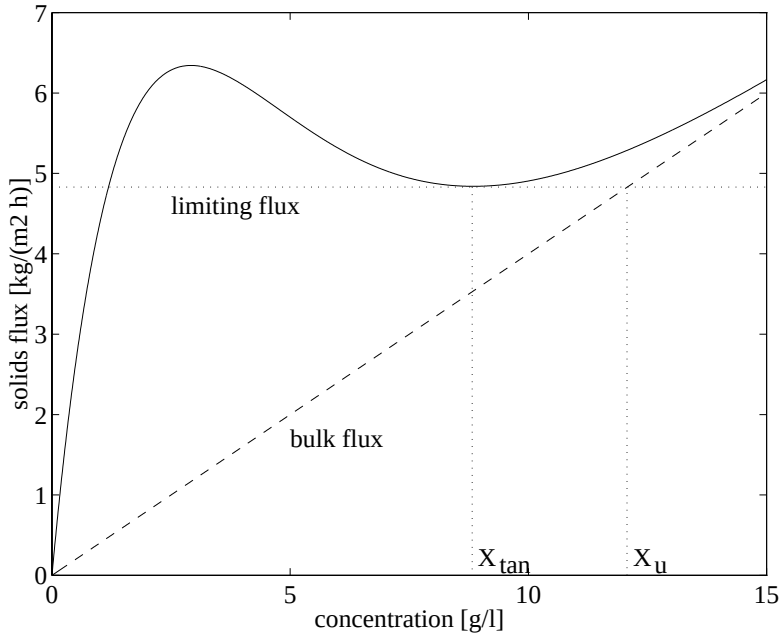


Figure 5.6 Continuous-flow sedimentation flux curve (critical loading).

In this case it is not critical whether $X_{\max}$ is finite or not. If the bulk flux is too large, then the settler will be underloaded and no stable discontinuity can be sustained. This situation can also be graphically represented using the same principle as described above. Similar plots can also be constructed for the clarification zone.

5.3 Traditional One-Dimensional Layer Models

Using the solids flux theory as the constitutive assumption and formulating the conservation law, yield equation (5.10). From this equation a dynamic model can be developed. Tracy and Keinath (1973) produced one of the first dynamical models using a mass balance and Kynch's sedimentation law to obtain a partial differential equation (PDE) which was then solved numerically by finite differences. Though their work solved the problem from a conceptual point of view, the resulting model was too complex and suffered from typical numerical shortcomings in terms of stability and boundary condition specifications. Stehfest (1984) proposed a numerical method to solve some of the problems by reducing the original PDE into a single ordinary differential equation by the so called method of lines solution technique (e.g., Schiesser, 1991).

General Model Description

The principles of the one-dimensional layer model used today are primarily based on the continuing work of Bryant (1972), Stenstrom (1975), Hill (1985) and Vitasovic (1985). Initially the work was focused on the thickening process. Bryant (1972) based the work on the continuity equation formulated as

$$\frac{\partial X}{\partial t} = \frac{\partial(D\partial X)}{\partial z^2} - \frac{\partial(vX)}{\partial z} - RX \qquad (5.20)$$

where X is the solids concentration, D is the dispersion coefficient, v is the settling velocity and R is a reaction rate affecting the solids concentration. Various simplifying assumptions, based on conditions known to exist in a continuous thickener, were applied to facilitate a solution, and initial and boundary conditions were established. The main assumptions were (Stenstrom, 1975):

- the continuous thickener does not exhibit vertical dispersion;
- the concentration of suspended solids is completely uniform within any horizontal plane within the settler;
- the bottom of the solids-liquid separator represents a physical boundary to separation and the solids flux due to gravitational settling is zero at the bottom;
- there is no significant biological reaction affecting the solids mass concentration within the separator;
- the mass flux into a differential volume cannot exceed the mass flux the volume is capable of passing, nor can it exceed the mass flux which the volume immediately below it is capable of passing;
- the gravitational settling velocity is a function only of the suspended solids concentration except when the assumption immediately above is violated.

It should be noted that the assumption concerning the largest possible mass flux into a differential volume is an assumption on the solution of (5.20). Instead, the numerical algorithm used to solve the continuity equation should deal with the possible mass flux into a specific layer.

Stenstrom divided the settler (only the thickening zone) into a number of horizontal layers and formulated a mass balance for each layer assuming

complete mixing within each layer. Based on (5.20) and the above assumptions, the following expression was obtained

$$\frac{\partial X_i}{\partial t} = \frac{v_{dn}\left(X_{i-1} - X_i\right) + \min\left(J_{s,i}, J_{s,i-1}\right) - \min\left(J_{s,i}, J_{s,i+1}\right)}{z_i} \quad (5.21)$$

where J_s is the settling flux defined in (5.1), z_i is the height of layer i. The downward bulk fluid velocity is

$$v_{dn} = \frac{Q_u}{A} \quad (5.22)$$

where Q_u is the underflow volumetric flow rate and A is the cross-sectional area of the settler. Finally, the boundary conditions for the top and bottom layers were established to permit simultaneous solution of the equations represented by (5.21). The upper boundary condition was obtained by the rate of addition of solids to the settler (because only the thickening zone was modelled) and the lower boundary condition by the rate of removal of thickened sludge. The equation for the top layer was

$$\frac{dX_1}{dt} = \frac{\frac{Q_f X_f}{A} - v_{dn} X_1 - \min\left(J_{s,1}, J_{s,2}\right)}{z_1} \quad (5.23)$$

and for the bottom layer

$$\frac{dX_n}{dt} = \frac{v_{dn}\left(X_{n-1} - X_n\right) - \min\left(J_{s,n-1}, J_{s,n}\right)}{z_n} \quad (5.24)$$

where Q_f is the feed volumetric flow rate to the settler, X_f is the suspended solids concentration of the feed and the subscript n denotes the bottom layer. The underflow concentration was defined to equal X_n, whereas no material escaped from the top layer as no upward bulk flow was modelled.

A major drawback of the thickening model was its inability to predict the behaviour in the zone above the feed layer. Due to the upper boundary condition, the model could only be applied to the regions below the feed. Therefore, the model was extended to include the clarification zone (Vitasovic, 1985). The settler was divided into n layers with the feed entering in layer m, as shown in Figure 5.7. It was assumed that the feed was instantaneously and completely distributed throughout the feed layer.

Fluid flows upward from the feed layer at the rate determined by the overflow and downward at the rate at which the thickened underflow is removed.

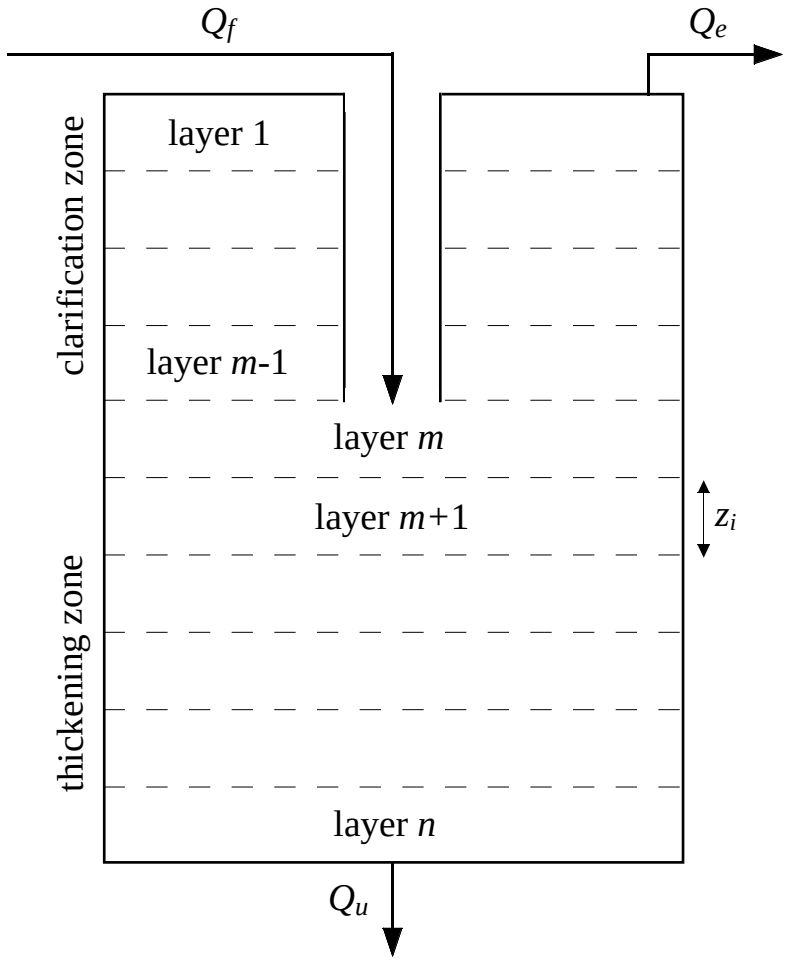


Figure 5.7 Schematic view of a settler.

The region below the feed level was modelled according to Stenstrom's approach. In the region above layer m, the solids were assumed to have a gravitational settling velocity greater than the upward movement of fluid in order to be separated from the overflow. An empirical threshold concentration, X_t, was defined in order to describe the behaviour in the upper section of the settler. Whenever the solids concentration is greater than X_t, it was assumed that the settling flux in that layer will affect the rate of settling within adjacent layers. It was presumed that the threshold concentration corresponded to the onset of hindered settling behaviour. The top of the sludge blanket was determined by the highest layer with a solids concentration equal to or greater than X_t.

The full set of equations constituting a traditional one-dimensional layer model of the secondary clarifier is presented below. In the clarification zone (layer 2 to $m-1$) the following equations are given

$$\frac{dX_i}{dt} = \frac{J_{up,i+1} + J_{clar,i-1} - J_{up,i} - J_{clar,i}}{z_i} \tag{5.25}$$

where

$$J_{up,i} = v_{up} X_i \tag{5.26}$$

$$v_{up} = \frac{Q_e}{A} \tag{5.27}$$

(we assume that the cross-sectional area of the clarifier, A, is the same) and the solids flux in the clarification zone is defined as

$$J_{clar,i} = \begin{cases} J_{s,i} & \text{if } X_{i+1} \le X_t \\ \min\left(J_{s,i}, J_{s,i+1}\right) & \text{if } X_{i+1} > X_t \end{cases} \tag{5.28}$$

In the feed layer (layer m) there is a bulk fluid movement upward at velocity v_{up} and downward with velocity v_{dn}. The resulting equation is

$$\frac{dX_m}{dt} = \frac{\frac{Q_f X_f}{A} + J_{clar,m-1} - \left(v_{up} - v_{dn}\right) X_m - \min\left(J_{s,m}, J_{s,m+1}\right)}{z_m} \tag{5.29}$$

For the layers below the feed level (layer $m+1$ to $n-1$), equation (5.21) still holds. The equation for the bottom layer is still described by (5.24) and the underflow concentration is defined to be equal to X_n. The modified equation describing the top layer becomes

$$\frac{\partial X_1}{\partial t} = \frac{J_{up,2} - J_{up,1} - J_{clar,1}}{z_1} \tag{5.30}$$

where the effluent concentration is defined to be equal to X_1. A complete description of Vitasovic's layer model is shown in Figure 5.8.

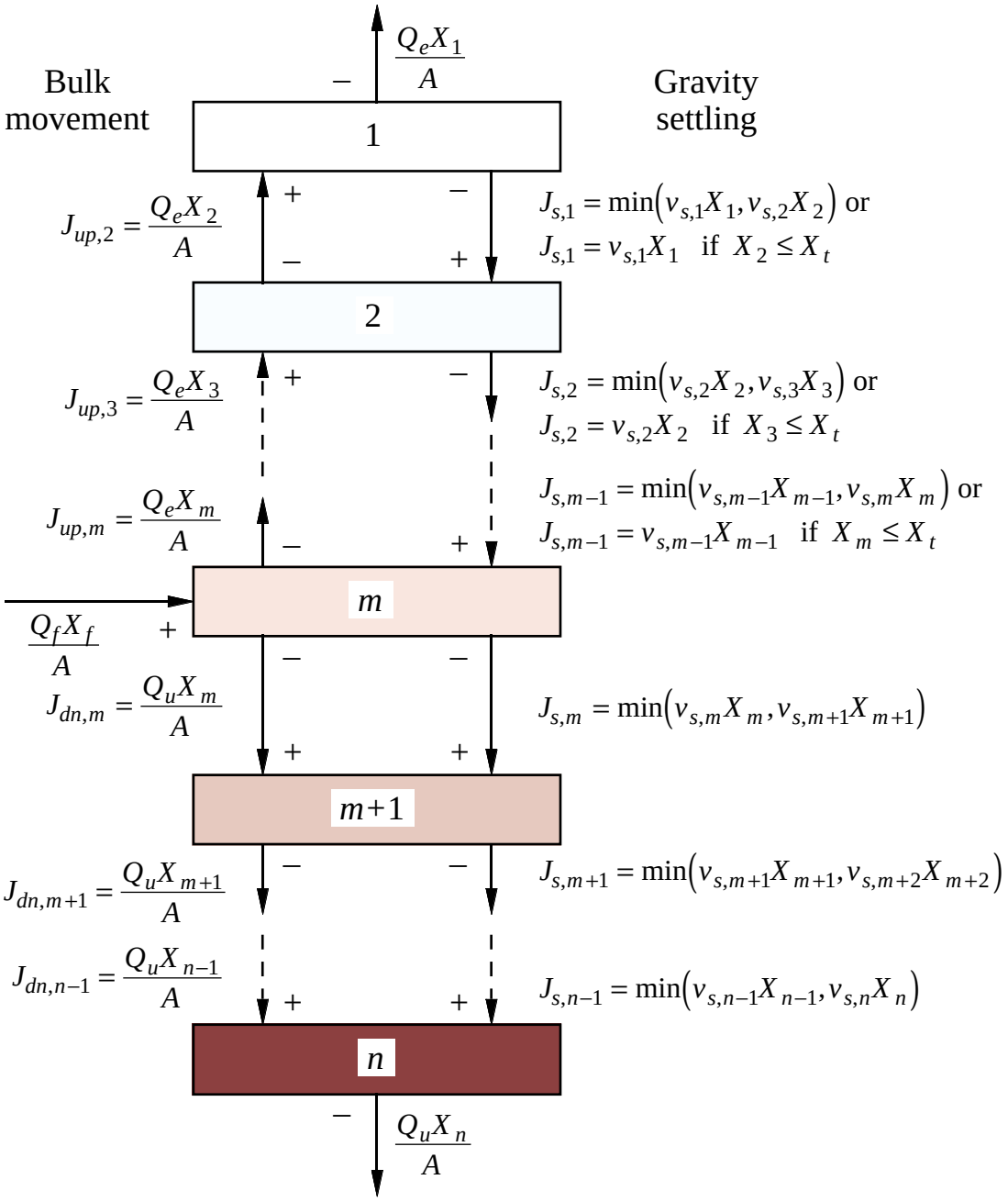


Figure 5.8 General description of the traditional one-dimensional layer model by Vitasovic (1985) with equidistant layers and a constant cross-sectional area (Takács *et al.*, 1991).

Settling Velocity Functions

The model description shown in Figure 5.8 have set the framework for practically all layer models today. However, in its original form it deals mainly with the underflow concentration, leaving realistic effluent suspended solids predictions to empirical or statistical models (e.g., Pflanz, 1966; Busby and Andrews, 1975; Chapman, 1984; Dupont and Henze, 1992).

This was partly because the *settling velocity function* used in the original model was of a type that could not predict a reasonable settling velocity for low concentrations of solids (usually found in the clarification zone). Vitasovic used the traditional *exponential* settling velocity function proposed by Thomas (1963) and Vesilind (1968b)

$$v_s = k\mathrm{e}^{-nX} \tag{5.31}$$

which predicts unreasonably high settling velocities for low concentrations of solids. The determination of an appropriate settling velocity model is indispensable for modelling the secondary clarifier using the solids flux theory. Therefore, a number of empirical functions of the settling velocity have been proposed. The majority of the functions are based either on the exponential function of (5.31) or the *power* function (Yoshioka *et al.*, 1955; Dick and Young, 1972)

$$v_s = kX^{-n} \tag{5.32}$$

Usually, the exponential function is considered to be more accurate but is sometimes considered to require more complex numerical procedures for the mathematical analysis (Smollen and Ekama, 1984). A few examples of different settling velocity functions found in the literature are given in Table 5.1.

A major difficulty is to calibrate the settling velocity function to the actual settling characteristics of the sludge. A common and practical approach is to correlate measurements of the Sludge Volume Index (SVI) to the parameters of the settling function. The SVI is achieved by a simple test and provides a rough estimate of the settleability of the sludge. Several attempts to incorporate the SVI in the settling velocity function can be found in the literature, e.g., Pitman (1985), Daigger and Roper (1985), Wahlberg and Keinath (1988), Sekine *et al.* (1989) and Härtel and Pöpel (1992). For example, the following empirical exponential settling velocity function was proposed by Härtel and Pöpel

$$v_s = \left(17.4\mathrm{e}^{-0.0113\cdot\mathrm{SVI}} + 3.931\right)\mathrm{e}^{-\left(-0.9834\mathrm{e}^{-0.00581\cdot\mathrm{SVI}}+1.043\right)X} \tag{5.33}$$

Less satisfactory attempts have also been made to describe the settling velocity as a function of the organic loading rate (F/M ratio) or solely of the mixed-liquor suspended solids concentration (MLSS) (Härtel, 1990).

$v_s = k(1-nX)^{4.65}$	Richardson *et al.* (1954)
$v_s = k\dfrac{(1-nX)^3}{X}$	Scott (1966)
$v_s = k\dfrac{(1-nX)^4}{X}$	Cho *et al.* (1993)
$v_s = k\dfrac{e^{-nX}}{X}$	Cho *et al.* (1993)
$v_s = kX(1-X)$	Scott (1968)
$v_s = k\dfrac{(1-n_1X)^4}{X}e^{-n_2X}$	Cho *et al.* (1993)
$v_s = k(1-nX)^2 e^{-4.19X}$	Steinour (1944)
$v_s = k\left(1-n_1X+n_2X^2+n_3X^3+n_4X^4\right)$	Shannon *et al.* (1963)
$v_s = k_1(1-n_1X)^{n_2}+k_2$	Vaerenbergh (1980)
$v_s = k(1-n_1X)^{n_2}$	Vaerenbergh (1980)

Table 5.1 Proposed settling velocity functions for activated sludge.

It should be recognized that none of the traditional settling velocity functions accounts for compression effects and channelling effects at the bottom of the clarifier, as described by Vesilind (1979). In order to partly compensate for this, Härtel and Pöpel (1992) imposed an empirical Ω-function which is a function of the settler depth, SVI, influent depth and influent solids concentration. By multiplying the settling flux with the correcting Ω-function during the numerical calculations, more accurate solids profiles were reported, especially in the transition and compression zones of the settler.

In order for the layer model to predict the effluent suspended solids concentration realistically, the settling velocity function must be modified further. According to Figure 5.1, the sludge can be divided into two distinct fractions, where the primary particles have a very low settling velocity (Li and Ganczarczyk, 1987) and the flocs settle according to gravity and concentration of solids. To compensate for this, Otterpohl and Freund (1992) extended the velocity function of Härtel–Pöpel to include a separate settling velocity for primary particles (set to a constant value of 0.01 m/h). The fraction of primary particles versus flocs was empirically

determined as a function of the solids concentration in the settler influent and an exponential function was calibrated to the data. A similar approach was proposed by Dupont and Dahl (1995), who defined the concentration of primary particles (X_{pp}) in the settler influent as a function of the effluent flow rate

$$X_{pp} = X_{pp,\min} + k_1\left(\frac{Q_e}{A}\right)^{k_2} \tag{5.34}$$

where $X_{pp,\min}$ is the minimum possible concentration of primary particles. This fraction was then assumed to have zero settling velocity, that is, the primary particles simply follow the upward and downward bulk flows. Furthermore, a different settling velocity model for gravity settling was suggested by Dupont and Dahl (1995)

$$v_s = v_0 \mathrm{e}^{-0.5\left(\frac{\ln(X/n_1)}{n_2}\right)^2} \tag{5.35}$$

where v_0 is the maximum settling velocity of flocs. A first attempt to compensate for density currents and short-circuiting in the settler was also proposed (usually only found in complex hydraulic models). Density currents will cause a vertical transport of the influent through the settling tank, up or down to the layer where the suspended solids concentration is closest to the concentration of the influent (Larsen, 1977). This effect is modelled by dynamically changing the inlet layer of the model to the layer where the solids concentration is closest to the influent concentration. Short-circuiting in the settler means that a part of the influent flow is transported directly to the return sludge pit without taking part in the actual settling process and correspondingly leading to a lower concentration of the sludge in the settler underflow than a traditional model will predict. This behaviour is modelled by introducing a constant empirical short-circuiting factor, Ω, and simply diverting the fraction Ω of the settler influent flow directly to the recirculation stream. Further research is carried out to determine a suitable model for dynamically updating the short-circuiting factor. (Note that the Ω-factor of Dupont and Dahl is not related to the Ω-function of Härtel and Pöpel.)

Another settling velocity function was proposed by Takács *et al.* (1991). This was the *double-exponential* settling velocity function defined as

$$v_s = \max\left(0, \min\left(v_0', v_0\left(\mathrm{e}^{-r_h(X-X_{\min})} - \mathrm{e}^{-r_p(X-X_{\min})}\right)\right)\right) \tag{5.36}$$

where v_0' and v_0 is the maximum practical and theoretical settling velocity, respectively, r_h is a settling parameter characteristic of the hindered settling zone and r_p is a parameter associated with the settling behaviour at low solids concentrations. $X_{\min}$ is the minimum attainable concentration of suspended solids in the effluent and is in turn a function of the settler influent concentration of solids

$$X_{\min} = f_{ns} X_f \tag{5.37}$$

where f_{ns} is the non-settleable fraction of X_f. The inclusion of (5.37) will directly influence the behaviour of the settler, especially within the clarification zone. The function (5.36) divides the settling velocity into four regions, schematically illustrated in Figure 5.9, in order to describe the behaviour of the different fractions of the sludge, i.e., the

- unsettleable fraction;
- slowly settling fraction;
- rapidly settling fraction.

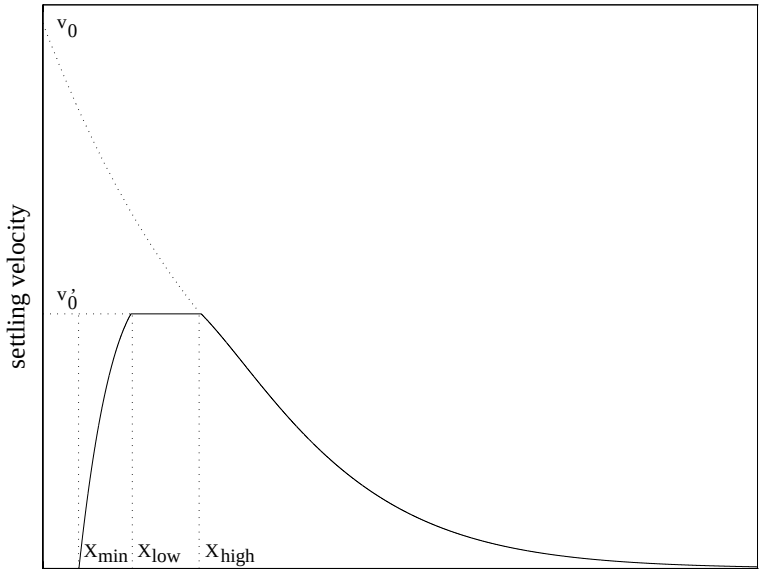


Figure 5.9 Schematic description of the double-exponential settling velocity model (5.39) suggested by Takács *et al.* (1991).

For $X < X_{\min}$ the settling velocity is zero, see (5.36). When $X_{\min} < X < X_{\text{low}}$, the settling velocity is dominated by the slowly settling particles. For low concentrations of suspended solids, Patry and Takács (1992) showed that the mean particle diameter increases as the solids concentration in the free settling zone of the clarifier gets higher. According to Li and Ganczarczyk

(1987), an increasing particle diameter implies a higher settling velocity and this effect is reflected in the behaviour of (5.36) within the region $X_{min} < X < X_{low}$. When $X_{low} < X < X_{high}$ (typically for concentrations ranging from 200 to 2000 g/m^3), the settling velocity is considered to be independent of the concentration as the flocs reach their maximum size. Finally, when $X > X_{high}$, the model reduces to the traditional exponential velocity function (5.31) describing the effects of hindered settling.

A recent evaluation (Grijspeerdt *et al.*, 1995) of six one-dimensional settler models (Laikari, 1989; Takács *et al.*, 1991; Otterpohl and Freund, 1992; Dupont and Henze, 1992; Hamilton *et al.*, 1992; a combination of Takács *et al.*, 1991 and Otterpohl and Freund, 1992), based on the solids flux theory, concluded that the model of Takács *et al.* (1991) provided the most realistic results when compared with ten sets of experimental data, both for steady-state and dynamic conditions. Therefore, the evaluation presented in Chapter 6 of the robust settler model proposed by Diehl (1996b) and Diehl and Jeppsson (1996), described in Section 5.4, will be performed using the Takács model as the reference model capturing the behaviour of traditional layer-model approaches. It should be noted that the Takács' model is actually identical to the model by Vitasovic (1985) but with the special settling velocity function (5.36) introduced by Takács *et al.* (1991).

5.4 Robust Modelling of the Settler

The behaviour of the secondary clarifier is very complex and exhibits very non-linear phenomena. Therefore, it is difficult to obtain a mathematical model that captures the behaviour in a satisfactory way. The aim of many papers presented in wastewater journals has been to obtain a good model fit to some set of experimental data, and to acquire this, *empirical* reasoning and *ad hoc* assumptions have been used. However, a model becomes more reliable if some of the *ad hoc* assumptions can be replaced by conditions that are derived rigorously from basic physical principles. In this section, a settler model that is *derived* using the knowledge of the analytical solution of the continuity equation (see Section 5.2) and Kynch's constitutive assumption only, will be presented. The continuity equation is a universal equation that must always be satisfied and the robust model guarantees this based on the mathematics from which the numerical solution technique is derived.

The conceptual approach of modelling the settler by dividing the settler into a fixed number of layers, was described in the previous section. Within each layer the concentration is assumed to be constant and the dynamical update is performed by imposing a mass balance for each layer. The numerical fluxes are defined by empirical reasoning and sometimes contain *ad hoc* assumptions. Often further *ad hoc* conditions are imposed in the layers at the inlet and outlets. Then the solution will depend on the number of layers and there is no guarantee that the method produces physically relevant solutions satisfying the continuity equation. Of course, any numerical method must divide the settler into a finite number of layers, but as the number of layers increases a natural claim is that the method should produce better and better approximations of the 'physically correct' solution (under the given assumptions). The proposed robust settler model uses numerical fluxes and formulae for the prediction of the effluent and underflow concentrations that are derived from basic physical principles without applying any *ad hoc* assumptions. The approach is based on the continuity equation written as a non-linear partial differential equation with a source term and a discontinuous flux function, modelling the inlet and outlets. The model is based on new mathematical results presented for this type of equation by Diehl (1995b, 1996b). The only assumption that needs to be specified is the batch settling flux curve.

Away from the inlet and outlets, that is, within the clarification and thickening zones, a numerical scheme converging to the 'physically correct' solution (under Kynch's assumption) was introduced by Godunov (1959). The application of the scheme to the thickening zone of the settler was introduced to the field of wastewater treatment by Diehl *et al.* (1990). The method of Godunov has then been generalized by Diehl to apply to the entire settler, including the prediction of the effluent and underflow concentrations. Barton *et al.* (1992) suggests a multi-step method to obtain even more accurate solutions within the thickening zone.

The settler model described below should be regarded as a 'first-order' model that captures the wave behaviour and the conservation of mass. The number of layers used in the model does not affect the actual results of the numerical solution, but only enhances the spatial resolution. Effects such as compression, hydrodynamics and biological activity in the settler are not modelled. Moreover, the cross-sectional area of the settler is assumed to be constant, though current research has provided theoretical results for extending the settler model to include a varying cross-sectional area as well (Diehl, 1996c).

It should be noted that the numerical results of the proposed model do not necessarily provide a better fit to real data than traditional layer models do. The model is more of an attempt to focus on the importance of model consistency. Possible discrepancies with real data are therefore an indication that the model assumptions are too rough and that the model assumptions should be extended in order to explain the true behaviour, rather than that the model is not basically correct. For example, phenomena due to compression and hydrodynamics will certainly have an effect on the true settler behaviour and, consequently, these effects need to be included in a mathematical model. However, the strict mathematical basis of the proposed model makes it a more reliable platform for future model refinements than a model using *ad hoc* and possibly erroneous assumptions in order to fit the model predictions to real data.

General Model Description

The mass per unit time entering the settler is $Q_f X_f$, and the feed inlet is modelled by the source function

$$s(t) = \frac{Q_f(t) X_f(t)}{A} \geq 0 \tag{5.38}$$

The only constitutive assumption used in the model is the one by Kynch, that is, the settling velocity of the particles is assumed to depend only on the local concentration; $v_s(X)$. The batch settling flux is then defined as $J_s(X) = v_s(X)X$ and is assumed to satisfy $J_s(X) \geq 0$ and to have one point of inflection with $J_s''(X) > 0$ for $X > X_{infl}$ (cf. Figure 5.4). These conditions hold for most traditional settling velocity functions, for example, the exponential velocity function (5.31). In continuous sedimentation the volumetric flows Q_u and Q_e also influence the flux of particles downwards. Thus, the total flux functions are defined as

$$\begin{aligned} f(X) &= v_s(X)X + \frac{Q_u X}{A} = J_s(X) + J_{dn}(X) \quad \text{in the thickening zone} \\ g(X) &= v_s(X)X - \frac{Q_e X}{A} = J_s(X) - J_{up}(X) \quad \text{in the clarification zone} \end{aligned} \tag{5.39}$$

where X is a function of the settler depth z and time t. In order to obtain a mathematical model of the settler enabling the prediction of the outlet concentrations X_e and X_u, one can introduce fictitious flux terms in the regions $z < -H$ and $z > D$, where H is the height of the settler above the feed

point, D is the depth below the feed point and the z-axis is defined positively in the downward direction. The mass per unit time leaving the settler through the outlets is the sum of $Q_e X_e$ and $Q_u X_u$. Assuming there is no sedimentation outside the settler, the conservation of mass at the outlet yields

$$\begin{cases} Ag(X_H) = -Q_e X_e \\ Af(X_D) = Q_u X_u \end{cases} \tag{5.40}$$

where X_H and X_D are the boundary concentrations at the top and bottom within the settler, respectively. The second equation of (5.40) is equivalent to the conclusions drawn from Figure 5.6 (X_D is equal to X_{tan}) following the discussion in Section 5.2. The full extended flux function including the modelling of the outlets is, consequently,

$$\Psi(X,z) = \begin{cases} -J_{up}(X), & z \le -H \\ g(X), & -H < z < 0 \\ f(X), & 0 < z < D \\ J_{dn}(X), & z \ge D \end{cases} \tag{5.41}$$

with the expressions defined in (5.39). The conservation of mass can be used to derive the partial differential equation

$$\frac{\partial X(z,t)}{\partial t} + \frac{\partial}{\partial z}\left(\Psi\left(X(z,t),z\right)\right) = s(t)\delta(z) \tag{5.42}$$

The source term on the right-hand side models the feed inlet (δ is the delta function). For example, in the thickening zone ($0 < z < D$) the concentration is governed by the equation

$$\frac{\partial X(z,t)}{\partial t} + \frac{\partial}{\partial z}\left(f\left(X(z,t)\right)\right) = 0 \tag{5.43}$$

This type of equation has been studied much and, generally, solutions contain discontinuities, as discussed in Section 5.2, see Oleinik (1959). Equation (5.42) is analysed in a general sense in Diehl (1995b) and specifically for the continuous-sedimentation case in Diehl (1996b). The main feature of this class of continuity equations is that for given initial data there are generally infinitely many possible solutions, which all satisfy the conservation of mass. In order to obtain a unique physically relevant solution an extra condition, called the entropy condition, must be imposed,

which picks out the stable discontinuities, see Section 5.2. One way to derive this entropy condition is to include some diffusion or viscosity in the model. For example, equation (5.43) then becomes

$$\frac{\partial X(z,t)}{\partial t} + \frac{\partial}{\partial z}\left(f\left(X(z,t)\right)\right) = \varepsilon \frac{\partial^2 X(z,t)}{\partial z^2} \tag{5.44}$$

where ε is a small positive parameter (cf. equation (5.18)). This equation has a unique smooth solution for given initial data. Letting $\varepsilon \to 0$ we get a limit solution with discontinuities solving (5.43) in the weak sense (cf. (5.16)). The method has been verified in several applications, see for example, Auzerais *et al.* (1988) and Davis and Russel (1989) in the case of batch sedimentation. The discontinuities of the total flux function $\Psi(\cdot, z)$ (at $z = -H$, $z = 0$ and $z = D$) make equation (5.42) even more difficult to solve. To handle these discontinuities, a generalized entropy condition is presented by Diehl (1995b), and its connection with the introduction of a diffusion term is analysed by Diehl (1996a) and Diehl and Wallin (1996).

The Numerical Algorithm

By using the theory of analytical solutions of the continuity equation (5.43), one can derive numerical algorithms that automatically take the entropy condition into account. Since the batch settling flux function J_s is non-convex (and thereby also f and g), the algorithms must work for this theoretically more complicated case as well. One such algorithm is the Godunov method. A proof of convergence of this method is presented by Le Roux (1976). The numerical implementation used in this work uses a generalization of Godunov's method to the case of point source and discontinuous flux function, that is, equation (5.42). The scheme is derived by averaging analytical solutions of the partial differential equations, see Diehl (1995b).

The numerical method is based on the division of the z-axis by n grid points equally distributed, such that $z = -H$ and $z = D$ are located half-way between the first two and the last two grid points, respectively, as shown in Figure 5.10. Let the index i stand for the space grid point (or, equivalently, layer), j for the time step and X_i^j for the corresponding concentration of suspended solids. The feed source is assumed to be located at the grid point closest to $z = 0$, denoted with the index m. The distance between two grid points is thus $\Delta = (H + D)/(n - 2)$ and the grid point $m = \text{round}(H/\Delta + 3/2)$ is closest to the feed level. The length of the time step is denoted by τ.

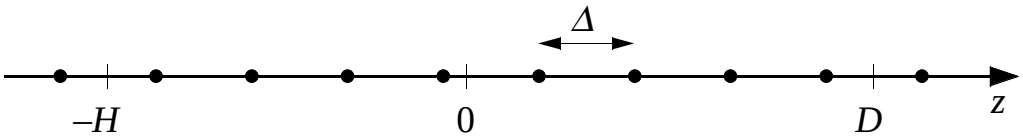


Figure 5.10 Locations of the grid points in the case $n = 10$.

Given data at time $j\tau$, first the grid points $i = 2,\ldots,n-1$ are updated according to

$$\begin{aligned} X_i^{j+1} &= X_i^j + \frac{\tau}{\Delta}\left(G_{i-1/2}^j - G_{i+1/2}^j\right) && i = 2,\mathsf{K}\ ,m-1 \\ X_m^{j+1} &= X_m^j + \frac{\tau}{\Delta}\left(G_{m-1/2}^j - F_{m+1/2}^j + s^j\right) && i = m \\ X_i^{j+1} &= X_i^j + \frac{\tau}{\Delta}\left(F_{i-1/2}^j - F_{i+1/2}^j\right) && i = m+1,\mathsf{K}\ ,n-1 \end{aligned} \tag{5.45}$$

where Godunov's numerical flux term for the clarification zone is

$$G_{i-1/2}^j = \begin{cases} \min\limits_{X\in\left[X_{i-1}^j, X_i^j\right]} g(X) & \text{if } X_{i-1}^j \le X_i^j \\ \max\limits_{X\in\left[X_i^j, X_{i-1}^j\right]} g(X) & \text{if } X_{i-1}^j > X_i^j \end{cases} \tag{5.46}$$

and the flux term for the thickening zone is defined as

$$F_{i-1/2}^j = \begin{cases} \min\limits_{X\in\left[X_{i-1}^j, X_i^j\right]} f(X) & \text{if } X_{i-1}^j \le X_i^j \\ \max\limits_{X\in\left[X_i^j, X_{i-1}^j\right]} f(X) & \text{if } X_{i-1}^j > X_i^j \end{cases} \tag{5.47}$$

and, finally, the source term is calculated as

$$s^j = \frac{Q_f^j X_f^j}{A} \tag{5.48}$$

with values at time $t = j\tau$. In order to guarantee stable and correct numerical solutions, the time step τ must be chosen so that

$$\frac{\tau}{\Delta} < \min\left(\frac{1}{\max\limits_{X\in[0,X_{\max}]}|f'(X)|}\ ,\ \frac{1}{\max\limits_{X\in[0,X_{\max}]}|g'(X)|}\right) \tag{5.49}$$

where $X_{\max}$ denotes the maximum possible packing concentration of particulate material. In wastewater applications $X_{\max}$ is normally not defined and should then be considered to be infinite.

The boundary values (grid points 1 and n) are then updated according to

$$X_1^{j+1} = \begin{cases} X_2^{j+1} & \text{if } g\left(X_2^{j+1}\right) \le 0 \\ 0 & \text{if } g\left(X_2^{j+1}\right) > 0 \end{cases}$$
$$X_n^{j+1} = \begin{cases} X_{n-1}^{j+1} & \text{if } X_{n-1}^{j+1} \in \left[0, X_m\right) \cup \left(X_M, X_{max}\right] \\ X_M & \text{if } X_{n-1}^{j+1} \in \left[X_m, X_M\right] \end{cases} \tag{5.50}$$

and the outputs X_e and X_u are calculated as

$$X_e^{j+1} = X_1^{j+1} - \frac{J_s\left(X_1^{j+1}\right)A}{Q_e^j}$$
$$X_u^{j+1} = X_n^{j+1} + \frac{J_s\left(X_n^{j+1}\right)A}{Q_u^j} \tag{5.51}$$

where X_M is the local minimizer of f. The constant X_m is the value strictly less than X_M satisfying $f(X_m) = f(X_M)$, see Figure 5.11. For stricter definitions see Diehl (1996d). The value $f(X_M)$ is the previously discussed limiting flux. This terminology is in good agreement with the algorithm above, where the updates of the boundary values are derived from the generalized entropy condition (Diehl, 1996b). For example, the most common steady-state solution has a discontinuity, the sludge blanket, in the thickening zone with the concentration X_m above and X_M below it.

The presented numerical method is stable and non-oscillatory near discontinuities and it is mass preserving. This implies that even if a discontinuity is smeared out by numerical diffusion (due to the discretization), it is located at the right position, that is, it will have the same speed as the discontinuity of the analytical solution of (5.42).

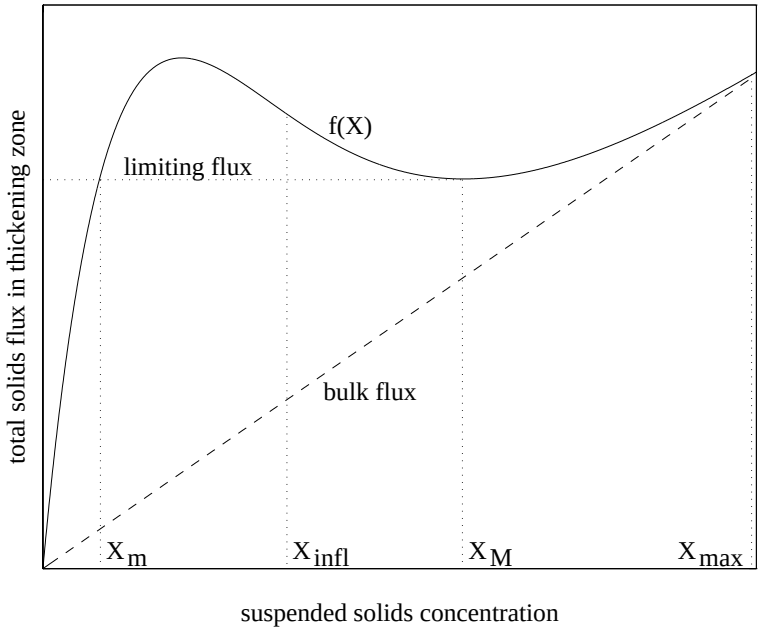


Figure 5.11 The constant X_M is the local minimizer of $f(X)$. The slope of the dashed line is Q_u/A.

As discussed in the previous section, the double-exponential settling velocity function (5.36) by Takács *et al.* (1991) is considered to provide the most realistic results of the various models presented. Therefore, this function was selected to be used in the presented settler model. For the Takács model, $v_s(X)$ is zero for concentrations $0 \le X \le X_{\min}$ (see Figure 5.9), although for higher concentrations the batch settling flux has the same qualitative shape as for the traditional exponential settling velocity model (5.31). This implies that $g(X)$ will have a local minimizer (when $Q_e/A > 0$) located somewhere in the neighbourhood of $X_{\min}$. Takács *et al.* suggest that $X_{\min}$ is proportional to the concentration entering the settler (5.37) and hence the batch settling flux will be a function of X_f as well. This can be included in the robust model by taking into account that the characteristic concentrations X_m and X_M will then depend on X_f and hence vary from one time step to another. For simplicity, $X_{\min}$ will be considered to be constant (set to 10 mg/l) in the simulations presented in Chapter 6 and, consequently, the dependence of X_f on the settling velocity will be neglected. Nevertheless, the local minimum of $g(X)$ must be taken into account. This is done in a very similar way as we defined the local minimizer X_M of $f(X)$ and the corresponding concentration X_m. Let us denote the local minimizer of $g(X)$ by X_G and then define X_g to be the strictly greater value satisfying $g(X_g) = g(X_G)$, as illustrated in Figure 5.12.

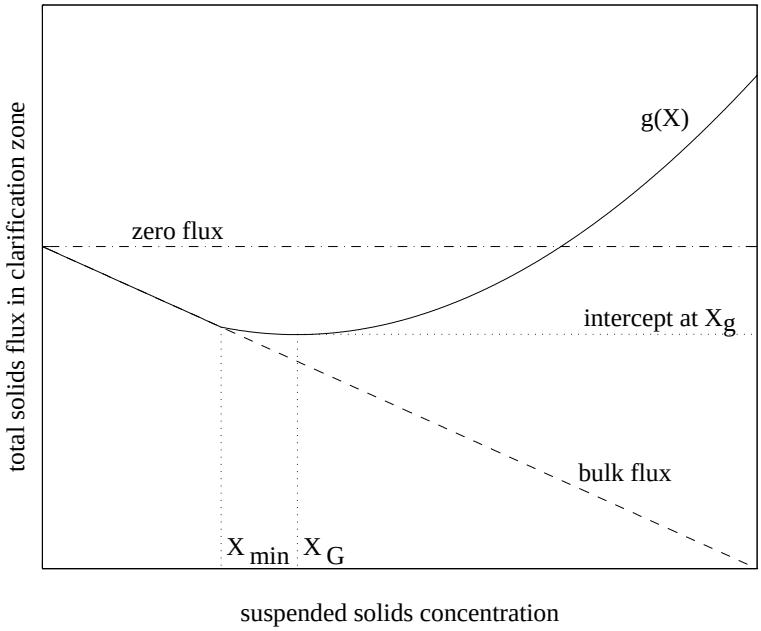


Figure 5.12 The constant X_G is the local minimizer of $g(X)$. Note that the plot only shows the flux curve for very low concentrations of solids. The concentration X_g is outside the range of this plot.

This leads to the following modification. The dynamic calculation of the first grid point (5.50), should be replaced by

$$X_1^{j+1} = \begin{cases} X_2^{j+1} & \text{if } X_2^{j+1} \in [0, X_G) \cup (X_g, X_{\max}] \\ X_G & \text{if } X_2^{j+1} \in [X_G, X_g] \end{cases} \tag{5.52}$$

The soluble material is not influenced by gravity settling but simply accompanies the bulk flow upward and downward from the feed point. For one-dimensional layer models there are basically four different approaches found in the literature of how to model the propagation of soluble material in the settler. The simplest is instantaneous propagation, that is, the concentration at the outlets is set identical to the current influent concentration. The second approach is to regard the whole settler volume as a completely mixed reactor with regard to the soluble material. Thirdly, delay variables can be used for modelling the propagation from the inlet to the outlets. The values of these delay variables are usually based on the hydraulic detention time of the settler, and can be either static (based on average flow rates) or, preferably, dynamically updated as the volumetric flow rates change. The last approach, and probably the most common today, is to use the basic layer structure already defined for the modelling of the particulate material for the soluble material as well. Every layer is then considered to be completely mixed and the transport between the layers is calculated dynamically based on the current volumetric flow rates.

For the robust settler model, the same numerical method as previously described for the particulate material can also be used for the soluble material if the batch settling flux is set to zero. This means that the updates become particularly simple. The scheme can be written as

$$\begin{aligned} S_i^{j+1} &= S_i^j + \frac{\tau}{\Delta}\frac{Q_e^j}{A}\left(S_{i+1}^j - S_i^j\right) && i = 2,\mathsf{K}\ ,m-1 \\ S_m^{j+1} &= S_m^j + \frac{\tau}{\Delta}\left(-\left(\frac{Q_e^j + Q_u^j}{A}\right)S_m^j + \frac{Q_f^j S_f^j}{A}\right) && i = m \\ S_i^{j+1} &= S_i^j + \frac{\tau}{\Delta}\frac{Q_u^j}{A}\left(S_{i-1}^j - S_i^j\right) && i = m+1,\mathsf{K}\ ,n-1 \end{aligned} \tag{5.53}$$

where S is the concentration of soluble material (cf. equation (5.45)). The soluble concentrations at the outlets are calculated as

$$\begin{aligned} S_e^{j+1} &= S_2^{j+1} \\ S_u^{j+1} &= S_{n-1}^{j+1} \end{aligned} \tag{5.54}$$

Note that grid points 1 and n are not used in the case of soluble material. They are only used to define the boundary conditions and, consequently, for the prediction of analytically correct particulate concentrations at the outlets. The algorithm (5.53) is identical to the fourth approach discussed above (every layer regarded as a completely mixed volume) combined with an Euler algorithm for the numerical updating.

Dynamic Propagation of the Biological Components

From a modelling point of view, the components of the wastewater are described differently for the biological reactor and the settler unit. For example, the IAWQ model differentiates between thirteen types of material in the water (six particulate, six soluble and alkalinity), whereas the robust settler model only divides the material into particulate and soluble matter. Therefore, all soluble components of the IAWQ model are lumped together into a single variable when entering the settler and the reversed process is performed at the outlets. The particulate material is treated analogously but the components are simultaneously transformed from mg COD/l into mg SS/l by individual transformation coefficients. The settler model must then be extended with a material propagation algorithm in order to unravel

the composite variables (soluble and particulate material) into their biological equivalents as they reach the outlets. Such an algorithm, especially designed for the robust settler model, is described below (Diehl, 1996d; Diehl and Jeppsson, 1996; Jeppsson and Diehl, 1996b). The algorithm is evaluated in Chapter 6.

Let $\boldsymbol{p}(z,t)$ denote the vector of percentages of a floc at depth z and time t in the settler. Since the concentration of suspended solids is denoted by $X(z,t)$, the vector $\boldsymbol{p}X$ contains the concentration of the different COD components in the unit mg SS/l. These values can then be retransformed into mg COD/l by dividing them by the conversion coefficients given in Table 6.1.

In order to obtain a numerical algorithm for the update of the percentage vector, we start by considering the thickening zone. If we denote the particle/floc velocity, $v_{\text{floc,dn}}$, where

$$v_{\text{floc,dn}}(X) = v_s(X) + \frac{Q_u}{A} \tag{5.55}$$

then we can describe the flux of the biological components by the vector $v_{\text{floc,dn}}(X)\boldsymbol{p}X = \boldsymbol{p}f(x)$. The conservation law in differential form is the system of partial differential equations

$$\frac{\partial(\boldsymbol{p}X)}{\partial t} + \frac{\partial(\boldsymbol{p}f(X))}{\partial z} = 0 \tag{5.56}$$

Since the sum of all components of the vector $\boldsymbol{p}$ is one, the sum of all equations in (5.56) gives back the conservation law for the total concentration (5.43), which can be solved numerically by the method in the previous subsection. It is not straightforward to turn a partial differential equation into a numerical algorithm. In the case of the conservation law the main difficulty is to find a numerical flux that is consistent with the analytical one. It is, however, possible to derive a numerical algorithm for the update of the percentage vector in a very similar way as the method in the previous subsection was derived. Since the analytical solution X is known, it is possible to combine equations (5.56) and (5.43) to obtain the following equation for $\boldsymbol{p}(z,t)$

$$\frac{\partial \boldsymbol{p}}{\partial t} + v_{\text{floc,dn}}(X)\frac{\partial \boldsymbol{p}}{\partial z} = 0 \tag{5.57}$$

This is a simple wave equation, which says that a floc with a certain percentage vector propagates with the speed $v_{\text{floc,dn}}(X(z,t))$ at the point (z,t). Without going into the details of the mathematics (see Diehl (1996d) for the full mathematical description) we can obtain the numerical updates for the percentage vector for the grid points in the thickening zone, that is, $i = m+1,\ldots,n-1$, according to

$$\boldsymbol{p}_i^{j+1} = \left(\boldsymbol{p}_i^j X_i^j + \frac{\tau}{\Delta}\left(\boldsymbol{p}_{i-1}^j F_{i-1/2}^j - \boldsymbol{p}_i^j F_{i+1/2}^j \right) \right) \Big/ X_i^{j+1} \tag{5.58}$$

Equation (5.58) holds as long as $X_i^{j+1} > 0$. If $X_i^{j+1} = 0$, then there are no flocs and the percentage vector is not interesting. F is defined according to (5.47). Similarly, the updates in the clarification zone ($i = 2,\ldots,m-1$) and at the feed level (layer m) are

$$\boldsymbol{p}_i^{j+1} = \left(\boldsymbol{p}_i^j X_i^j + \frac{\tau}{\Delta}\left(\boldsymbol{p}_{i-1/2}^j G_{i-1/2}^j - \boldsymbol{p}_{i+1/2}^j G_{i+1/2}^j \right) \right) \Big/ X_i^{j+1} \tag{5.59}$$

$$\boldsymbol{p}_m^{j+1} = \left(\boldsymbol{p}_m^j X_m^j + \frac{\tau}{\Delta}\left(\boldsymbol{p}_{m-1/2}^j G_{m-1/2}^j - \boldsymbol{p}_m^j F_{m+1/2}^j + \boldsymbol{p}_f^j s^j \right) \right) \Big/ X_m^{j+1} \tag{5.60}$$

where

$$\boldsymbol{p}_{i-1/2}^j = \begin{cases} \boldsymbol{p}_i^j & \text{if } G_{i\text{-}1/2}^j \le 0 \\ \boldsymbol{p}_{i-1}^j & \text{if } G_{i\text{-}1/2}^j > 0 \end{cases} \tag{5.61}$$

and $\boldsymbol{p}_f$ is the percentage vector associated with the suspended solids entering the settler. G and s are defined according to (5.46) and (5.48), respectively. The reason for the slightly more complicated scheme in these cases is that the total velocity of a floc in the clarification zone may be directed both upwards and downwards, depending on the concentrations and the flow conditions. Finally, the percentage vectors at the outlets are defined as

$$\begin{aligned} \boldsymbol{p}_e^{j+1} &= \boldsymbol{p}_2^{j+1} \\ \boldsymbol{p}_u^{j+1} &= \boldsymbol{p}_{n-1}^{j+1} \end{aligned} \tag{5.62}$$

Once again, note that grid points 1 and n are not explicitly used for these calculations. They are only used to define the boundary conditions when determining the concentration of the total suspended solids at the outlets.

The definition of G will guarantee that a value for $\boldsymbol{p}_1$ will never be required when using the above algorithms as $G^j_{2-1/2}$ will always be ≤ 0 with equality only when the effluent concentration equals 0.

For the soluble material the same percentage propagation algorithm can be directly applied, though in a much simpler form as the soluble material will always follow the bulk flows (upwards and downwards) in the settler. The following equations are obtained

$$\begin{aligned}
\boldsymbol{r}_i^{j+1} &= \left(\boldsymbol{r}_i^j S_i^j + \frac{\tau}{\Delta} \frac{Q_e^j}{A} \left(\boldsymbol{r}_{i+1}^j S_{i+1}^j - \boldsymbol{r}_i^j S_i^j \right) \right) \Big/ S_i^{j+1} && i = 2,..,m-1 \\
\boldsymbol{r}_i^{j+1} &= \left(\boldsymbol{r}_i^j S_i^j - \frac{\tau}{\Delta} \left(\frac{Q_e^j + Q_u^j}{A} \boldsymbol{r}_i^j S_i^j - \boldsymbol{r}_f^j \frac{Q_f^j S_f^j}{A} \right) \right) \Big/ S_i^{j+1} && i = m \qquad (5.63) \\
\boldsymbol{r}_i^{j+1} &= \left(\boldsymbol{r}_i^j S_i^j + \frac{\tau}{\Delta} \frac{Q_u^j}{A} \left(\boldsymbol{r}_{i-1}^j S_{i-1}^j - \boldsymbol{r}_i^j S_i^j \right) \right) \Big/ S_i^{j+1} && i = m+1,..,n-1
\end{aligned}$$

where $\boldsymbol{r}$ is the percentage vector for the soluble components. Finally, at the outlets the percentage vectors are defined in the same way as in (5.62), i.e.,

$$\begin{aligned}
\boldsymbol{r}_e^{j+1} &= \boldsymbol{r}_2^{j+1} \\
\boldsymbol{r}_u^{j+1} &= \boldsymbol{r}_{n-1}^{j+1}
\end{aligned} \qquad (5.64)$$

This algorithm will produce identical results as if each soluble component were applied to (5.53) separately, that is, it models the transport of soluble material as a flow through a series of completely mixed layers upwards and downwards from the feed layer.

The presented algorithm for describing the propagation of the individual components in the settler while maintaining the physically relevant basis for the settler model, that is, the settling velocity and numerical flux terms are defined from the true concentration of suspended solids, is considered to be more robust than other methods based on various *ad hoc* assumptions. The way the algorithm can be derived from a stringent mathematical analysis further emphasizes this fact.

Chapter 6

Model Evaluation

In this chapter we will examine the behaviour of the robust settler model and compare the results to those obtained by a well-known traditional model by means of numerical simulations. The models are investigated both with regard to their steady-state and dynamic behaviour. Moreover, the models are investigated both as stand-alone models and when coupled to a biological reactor model, that is, as an integrated part of the activated sludge process. Some aspects concerning the coupling of the settler model and the bioreactor model are also discussed together with a method for obtaining steady-state solutions for such an integrated process. Finally, some possible future extensions of the robust settler model is presented. The chapter covers material from [90], [95], [186], [188] and [189].

6.1 Bioreactor–Settler Interactions

A mathematical model describing the settler behaviour is useful as a stand-alone model in many applications, for example, in chemical and mineral applications. However, within the field of wastewater treatment there is usually a need to combine the settler model with other models (for example, bioreactor models) in order to describe and predict the behaviour of an entire WWT process. In Chapter 3, the most commonly used model describing the behaviour of the bioreactor, that is, the IAWQ Activated Sludge Model No. 1 (Henze *et al.*, 1987), was thoroughly described. In this section, we will discuss some important aspects when combining this bioreactor model with a model of the secondary clarifier.

A general problem is that settler models and bioreactor models appear to be developed by different people. In scientific papers usually only one or the other is discussed and the interactions are seldom commented upon. Moreover, simulations of the entire AS process are often presented without any description of how the settler and bioreactor models have been integrated. Of the quite extensive number of papers listed in the Bibliography, the only papers (disregarding those by Jeppsson and Diehl) that provide a slightly more detailed description of how the interactions have been modelled are Dupont and Henze (1992) and Otterpohl and Freund (1992).

Propagation of Individual Components

From a modelling point of view, the components of the wastewater are described differently for the biological reactor and the settler unit. The IAWQ model differentiates between thirteen types of material in the water, whereas a 'normal' settler model only divides the material into particulate and soluble matter. This was discussed in Section 5.4. In order to separate the composite variables of the settler (soluble and particulate material) into their biological equivalents at the outlets, it is necessary to keep track of the individual flocs as they move through the settler or model the propagation of the different fractions of the material making up the sludge. Such an algorithm was described for the robust settler model in Section 5.4.

Attempts have been made to avoid the above problem arising from the different description of the components in the IAWQ model and settler models. In Otterpohl and Freund (1992), the settler was described as a traditional one-dimensional layer model (see Section 5.3), and they modelled the flux of each single particulate component of the IAWQ model separately, thereby avoiding many problems. The structure of the model is exemplified by the following equation for a layer *i* below the feed level

$$\frac{dX_{i,j}}{dt} = \frac{v_{dn}\left(X_{i-1,j} - X_{i,j}\right)}{z_i} + \frac{\min\left(v_{s,i-1}X_{i-1,j}\ ,\ v_{s,i}X_{i,j}\right)\Omega_{i-1}}{z_i} - \frac{\min\left(v_{s,i}X_{i,j}\ ,\ v_{s,i+1}X_{i+1,j}\right)\Omega_i}{z_i} \tag{6.1}$$

The basic model structure is identical to (5.21). One difference is the inclusion of the earlier discussed Ω-function (see Section 5.3), but more importantly in this context is the extra subscript j, which means that $X_{i,j}$ represents the specific particulate component j of the IAWQ model, that is,

X_S, X_I, X_P, $X_{B,H}$, $X_{B,A}$ and X_{ND}, in a specific layer *i*. On the other hand, the settling velocity is a function of the true suspended solids concentration, which means that a conversion is still performed to transform COD into SS. Otterpohl and Freund also claim that the above approach is more computationally efficient than keeping track of delay variables and fractions for the separate components. However, applying Otterpohl and Freund's full model to a settler with 50 layers connected to an IAWQ model will require 950 non-linear differential equations (19 equations for each layer) to be solved on each time step in the settler alone. This can hardly be considered to be very efficient.

At first, the above approach may appear to be wise (disregarding the computational complexity). However, there are some obvious drawbacks. Firstly, the coefficients for transforming the COD components of the IAWQ model into a total suspended solids concentration must still be determined in order to define the settling velocity function. Secondly, trying to apply the above approach to the Vitasovic model (see Figure 5.8) means that the empirical threshold parameter in the clarification zone is no longer valid as it is based on the solids flux of suspended solids and not on individual particulate components. For the layers above the feed level, Otterpohl and Freund have in fact assumed that X_{i+1} is always smaller than X_t and modified their model accordingly, see equation (5.28). This will lead to somewhat different results when compared with those of the original Vitasovic model, especially when the system approaches a sludge overflow and the concentration in the clarification zone is high. However, X_t may still be applicable if the COD components in every layer are transformed dynamically into a suspended solids concentration and then used to determine the correct flux term for the individual components. Thirdly, the *ad hoc* formulation of the numerical flux terms in the Vitasovic model becomes even more heuristic when applied to each individual particulate component of the IAWQ model as in (6.1). For example, at a specific time and a specific layer, the downward flux of one particulate component (*a*) from that layer may be limited by the flux term in that layer (the first part of the last min-term in (6.1), i.e., $v_{s,i}X_{i,a}$), whereas another component (*b*) may be limited by the flux term of the layer below (the second part of the min-term, i.e., $v_{s,i+1}X_{i+1,b}$). This is not physically relevant as all the different components are assumed to be present as uniform flocs (made up of all the particulate components) and the solid flux of the individual components should consequently be that of the flocs and not differ for separate components. Numerical simulations in Section 6.3 will demonstrate a non-physical oscillating behaviour, especially in the layers from the feed level and down to the sludge blanket. It should be noted that

the above approach works perfectly well for the soluble components, as they simply follow the upward or downward bulk flows in the settler and do not require any complex numerical flux terms to describe the transport between the layers.

The principal analytical model available today, and hence a reliable numerical algorithm, is built on Kynch's assumption that the particles in the settler are uniform in size and density, that is, it is a method for predicting only the total concentration of suspended solids. In order to follow the propagation of the different COD components of the wastewater through the settler, the algorithm used in the robust model makes use of the reasonable assumption that the components are lumped together into larger particles or flocs. When the system is in a perfect steady state, the relative amount of each component is identical at the settler inlet, outlets and throughout the entire settler (though the absolute concentrations will be different). During dynamic conditions the relative amount of each COD component entering the settler is a function of time due to several factors, for example, changing influent wastewater characteristics and changing behaviour of the bioreactors. Consequently, the percentages of the components of the suspended solids entering the settler must be updated dynamically as the flocs are transported through the settler. This approach (described in Section 5.4) will also be evaluated in Section 6.3 and compared with the behaviour of the Otterpohl-Freund approach.

Unit Transformations

A problem when integrating the bioreactor model (i.e., the IAWQ model) and the settler model is due to the different units used. The soluble material can be added together in a straightforward way with no regard to the different units used (mg COD/l and mg N/l) as the soluble material only follows the bulk flows, and no biological reactions are assumed to take place in the settler. However, the particulate components of the IAWQ model should be converted into one variable with the unit mg SS/l, since gravity settling models are based on the actual mass when defining the settling velocity and the solids flux function. A set of transformation coefficients suggested by Henze *et al.* (1995) are given in Section 6.2 (Table 6.1), and will be used in this chapter. These values are based on averages from measurements of many municipal WWT plants. However, when applied to real data, the values should be confirmed by laboratory experiments on the actual sludge. The conversion means that the true suspended solids concentration is calculated according to the expression

$$X = \left(X_S + X_P + X_I\right) \cdot 0.75 + \left(X_{B,H} + X_{B,A}\right) \cdot 0.9 \qquad (6.2)$$

Note that X_{ND}, which is expressed as mg N/l, should not be included in the transformation into SS concentration although it is a particulate material. This is because it is a subset of the other particulate components, expressed as mg COD/l, and is already included in their concentrations (Henze *et al.*, 1987). Because of this special feature we run into a non-invertible problem, that is, we can calculate the suspended solids concentration from the particulate components of the IAWQ model according to (6.2) but we cannot recalculate the X_{ND} concentration at the outlets even if we know how large the fraction of each COD component of the SS are. Some possibilities, in order to get reasonable X_{ND} concentrations in the settler underflow are to assume that the

- proportion of X_{ND} to total COD is the same in the last bioreactor and at the settler outlets;
- proportion of X_{ND} to biomass COD is the same in the last bioreactor and at the settler outlets;
- X_{ND} concentration is in steady state at the outlets, that is, apply the IAWQ differential equation (3.16) with all other variables defined by the concentrations at the outlets and the parameters set to the same values as in the last bioreactor;
- dynamic state of X_{ND} is the same as in the last bioreactor, that is, apply the above method but set dX_{ND}/dt equal to the value in the last bioreactor.

For the simulations presented in this chapter, the third approach have been applied. It should, however, be noted that the concentration of X_{ND} is usually small and will only affect the behaviour of the activated sludge process simulations to a very small extent, regardless of which approach is applied. It is much more important to use a good algorithm for describing the dynamic propagation of the various components of the IAWQ model through the settler. This is necessary to describe the delay and change in composition of the suspended solids from the time when the particulate material enters the settler until it reaches the outlets.

Obtaining Steady-State Solutions

When working with complex mathematical models it is always a great advantage if steady-state solutions can be calculated fairly easily. Steady-state solutions can reveal a great deal of information about the general

model structure, robustness and validity. Moreover, it is important to be able to initiate dynamic simulations from a well-defined steady-state situation in order to investigate the impact of individual dynamic variations. In this section, a straightforward scheme to determine the steady state of the bioreactor model combined with the robust settler model, is outlined.

A steady-state solution means that all concentrations and flow rates are constant as functions of time. Many commercial tools for WWT modelling and simulation provide no special means for easily determining a steady-state solution. Instead they often rely on fast computers and determine the steady state by letting the dynamic algorithm run for a long time, applying different methods of relaxation. As the model complexity increases this is not the most efficient way to proceed.

In this work a different approach has been used. Since the IAWQ model, which describes the behaviour of the bioreactor, consists of ordinary differential equations, stationary concentrations are obtained by solving a system of non-linear equations. This is accomplished by a routine for constrained optimization which restricts the state derivatives to zero and solves a minimax problem. The underlying method is sequential quadratic programming (MathWorks, 1995). Other methods could also be applied, for example, the simplex search method. The input variables to this computation (with connection to the settler model) are the concentration of suspended solids in the underflow X_u, the concentration of soluble material S_u and all the percentages (both for particulate and soluble components) used to unravel the lumped variables into the individual biological components, which are stored in the vectors $\boldsymbol{p}_u$ and $\boldsymbol{r}_u$ (see Figure 6.1) Analogously, the output variables from the bioreactor are X_f, S_f, $\boldsymbol{p}_f$ and $\boldsymbol{r}_f$.

For the settler, the concentrations of soluble material at the inlet and outlets (and throughout the entire settler) are identical when the system is in steady state, that is, $S_f = S_u$, and the same holds for the percentage vectors; $\boldsymbol{p}_f = \boldsymbol{p}_u$ and $\boldsymbol{r}_f = \boldsymbol{r}_u$. For the total particulate concentration there are several different steady-state solutions depending on the actual values of Q_f, Q_u, Q_e and X_f. All possible cases for a settler with a constant cross-sectional area have been thoroughly investigated by Diehl (1996b). The steady-state solution we are primarily aiming for has a very low or zero concentration in the clarification zone (depending on the applied settling velocity function) and a sludge blanket (discontinuity) in the thickening zone with the constant concentrations X_m above and X_M below it, see Figure 5.11.

Using a Simple Settling Velocity Function

Assume at first that we are using a simple settling velocity function such as the exponential one (5.31). Then the following relations must hold for such a steady-state solution (Diehl, 1996b)

$$\begin{aligned} X_e &= 0 \\ \frac{Q_f X_f}{A} &= f(X_m) = f(X_M) = \frac{Q_u X_u}{A} \\ Q_f &= Q_u + Q_e \\ X_M &= M(Q_u) \\ X_m &\le X_f \le X_M \end{aligned} \tag{6.3}$$

where M denotes a function that computes the local minimizer of $f(X)$ given Q_u. When X_M is known, X_m can be calculated as shown in Figure 5.11. Naturally, the conditions in (6.3) are not always satisfied for any given values of X_f, Q_f and Q_u. Therefore, Q_u is considered to be a variable that can be manipulated in order to find a suitable steady state. The values of Q_{in}, Q_{intr}, Q_f and Q_r are assumed to be fixed by the initial settings (Q_{intr} represents the internal recirculation from the last aerobic reactor to the anoxic reactor). Since $Q_w \ge 0$ and $Q_e \ge 0$, the bounds on Q_u are obviously

$$Q_r \le Q_r + Q_w = Q_u \le Q_u + Q_e = Q_f \tag{6.4}$$

A steady-state solution for the combined bioreactor-settler system can be obtained by an iterative process with regard to the equations of the IAWQ model and the equations for the settler. The iterations are schematically outlined in Figure 6.1 where the equations of the settler are described by

$$\begin{aligned} Q_u^k &= \frac{Q_f X_f^k}{X_u^k} \\ Q_e^k &= Q_f - Q_u^k \\ X_M^k &= M(Q_u^k) \\ X_u^{k+1} &= \frac{A f(X_M^k)}{Q_u^k} \end{aligned} \tag{6.5}$$

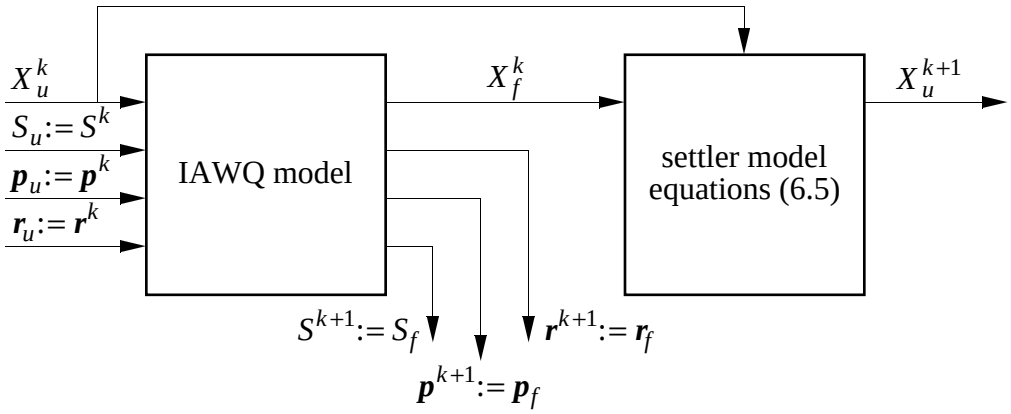


Figure 6.1 Schematic description of the iterative process to obtain a steady-state solution with a sludge blanket for the combined bioreactor-settler model.

The iterations terminate when the differences $\left|X_u^{k+1} - X_u^k\right|$, $\left|S_u^{k+1} - S_u^k\right|$, $\left\|\boldsymbol{p}^{k+1} - \boldsymbol{p}^k\right\|$ and $\left\|\boldsymbol{r}^{k+1} - \boldsymbol{r}^k\right\|$ are sufficiently small. Because of the large non-linear system of equations in the IAWQ model, it is difficult to analyse the convergence of the algorithm. However, computer simulations show convergence for a very wide range of initial values.

Using the Takács Double-Exponential Settling Velocity Function

If instead the Takács double-exponential settling velocity function (5.36) is used in the settler model, the iterative procedure above must be modified slightly to account for the fact that the flux function in the clarification zone, $g(X)$, then has a local minimizer, as illustrated in Figure 5.12. In the thickening zone, the steady-state solution is the same as discussed above, but the concentration of suspended solids in the clarification zone is now equal to X_G (> 0). The equations (6.3) should in this case be replaced by

$$\begin{gathered}
\frac{Q_f X_f}{A} + g(X_G) = f(X_m) = f(X_M) = \frac{Q_u X_u}{A} \\
Q_f = Q_u + Q_e \\
X_M = M(Q_u) \\
X_G = Z(Q_e) \\
X_e = X_G - \frac{J_s(X_G)A}{Q_e} \\
X_m \le X_f \le X_M
\end{gathered} \tag{6.6}$$

where Z is a function that computes X_G given Q_e. Consequently, the iterative equations (6.5) given above should be replaced by

$$\begin{aligned} Q_u^k &= \frac{Q_f X_f^k + Ag\left(X_G^k\right)}{X_u^k} \\ Q_e^k &= Q_f - Q_u^k \\ X_G^{k+1} &= Z\left(Q_e^k\right) \\ X_M^k &= M\left(Q_u^k\right) \\ X_u^{k+1} &= \frac{Af\left(X_M^k\right)}{Q_u^k} \end{aligned} \tag{6.7}$$

The procedure above yields steady-state values of X_f, X_u, X_e, $S_f = S_u = S_e$, X_M, X_G, $\boldsymbol{p}_f = \boldsymbol{p}_u$, $\boldsymbol{r}_f = \boldsymbol{r}_u$ and the control variable Q_u. Then the bounds (6.4) must be checked. For a given X_u the bioreactor model produces a settler feed concentration that is much lower than X_u under normal operating conditions. Since $g(X_G)$ is a very small number we assume that

$$X_f + \frac{Ag\left(X_G\right)}{Q_f} \le X_u \tag{6.8}$$

is fulfilled. Note that $g(X_G)$ will be zero if a traditional exponential settling velocity function is used, and is of interest only for the Takács velocity function. Furthermore, all steady-state solutions in the settler also satisfy (6.8), because

$$\begin{aligned} Q_f X_f = Q_e X_e + Q_u X_u &= -Ag\left(X_G\right) + Q_u X_u \\ Q_u &\le Q_f \end{aligned} \tag{6.9}$$

(6.8) says that the right inequality of (6.4) is satisfied. In other words, Q_e is always ≥ 0 during the iteration. The left inequality of (6.4) is equivalent to $Q_w \ge 0$.

If (6.4) and (6.6) are fulfilled (set X_G to zero if not using the Takács velocity function), then there exists a steady-state solution with a sludge blanket somewhere in the thickening zone. The depth of the sludge blanket is arbitrary, because the cross-sectional area is constant as function of depth and should be defined by the user. If the area was decreasing, then

the depth of the sludge blanket would be uniquely determined by the value of $Q_f X_f$ (Diehl, 1996c).

If either (6.4) or (6.6) is violated then there exists no steady-state solution with a well-defined sludge blanket level in the thickening zone. In this case we suggest some alternative ways to proceed. However, there is no theoretical guarantee that a satisfactory steady-state solution will be found.

1. Assume that the variables satisfying (6.6) have been found, but that $Q_w = Q_u - Q_r < 0$ holds. This is a clear indicator of an erroneous solution and means that Q_u is too low ($< Q_r$), i.e., the IAWQ model probably cannot produce a value of X_f large enough for a steady-state solution with a sludge blanket. Two possible ways to proceed are

   A. To obtain another steady-state solution in the thickening zone, which then simply consists of a constant concentration profile with the value X_c, Q_u is kept fixed ($\geq Q_r$) and the iterations (6.7) are replaced by

$$\begin{aligned} Q_u^k &= \frac{Q_f X_f^k + Ag\left(X_G^k\right)}{X_u^k} \\ Q_e^k &= Q_f - Q_u^k \\ X_G^{k+1} &= Z\left(Q_e^k\right) \\ X_u^{k+1} &= \frac{Q_f X_f^k + Ag\left(X_G^{k+1}\right)}{Q_u^k} \end{aligned} \tag{6.10}$$

   At termination we have a new pair of X_f and X_u that will give a steady state in the bioreactor, and the steady-state constant concentration X_c in the entire thickening zone is uniquely determined by

$$f\left(X_c\right) = \frac{Q_f X_f}{A} + g\left(X_G\right) \tag{6.11}$$

   provided that $Q_f X_f / A + g(X_G) < f(X_m)$. Then $X_c < X_m$.

   B. Modify Q_r and restart the iterative procedure according to Figure 6.1. However, a suitable size of the change in Q_r is difficult to anticipate.

2. If (6.4) and (6.6) are satisfied except that $X_f < X_m$, then Q_e is large. To prevent a steady-state solution with a sludge overflow, try finding a constant X_c according to procedure 1A.
3. If (6.4) and (6.6) are satisfied except that $X_f > X_M$, then a constant X_c could be found by the procedure outlined in 1A, but now with the requirement $Q_f X_f / A + g(X_G) > f(X_m)$ instead. Then $X_M \leq X_c \leq X_f$ holds.
4. A special case of 2 and 3 occurs if Q_u is so large that $f(X)$ is strictly increasing. Then $X_m = X_M = X_{infl}$ and the procedure of 1 applies, and X_c is determined by (6.11) without any further restrictions.

It is often of interest to find a steady state that fulfils a requirement on the sludge age of the system, θ_X, as the sludge age is the most important control variable for WWT plants in practice and has a major impact on the behaviour of the plant. This can be included in the above algorithm by an additional requirement on Q_w. The sludge age is defined as the ratio of the sludge mass in the bioreactor and the rate of the wasted sludge, i.e.,

$$\theta_X = \frac{X_f V_{\text{bioreactor}}}{X_u Q_w + X_e Q_e} \tag{6.12}$$

Using the steady-state mass balance $X_f Q_f = X_u Q_u + X_e Q_e$ of the settler and the plant flow conditions (6.14), (6.12) can be rewritten as

$$\theta_X = \frac{V_{\text{bioreactor}}}{Q_{in} - (X_u / X_f - 1) Q_r} \tag{6.13}$$

see Olsson and Andrews (1978). Since the reactor volume and Q_{in} are considered to be constants, it is the compaction ratio X_u / X_f and Q_r that influence the sludge age in steady state. By including the above equation in the iteration routine for finding a steady-state solution, it may be possible to determine a value of Q_r corresponding to a required sludge age. The procedure will become more complicated as Q_r and, consequently, Q_f will have to be updated on every iteration, cf. equation (6.7).

The method for finding a steady-state solution demonstrates one of the advantages of using a model that is founded on a stringent mathematical analysis. Fairly easily, it allows us to define analytical conditions that must hold for a specific type of solution and use them in a procedure for calculating a steady-state solution of an integrated bioreactor-settler system. The advantage of this two-step algorithm is that first an optimization algorithm is used to find a steady state for the bioreactor, which is quite easy when

the recirculated sludge concentration is constant. The steady-state output is then fed into the analytical equations describing the settler in steady state and a new return sludge concentration is defined. This iteration is then repeated until a perfect steady state is achieved. Practical experience have shown that usually only 5–10 such iterations are required to find a steady state with a tolerance in the state variables of 10^{-7} mg/l. Moreover, the complexity of finding a steady-state solution is not affected at all whether we want to use a settler model with 10 or 100 layers. The traditional layer models in combination with a bioreactor have to rely on either lengthy numerical simulations (relaxation) or complex optimization algorithms (as the settler and bioreactor models must be coupled during the calculations, i.e., there is an active, changing feedback within the system) to determine a steady state. The major drawbacks of such optimization routines are the rapidly increasing computational effort as the systems increase in complexity and the fact that a possibly successful outcome often depends heavily on the initial conditions.

6.2 Steady-State Behaviour

To demonstrate the behaviour of the integrated bioreactor-settler system (coupled case), a number of numerical simulations will be performed where the results when using the robust settler model is compared with results of the traditional Vitasovic layer model. Both models use the same Takács double-exponential settling velocity function (5.36). Apart from the basic structure of the settler model all other conditions are identical for the simulated process. The simulations will demonstrate the importance of a good settler model when it is included as a part of an entire activated sludge process. Later in this section we will also study the behaviour of the two settler models when simulated as separate units without any bioreactor interactions (decoupled case).

Plant Configuration and Simulation Conditions

For the simulations presented in this and the next section, we assume the plant configuration schematically outlined in Figure 6.2. The simulated plant is a low-loaded predenitrification-nitrification AS system with one anoxic reactor and two aerobic reactors in series. All reactors are assumed to be completely mixed and the biological mechanisms are modelled using

the default set of parameters for the IAWQ model for 20 ˚C, suggested by Henze *et al.* (1987), see Table 3.1. Note that effects due to alkalinity changes have been removed from the original model. The influent wastewater characteristics describe a typical presettled wastewater in Sweden, see Aspegren (1995). The entire influent flow, Q_{in}, is directed to the anoxic zone (i.e., no step feed) and the wastage sludge flow, Q_w, is withdrawn from the settler underflow. Data describing the details of the simulated WWT plant including the influent wastewater characteristics are given in Table 6.1.

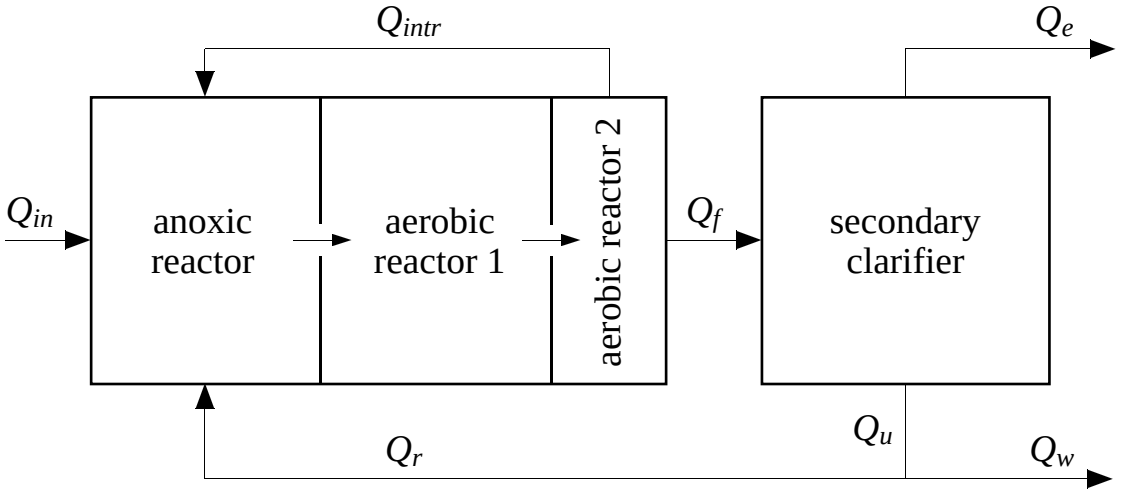


Figure 6.2 Schematic view of the simulated activated sludge plant showing the principal volumetric flows.

The obvious flow relations in Figure 6.2 are

$$\begin{cases} Q_{in} = Q_e + Q_w \\ Q_f = Q_e + Q_u \\ Q_u = Q_r + Q_w \end{cases} \tag{6.14}$$

This means, for example, that a change of Q_{in} will instantaneously result in a similar change of Q_f and Q_e (if no special control action in the recycle flow is imposed). In reality, a change of flow will be delayed in time and smeared out as it propagates through the system (Olsson *et al.*, 1986). It should be noted that no attempt has been made to model this type of flow propagation in the system.

Design and operational variables		
anoxic reactor volume	=	1000 m^3
aerobic reactor 1 volume	=	1000 m^3
aerobic reactor 2 volume	=	500 m^3
settler surface area	=	500 m^2
settler volume	=	2000 m^3
settler depth	=	4 m
settler inlet depth	=	1.8 m
influent flow rate	=	250 m^3/h
recycle flow rate	=	200 m^3/h
internal recycle flow rate	=	1000 m^3/h
wastage flow rate	=	1.163 m^3/h
hydraulic load to settler	=	0.5 m/h
sludge load to settler	=	4 kg SS/(h m^2)
anoxic sludge retention time	=	13.5 days
aerobic sludge retention time	=	20.2 days
hydraulic retention time	=	10 hours
aerobic oxygen conc.	=	2.0 mg/l

Transformation coefficients		
SS to X_S ratio	=	0.75 g_{SS}/g_{COD}
SS to X_I ratio	=	0.75 g_{SS}/g_{COD}
SS to X_P ratio	=	0.75 g_{SS}/g_{COD}
SS to biomass ratio	=	0.90 g_{SS}/g_{COD}

Wastewater characteristics		
S_S	=	80 mg COD/l
X_S	=	105 mg COD/l
S_I	=	35 mg COD/l
X_I	=	40 mg COD/l
S_{NO}	=	0 mg N/l
S_{NH}	=	20 mg N/l
S_{ND}	=	2 mg N/l
X_{ND}	=	6 mg N/l
S_O	=	0 mg (-COD)/l
$X_{B,H}$	=	25 mg COD/l
$X_{B,A}$	=	0 mg COD/l

Settling parameters		
v_0	=	145 m/d
v_0'	=	100 m/d
r_p	=	0.005 m^3/g
r_h	=	0.00042 m^3/g
X_{min}	=	10 mg/l

Threshold value		
X_t	=	3000 mg/l

Table 6.1 The simulated WWT plant.

A significant problem in the evaluation of the entire activated sludge process is the difficulty of separating the dynamics of the biological reactors from the settler, due to the recycle flow, Q_r, from the clarifier to the bioreactor. As the main purpose of this chapter is to demonstrate the behaviour of the robust settler model and to compare it with traditional layer models, the simulated plant is deliberately operated with a very high sludge age. This implies that differences that appear as a result of which settler model is currently being tested will mainly affect the concentration of inert solids in the system and, consequently, only have a limited effect on the biological mechanisms in the bioreactors. Moreover, the simulations have been performed with the oxygen concentration in the anoxic zone fixed at 0 mg/l (i.e., perfectly anoxic) during the whole duration of the simulations. This is mainly done to speed up the simulations and will only have a limited influence on the results. For a comparative study of various settler models this simplification is of no importance. However, when calibrating an AS model to real data, such an assumption should not be used.

Coupled Case

A steady state of a WWT plant can be defined by different criteria. For example, a steady-state solution can be based on certain flow rates, a required sludge retention time (SRT), a specific sludge blanket level, etc. In this section, the flow rates of the plant are set to values based on an analytical solution of the continuity equation that has a well-defined sludge blanket in the thickening zone. Recall from the previous section that if certain conditions are fulfilled then the sludge blanket can be positioned anywhere in the thickening zone (if the cross-sectional area of the settler is constant). In the simulations below, we have defined the sludge blanket level to be one meter below the feed level.

The steady state of the entire activated sludge plant, based on the defined flow conditions and influent wastewater characteristics, presented in Table 6.1, are calculated for the IAWQ model coupled to the robust settler model and the traditional Vitasovic model, respectively, using the procedure described in the previous section. The solutions are determined for different number of layers in the settler, to demonstrate the most important differences between the two models. In Figures 6.3 and 6.4, the obtained steady-state concentrations profiles in the settler are shown for the two models. It is evident that the robust model provides results closest to the analytical solution. Moreover, the number of layers has a significantly larger effect on the solution of the traditional model than on the robust model, where the differences depend on the spatial resolution only. Note once again that, apart from the basic structure of the two settler models and the number of layers used, all plant conditions and model parameters are identical during the simulations.

The results in Table 6.2 confirm the consistency of the robust model with respect to the number of layers used. Only the mass of solids in the settler differs from the analytical solution, and it is clear that the model prediction is approaching the analytical result as the number of layers increases. Note that the underflow and effluent SS concentrations of the robust model are not the same as the concentrations in the bottom and top layers, as is the case for the traditional model. Instead they are calculated according to the analytically derived expression (5.51). The underflow and effluent concentrations of the traditional model appear to be approaching a solution close to the analytically calculated concentrations. However, even the small difference for the 50-layer case has a dramatic influence on the steady-state solution with regard to the mass of solids in the settler and level of the sludge blanket.

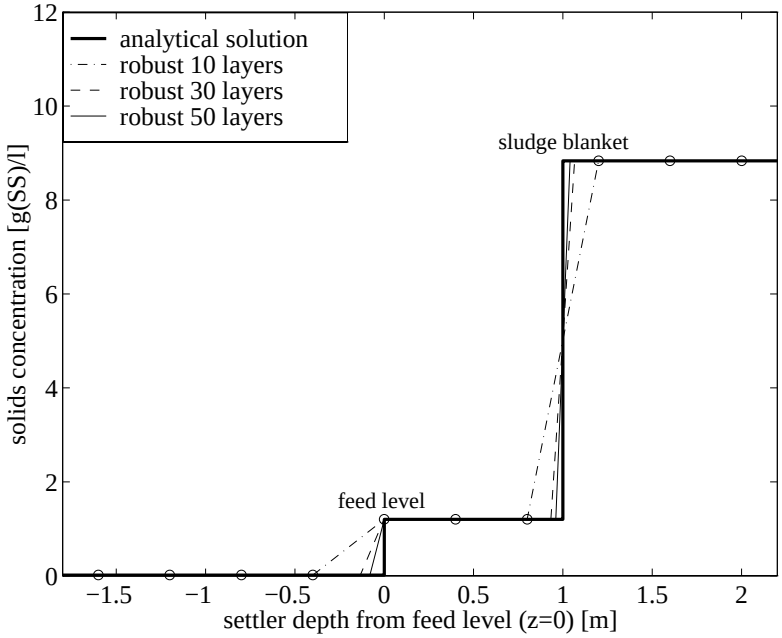


Figure 6.3 Settler concentration profiles of the suspended solids in steady state for the robust model. The solutions for different number of layers are compared with the analytical solution. The positions of the layer's midpoints are marked (o) only for the 10-layer case.

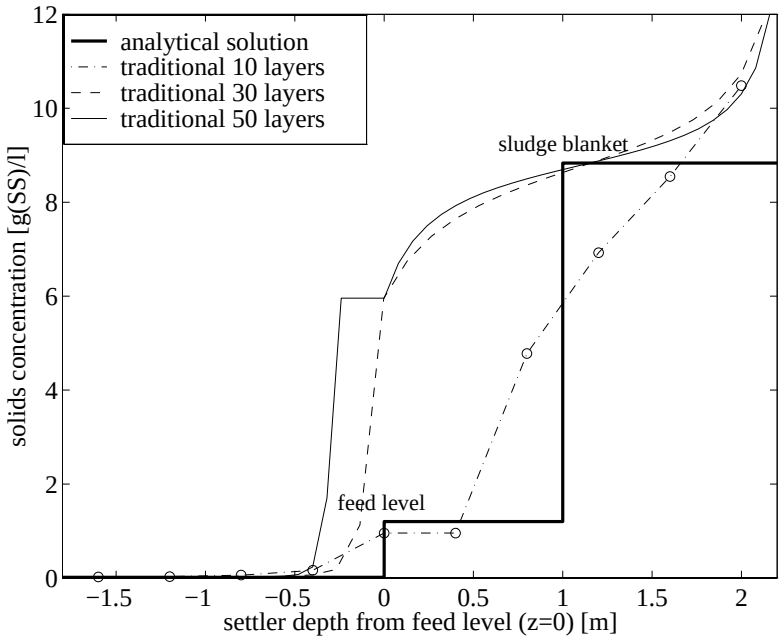


Figure 6.4 Settler concentration profiles of the suspended solids in steady state for the traditional Vitasovic model. The solutions for different number of layers are compared with the analytical solution. The positions of the layer's midpoints are marked (o) only for the 10-layer case.

Variables	Traditional model			Analytical solution	Robust model		
	10 layers	30 layers	50 layers		10 layers	30 layers	50 layers
X_u [g/l]	10.5	11.9	12.0	12.1	12.1	12.1	12.1
X_e [mg/l]	20.7	12.0	11.4	10.9	10.9	10.9	10.9
X_f [g/l]	4.70	5.34	5.38	5.41	5.41	5.41	5.41
mass [ton]	6.58	10.0	10.6	5.91	6.03	5.95	5.94
SRT [d]	28.3	33.0	33.4	33.6	33.6	33.6	33.6

Table 6.2 The key variables for the settler in the coupled case.

A particularly interesting observation of the traditional model is the significant difference between the 10 and 30/50-layer solution. A comparison to the analytical solution shows that a 10-layer model is too crude. Note that 10-layer models are the most common ones used in simulation programs of WWT plants available today. The dependence of the numerical solution on the number of used layers is really a problem for almost all one-dimensional layer models as they are all based on the same concept as the traditional Vitasovic model. The problem has in most cases been overlooked and many publications dealing with settler modelling do not even mention the number of layers used in the presented modelling approaches although the impact is significant. A few recent publications have begun to realize the problem. For example, in Krebs (1995) the importance of investigating the sensitivity of settler models to the number of layers is briefly discussed, and in Vanrolleghem *et al.* (1995) a 50-layer settler model was suggested as a good trade-off between prediction accuracy and computational burden. Sometimes it is suggested that the number of layers could be used as a tool for calibrating the model to real data. However, a reliable model should naturally predict more accurate solutions as the number of layers increases, and the idea to use the number of layers for calibration purposes, that is, use a built-in error of the model to fit the model predictions to real data, is not a good way to proceed. Such an approach is more of an indication that there are fundamental problems with the traditional layer models that are not fully understood. The robust settler model eliminates many of these problems.

The extremely long sludge age of the simulated process was deliberately chosen so that the behaviour of the bioreactors has a limited influence on the results shown above. The differences that can be observed depending on which settler model is used, are mainly due to different concentrations of inert material in the system. In Table 6.3, the steady-state values of all the state variables in the IAWQ model are shown for the analytically calculated case. Apart from the inert material concentration (and conse-

quently the suspended solids concentration), the steady-state concentrations for all the tested cases above are within a few percent of these values. Therefore, the discrepancies and effects of the traditional Vitasovic model on the entire AS process would be even more dramatic for an AS system with a low sludge age (as the relative amount of active biomass in the recycled sludge would be higher).

Variables	Anoxic zone	Aerobic zone 1	Aerobic zone 2
S_S [mg COD/l]	3.965	2.960	2.413
S_{NO} [mg N/l]	0.634	3.334	4.339
S_{NH} [mg N/l]	4.332	1.662	0.894
S_{ND} [mg N/l]	0.529	0.953	0.945
S_O [mg (-COD)/l]	0.000	2.000	2.000
X_S [mg COD/l]	45.91	25.77	19.30
$X_{B,H}$ [mg COD/l]	1346	1351	1351
$X_{B,A}$ [mg COD/l]	60.42	60.80	60.90
X_I+X_P [mg COD/l]	5499	5501	5502
X_{ND} [mg N/l]	3.404	2.030	1.576
SS [mg SS/l]	5425	5416	5412

Table 6.3 Steady-state values for the components of the bioreactor for the analytical case. The suspended solids concentration is calculated according to (6.2).

Decoupled Case

For a settler model integrated with an entire AS process, the simulated bioreactor act as an 'equalizer', that is, many differences are smoothed and, consequently, more difficult to detect and verify. A good model of the settler should naturally produce reliable and consistent results when simulated separately as well. Table 6.2 clearly shows that the robust settler model is consistent with regard to the number of layers used and this will be demonstrated even more thoroughly in the next section. The results shown in the previous subsection clearly indicated that this is not the case for the traditional settler model. In order to further demonstrate the limitations of this model with regard to the number of layers used, its behaviour is simulated for a decoupled case (no coupling to the bioreactor, stand-alone model) and the largely different steady-state solutions are shown in Figure 6.5. The conditions and parameters used for the simula-

tions are identical to those given in Table 6.1 except that the settler feed concentration of suspended solids is *fixed* at 5200 g SS/m^3, that is, a value within the same region as the X_f values found for the 10, 30 and 50-layer traditional models seen in Table 6.2. Note that in the previous subsection the settler feed concentration was determined as a result of the model equations and the settler-bioreactor interaction. Figure 6.5 clearly demonstrates the highly different steady-state concentration profiles although all inputs and parameters are identical, *only* the number of layers differ. The results are a clear indication that the applied numerical fluxes and the prediction of the concentrations at the outlets are not mathematically sound. The key variables for the steady state of the Vitasovic model are given in Table 6.4.

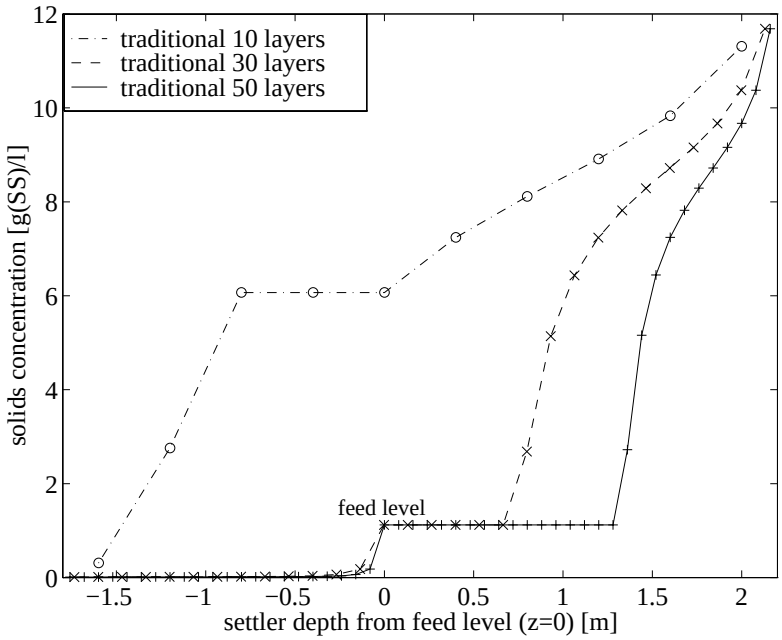


Figure 6.5 Settler concentration profiles of the suspended solids in steady state for the decoupled traditional Vitasovic model for different number of layers using identical model inputs and parameters.

	Traditional model		
Variables	10 layers	30 layers	50 layers
X_u [g/l]	11.3	11.7	11.7
X_e [mg/l]	309	11.9	11.3
mass [ton]	13.3	6.29	4.28

Table 6.4 The key variables for the settler in the decoupled case.

Figure 6.6 demonstrates the problem of using *ad hoc* conditions in models. The threshold value (X_t) included in the traditional Vitasovic model is an attempt to describe the different behaviour of hindered and non-hindered settling in the clarification zone, as discussed in Section 5.3 (see equation (5.28)). Such a description should instead be included in the constitutive assumptions of the model. X_t has a direct influence on the behaviour of the numerical algorithm used for solving the model equations. A suitable value of X_t has to be determined empirically or estimated and varies with time and quality of the sludge. Therefore, an erroneous choice may have significant effects both when calculating a steady-state solution and especially during dynamic simulations, as X_t will affect the velocity of a shock wave moving through the clarification zone (when approaching a sludge overflow). In Figure 6.6, the previous simulation for the 10-layer traditional model is repeated using different values of X_t. Note that because of the set of parameters used in the settling velocity function, the value of X_t must be increased significantly to show any major differences between the steady-state solutions (X_t must be larger than the steady state solution in the clarification zone, i.e., $X_t > 6$ g SS/l, see Figure 6.5). In the dynamic case, however, a fairly small change in X_t will also have significant effects when the settler process is approaching a sludge overflow. The key variables for this simulation is given in Table 6.5.

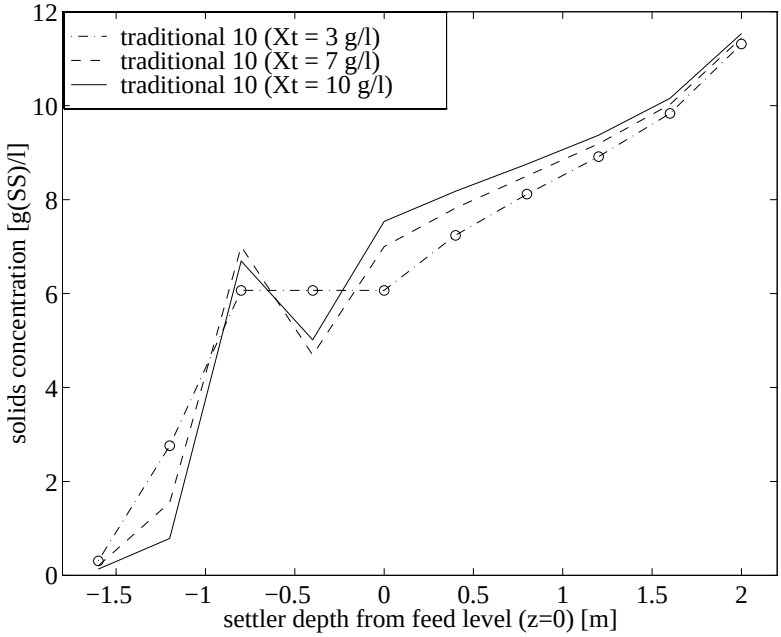


Figure 6.6 Settler concentration profiles of the suspended solids in steady state for the decoupled traditional 10-layer Vitasovic model for different values of the threshold coefficient X_t.

Variables	Traditional model, 10 layers		
	X_t=3 g/l	X_t=7 g/l	X_t=10 g/l
X_u [g/l]	11.3	11.4	11.5
X_e [mg/l]	309	205	133
mass [ton]	13.3	13.5	13.6

Table 6.5 The key variables for the settler in the uncoupled case.

It is clear that the prediction of mass in the settler is almost the same for all cases but the effluent solids concentration and the profile in the clarification zone are quite different. In the example above it should also be noted that for a certain set of values of X_t (≈6–8 g/l, in these simulations) there exists no perfect steady-state solution. Instead, the solution in some layers in the clarification zone oscillates around the value of X_t as the numerical flux terms change when the concentration in a layer is slightly above or below the value X_t (cf. Figure 5.8). In this case an optimization algorithm cannot be used to determine the steady-state solution (as it does not exist). The graph in Figure 6.6 for X_t=7 g/l was instead determined by relaxation and is not a real steady state (although the oscillations are small). Moreover, for large values of X_t (>8 g/l), the calculated steady-state solution depends on the chosen initial concentration profile in the settler and is not unique with regard to the inputs. Such a large value of X_t implies that the numerical flux expressions will be controlled by the conditions in the actual layer, regardless of the situation in the receiving layer below (X_i will always be smaller than X_t, see Figure 5.8). Figure 6.7 shows two possible steady-states profiles for X_t=10 g/l. The solid one is identical to the one in Figure 6.6 when the initial concentration profile was set to the steady-state profile found for X_t=3 g/l. The other steady-state solution was found when setting the initial concentration profile to a more normal one, with a well-defined sludge blanket in the thickening zone and only a very low concentration of solids in the clarification zone.

Note that the specific values of X_t discussed above only hold for the conditions used in this specific simulation. The same type of behaviour can, however, be found for any set of model parameters and conditions. In this subsection only a few examples were given to demonstrate the influence of the threshold parameter on the predictions given by the traditional layer model. It appears that X_t has been included to provide more reasonable predictions when compared with real data and as a tool for model calibration without fully investigating the physical violations it may lead to. It

is recommendable to remove this threshold value (together with the *ad hoc* flux terms) from the model and instead apply the analytically derived Godunov flux terms discussed in Section 5.4.

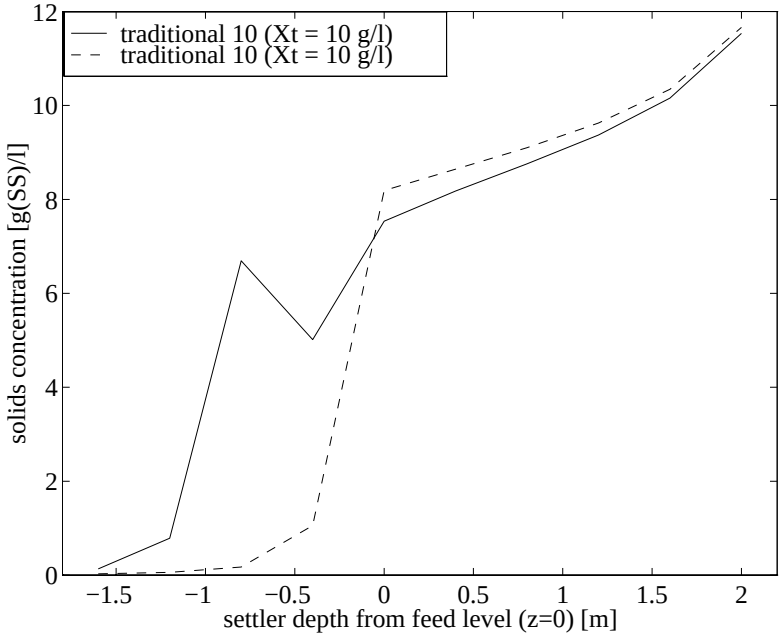


Figure 6.7 Different steady-state concentration profiles for the decoupled traditional Vitasovic model depending on the initial concentration profile.

6.3 Dynamic Behaviour

By imposing a number of disturbances on the process, the behaviour of an entire AS plant will be investigated, both when using the robust settler model and also with a traditional layer model. The importance of a proper algorithm for describing the propagation of the individual biological components in the settler will also be demonstrated. However, to demonstrate the consistency of the robust settler model with regard to the number of layers used we will first present results of simulations of a decoupled robust settler during dynamic conditions. A similar study is not carried out for the traditional layer model, as we have already shown that this model is not even consistent for steady-state conditions (see Section 6.2).

Decoupled Case

In order to exemplify the behaviour of the robust settler model, a number of simulations are demonstrated in Figures 6.8–6.11. In this case the settler model is used as a separate unit (no coupling to a biological reactor) and the robustness with regard to the number of grid points and the mass preservation is demonstrated. The applied settling velocity model is also in this case the one by Takács (5.36) with the parameters given in Table 6.6. The settler design variables and the initial conditions are also presented in Table 6.6.

A	500 m^2	X_f	5390 g/m^3
H	1.8 m	v_0	145 m/d
D	2.2 m	v_0'	100 m/d
Q_f	450 m^3/h	r_p	0.005 m^3/g
Q_e	250 m^3/h	r_h	0.00042 m^3/g
Q_u	200 m^3/h	$X_{\min}$	10 g/m^3

Table 6.6 Parameters for settler model simulations.

The simulations are performed for four different cases – 12, 32, 52 and 72 grid points (note that the first and last grid point are only used to define the boundary conditions, see Figure 5.10). For the initial data a steady-state solution is calculated with a sludge blanket level one meter below the feed level. After half an hour the settler feed concentration, X_f, is increased to 7000 g/m^3 and at $t = 2$ h, Q_f is increased to 600 m^3/h (which directly affects Q_e as Q_u is constant). At $t = 2.5$ h, Q_f and X_f are set back to the values given in Table 6.6 and the simulation is continued until $t = 5$ h.

The three-dimensional plots presented in Figures 6.8–6.11 show how the sludge blanket starts to move upwards from $z = 1$ m when X_f is increased. The higher flow rate imposed at $t = 2$ h causes a shock wave to move upwards in the clarification zone and at $t = 5$ h the system has almost reached a new steady state. The courser the grid mesh is, the more the shocks are smeared out and with only 12 grid points, see Figure 6.8, the shocks can hardly be located at all. However, the predictions of mass in all four simulations agree well, as shown in Table 6.7.

n	initial mass [kg]	final mass [kg]	$\lvert\Delta$mass$\rvert$ [kg]
12	6042	8039	1997
32	5963	7961	1998
52	5947	7945	1998
72	5940	7938	1998

Table 6.7 Mass preservation for different number of grid points.

Although the initial steady state is identical for all four cases, it is clear that there is a difference with regard to the initial amount of mass in the settler. This is due to the distance between the grid points. Just above the feed point (from grid point m to $m-1$) the concentration drops sharply and since the concentration is considered to be constant around each grid point (i.e., within each layer) this will lead to the prediction of a slightly larger mass as the grid mesh becomes coarser. Apart from this initial difference, Table 6.7 shows that the preservation of mass during dynamic conditions is good and independent of the grid mesh. Although not shown, the predictions of the effluent and underflow concentrations agree perfectly for all cases.

A problem that may affect the prediction of mass slightly more may be anticipated by reviewing Figure 5.10. Here it was stated that $z=-H$ and $z=D$ are located exactly half-way between the first two and the last two grid points, respectively. Depending on the values of H and D and the number of grid points used, it is impossible to guarantee that one grid point can be positioned exactly at $z=0$. The small round-off error produced will be more prominent when the number of grid points is small and insignificant for a narrow grid mesh. For the cases shown in Figure 6.8–6.11, the size of the settler (H and D) and the number of grid points were chosen in such a way that the above problem did not occur, that is, the value of $(H/\Delta+3/2)$ is an integer for all cases shown.

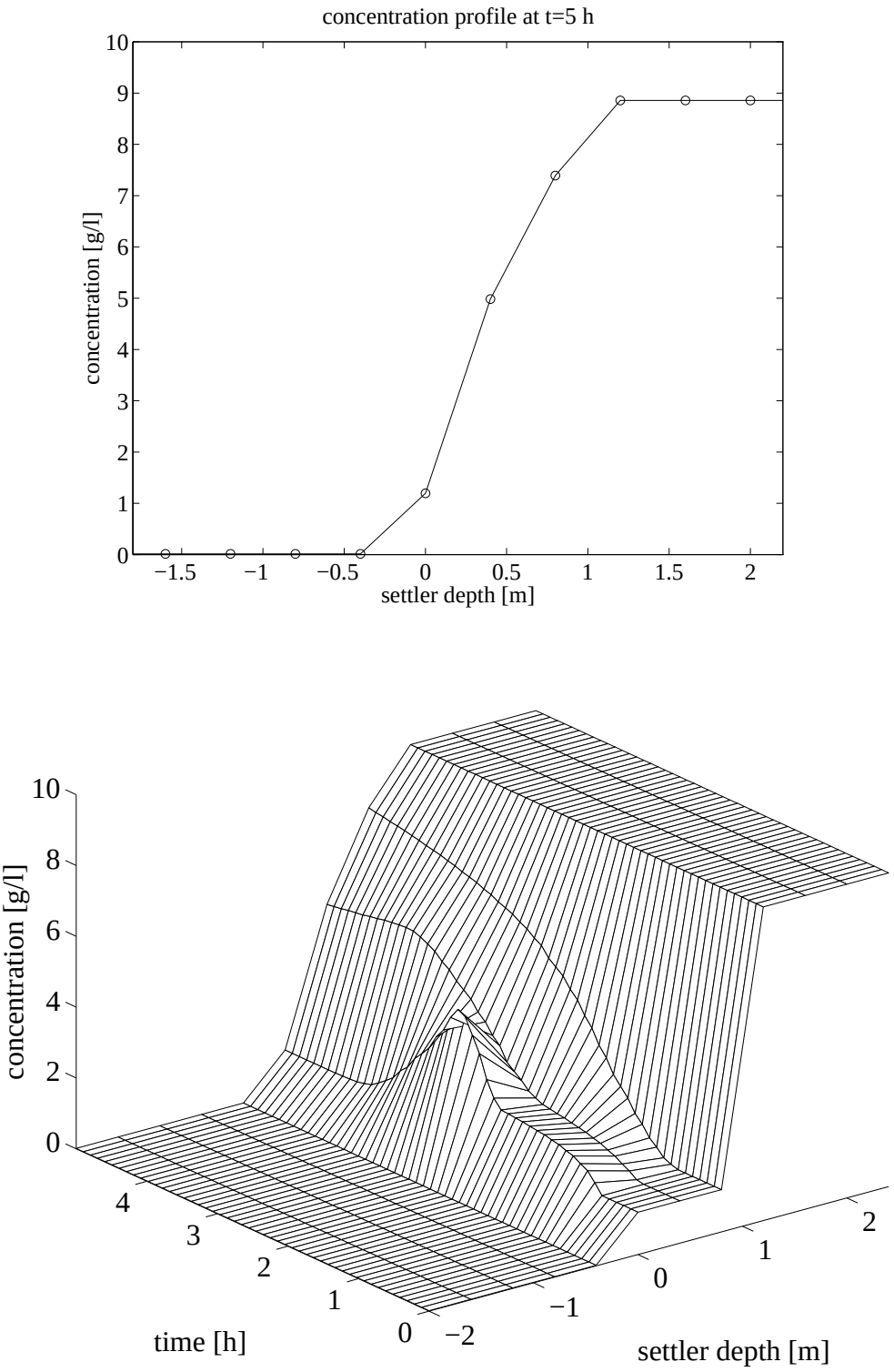


Figure 6.8 A dynamic simulation of the robust settler model starting at steady state according to Table 6.6. The number of grid points $n = 12$. The initial mass of solids in the settler is 6042 kg and at the end of the simulation the mass in the settler is 8039 kg.

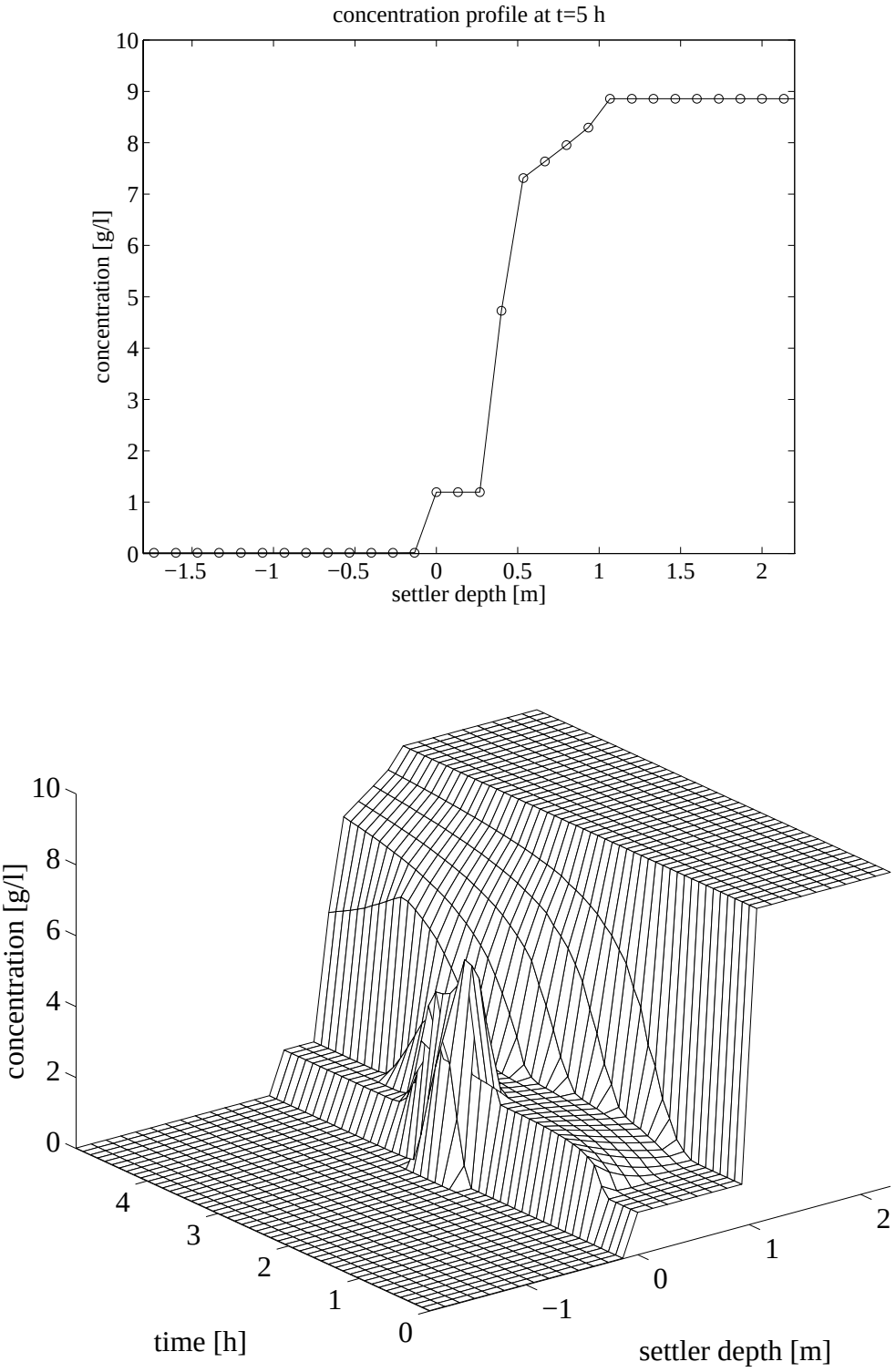


Figure 6.9 A dynamic simulation of the robust settler model starting at steady state according to Table 6.6. The number of grid points $n = 32$. The initial mass of solids in the settler is 5963 kg and at the end of the simulation the mass in the settler is 7961 kg.

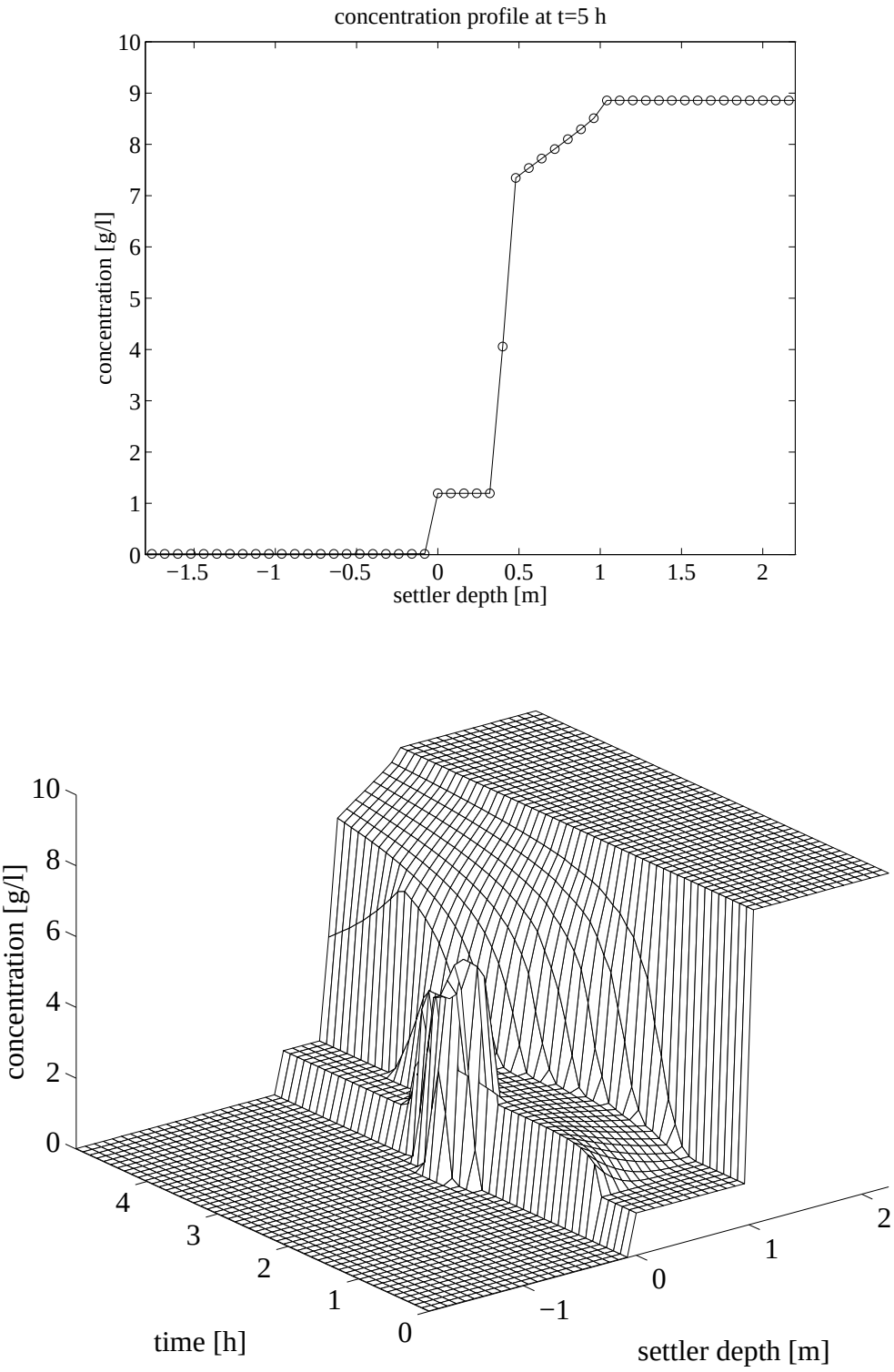


Figure 6.10 A dynamic simulation of the robust settler model starting at steady state according to Table 6.6. The number of grid points $n = 52$. The initial mass of solids in the settler is 5947 kg and at the end of the simulation the mass in the settler is 7945 kg.

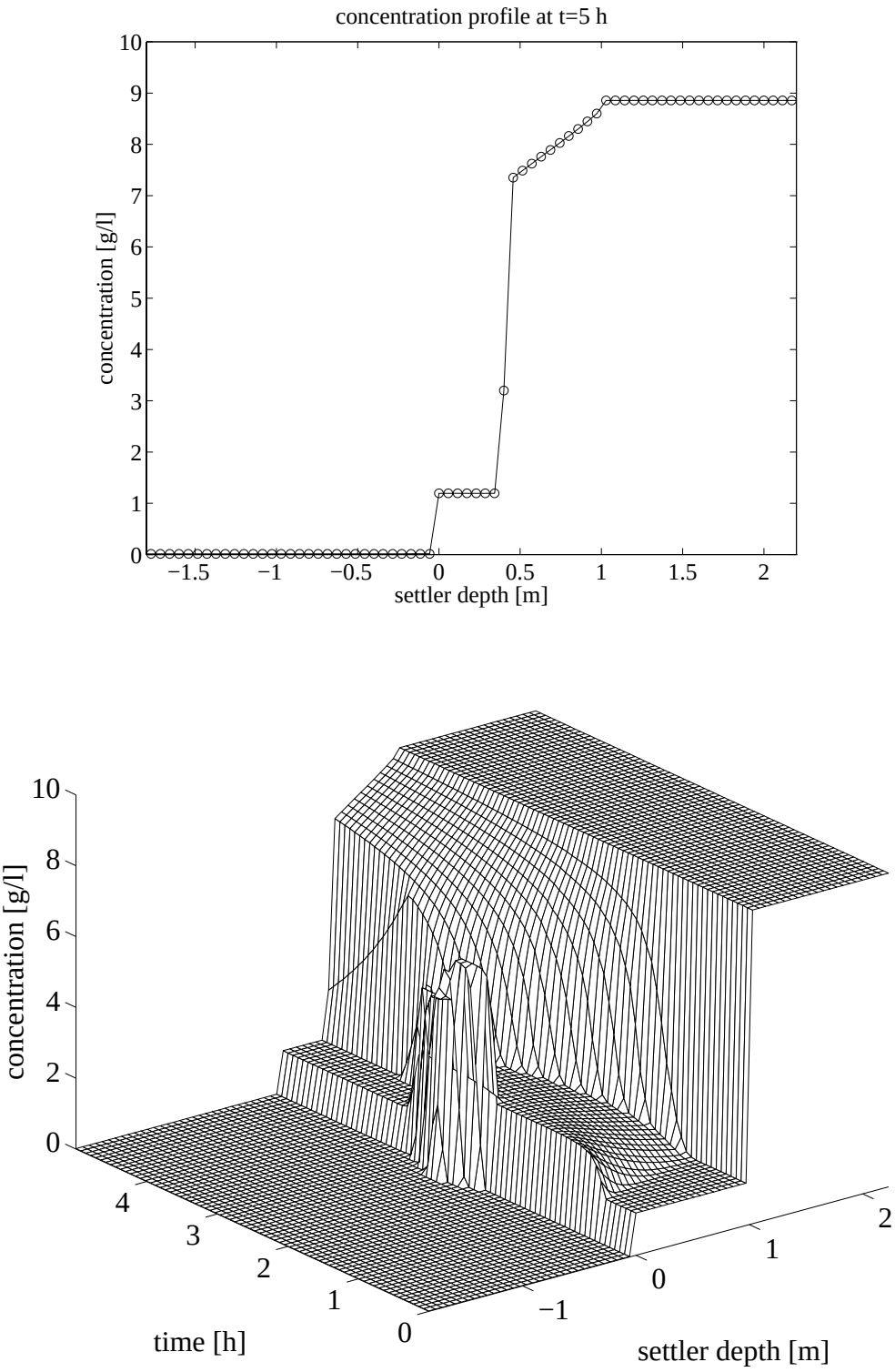


Figure 6.11 A dynamic simulation of the robust settler model starting at steady state according to Table 6.6. The number of grid points $n = 72$. The initial mass of solids in the settler is 5940 kg and at the end of the simulation the mass in the settler is 7938 kg.

Coupled Case

By imposing a number of disturbances on the process, the behaviour of an entire AS plant will be demonstrated, both when using the robust settler model and the traditional layer model. All conditions and model parameters are identical to those used in the coupled case, described in Section 6.2, with the exception of a few necessary modifications. In order to make a fair comparison of the different models it is important that the initial conditions are as similar as possible. Figure 6.4 showed the very different steady states for the traditional Vitasovic model depending on the number of layers used. In order to compensate for this, the wastage flow rate (Q_w) is slightly changed for the 30 and 50-layer cases to allow for an initial steady state with almost the same amount of mass in the settler and the same sludge blanket height as for the 10-layer case. These solutions are also similar to the steady state of the robust model. For the 30-layer case, Q_w is increased from 1.163 m^3/h to 1.20 m^3/h and for the 50-layer case the new Q_w is set to 1.183 m^3/h (Q_e is modified accordingly). Note that the settling process is very sensitive to Q_w, although the observable effects may be slow. For example, starting from the steady state defined in Figure 6.4 for the 50-layer case and setting Q_w to the new value given above, will require the process to be simulated several years (simulated time, not CPU time) before a new steady state is obtained. This further emphasizes the need for good algorithms to calculate a steady-state solution. The above modifications lead to new steady-state concentration profiles using the traditional settler model, which are shown in Figure 6.12, and the new key variables are given in Table 6.8.

The second modification deals with the propagation of the individual components of the biological model through the settler. As discussed in Section 6.1, this is not an important issue when investigating the steady-state behaviour of the settler, since the relative amounts of all soluble and particulate components in this case are the same throughout the settler (identical to the relative amounts in the last bioreactor since we assume that no biological reactions occur in the settler). However, during dynamic conditions we must be able to calculate how the relative amounts of the components vary (especially in the settler underflow). For the robust settler model this is done in a straightforward manner by applying the equations (5.58)–(5.65) in Section 5.4, and it is assumed that the concentration of X_{ND} is always in steady state in the settler underflow. In the traditional model we describe the propagation of the various soluble components in a similar way as in the robust model, that is, the propagation of each soluble component is modelled separately and every layer is completely mixed.

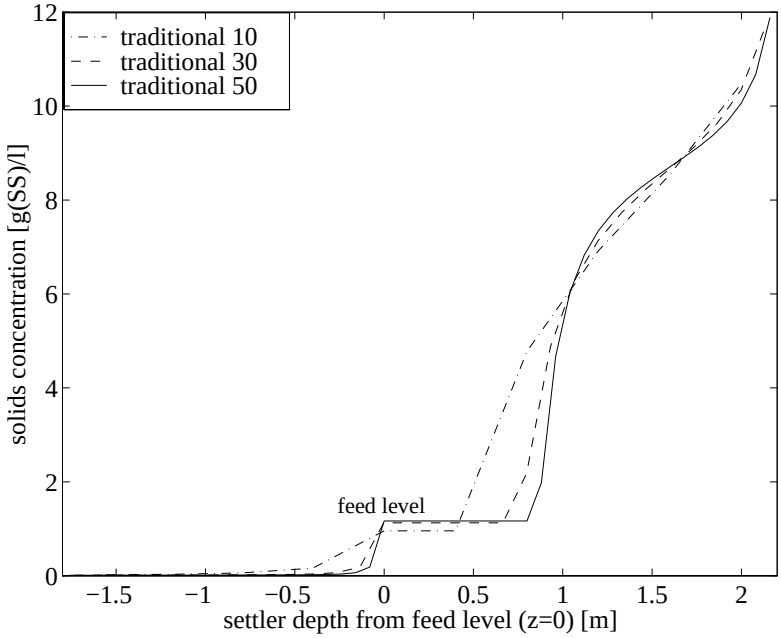


Figure 6.12 Initial steady-state concentration profiles for the coupled traditional Vitasovic model used for investigating the dynamic behaviour.

	Traditional model		
Variables	10 layers	30 layers	50 layers
X_u [g/l]	10.5	11.7	11.9
X_e [mg/l]	20.7	11.8	11.3
X_f [g/l]	4.70	5.22	5.32
mass [ton]	6.58	6.20	6.05
SRT [d]	28.3	32.1	32.9

Table 6.8 The key variables for the initial steady state.

For the particulate components there is no easy way to model the material propagation. We will apply the method described in Otterpohl and Freund (1992), which was discussed in Section 6.1, see equation (6.1). This means that the traditional Vitasovic model has to be modified slightly. The basic structure illustrated in Figure 5.8 is still valid but all the flux terms are now calculated for each particulate component of the IAWQ model, that is, all symbols X_i in Figure 5.8 are replaced by $X_{i,j}$, except in the if-conditions for the clarification zone. Consequently, the flux terms are no longer defined as g SS/(h m^2) but instead as g COD/(h m^2) and g N/(h m^2). We maintain the threshold coefficient based on the total suspended solids concentration in the clarification zone, as discussed in Section 6.1, calculated according to

(6.2). Note that the settling velocity is also based on the total concentration of suspended solids (also calculated according to (6.2)) and not defined for each individual particulate component. We are neither applying the Ω-function of Otterpohl and Freund nor their special modelling of primary particles, but only the method for material propagation in the settler. Some numerical drawbacks of the approach will be demonstrated in the next subsection. It should also be observed that the steady-state solutions presented earlier are not affected by the above model modifications, as the propagation algorithm only influences the dynamic behaviour.

The dynamic behaviour of the robust settler model is investigated and compared with the behaviour of the modified Vitasovic model (still referred to as the traditional model) by imposing some disturbances during a 20-hour simulation of the entire AS process. This is done so that an extreme situation is obtained where the settler becomes overloaded. All disturbances are imposed as step functions since the purpose is to test the models and not to describe a true situation. Moreover, step changes produce more rapid model responses and model discrepancies are easier to detect than if more realistic smooth variations are used. The simulations are initiated to the steady-state conditions described in the previous section combined with the small modifications discussed above. At time $t = 1$ h, Q_{in} is increased from 250 to 750 m^3/h and the influent S_S is increased from 80 to 160 mg COD/l. The hydraulic shock leads to a high load of solids leaving the bioreactor and increases the load on the settler. At $t = 6$ h, Q_r is doubled to 400 m^3/h in an attempt to lower the amount of solids in the settler. At $t = 10$ h, Q_{in} and S_S are set back to their initial values (250 m^3/h and 80 mg COD/l) and, finally, at $t = 16$ h, Q_r is reduced to 200 m^3/h in order to return the process to a situation similar to the initial one. The imposed flow disturbances are shown in Figure 6.13. The disturbances will naturally also cause a smooth variation of the suspended solids concentration entering the settler. The flow-rate disturbances are deliberately chosen to create a troublesome situation with a discontinuity at the very top of the settler. Note that all simulations are performed with so small time steps that the numerical errors in this respect are negligible, that is, performing simulations with even smaller time steps will produce results that cannot be distinguished from those shown in this subsection.

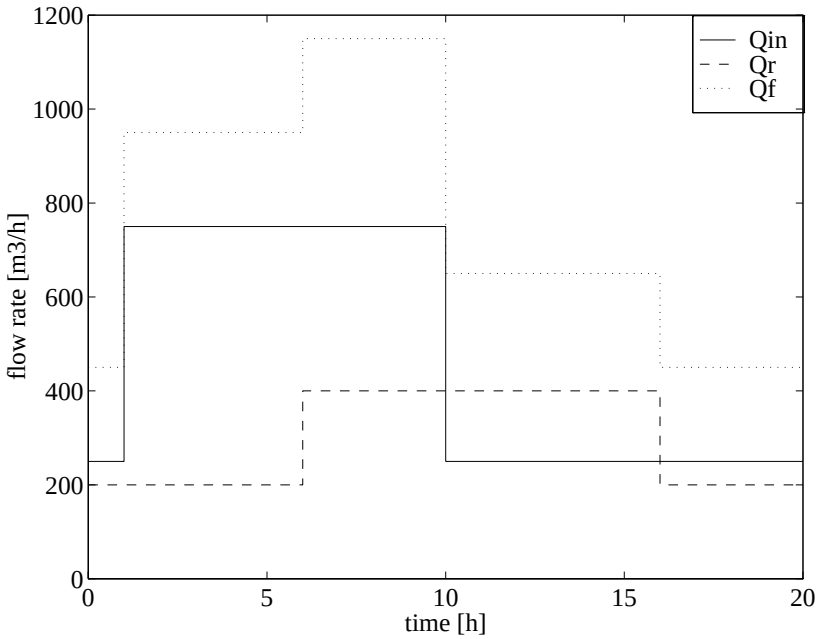
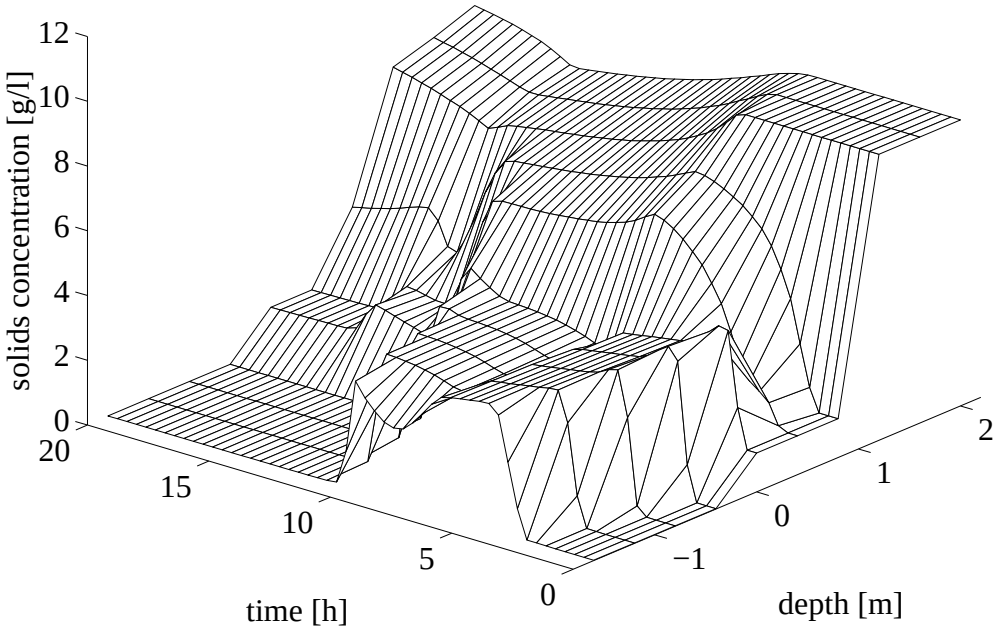
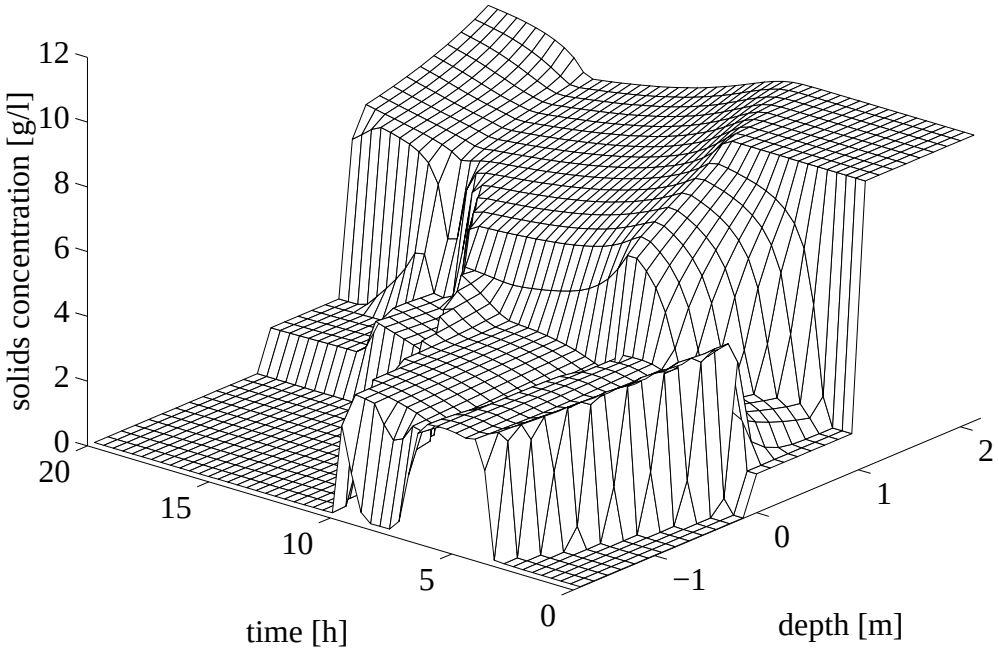


Figure 6.13 The imposed flow disturbances of Q_{in} and Q_r for the dynamic simulations and the resulting disturbance of Q_f.

In Figures 6.14 and 6.15, the dynamics of the SS concentration profile is shown for the two settler models, respectively. The SS concentration is calculated according to (6.2) for the modified Vitasovic model. Again, the consistency of the robust model is apparent, whereas the result of the 10-layer Vitasovic model is quite different from the 30 and 50-layer results. The speeds of the shock waves (for 30 and 50 layers) seem to be quite similar for both models. For example, the shock wave in the clarification zone rises with the same speed in both models between $t \approx 1$ and $t \approx 3$ h. This is because the numerical fluxes for the two models produce the same values in this particular case. However, the Vitasovic model makes the rising shock wave in the clarification zone stop before the effluent level. This is observed in more severe overflow situations as well. There is no physical reason why a rising shock wave should not reach the effluent level. When this happens (at $t \approx 4$ h), the mass in the settler predicted by the Vitasovic model becomes lower than the one by the robust model, which is seen in Figure 6.16. A similar effect is seen for the 10-layer robust model. In this case, however, it is a result of the rough discretization, which leads to poor prediction of the effluent concentration of SS.


solids concentration [g/l]
12
10
8
6
4
2
0
20
15
10
5
0
time [h]
2
1
0
−1
depth [m]


solids concentration [g/l]
12
10
8
6
4
2
0
20
15
10
5
0
time [h]
2
1
0
−1
depth [m]

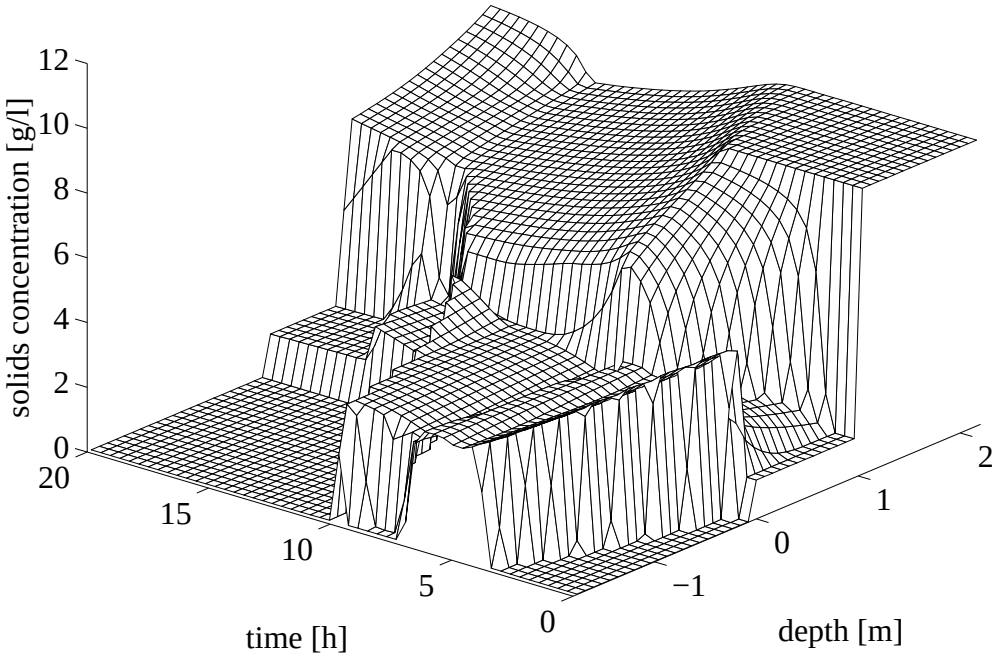
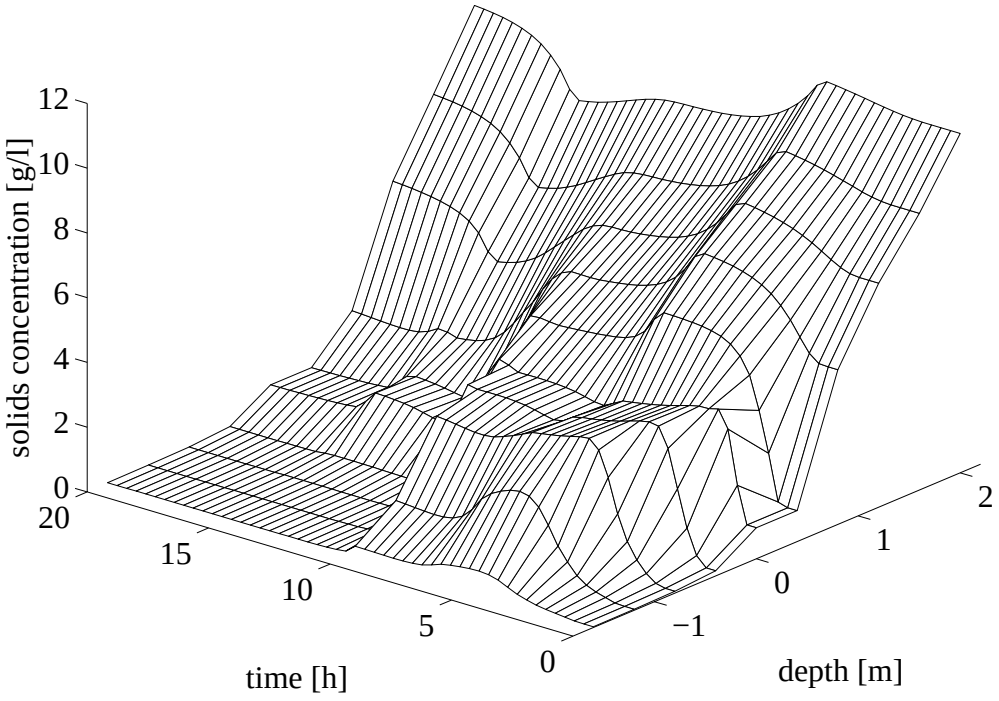


Figure 6.14 (This and the previous page) The suspended solids concentration in the settler as a function of time and depth for the robust model (first plot: 10 layers; second: 30 layers; third: 50 layers).



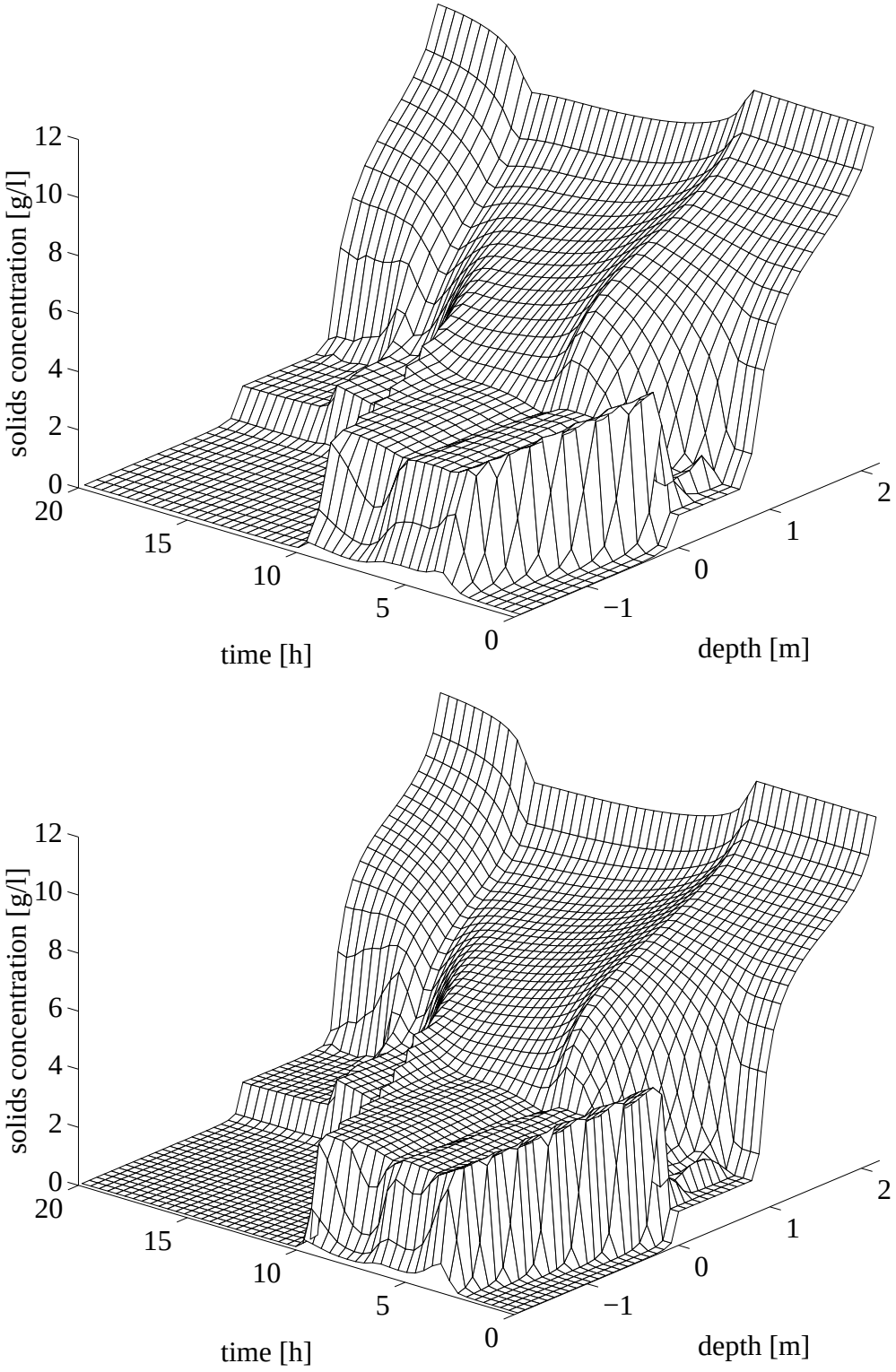


Figure 6.15 (This and the previous page) The suspended solids concentration in the settler as a function of time and depth for the modified Vitasovic model (first plot: 10 layers; second: 30 layers; third: 50 layers).

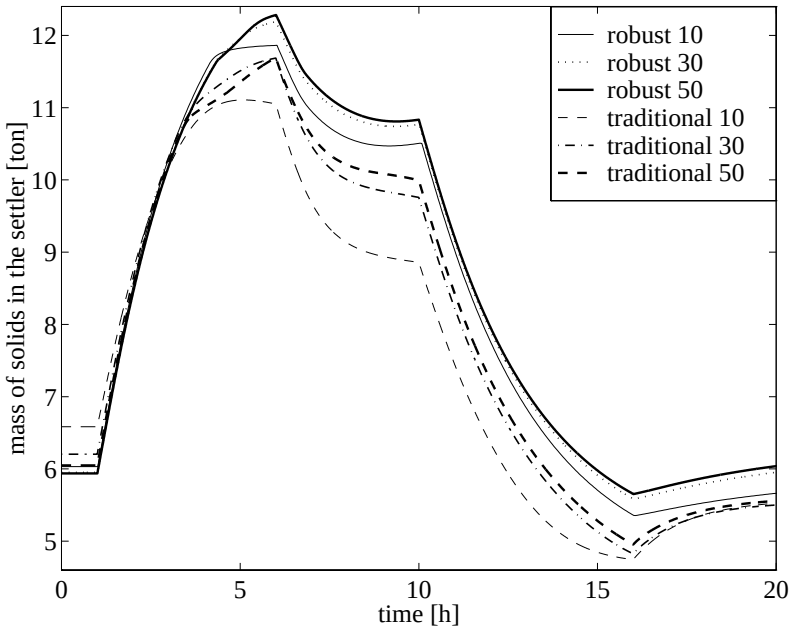


Figure 6.16 The mass of solids in the settler when using the robust and the modified Vitasovic settler models with different number of layers during a dynamic 20-hour simulation.

The settler feed and underflow concentrations predicted by the robust model and shown in Figures 6.17 and 6.18, are quite similar, especially for the 30 and 50-layer cases. The results of the modified Vitasovic model are more scattered (taking the different initial concentrations into account).

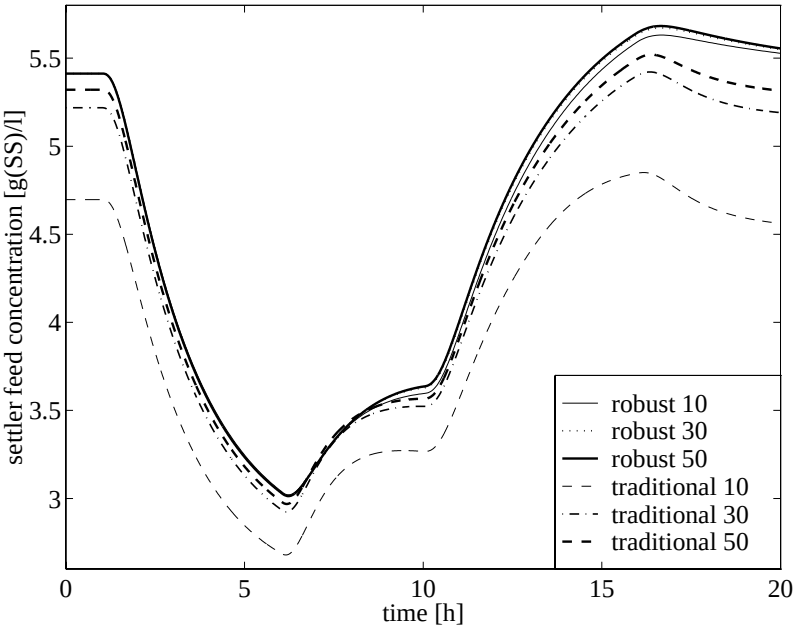


Figure 6.17 The settler feed concentration of SS when using the robust and the modified Vitasovic settler models with different number of layers during a dynamic 20-hour simulation.

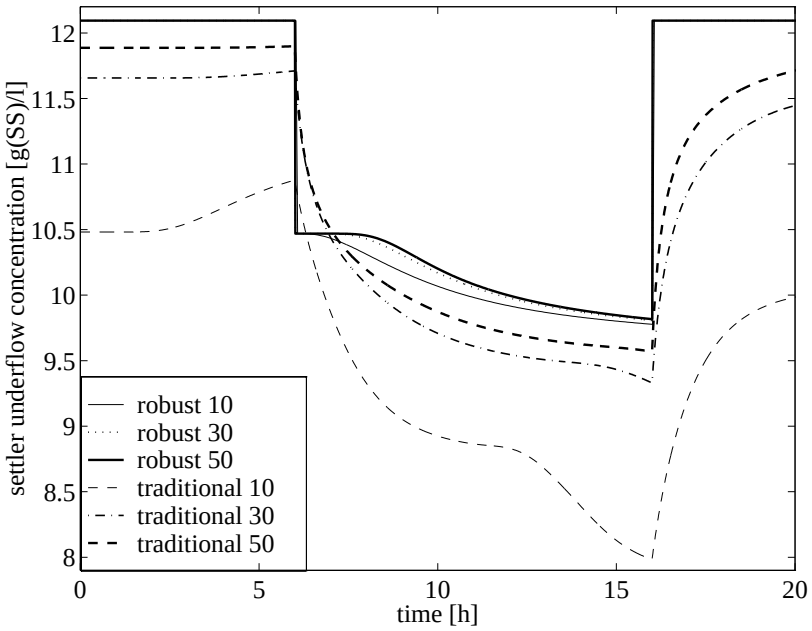


Figure 6.18 The settler underflow concentration of SS when using the robust and the modified Vitasovic settler models with different number of layers during a dynamic 20-hour simulation.

For the effluent concentration the differences between the models are especially prominent, which is a result of the troublesome conditions chosen for the simulations. This is shown in Figures 6.19 and 6.20 (the results of the two models have been separated to improve the readability). The large differences *between* the two models come from the different predictions of the effluent concentration discussed above. The large differences *within* the respective model come from the fact that there is a discontinuity at the very top of the settler between $t \approx 4$ and $t \approx 10$ h. To resolve the dynamics of the effluent solids concentration for such a difficult situation, one needs to use an even larger number of layers. Therefore, the result when using a 70-layer robust model has been included in Figure 6.19, which makes it easier to see the convergence of the numerical solution.

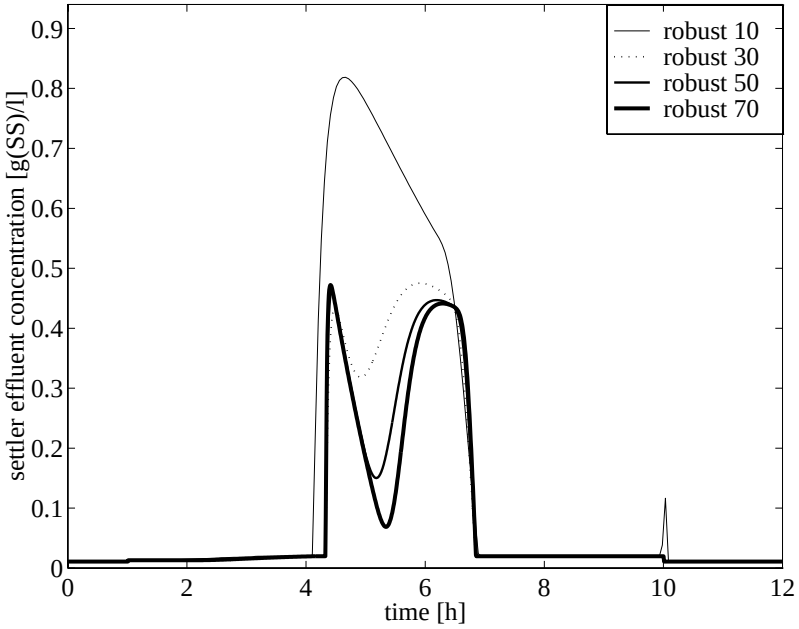


Figure 6.19 The settler effluent concentration of SS when using the robust settler model with different number of layers during a dynamic 20-hour simulation. Note the different time scale in the plot.

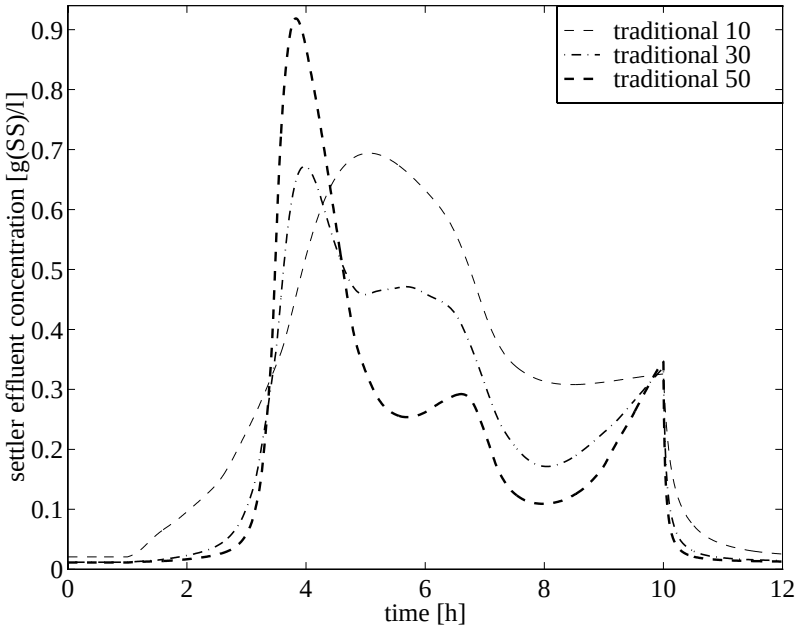


Figure 6.20 The settler effluent concentration of SS when using the modified Vitasovic model with different number of layers during a dynamic 20-hour simulation. Note the different time scale in the plot.

In order to provide a complete description of the process, the dynamic behaviour of the biological reactor is illustrated in Figures 6.21 and 6.22. Figure 6.21 shows the concentrations of the soluble components in the bioreactor and Figure 6.22 shows the variations of the major particulate components. Note that the concentration of the inert particulate material is the sum of X_P and X_I. The shown graphs are from the case where the bioreactor is simulated in combination with the 50-layer robust settler model. However, the qualitative behaviour of the variations in the bioreactor is the same for all cases presented in this subsection.

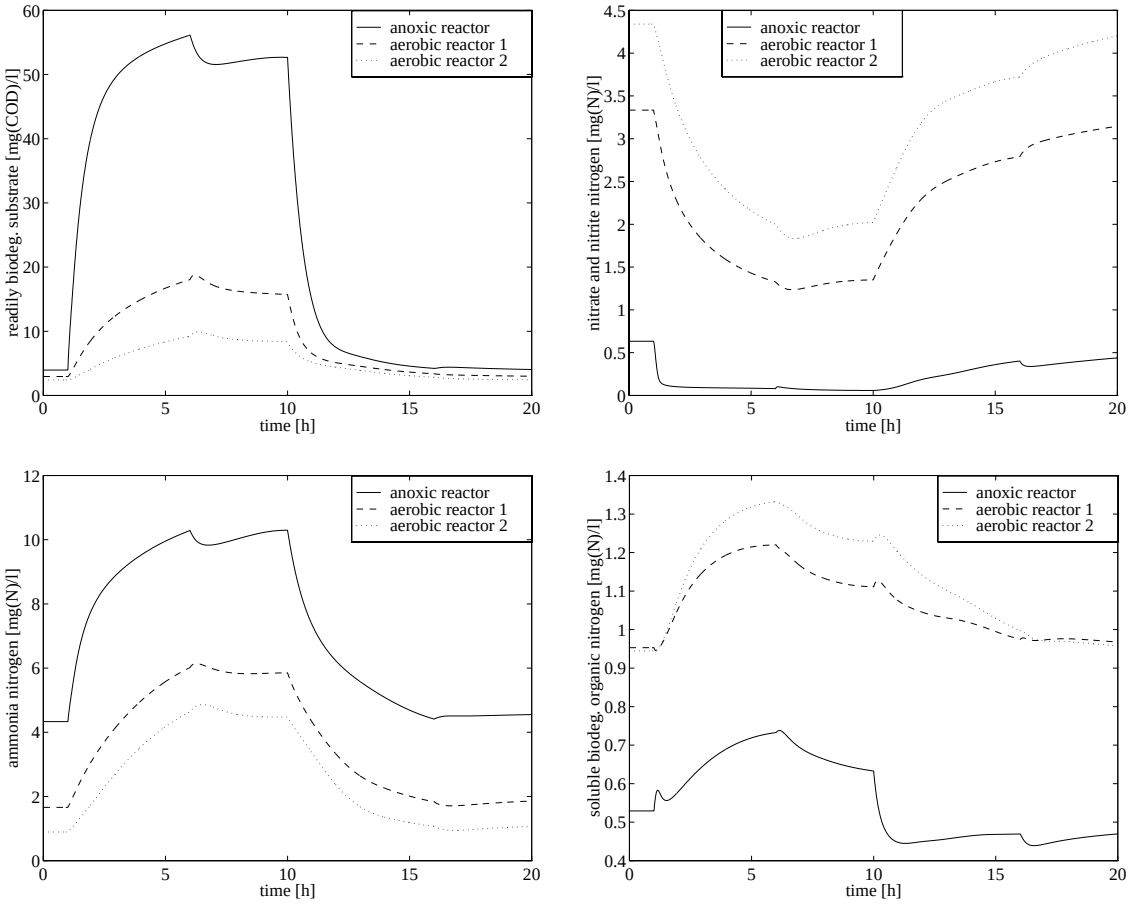


Figure 6.21 Behaviour of the soluble components in the bioreactors during a dynamic 20-hour simulation using the 50-layer robust settler model.

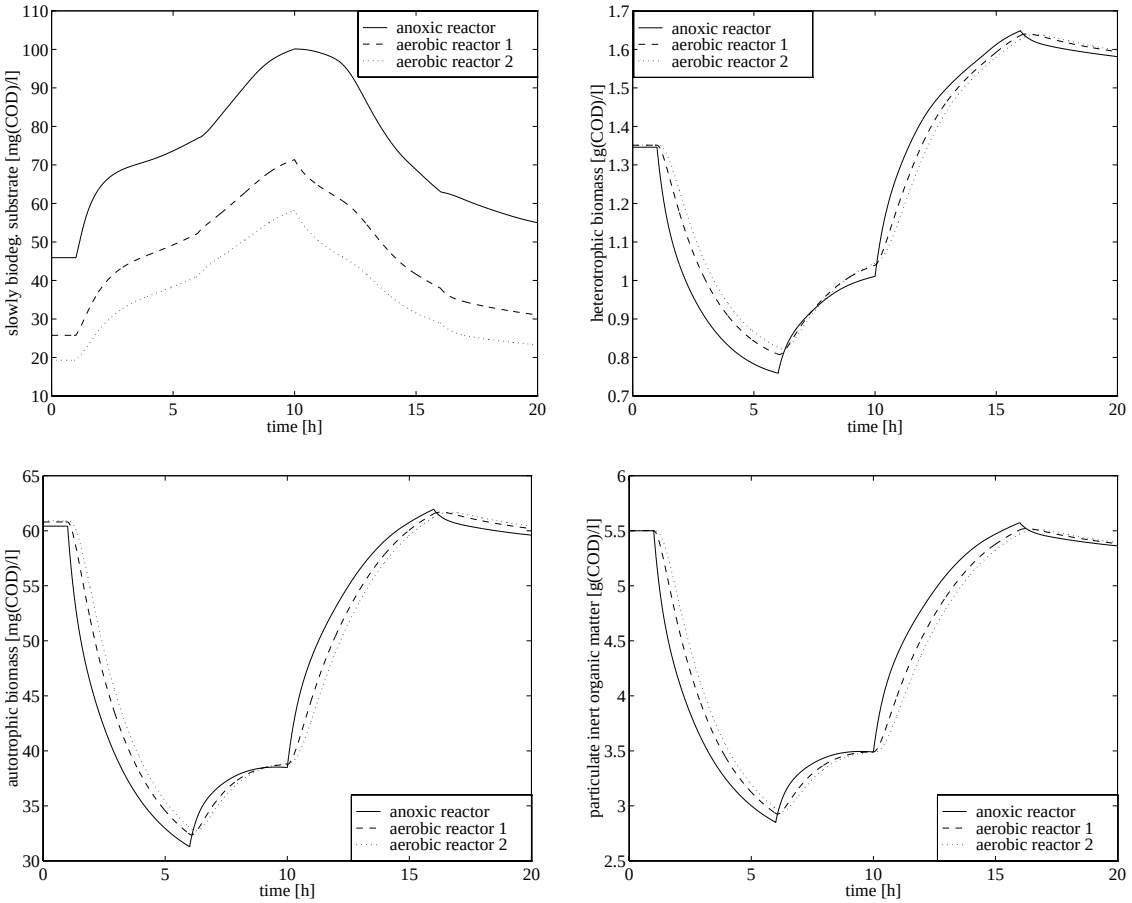


Figure 6.22 Behaviour of the major particulate components in the bioreactors during a dynamic 20-hour simulation using the 50-layer robust settler model.

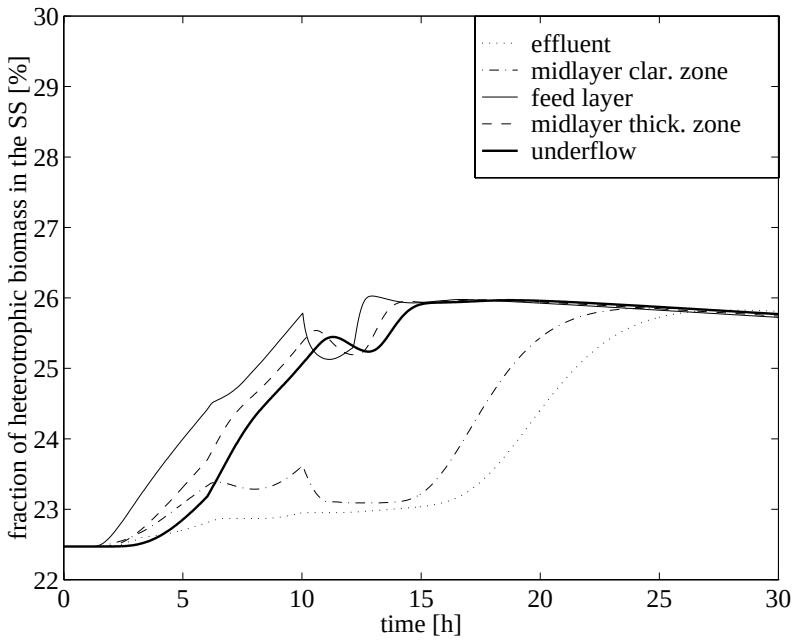
Dynamic Propagation of Biological Components

In this subsection we will demonstrate the behaviour of the two different algorithms for the dynamical update of the concentrations of the individual biological components as the solids are transported through the settler. All simulations are identical to those shown in the previous subsection except that the simulations are continued for another 10 hours (with constant model inputs), that is, a total simulation period of 30 hours.

The positive effects of the material propagation algorithm used by the robust settler model (see Section 5.4) are difficult to observe in the graphs shown in the last subsection. A percentage vector, which describes how many percent of the suspended solids are actually X_S, $X_{B,H}$, $X_{B,A}$ and $X_I + X_P$, is kept for each layer in the settler and is updated on every time step. As this update is based on the analytically correct concentration of SS

and the Godunov flux terms, the percentage vectors are also analytically correct. This means that we can easily calculate the concentration of the particulate components used in the biological model for every layer of the settler at every time instance. For the simulations in this section it is only of importance to know the different concentrations in the settler underflow, as we are recycling sludge to the bioreactor. However, if we want to extend the settler model by taking into account various biological reactions within the settler, it is almost of equal importance to know the composition of the sludge within every separate layer. This is accomplished by using the material propagation algorithm in Section 5.4. The algorithm is stable and computationally efficient in the sense that the extra required CPU time is hardly noticeable when compared with the time required to solve the basic robust model equations discussed in Section 5.4 numerically.

To exemplify the algorithm, Figure 6.23 shows the percentages of the suspended solids which constitute the heterotrophic biomass in the effluent and underflow as well as in the feed layer, a layer in the middle of the clarification zone and a layer in the middle of the thickening zone for the same dynamic simulation as described earlier in this section, using the 10, 30 and 50-layer robust settler model.



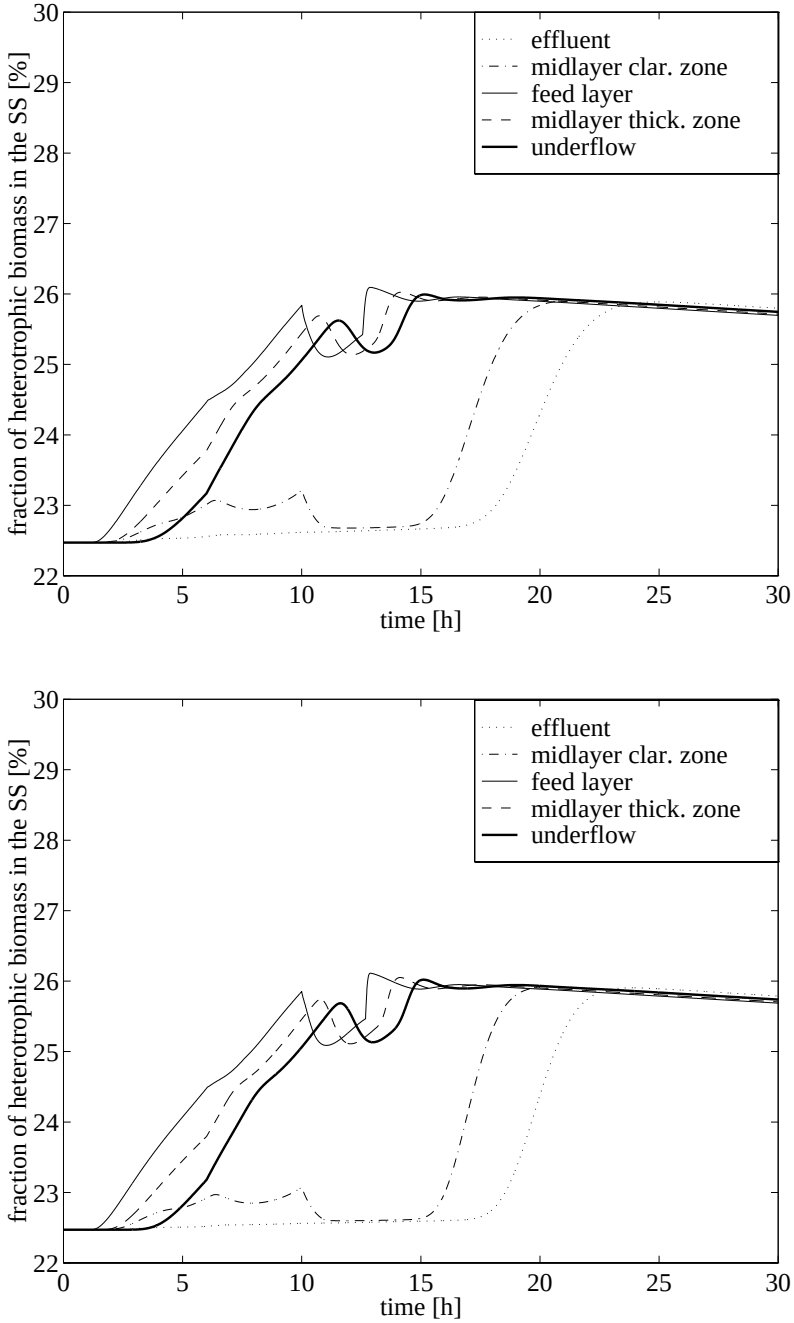


Figure 6.23 (This and the previous page) The fraction of heterotrophic biomass of the suspended solids in different layers of the settler when using the robust model during a 30-hour dynamic simulation (first plot: 10 layers; second: 30 layers; third: 50 layers).

It is clear that the results are practically independent of the number of layers. We also see how the fraction of $X_{B,H}$ increases rapidly in the feed layer as an effect of the higher influent flow rate and the extra added S_S (at $t = 1$ h), and how this 'dynamic' sludge with a different composition than the initial steady state rapidly propagates downwards in the settler. The

upward propagation of sludge from the feed layer looks quite differently and it is apparent that practically no part of the new sludge reaches the effluent, as the fraction of $X_{B,H}$ is almost constant here until $t \approx 17$ h. This means that the high concentration of SS in the effluent seen in Figure 6.19 is actually made up of the sludge that was in the clarification zone from the initial steady state and this has propagated upwards. At $t \approx 13$ h most of the sludge, which had accumulated in the clarification zone during the first hydraulic shock, has moved downwards into the thickening zone and new sludge (of very low concentration) starts to move upwards according to the double-exponential settling velocity function. It reaches the middle of the clarification zone at $t \approx 15$ h and the effluent at $t \approx 17$ h. At $t \approx 23$ h, the fraction of heterotrophic biomass is almost constant within the entire settler but it is not a perfect steady state. We see how the fraction of $X_{B,H}$ is slowly decreasing and after a few days the system will have returned to the initial steady state in every respect. The reason for this slow variation is, naturally, that it is controlled by the decay processes in the bioreactor, which gradually transform the additional fraction of $X_{B,H}$ created by the dynamic disturbances into inert material.

In a similar way we can construct three-dimensional graphs that show the fraction of any biological component as a function of settler depth and time. Such a graph is given in Figure 6.24, showing the fraction of heterotrophic biomass of the suspended solids in the settler using the 50-layer robust settler model.

By multiplying the percentage vectors for each layer with the corresponding concentration of SS and transforming the result into g COD/l (by dividing the result with the appropriate transformation coefficient from Table 6.1) we can obtain dynamic concentration profiles for every particulate component of the IAWQ model within the settler. In Figure 6.25, this is illustrated by the concentration profile for the heterotrophic biomass using the 50-layer robust settler model (only shown for the first 20 hours of the simulation). Naturally, the plot is qualitatively very similar to the one in Figure 6.14 showing the SS concentration, but detailed studies reveal some small differences.

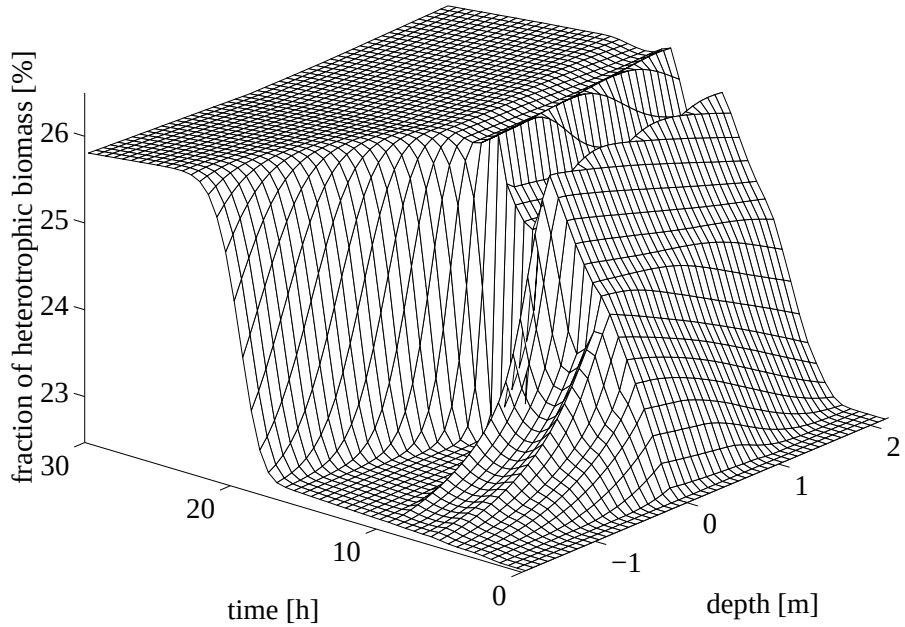


Figure 6.24 The fraction of heterotrophic biomass of the suspended solids in the settler when using the 50-layer robust model during a 30-hour dynamic simulation as a function of settler depth and time.

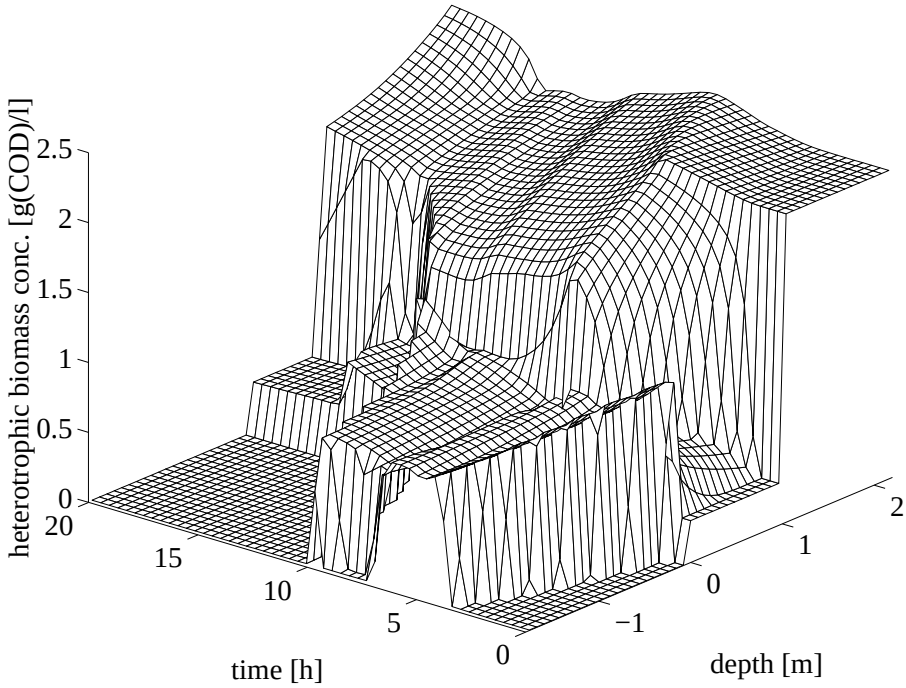
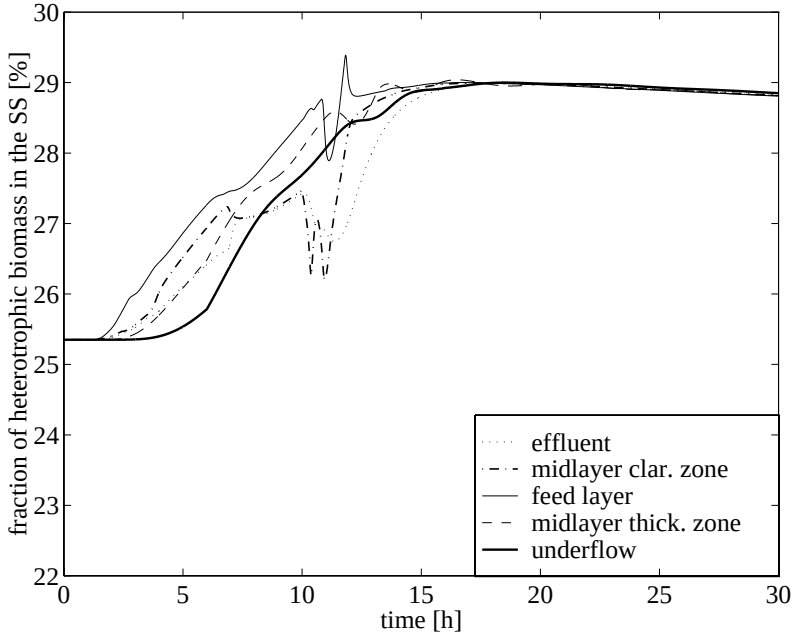


Figure 6.25 The concentration of heterotrophic biomass in the settler when using the 50-layer robust model during a 30-hour dynamic simulation as a function of settler depth and time (only shown for the first 20 hours).

The results for the other particulate components of the SS are similar and not shown here. No example is given for the propagation of the different soluble components. This algorithm is quite simple and straightforward, as the soluble material only follows the bulk flows (upwards in the clarification zone and downwards in the thickening zone). The particulate material propagation algorithm is believed to improve the dynamic behaviour of the robust settler model. However, it is possible that a one-dimensional model may be too rough an approximation to describe this type of detailed behaviour in reality, since other effects such as hydrodynamic phenomena may have a significant impact on the true behaviour.

The alternative method for modelling the propagation of individual particulate components in the settler (Otterpohl and Freund, 1992), applied to the traditional Vitasovic model in this work, was discussed in Section 6.1 and some drawbacks were pointed out. Although the method is quite straightforward, detailed studies of its behaviour reveal some drawbacks. The algorithm calculates the flux of each particulate component instead of combining the components into one variable describing the suspended solids concentration. We can easily transform the individual concentration for each layer into a SS concentration by equation (6.2) and, consequently, determine the percentage of each component of the total SS concentration and compare the results with the ones shown in Figure 6.23 when using the robust settler model. In Figure 6.26 the results are presented, showing the fraction of $X_{B,H}$ in the settler (for the same layers as in Figure 6.23) using a 10, 30 and 50-layer modified Vitasovic model. When comparing the results to those in Figure 6.23, the drawbacks of the algorithm become apparent.



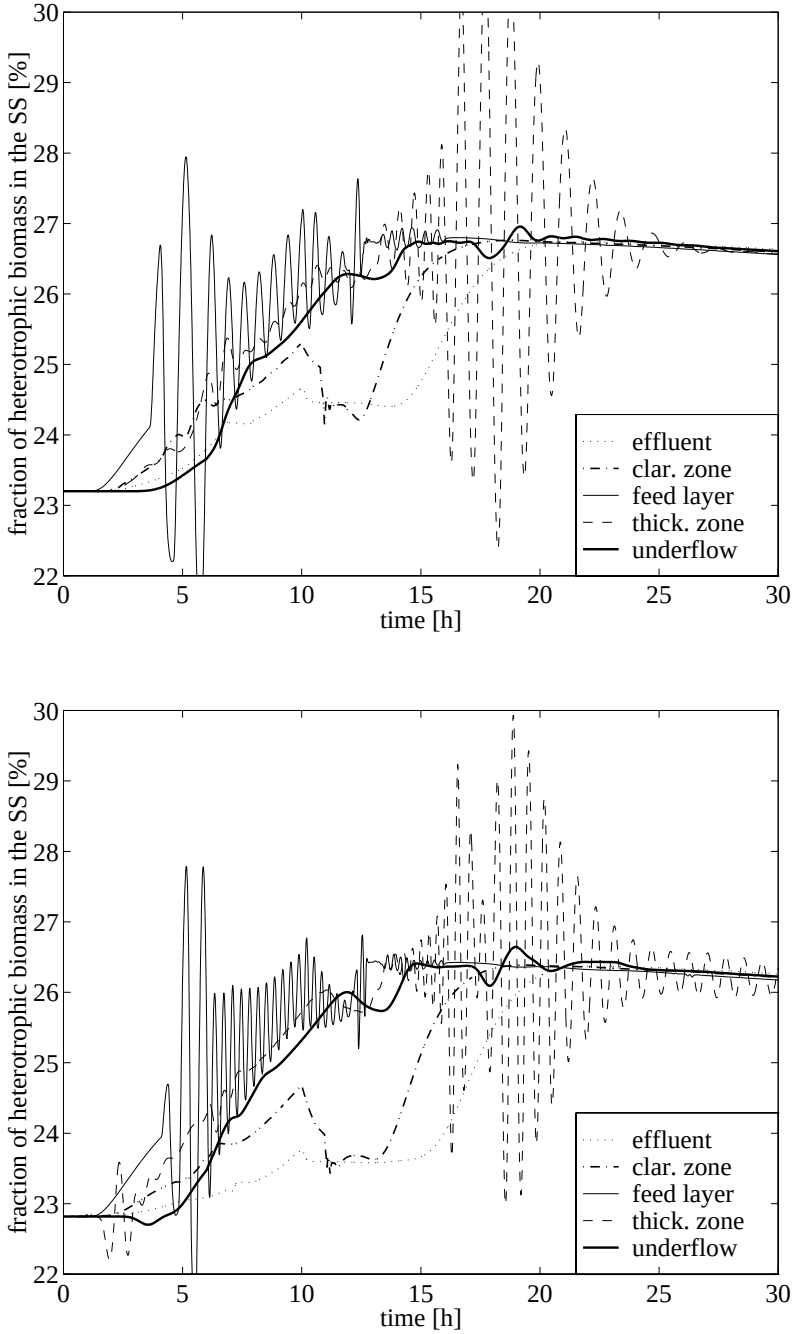


Figure 6.26 (This and the previous page) The fraction of heterotrophic biomass of the suspended solids in different layers of the settler when using the modified Vitasovic model during a 30-hour dynamic simulation (first plot: 10 layers; second: 30 layers; third: 50 layers).

The same type of oscillation is found for all the particulate components. For the fractions in the effluent, underflow and above the feed layer the results are not that much different from the ones in Figure 6.23 (qualitatively). From the feed layer and downwards the oscillating behaviour is

much more prominent. The problem is only to a small part a numerical one. Otterpohl and Freund (1992) suggest that the volume of a layer must be at least one order of magnitude larger than the flow in one time interval, to avoid numerical problems, but this does not suffice. The oscillations can be somewhat reduced by decreasing the tolerance of the numerical algorithm. In the case above, a Gear algorithm with a low error tolerance was used, that is, an accurate numerical solution was obtained (e.g., Kahaner *et al.*, 1989). Other algorithms (Euler, Runge-Kutta and Adams) with different error tolerances and time steps have been tested with similar results. An advantage of an analytically derived model (such as the robust settler model) is that an exact upper bound of the time-volume ratio can be given (see equation (5.49)).

The basic problem lies within the model structure itself. First of all, the numerical flux terms of the Vitasovic model are not the analytically correct ones, as discussed in previous sections. However, the oscillating behaviour is only apparent when the numerical flux terms are defined for the individual particulate components and not when we model all the particulate material as one variable. To explain this we look at the boundary between the feed layer and the layer below it (layer m and $m+1$) and consider only the heterotrophic biomass (index a) and the inert material (index b), as these two components make up approximately 99% of the suspended solids in the simulated case. The downward settling flux from layer m to layer $m+1$ is calculated for the two components as

$$\begin{aligned} J_{s,m,a} &= \min\left(v_{s,m}(X_m)X_{m,a}\ ,\ v_{s,m+1}(X_{m+1})X_{m+1,a}\right) \\ J_{s,m,b} &= \min\left(v_{s,m}(X_m)X_{m,b}\ ,\ v_{s,m+1}(X_{m+1})X_{m+1,b}\right) \end{aligned} \tag{6.15}$$

according to the algorithm by Otterpohl and Freund (1992). From $t=0$ to $t=1$ h, the system is in steady state and all concentrations in the layers from the feed level down to the beginning of the sludge discontinuity are identical. At $t=1$ h, disturbances are imposed on the system and the settler feed concentrations and flow conditions start to change. At some time instances the first part of the min-terms will determine the settling flux and at other times the second part will be the minimum. There is no guarantee that both flux terms in (6.15) will change at the same time, that is, at some time the settling flux of heterotrophic biomass will be controlled by the situation in layer m whereas the settling flux of the inert material will be controlled by the current situation in layer $m+1$. This means that the particulate material is no longer modelled to settle as a uniform floc but as individual components, which contradicts the fundamental assumption that

the biological components are present as flocs. Figure 6.27 illustrates this by showing which part of the min-terms in (6.15) that are active during the first 15 hours of the simulation at the boundary between layers m and $m+1$. A value of 1 means that it is the conditions in layer m that determine the settling flux and a value of 0 implies that the conditions in layer m+1 are controlling the flux at the specific time instance. On several occasions, the settling flux of the biomass and the inert material are determined by conditions in different layers. Even during steady-state conditions ($t < 1$ h) some small variations can be detected as a result of the limited numerical tolerance. This implies that oscillations can never be avoided completely when the above algorithm is used.

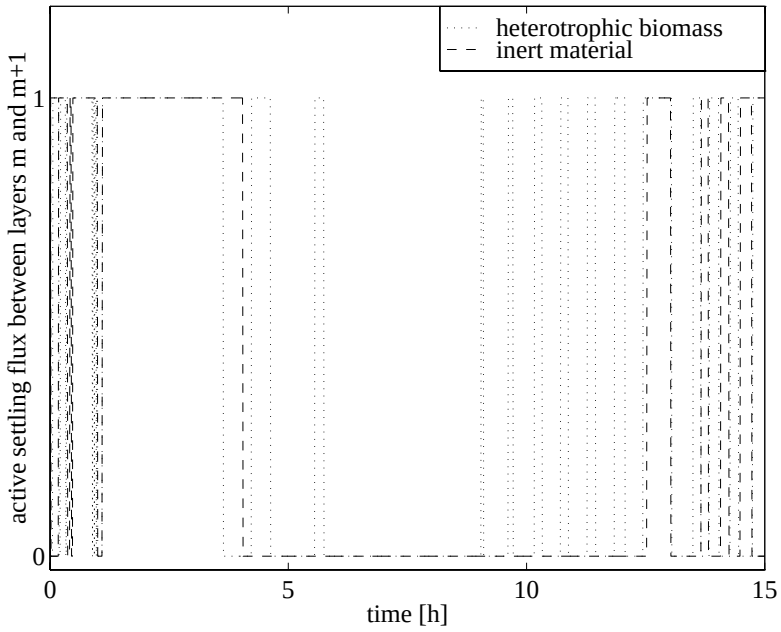


Figure 6.27 Graphs showing which parts of the min-terms in (6.15) are controlling the settling fluxes of $X_{B,H}$ and X_I+X_P at the boundary between the feed layer (m) and the layer below when using the 30-layer modified Vitasovic model (1 = first part; 0 = second part).

The system does not seem to become completely unstable, and as the external disturbances cease, the oscillations are reduced. It should also be noted that the oscillations are correlated so that an increasing concentration of heterotrophic biomass in one layer coincides with a decreasing concentration of inert material and vice versa. The bulk flows will in most cases also have a smoothing effect on the oscillations as it will transport components between the layers regardless of the settling fluxes. Therefore, the SS concentration in a layer does not oscillate as much as the concentrations of the individual components. This is illustrated in Figure 6.28, showing the

concentration of heterotrophic biomass, inert material and suspended solids in a layer 0.4 m below the feed level when using the 30-layer modified Vitasovic model (that is, the traditional layer model combined with the Otterpohl-Freund algorithm).

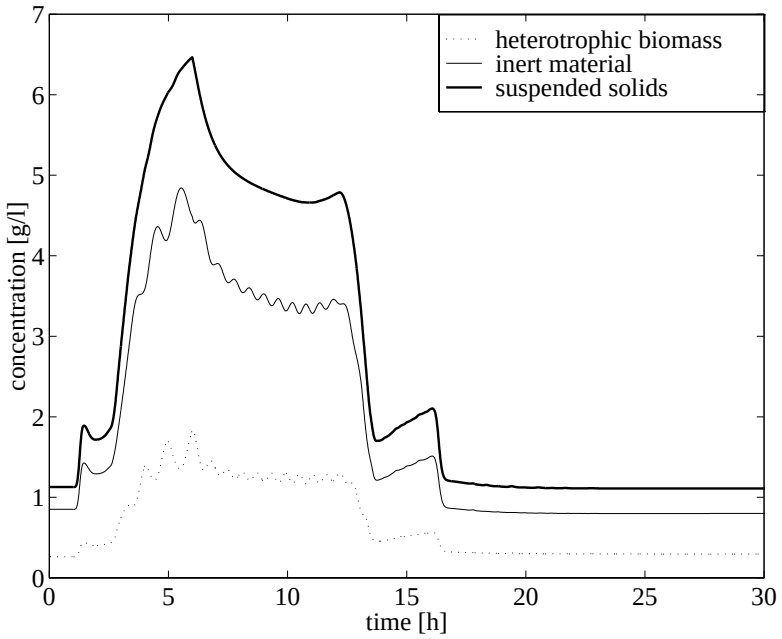


Figure 6.28 The concentrations of heterotrophic biomass, inert material and suspended solids in a layer 0.4 m below the feed level using the 30-layer modified Vitasovic model during a 30-hour dynamic simulation.

This is not a physically relevant situation unless there are oscillations in the settler feed concentrations. In the case shown here, the variations of the settler feed concentrations are smooth and change according to Figure 6.22. Note that oscillations in the clarification zone often occur as secondary effects of the oscillations in the feed layer caused by the upward bulk movement from the feed layer carrying sludge into this zone. Only when the SS concentration in the clarification zone reaches above the value of the threshold coefficient (3 g SS/l in our case) will the settling fluxes not be uniquely determined. This is one reason why no oscillations occur in the effluent in Figure 6.26. Oscillations may also occur in layers below the beginning of the sludge discontinuity, though most of them are also due to secondary effects from the layers above. The reason why the oscillations are reduced in this region is because of the large concentration differences between adjacent layers, which increase the probability (but does not guarantee) that the settling fluxes will be uniquely defined, that is, determined by the situation in the same layer for the different components. Therefore, the oscillating behaviour is mainly a problem in the layers

between the feed layer and the beginning of the sludge discontinuity, where the concentrations are similar. However, secondary effects will influence the behaviour in other layers as well. The model is also more sensitive to the oscillations as the number of layers increases (see Figure 6.26). It is quite difficult to detect any oscillations for the 10-layer modified Vitasovic model during the dynamic simulation in this section. However, it is advantageous to use a large number of layers, as this will improve the reliability and resolution of the solution. Therefore, we require a settler model where this can be accomplished without producing results that exhibit the above type of undesired behaviour.

A complete investigation of the oscillating behaviour discussed above is beyond the scope of this work. However, we can easily demonstrate that the oscillations are due to the switching of the flux terms. In Figure 6.29, the simulations of Figure 6.26 are repeated for the 30-layer model, but with a slight model modification. During this simulation the settler model is modified so that the flux terms in every layer are *always* defined by the situation in the current layer, that is, the first part of the min-terms (see equation (6.15)). All oscillations are removed by such a model modification.

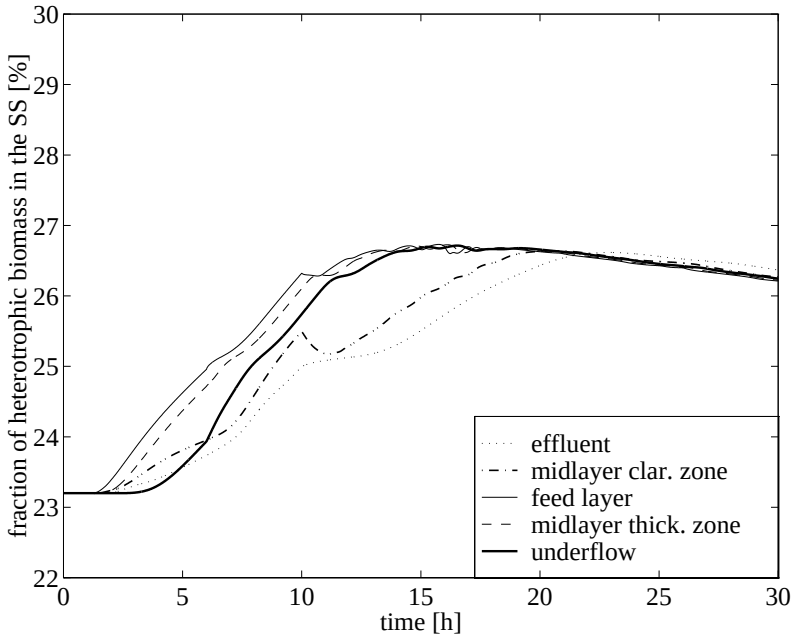


Figure 6.29 The fraction of heterotrophic biomass of the suspended solids in different layers of the settler when using the 30-layer modified Vitasovic model with modified numerical flux terms during a 30-hour dynamic simulation.

The above model modification will, naturally, influence the complete behaviour of the model and it is therefore not relevant to make any detailed comparisons with the previous results. Another possible approach to reduce the oscillations is to use low-pass filters, but this will also alter the dynamic behaviour of the model. A better method is to apply the material propagation algorithm presented in Section 5.4 in combination with the robust settler model.

6.4 Future Model Development

The future potential of the robust settler model is believed to be good, primarily for operational applications. The model is easily combined with other types of models, such as the IAWQ model, and is a useful tool when developing control strategies for the AS process. However, the model must be refined further to describe the behaviour of a real settler.

In this chapter, the behaviour of the robust settler model was compared with an established one-dimensional layer model. It must be emphasized that this comparison only reveals various modelling and numerical problems based on a mathematical analysis. Although the numerical solution of the model has the ability to approximate the analytical solution of the conservation law, this does not guarantee that the model will predict reality well. For example, the steady-state concentration profiles do not look like a 'normal' profile found in a real settler. The reason for this is, naturally, that the model is a simplification because several effects are not included in the constitutive assumptions. On the other hand, the model is consistent and the previously used *ad hoc* conditions are in the robust model replaced by formulae that are derived with a strict mathematical analysis by using basic physical principles and Kynch's assumption. This means that it is a reliable platform for future model refinements.

One obvious improvement is to extend the model so that a settler with a varying cross-sectional area can be described. The theory for this has been developed by Diehl (1996c). If the settler area is decreasing towards the bottom then the height of the steady-state sludge blanket level is uniquely defined within the thickening zone (if there exists a discontinuity in the thickening zone) based on the feed concentration and the bulk flow rates in the settler. Moreover, the predicted concentration profile will assume a more realistic appearance, as illustrated in Figure 6.30.

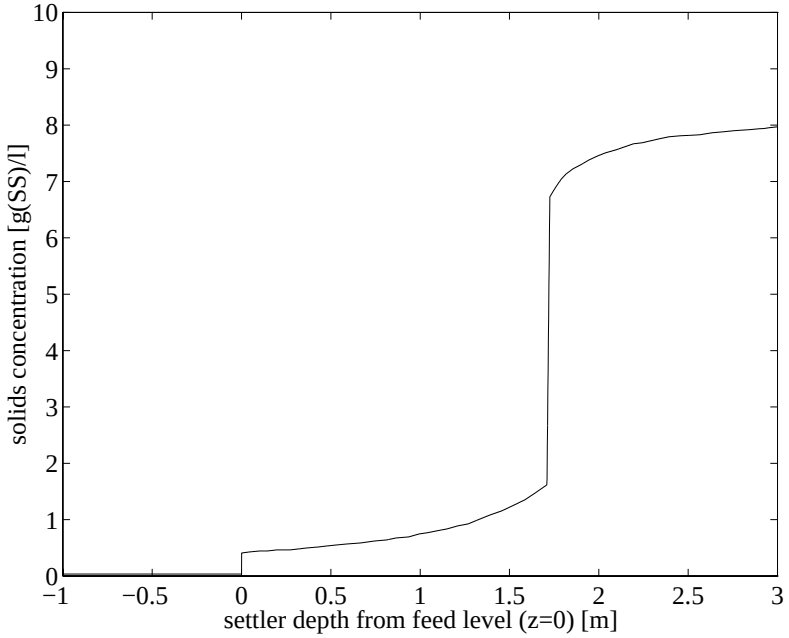


Figure 6.30 A steady-state concentration profile in a conical settler using the robust settler model, from (Diehl, 1996c). The simulated conditions are: $A(0) = 1250$ m^2, $A(D) = 100$ m^2, $H = 1$ m, $D = 3$ m, $Q_f = 1300$ m^3/h, $Q_u = 500$ m^3/h, $Q_e = 800$ m^3/h and $X_f = 3.08$ g SS/l.

A second improvement, which can be implemented, is to include some biological reactions in the settler. In principle, a complete IAWQ model (preferably including the production of nitrogen gas) can be used to model the biological behaviour within every layer of the settler. The inclusion of biological reactions in the settler model stresses the need for a good material propagation algorithm as discussed earlier in Sections 5.4 and 6.3. In particular, extensive denitrification in the settler has been reported in the literature (e.g., Siegrist *et al.*, 1995), which may also give rise to secondary problems, for example, rising sludge. These types of secondary effects are, however, much more difficult to model accurately.

Hydrodynamic phenomena, such as turbulence, short-ciruiting, density currents and horizontal flow conditions are obviously impossible to include correctly in a one-dimensional model. However, it is possible to include some of the *effects* of these phenomena in the model, for example, as suggested by Dupont and Dahl (1995). The approach was discussed in Section 5.3. On the other hand, this would mean imposing a number of *ad hoc* conditions on an otherwise analytically derived model, which is something that should be avoided if possible. Flocculation and compaction are other important processes that affect the behaviour of a real settler but are difficult to model in an accurate and fairly simple manner, without imposing new *ad hoc* and empirical conditions.

In order to improve the computational efficiency of the robust model, it is possible to allow for larger time steps in the numerical algorithm. The time step is currently defined by equation (5.49) and calculated when a dynamic simulation is initiated and then fixed to this value, based on a worst case scenario. Instead the time step could be adjusted on-line by an adaptive algorithm based on the actual conditions in the settler. In this way the computational effort would be reduced, especially when the process is not exposed to large dynamic disturbances.

Finally, the model must naturally be verified with real data. However, the purpose of this primary investigation was to investigate what a mathematical analysis of the settler model might reveal.

PART IV

Modelling the Biofilm Process

Chapter 7

Modelling Principles – a Review

In this chapter we describe the basic principles of biofilm processes. Various process configurations are introduced together with a description of the fundamental process mechanisms. The influence of higher order organisms on biofilm systems is discussed. A short review of recent progress within the field of biofilm modelling is also provided. The principles applied in a state-of-the-art biofilm model are presented in greater detail (especially the transport mechanisms) together with a numerical algorithm for solving the resulting system of stiff non-linear partial differential equations. Parts of the material in this chapter are covered in [187].

7.1 Fundamentals

A thin layer of microorganisms attached to a solid surface is called a biofilm. Biofilms can develop on almost any kind of surface exposed to an aqueous environment. In wastewater treatment they are used to eliminate and oxidize organic and inorganic components from the wastewater. The basic feature of the biofilm reactor is the heterogeneous nature of the processes. The reactor involves a solid medium to which the bacteria are fixed as a matrix, the surface of which is exposed to water passing through the reactor. The essential feature of this configuration is the need for the substrates and the resulting products to diffuse through the biofilm. This purely physical phenomenon has turned out to be crucial to the understanding of the performance of water purification in biofilm reactors. The biological mechanisms occurring within the biofilm are quite similar to the ones described for the activated sludge process in Chapter 3.

Process Configurations

Biofilm techniques are generally used in small sewage works, serving populations of less than 20000 (Halling-Sørensen and Jørgensen, 1993). They tend to be higher in capital costs but lower in running costs than activated sludge plants.

Many different types of reactors have been developed over the years (see e.g., Arvin and Harremoës, 1990). The oldest is the traditional filter, which was employed already before the turn of the century – initially as a screening device, but it was soon realized that the dominating mechanism was biological degradation more than simple screening. That led to the development of the trickling filter, where the wastewater flows by gravity as a free surface stream over a porous medium (today usually a plastic material) packed in a reactor. The rotating biological contactor process can be dated back to the turn of the century. In this process the rotating discs, covered with biofilm, are partially submerged in wastewater. More recent process types include submerged filters with either upflow or downflow and the so called fluidized beds, where spherical particles coated with a biofilm are fluidized by upflowing wastewater. The sizes of the particles used in fluidized beds typically range from 0.1 to 1 mm and the carrier material can be sand, glass beads, coal, activated carbon, etc. The different process configurations are schematically described in Figure 7.1.

The main advantage of the fixed film processes is that high volumetric densities of microorganisms can be accumulated by natural attachment as biofilms. This high density of biomass accumulation allows excellent treatment performance in fairly small reactor volumes, which is economically beneficial. Other important factors are that – for most processes – the need for sedimentation is very limited, normally there is no need for sludge recirculation (Henze *et al.*, 1992) and sometimes there is no need for recirculation of the effluent (Rittmann, 1989a). Moreover, the traditional trickling filter process is practically self-controlled and may in some cases also have an ability to survive shock loads of toxic wastes due to the relatively short retention time of the wastewater in the reactor, which means that only organisms on the surface of the film will be affected (Grady and Lim, 1980). This only holds if the duration of the toxic shock load is very limited. However, the limited flexibility and controllability of the process may also be regarded as a disadvantage. The more recent process configuration, for example fluidized beds, are more sensitive to short-term changes in the influent concentrations due to the extremely short hydraulic retention times (a few minutes). Therefore, effluent recircu-

lation is usually practised to control the flow velocity, the degree of bed expansion and the inlet substrate concentration (Rittmann, 1989a).

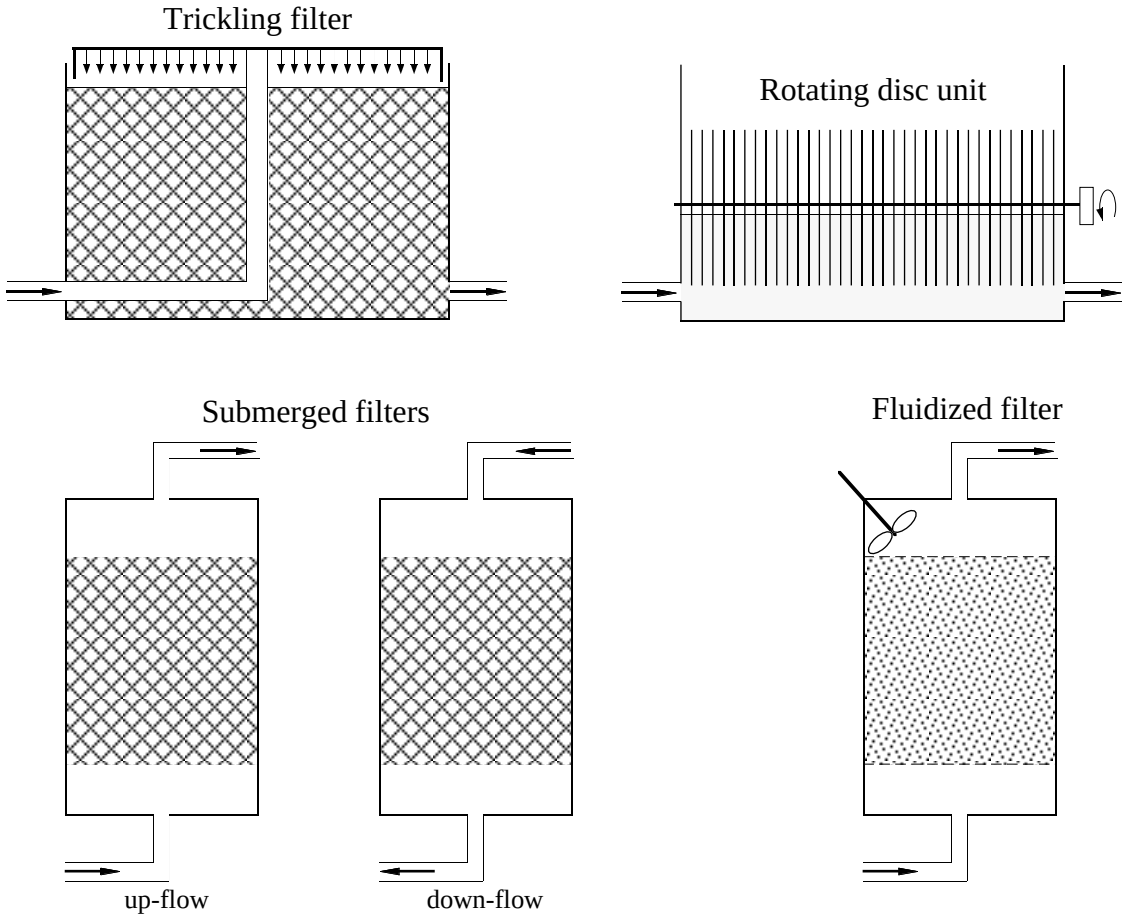


Figure 7.1 Biofilm reactors used in wastewater treatment.

Applications for most biofilm processes extend far beyond the traditional aerobic treatment of sewage and industrial wastewaters. Some of the most interesting applications include methanogenic treatment of wastewaters, nitrification and denitrification of many wastewaters and drinking waters and detoxification of waters containing hazardous organic chemicals. The excellent biomass retention and relatively short hydraulic detention times of biofilm processes make them attractive when bacterial growth rates are slow or when the compounds are inhibitory or slowly degraded.

Process Mechanisms

Typical biofilms are only a fraction of a millimeter thick. Over this short distance the physical and chemical conditions in the biofilm may drastically change, for example, from aerobic to anoxic conditions. These

changes lead to the formation of microenvironments, which may house completely different types of microbial species. Different zones may develop as a function of the loading of substrate to the biofilm. According to Kinner (1983) the most varied biofilm induced by a heavily loaded wastewater can have four dramatically different layers, composed as follows:

- an outer layer with heterotrophic oxidation of organic carbons, nitrification, denitrification and sulphide oxidation;
- a second microaerophilic layer with denitrification and fermentation;
- an anaerobic layer with sulphate respiration and fermentation;
- an anaerobic layer adjacent to the support material with methanogenesis and fermentation.

If the wastewater becomes less heavily loaded, or possibly acquires a different composition, the biofilm will be built up of the two upper layers or consist of the top layer only.

The behaviour of a biofilm is determined by a variety of biological, chemical and physical processes internal to the biofilm as well as by interactions between the biofilm and its environment. The biofilm and its environment form a very complex system, which is often difficult to analyse experimentally due to its heterogeneity and small dimensions. Thus, mathematical models represent important tools in biofilm research and applications. Models aid the researcher to state and test hypotheses, as well as represent and interpret data. In practical applications, models provide means for prediction of biofilm behaviour and for failure analysis.

The major difference when modelling biofilm processes compared with suspended-growth processes (e.g., the AS process) is the necessity to include equations describing some of the material transport processes in the biofilm and at the interface between the film and the liquid phase, e.g., molecular diffusion, turbulent diffusion, advection, attachment and detachment. The biological reactions within the biofilm can on the other hand be modelled in a similar manner as in a suspended-growth bioreactor, thoroughly described in Chapter 3, although concentrations may vary significantly as a function of the biofilm depth. Therefore, biofilm models are generally more complex and contain a very large number of model parameters compared with activated sludge models, which often only deal with the biological reaction mechanisms.

Molecular diffusion is driven by a potential energy gradient related to Brownian motion (Gujer and Wanner, 1989). Within the biofilm, two forms of potential energy are of prime importance: (1) the chemical potential and (2) an electrical potential for electrically charged particles. The molecular diffusion is governed by three forces:

- a concentration gradient expressed in the form of Fick's first law of diffusion;
- a gradient of the so called activity coefficient, which depends on the local chemical environment, usually considered negligible in biofilms;
- a gradient in the electrostatic potential, which may be due to electrical interactions between charges in the solution or between charges in solution and charges fixed to the solids matrix.

Moreover, the diffusion is usually considered to be affected by a stagnant liquid layer adjacent to the biofilm, through which the soluble matter must diffuse before reaching the actual biofilm. A method for determining the thickness of this boundary diffusion layer is described by Bouwer and McCarty (1985). An example of a typical oxygen concentration profile in a biofilm is shown in Figure 7.2.

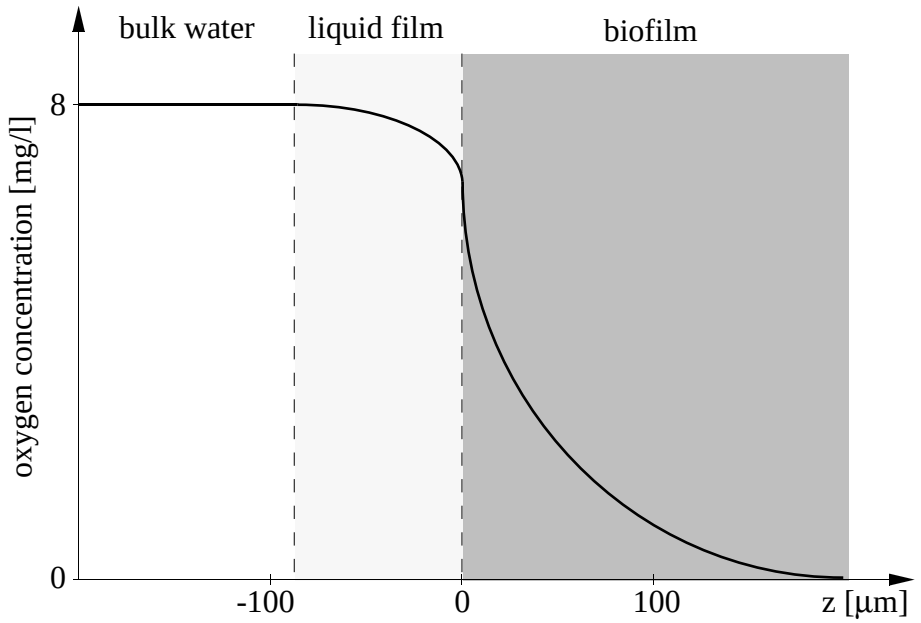


Figure 7.2 Example of an oxygen concentration profile in a biofilm including boundary layer diffusion.

Turbulent (or eddy) diffusion is driven by eddies of different scale and results in a net transport of matter in the direction of the negative concentration gradient. If the scale of the eddies are considerably smaller than the

scale of the biofilm system over which these gradients are of interest, turbulent diffusion is described in analogy to Fick's first law.

All solid biofilm phases are subject to a common advection process. If any one attached particle moves, this causes displacement of its neighbouring particles. Therefore, the advection velocity is identical for all solid materials due to their production (e.g., biomass growth) or of a volume change of the liquid phase. This process may cause local change in biomass composition.

The so called interfacial transfer processes are those specific for the interface between the two bulk compartments in a biofilm system; bulk water and biofilm matrix. These processes include attachment (deposition of particles on the biofilm) and detachment, which may be separated into four different mechanisms: erosion, grazing, abrasion and sloughing (Rittmann, 1989b). They are physical processes, but still no generally accepted rate expressions are available (Gujer and Wanner, 1989). According to Arvin and Harremoës (1990), the removal of particulate matter from the bulk water phase depends on aspects such as the

- size and the charge of the particulate matter;
- size, shape and chemical composition of the support media;
- surface configuration of the biofilm;
- hydraulic flow regime.

Bouwer (1987) suggests an attachment model based on Stoke's law to predict the flux of particulate material towards the surface of the film and a sticking efficiency parameter to describe the fraction of collisions between suspended particles and the biofilm interface that result in particle attachment. Erosion has been studied by Trulear and Characklis (1982), who indicate that nutrient supply, biofilm thickness and hydraulic shear are important factors. Furthermore, the shear stress history of the biofilm apparently influences the rate and extent of erosion. Grazing results from protozoa and other higher order organisms feeding on the outer surface of the film and abrasion is caused by the collision or rubbing together of particles (e.g., in fluidized beds), some of which are covered with biofilm. The cause of biofilm sloughing, that is, the periodic loss of large patches of biofilm, is also poorly understood. Harremoës *et al.* (1980) report that sloughing is due to the development of nitrogen gas bubbles in the lower layers of a denitrifying biofilm. Other researchers speculate that sloughing occurs as a consequence of nutrient deficiency in the depth of the biofilm. No single mechanism for this process has been accepted (Gujer and

Wanner, 1989). Some of the most essential phenomena occurring in a biofilm process are illustrated in Figure 7.3.

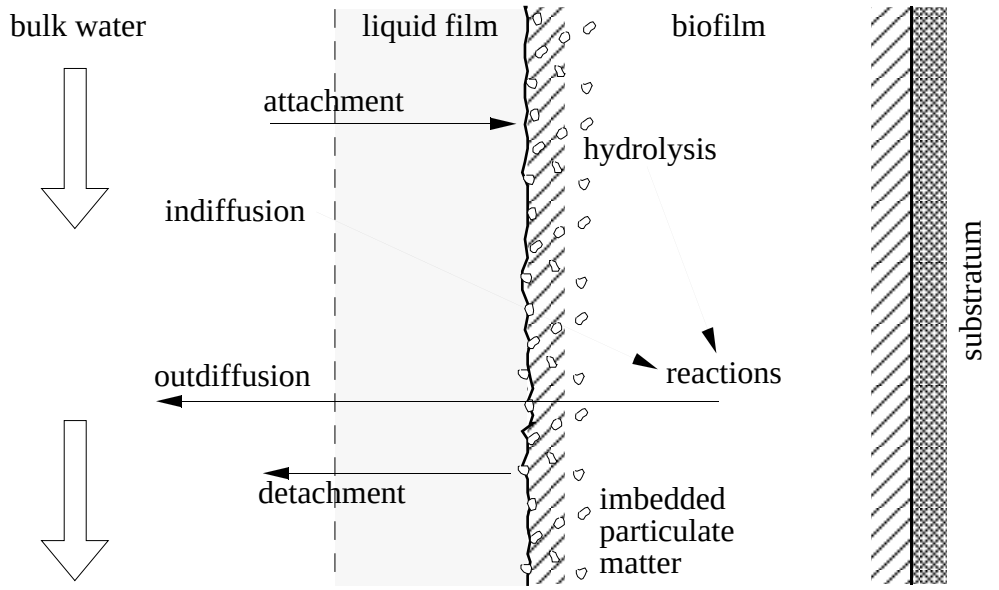


Figure 7.3 Illustration of some essential phenomena in a biofilm process.

The density and thickness of the biofilm are other important factors when modelling attached-growth systems. There is generally very little even qualitative information available on the mechanisms affecting the density and internal structure of the biofilm (Arvin and Harremoës, 1990). Therefore, it is common to assume a homogeneous biofilm with a constant density. Arvin and Harremoës also report that the thickness of the biofilm is primarily controlled by the following factors:

- growth of active biomass as a result of influx of substrate;
- decay of active biomass;
- accumulation of inert organic material from the decay of active bacteria;
- accumulation of polymers from the metabolism of the substrate;
- attachment of suspended particles from the bulk liquid;
- erosion of small particles from the surface of the biofilm;
- sloughing of large patches of the biofilm.

The influence of higher order organism through grazing will also affect the thickness of the film although the mechanisms are not well understood. Due to the large number of involved processes the ability to predict the thickness of a biofilm is relatively poor (Halling-Sørensen and Jørgensen, 1993).

Most theories for substrate removal in biofilm reactors suffer from the fact that they deal with soluble substrate alone. As most wastewaters contain a significant fraction of non-diffusible matter, this is a major drawback. The difficulties of modelling the attachment process was discussed above, but also the degradation mechanisms for the attached particulate matter is not known in detail (Arvin and Harremoës, 1990). There are two primary hypotheses for the hydrolysis process affecting attached particulate matter.

1) The degradation takes place by hydrolysis through the action of extra-cellular enzymes, either on the membranes of the bacteria (Banerji *et al.*, 1968) or released to the water (Larsen and Harremoës, 1994). This may lead to the release of soluble products to the liquid phase, which is then transported into the biofilm by diffusion.

2) The hydrolysing enzymes are attached to the surface of the bacteria and the products are absorbed directly by the bacteria as the hydrolysis takes place (Takahashi *et al.*, 1969; Sprouse and Rittmann, 1990).

In AS systems, the hydrolysis mechanism is often found to be the rate limiting process of the system and it is an important feature of the IAWQ model. Consequently, it is not unrealistic to assume that the fundamental function of the hydrolysis process is equally important in biofilm systems as well. The fate of the particulate organic material will also influence how to model the decay mechanism of microorganisms within the biofilm.

Higher Order Organisms

The complexity of the biofilm process was clearly indicated in the previous subsection. However, one important factor, which was not discussed in detail, deals with the influence of predators on the behaviour of the biofilm process. Protozoa and microscopic metazoa – in this work, for simplicity, defined as microfauna – as well as fungi and algae are normally found in large numbers in biofilm processes (Bishop and Kinner, 1986). Although their specific role when describing the mechanisms of the biofilm is usually neglected, it is often suggested that they strongly affect the overall process result (Stahl *et al.*, 1989; Arvin and Harremoës, 1990). High grazing pressure may increase the turnover rate of biomass, affect nutrient recycling and growth rate, significantly shorten the cell retention time, and influence the structure of the biofilm. Moreover, the grazing may be either selective (i.e., only affect certain types of organisms) or non-selective. Curds (1992) gives an excellent description of the higher order organisms found in wastewater treatment plants and their role in the processes. In this subsection, a short review of Curds (1992) is provided.

The most common group of higher organisms found in wastewater treatment plants are probably protozoa. They are normally single-celled eucaryotic (i.e., have a nucleus bounded by a nuclear envelope and are elaborately differentiated by a series of membrane systems) organisms. Protozoa are capable of reproduction, feeding, movement, excretion and respiration. The majority of protozoa are within the size range of 5 and 250 μm in diameter. At present more than 65000 protozoan species have been described and they are found in all moist habitats.

Protozoa are more abundant in aerobic sewage-treatment processes than in anaerobic processes. For example, it is common to find numbers in the order of 50000 cells per millilitre in the mixed liquor of an AS plant. Calculations based on such numbers indicate that protozoa can constitute approximately 5 % of the dry weight of the suspended solids in the aeration tank. According to Curds (1992) the most important protozoa found in both biofilm and AS processes are ciliates. Most of them are sedentary, attaching themselves directly to the microbial films by means of a stalk. A few are crawling forms and all the ciliates are known to feed on bacteria. Some of the most commonly found ciliates in sewage-treatment processes are shown in Figure 7.4.

In processes where the microbial film is held in a static position with respect to the flow of liquid (e.g., trickling filters and rotating biological contactors) different organisms thrive in different positions. For example, as sewage passes through the depth of a biological filter it is purified by microbial actions so that conditions at the top are very different from those at the bottom. This leads to different protozoan populations at the top of the filter compared with those below. This vertical stratification of organisms in a filter depends on many interrelated factors which are extremely difficult to unravel. However, generally there is a change from those species in the surface layers which utilize soluble organic substrates, to those in the middle which feed on bacteria, to those at the bottom which feed upon ciliates.

In an activated sludge plant the situation is different. Both the sewage and the microbial film or sludge floc flow down the aeration tank together and the organisms live in an ever-changing environment. After a period of low dissolved oxygen concentration in the settling tank, they are recirculated to be mixed with raw influent wastewater in the aeration tank. If a fixed medium is introduced into the aeration tank at various intervals along its length different protozoa colonise the medium and some of them are different to those growing upon the sludge floc.

Note! No figure available.

Figure 7.4 Some ciliates commonly found in wastewater treatment processes, from (Curds, 1992). A: *Trachelophyllum pusillum*; B: *Chilodonella uncinata*; C: *Cinetochilum margaritaceum*; D: *Aspidisca cicada* (vental and rear views); E: *Euplotes affinis*; F: *Vorticella convallaria* (extended and contracted); G: *Carchesium polypinum*; H: *Opercularia coarctata.*

While there is a significant amount of information concerning the role of protozoa in the AS process, little work has been carried out on the role of these organisms in biofilm systems. However, as there is a great deal of similarity between these aerobic processes it is reasonable to assume that the protozoa play a similar role in each. To the author's knowledge no studies have been made on anaerobic processes.

Protozoa were originally thought to be harmful to the AS process. It was argued that, as they feed on the bacterial populations, they inhibit microbial degradation processes. This has been shown to be untrue. Today, most

authorities agree that protozoa play a secondary but nonetheless important role in wastewater treatment processes. During protozoa-free conditions, AS plants produce highly turbid effluents of inferior quality and the turbidity is significantly related to the presence of very large numbers of bacteria suspended in the effluent (hundreds of millions per millilitre) and, consequently, the suspended solids concentration in the effluent is very high. By adding cultures of ciliated protozoa a dramatic improvement in the effluent quality can be detected within a few days (when the protozoan populations become properly established).

There are at least two ways in which protozoa might cause the improvement in effluent quality; either by flocculation or by predation. There is a considerable amount of published evidence showing that pure cultures of protozoa can flocculate suspended particulate matter and bacteria. In many species flocculation is thought to be brought about by the secretion of certain substances by the organisms. It is evident that the vast majority of ciliates found in AS processes can feed upon a whole variety of bacteria likely to be present in the process. Studies using batch and continuous-culture methods have supplied sufficient data to be able to assess the quantities of bacteria likely to be removed by protozoa. It seems that protozoa could, by predation alone, easily account for the removal of dispersed bacteria in the activated sludge of the experiments. It is therefore likely that the major role of the ciliated protozoa in aerobic wastewater treatment processes is the removal of dispersed growths of bacteria by predation and that protozoa-induced flocculation is not of any real importance. Flagellated protozoa and amoebae also feed upon bacteria and they play a similar role. Furthermore, the amoebae may also have the ability to ingest flocculated bacteria, which would have the effect of reducing sludge production.

Obviously, protozoa play a favourable role in the AS process. However, it is questionable how much of the above description is applicable to biofilm systems. In the next chapter, the results of an experimental investigation on the influence of higher order organisms on the behaviour of an aerobic, nitrifying biofilm system is presented. These results indicate that the higher order organisms may also have a significant negative effect on the overall performance of the biofilm process.

Model Development

The modelling of biofilm processes has been an on-going and advancing enterprise since the early 1970s. Early work (e.g., Vaughan *et al.*, 1973; Atkinson and Davies, 1974; Williamson and McCarty, 1976, Harremoës, 1976; La Motta, 1976; Harremoës, 1978) focused on the most fundamental phenomena of soluble substrate utilization and transport into and within the biofilm (mainly due to molecular diffusion). These investigations have shown that the removal of soluble substances is governed by diffusion resistance to the movement of the substrate into the biofilm, before degradation in the interior of the biofilm. The models were either simple steady-state models or a dynamic model was used to describe the soluble components whereas a steady-state model was used to describe the biomass. The rate of reaction in a biofilm was based on the concept of limiting substrate. If the wastewater is aerobic, the limiting substrate will consist of oxygen, organic carbon or ammonia. The intrinsic reaction rate of a limiting substrate is described, depending on the authors, as a Monod-type, first or zero-order equation (see Section 3.2). Subsequent work (Rittmann and McCarty, 1980; Rittmann, 1982) addressed the growth and loss of the biofilm and the establishment of a steady-state biofilm, or one in which biomass growth just balanced the losses to yield a biofilm with no net change of biomass over time.

In the mid 1980s, the first models appeared, which allowed for transient-state modelling of biofilms (Rittmann and Brunner, 1984; Chang and Rittmann, 1987a; Chang and Rittmann, 1987b). At the same time the first dynamic models to describe the distribution of different competing or complementary bacterial species were presented (Kissel *et al.*, 1984; Wanner and Gujer, 1984; Wanner and Gujer, 1986).

The model of Wanner and Gujer is based on the continuum approach, that is, a particulate component is not characterized by the shape, size and location of its cells in the biofilm, but by quantities which are spatial averages over a small biofilm volume element. It is also a one-dimensional description, i.e., only the space coordinate perpendicular to the substratum is considered. These two concepts are still the basis of practically all biofilm models used today (Wanner, 1996). During the last decade the model has been further refined and is today the most generally accepted mathematical model for describing the behaviour of the biofilm process. It is further discussed in Section 7.2.

Until recently, it was believed that all biofilms consist of a continuous, gel-like matrix in which microbial cells are embedded at random. New sophisticated experimental observations (Costerton *et al.*, 1994) have now revealed that biofilms exist in which there are clusters of microbial cells and pore channels of the size of a hundred micrometers, which form more or less independently. A hypothetical biofilm structure is shown in Figure 7.5. Modelling this type of structure may require two or three-dimensional models though this is not feasible today, as the mechanisms are not known.

Note! No figure available.

Figure 7.5 Structure of a hypothetical biofilm drawn from a large number of Confocal Scanning Laser Microscopy (CSLM) examinations of different biofilms (Lewandowski *et al.*, 1995). The arrows indicate convective flow within the water channels.

The effects of the biofilm structure and porosity on the microbial distribution, mass transport and biodegradation are today some of the most interesting fields of biofilm research (e.g., Fu *et al.*, 1994; Zhang and Bishop, 1994a; Zhang and Bishop, 1994b; Bishop *et al.*, 1995; De Beer and Stoodley, 1996; Lewandowski and Stoodley, 1996). New methods for investigating the biofilm structure have also been proposed (e.g., Gibbs and Bishop, 1996) and even methods based on fractal analysis are suggested (Hermanowicz *et al.*, 1996). As more knowledge and data about the biofilm structure become available, it may be possible to include these new concepts in a general biofilm model.

The importance of hydraulic phenomena on the biofilm process is also acknowledged. New methods are being developed to investigate the hydrodynamic conditions at the biofilm-water interface (e.g., Schindler *et al.*, 1995). Some initial attempts have also been made to couple biofilm models to models describing detailed microscopic hydrodynamic pheno-

mena in two dimensions (Chen *et al.*, 1994; Cunningham *et al.*, 1996). These types of models require extremely powerful computers for performing simulations and the approach is definitely not yet applicable for general biofilm modelling.

The influence of higher order organisms on the behaviour of biofilm systems is practically always neglected in available mathematical models. Several attempts have been made to include protozoa into biological models, but only in models describing suspended-growth systems (e.g., Curds, 1973; Curds, 1992; Ratsak *et al.*, 1993; Ratsak *et al.*, 1994). It is also realized that problems may occur when modelling 'predator-prey' systems, e.g., models of Volterra, Leslie-Gower, Holling-Tanner and Lotka-Volterra, which often lead to chaotic solutions. Such systems can be studied in, e.g., Renshaw (1993) and Sabin and Summers (1993).

7.2 A State-of-the-Art Biofilm Model

The model we will discuss in this section is based on the formulation by Wanner and Gujer (1984; 1986) and further extended in Wanner and Reichert (1996). The model is included in the simulation program AQUASIM (Reichert, 1994a; Reichert, 1994b; Reichert and Ruchti, 1994; Reichert *et al.*, 1995) developed at the Swiss Federal Institute for Environmental Science and Technology (EAWAG). AQUASIM is a general simulation and data analysis tool for laboratory, technical and natural aquatic systems (the current version contains models for mixed reactors, biofilm reactors and river sections), although we will only discuss it in connection with biofilm modelling (see also Appendix F). The program allows the user to define an arbitrary number of substances to be modelled and it is extremely flexible in the formulation of transformation processes. It not only offers the possibility of performing simulations but it also provides methods for system identification (sensitivity analysis and automatic parameter estimation) and for estimation of the uncertainty of calculated results. In this section we will primarily discuss the major modifications of the original model by Wanner and Gujer (1986).

Model Concepts

The biofilm model is basically a 'mixed' model, which includes a mechanistic core (e.g., mass balance equations) and empirical expressions (e.g., mathematical functions to describe attachment and detachment). The extensions of the original biofilm model is primarily due to several recent experimental findings (Wanner, 1996):

- transport of dissolved components in the biofilm is not always due to molecular diffusion only;
- transport of particulate components cannot be exclusively related to the net growth rates of the microbial species in the biofilm;
- the liquid phase volume fraction (porosity) of the biofilm is not constant;
- simultaneous detachment and attachment of cells and particles at the biofilm surface is an essential process.

As a consequence of applying the continuum concept, information about the biofilm structure at a micrometer scale is lost. However, this is not a problem unless the objective of the modelling is to investigate biofilm processes at the scale of individual cells, in which case the continuum approach is not applicable. In a similar way, some features of the biofilm structure are lost due to fact that the model only considers one spatial dimension, perpendicular to the substratum.

Model Structure

The biofilm model is based on the conservation law, written as a partial differential equation (cf. Sections 5.2 and 5.4)

$$\frac{\partial \hat{\rho}(z,t)}{\partial t} + \frac{\partial \hat{\boldsymbol{J}}(z,t)}{\partial z} = \hat{\boldsymbol{r}}(z,t) \tag{7.1}$$

where $\hat{\boldsymbol{\rho}}$ is an array of one-dimensional densities describing the quantities of the properties per unit length, $\hat{\boldsymbol{J}}$ is an array of one-dimensional fluxes describing the quantities of the properties transported per unit time relative to the resting frame of reference and $\hat{\boldsymbol{r}}$ are defined as the production rates of the properties per unit time and unit length of the system. The symbol '^' is used to distinguish one-dimensional variables from variables with traditional dimensions. The space coordinate z, defined perpendicular to the substratum, is calculated positively from the substratum and $z = 0$ at the

film-substratum interface (see Figure 7.3). Note that the mathematical model in AQUASIM is actually based on (7.1) written in the integral form (Reichert, 1994b), cf. equation (5.2).

The dynamic state variables are naturally divided into various particulate (X_j) and dissolved substances (S_j). They are represented by concentrations averaged over planes parallel to the substratum. For X_j, the concentration is defined as mass per unit of total volume (including both water and particles) and for S_j as mass per unit water phase volume (excluding the volume of the particles). If ρ_j denotes the density of particles of type j (as mass per unit of particle volume, defined to be constant over time and space in AQUASIM), then the volume fraction of the water phase (ε_l) of the total reactor volume is given by

$$\varepsilon_l = 1 - \sum_{k=1}^{n_X} \frac{X_k}{\rho_k} \tag{7.2}$$

where n_X is the number of modelled particulate variables. Note that the soluble concentrations per unit of total volume instead of concentration per unit water phase volume, are easily calculated by multiplying ε_l by S_j.

While the particulate substances are suspended in the bulk water volume, they form the solid matrix of the biofilm. Therefore, growth of particles in the depth of a biofilm leads to an advective displacement of the biofilm, if the additional particulate volume is not increased only at the expense of the volume of the water phase in the biofilm. The displacement velocity, v_F, at the location z is given as the total volume production between the subtratum-biofilm interface at $z = 0$ and the position z

$$v_F = \frac{1}{A} \int_0^z \left(\sum_{k=1}^{n_X} \frac{r_{X,k}}{\rho_k} + r_{\varepsilon_{l,F}} \right) A dz \tag{7.3}$$

where A is the surface area of the biofilm and $r_{X,j}$ is the net production rate of the particulate substance of type j.The first term in the integrand is the volume production due to growth of particulate substances and the second term is the volume production of free water volume between the particles (subscript *'F'* implies within the biofilm). It is a common assumption that the volume fraction of the water phase within the film remains constant. In this case, free water volume production is proportional to particulate volume production. The volume production term is therefore described as a sum of this generic term and an excess rate

$$r_{\varepsilon_{l,F}} = \frac{\varepsilon_{l,F}}{1-\varepsilon_{l,F}} \sum_{k=1}^{n_X} \frac{r_{X,k}}{\rho_k} + r'_{\varepsilon_{l,F}} \tag{7.4}$$

If the last term is zero then the volume fraction of the water phase in the biofilm is constant, as in the original model (Wanner and Gujer, 1986). On the other hand, $r'_{\varepsilon_{l,F}}$ may now be defined as a mathematical function of time, distance from the substratum, growth rate or any other variable in the model and can consequently be used to model the fact that the water volume fraction of the film may not be constant. This is one of the four important improvements of the modified biofilm model, listed at the beginning of this section. By combining (7.3) and (7.4), the advective velocity of the biofilm solid matrix can be rewritten as

$$v_F = \frac{1}{A} \int_0^z \left(\frac{1}{1-\varepsilon_{l,F}} \sum_{k=1}^{n_X} \frac{r_{X,k}}{\rho_k} + r'_{\varepsilon_{l,F}} \right) A dz \tag{7.5}$$

The flux due to diffusion of soluble material is modelled according to Fick's first law, i.e.,

$$J_{diff,S,j} = -\varepsilon_{l,F} D_{S,j} \frac{\partial S_j}{\partial z} \tag{7.6}$$

where $D_{S,j}$ is the diffusion coefficient for the soluble substance of type *j*. However, to describe some of the recently found phenomena where the molecular diffusion appears to be influenced by other processes as well (e.g., tortuosity and turbulence into the film), the diffusivity (D_S) is regarded as an *effective* diffusivity. It may be modelled as a function of biofilm density, thickness, depth, time or any other model variable, instead of describing D_S as being directly proportional to the molecular diffusivity in pure water in the traditional manner. This is the second important modification of the new biofilm model. In a similar way the boundary layer resistance, which makes it possible to limit the mass transfer between the film and the bulk water phase, is included in the model.

In order to account for transport of cells and particles in the direction opposite to that of the displacement velocity v_F, the original model is extended by an additional transport process, which is independent of microbial growth and is described as an effective diffusion process of particulate substances according to

$$J_{diff,X,j} = -D_{X,j} \frac{\partial X_j}{\partial z} \tag{7.7}$$

Note that $D_{X,j}$ is *not* the molecular diffusion of particles in water but a way of modelling changes of the biofilm due to detachment and attachment of particles within the film. In AQUASIM this process is only active when a so called *diffusive* biofilm is modelled. When using a *rigid* biofilm, changes of biofilm structure at a given distance from the substratum is only due to advection caused by growth or decay in the layers below or at the actual location, as in the original model (Wanner and Gujer, 1986). The effective diffusion of particulate material is a purely empirical description and there are still experimental observations that remain unexplained by this model extension (Wanner, 1996).

The biological reaction mechanisms are in AQUASIM provided by the user, depending on the type of process which is to be modelled. In a normal case with normal influent wastewater, these mechanisms are often modelled in a similar fashion as suggested in the IAWQ model (e.g., Gujer and Boller, 1990). Throughout this section we only describe the transformations by $r_{S,j}$ and $r_{X,j}$, which are the net production rates for a specific soluble and particulate substance, respectively.

We can now define the three functions in (7.1) in order to derive a complete model of the biofilm process. The one-dimensional densities of the conservation law are given by

$$\hat{\rho} = \begin{bmatrix} X_{F,j} \\ \varepsilon_{l,F} S_{F,j} \\ \varepsilon_{l,F} \end{bmatrix} \cdot A \tag{7.8}$$

The fluxes corresponding to the above densities are

$$\hat{\boldsymbol{J}} = \begin{bmatrix} v_F X_{F,j} - D_{X,j} \dfrac{\partial X_{F,j}}{\partial z} \\ -\left(1-\varepsilon_{l,F}\right) v_F S_{F,j} - \varepsilon_{l,F} D_{S,j} \dfrac{\partial S_{F,j}}{\partial z} \\ \varepsilon_{l,F} v_F + \displaystyle\sum_{j=1}^{n_X} \frac{D_{X,j}}{\rho_j} \frac{\partial X_{F,j}}{\partial z} \end{bmatrix} \cdot A \tag{7.9}$$

The first element of (7.9) represents the flux of particulate substances within the biofilm. The first term of this element describes the advective motion of the biofilm matrix due to growth processes, cf. (7.5). The second term describes changes in a diffusive biofilm solids matrix by an effective diffusion process, see (7.7). The second element of (7.9) represents the flux of dissolved substances within the biofilm. The first term describes advective transport due to water flowing into the film to compensate for volume changes produced by growth processes. It can often be neglected due to the slowness of microbial growth compared with diffusive processes. The second term describes the diffusive flux of dissolved substances in the film, see (7.6). The last element in (7.9) represents the flux of free water volume (not the water flux) within the biofilm. The first term describes advective motion and the second describes volume changes due to the effective diffusion process of particulate substances. To complete the definition of the conservation law (7.1), the transformation rates are simply

$$\hat{\boldsymbol{r}} = \begin{bmatrix} r_{X,j} & r_{S,j} & r_{\varepsilon_{l,F}} \end{bmatrix}^T \cdot A \tag{7.10}$$

The actual transformation rates will be defined by the biological processes included in the model by the user. All these rates specify transformation of mass (or volume) per unit of total biofilm volume including the volume of the particles. Based on (7.8), (7.9) and (7.10), we can expand (7.1) and rewrite the equations by using (7.4) and (7.5) into the form given below. The conservation of the different particulate substances is given as

$$\begin{aligned} \frac{\partial X_{F,j}}{\partial t} = & -v_F \frac{\partial X_{F,j}}{\partial z} + \frac{1}{A}\frac{\partial}{\partial z}\left(A D_{X,j} \frac{\partial X_{F,j}}{\partial z} \right) \\ & - r'_{\varepsilon_{l,F}} X_{F,j} + \left(r_{X,j} - \frac{X_{F,j}}{1-\varepsilon_{l,F}} \sum_{k=1}^{n_X} \frac{r_{X,k}}{\rho_k} \right) \end{aligned} \tag{7.11}$$

the conservation of the different soluble substances as

$$\begin{aligned} \frac{\partial S_{F,j}}{\partial t} = & \frac{1-\varepsilon_{l,F}}{\varepsilon_{l,F}} v_F \frac{\partial S_{F,j}}{\partial z} + \frac{1}{\varepsilon_{l,F}} \sum_{k=1}^{n_X} \frac{r_{X,k}}{\rho_k} S_{F,j} \\ & + \frac{1}{\varepsilon_{l,F}} \frac{1}{A} \frac{\partial}{\partial z}\left(A \sum_{k=1}^{n_X} \frac{D_{X,k}}{\rho_k} \frac{\partial X_{F,k}}{\partial z} \right) \\ & + \frac{1}{\varepsilon_{l,F}} \frac{1}{A} \frac{\partial}{\partial z}\left(\varepsilon_{l,F} D_{S,j} \frac{\partial S_{F,j}}{\partial z} \right) + \frac{1}{\varepsilon_{l,F}} r_{S,j} \end{aligned} \tag{7.12}$$

and for the conservation of the free water volume as

$$\frac{\partial \varepsilon_{l,F}}{\partial t} = -v_F \frac{\partial \varepsilon_{l,F}}{\partial z} - \frac{1}{A}\frac{\partial}{\partial z}\left(A \sum_{k=1}^{n_X} \frac{D_{X,k}}{\rho_k} \frac{\partial X_{F,k}}{\partial z} \right) + \left(1 - \varepsilon_{l,F}\right) r'_{\varepsilon_{l,F}} \quad (7.13)$$

Equation (7.11) describes the time evolution of the concentrations of particulate substances at a fixed position in the biofilm. The first term on the right hand side is the change in concentration due to advective motion of the biofilm matrix, the second term corresponds to changes due to a possible effective diffusion process of particulate material in the biofilm, the third term models dilution of particles caused by growth of free water volume and the last term describes concentration changes due to growth processes of particulate components. Equation (7.12) describes the time evolution of the concentration of soluble substances at a fixed position in the biofilm. The first term corresponds to the advective flux of water transported into the biofilm to compensate for volume changes produced by growth. The second term describes growth of concentration due to growth of particulate species (movement of a cell membrane due to growth increases the concentration of a dissolved substance not consumed in the growth process, i.e., not crossing the membrane). The next two terms describe changes in concentration of dissolved substances due to water flow caused by effective solid matrix diffusion of particles and due to diffusion of dissolved substances in the water phase, respectively. The last term is due to transformation processes consuming or producing the dissolved substance under consideration. Equation (7.13) describes the changes of free volume within the biofilm solid matrix. The first term is due to advection of the solid matrix, the second term is due to effective diffusion of particles and the last term is due to a user-defined empirical expression of how changes occur in the free volume.

In order to completely define equations (7.11), (7.12) and (7.13), the geometry of the biofilm must be defined by specifying the surface area of the film. The area (A) is described as a function of the distance from the substratum, z, which means that many different geometries can be modelled (e.g., plane film, film on a cylinder, film inside a cylinder, film on spheres).

Time evolution of the biofilm thickness, L_F, is given by

$$\frac{dL_F}{dt} = v_L \quad (7.14)$$

where due to attachment and detachment processes, the interface velocity v_L is not exactly the same as $v_F(L_F)$. Instead the relationship between the two velocities are

$$v_L = v_F(L_F) - v_{de} + v_{at} \tag{7.15}$$

where v_{de} and v_{at} are the detachment and attachment velocities, respectively. In AQUASIM, the detachment velocity can be defined as a global velocity equal for all particulate components. In this case it may be a function of any model variable (e.g., biofilm thickness, time, growth velocity). The second possibility is to define individual detachment coefficients ($k_{de,j}$) for each variable X_j. In this case the detachment velocity is given by

$$v_{de} = \frac{1}{1-\varepsilon_{l,F}} \sum_{k=1}^{n_X} \frac{k_{de,k} X_{F,k}}{\rho_k} \tag{7.16}$$

The attachment velocity is modelled in a similar fashion as

$$v_{at} = \frac{1}{1-\varepsilon_{l,F}} \sum_{k=1}^{n_X} \frac{k_{at,k} X_{L,k}}{\rho_k} \tag{7.17}$$

where $k_{at,j}$ is the attachment coefficient for the variable X_j. The subscript *'L'* indicates that $X_{L,j}$ is the concentration of a specific particulate substance at the interface between the biofilm and the liquid boundary layer.

In order to successfully model simultaneous attachment and detachment, a diffusive biofilm must be assumed. Otherwise, the attached cells would form the outer layer of the biofilm and would consequently be removed by the detachment process, instead of allowing attached cells to migrate through the biofilm. This is achieved by the effective diffusion process for particulate material, see equation (7.7).

Furthermore, in AQUASIM the biological reactions in the bulk phase are simulated in parallel with the reactions within the biofilm, if the user requests it. This makes it easy to model situations where the hydraulic retention time is so large that reactions in the bulk phase cannot be neglected. The equations governing the dynamic behaviour of bulk water concentrations of particulate and dissolved substances outside the biofilm is modelled as a mixed reactor. The only complication is that the factor ε_l has to be introduced into the equations for dissolved variables to correct for the water phase volume. Also note that the model describing the biological

transformations within the biofilm may be completely different from the model describing the reactions in the bulk phase.

Numerical Algorithm

In order to numerically solve the stiff and non-linear system of partial differential equations (PDE), AQUASIM uses the *method of lines solution technique* (e.g., Schiesser, 1991), as described in Reichert (1994b). This means that the PDEs are converted into a set of ordinary differential equations (ODE) and algebraic equations (due to the boundary conditions) by discretization in space only. Based on this discretization, consistent initial conditions must be defined for the model (as close as possible to the conditions specified by the user). This is done by a modified Newton algorithm, where the global convergence behaviour is improved by a flexible step-reduction mechanism guaranteeing monotonic decrease of the error term of the equations, and also the calculation of the matrix of partial derivatives at each iteration is simplified by a special updating algorithm, compared with the original Newton algorithm.

The numerical flux terms at the boundaries are calculated either by a first-order approximation, which leads to smooth solutions due to numerical diffusion, or a second-order approximation, which must be combined with a flux-limiter method to avoid oscillations in the numerical solution close to sharp density changes or discontinuities (cf. the Godunov algorithm described in Section 5.4), see LeVeque (1990). The new set of ODEs and algebraic equations are then numerically solved by a modified Gear algorithm (Petzold, 1983). This algorithm is a variable-step, variable-order method with a special error control criterion suited for solving stiff systems. Moreover, the positions of the grid points are modified dynamically as the biofilm thickness varies, although the relative distances between the grid points are maintained. The algorithm could be modified to dynamically resolve areas of high curvature of the solutions, by shortening the distances between the grid points in such regions, thereby improving the numerical results (Reichert, 1994b). Finally, steady-state solutions for biofilm processes in AQUASIM are obtained by simulating the system forward in time with constant boundary conditions (relaxation).

The description of a state-of-the-art biofilm model given in this section is primarily a review of Reichert (1994b), where more details can be found. The complete set of necessary boundary conditions for numerically solving the system is also provided by Reichert.

Chapter 8

Modelling Microfauna Influence

In this chapter the results from an experimental study are presented. The effects of microfauna influence on the nitrification capacity of an aerobic suspended-carrier biofilm system were investigated. From the experimental results a few hypotheses on the role of microfauna are formulated. A number of different modelling approaches and model extensions are suggested, based on these hypotheses. The capability of the models to predict the influence of the higher order organisms is investigated, focusing on the steady-state behaviour of the models. The numerical results are validated against the experimental data and various problems are discussed. Parts of the material in this chapter are covered in [187].

8.1 Experimental Study

In this section, the results from an experimental study are presented. The purpose of the experiment was to determine the influence of higher order organisms on the behaviour of an aerobic, nitrifying biofilm in a suspended-carrier system. More details about the study can be found in Lee and Welander (1994).

Hybrid Systems

A process scheme tested at many municipal WWT plants during the late 1980s was based on a combination of a biofilm system and a suspended-sludge system (Schlegel, 1988; Middeldorf, 1989; Andersson, 1990; Bonhomme *et al.*, 1990). The reasons for combining the processes were to

- increase the reactor capacity;
- increase the biomass content in the system without an additional loading of the unit process;
- achieve a better and more stable nitrification in existing AS plants.

The principle was to immerse a biofilm carrier material into the aerobic reactor. Based on this the idea was that removal of organic material primarily was to be accomplished by the suspended biomass. This in turn would enable the slow-growing nitrifying bacteria to dominate the biofilm developing on the carrier material, thus securing a stable nitrification capacity. However, it could not be shown that the hybrid process had a higher nitrification capacity in comparison to a conventional AS system. Furthermore, the observed sludge yield in the hybrid process was significantly lower than usually found in an AS system. Microscopic examination of the biofilm (Andersson, 1990) revealed that several species of microfauna were present in large numbers. Although it could not be shown to what extent the microfauna had affected the process, it was suspected that the higher order organisms were responsible for both the limited nitrification capacity and the low sludge yield.

A reduced nitrifying capacity has also been ascribed to grazing organisms in other types of nitrifying biofilm processes, such as trickling filter reactors (Boller and Gujer, 1986; Parker *et al.*, 1989). On the other hand, protozoa and metazoa have generally been considered to play a beneficial role in aerobic treatment processes. In the AS process by making the effluent water clear by consuming dispersed bacteria and by improving flocculation (Curds, 1975; WPCF, 1990; Curds, 1992), and in trickling filters by preventing excessive growth of biofilms, as well as facilitating the recycling of nutrients and improving the transport of substrates (Gray, 1989), see Section 7.1. However, these studies dealt primarily with systems aimed at BOD removal and not with nitrifying systems. Theoretically, non-selective grazing should have a more negative effect on the slow-growing nitrifiers than on the fast-growing heterotrophs in a homogeneous biofilm due to the slower recovery of the nitrifying population. However, the interactions between predator and prey in a microbial population are extremely complex and the mechanisms are not very well known, as discussed in the previous section.

Nitrification in Aerobic Biofilms

The capacity of an aerobic biofilm process to nitrify is determined by the amount of nitrifying bacteria in the system and the activity of these bacteria. Thus, to increase nitrification, measures must be taken to increase either the nitrifying population, the activity of the nitrifiers, or both. The amount of nitrifiers can be increased either by increasing the total amount of biomass in the system or by increasing the fraction of nitrifying bacteria in this biomass. The above naturally also holds for nitrification in suspended sludge systems.

The total amount of biomass retained in a biofilm process depends largely on the reactor design and the hydraulic conditions in the process, while the fraction of nitrifiers in the biomass is known to be strongly affected by the biochemical oxygen demand (BOD) to nitrogen ratio in the influent wastewater (Barnes and Bliss, 1983). The activity of the nitrifiers can be affected by a large number of factors, such as temperature, pH, substrate limitation (oxygen or ammonium), and inhibitory substances. In addition to being influenced by these abiotic factors, the capacity of a biofilm process to nitrify may also be affected by biological factors, as exemplified by grazing of predators on the nitrifying bacteria in the biofilm. Although many factors may influence nitrification in biofilm systems, it is often unclear which are the main factors limiting the capacity of the processes. Therefore, in order to maximize the nitrification capacity in aerobic biofilm processes, it is important to identify factors limiting nitrification in such systems under different operating conditions. While much attention has been focused on the effects of abiotic parameters on nitrification, biological factors have been much less studied.

The experimental study discussed in this section was carried out to determine whether predators have a negative effect on nitrification in aerobic biofilm processes, and if so, to what extent nitrification can be enhanced by suppressing these organisms. The experiment was performed under controlled conditions in a laboratory model system, using a synthetic wastewater and employing inhibitors specific to eucaryotic cells, to make sure that no factors affecting nitrification other than the amount of microfauna would vary.

Experimental Set-Up

The laboratory model system consisted of two aerobic 800 ml continuous-flow suspended-carrier biofilm reactors, operating in parallel on the same sterilized synthetic wastewater based on acetate and peptone (300 mg/l total COD, 44 mg/l total N, 6 mg/l total P and all essential trace metals). The double-walled glass reactors were filled to 50 % (volume/volume) with polyethylene carrier particles (10 x 7 mm), possessing a surface area for microbial growth of 400 m^2/m^3, see Figure 8.1. The carrier particles were kept suspended in the reactors by aeration and mechanical stirring, ensuring good contact between the biofilm and the bulk water. The temperature in the reactors was maintained at 15 °C and the pH at 7, by automatic controllers. The bulk water phase was kept saturated with regard to dissolved oxygen (≈ 9 mg O_2/l).

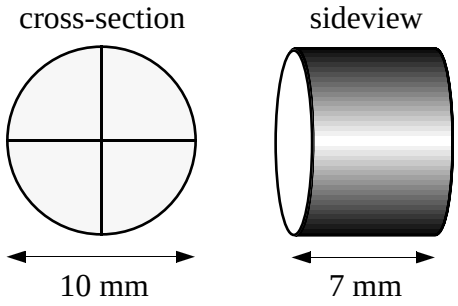


Figure 8.1 Schematic view of the carrier particles.

The experiment was carried out over a period of 450 days. Initially, the reactors were inoculated with activated sludge from the Sjölunda municipal treatment plant, Malmö, Sweden, and started up at a hydraulic retention time (HRT) of 7 hours. After 75 days, the HRT was shortened in small steps to 3 hours during a period of another 75 days. This was the HRT at which the inhibition study was carried out. After 50 days more, a good nitrification was established in both reactors and stable operating conditions were attained. In order to selectively inhibit the microfauna, nystatin and cycloheximide (8 mg of each added once every 8 h), substances specifically inhibitory to eucaryotic cells, were added to one of the reactors. The other reactor was operated as a reference system. After approximately 80 days, the addition of inhibitors was switched between the reactors (the previous reference reactor became the test reactor and vice versa) and the experiment was repeated for another 110 days. Then another switch of the addition of the inhibitors between the reactors was carried out for 60 days. At the end of the experiment (for a period of 14 days), the pH in both reactors was increased from 7 to 8.

Measurements and Analyses

The performance of the reactors was followed by daily analyses of ammonium, nitrite and nitrate concentrations in the bulk water phase according to standard methods (DIN, 1993). The COD concentrations – filtered and unfiltered samples – in the influent and effluent were measured regularly and analysed according to standard methods (APHA *et al.*, 1985a). Total suspended solids (TSS) concentration and volatile suspended solids concentration were determined in the reactor effluent, using methods described in APHA *et al.* (1985b). At the end of the experiment, the total amount of attached biomass in the reactors was estimated by determination of TSS and VSS on ten randomly picked carrier particles from each reactor. Moreover, in order to follow changes in the composition of the microbial populations, microscopical analyses of randomly picked carrier particles were carried out regularly. This made it possible to qualitatively study (very roughly) the fraction of protozoa and metazoa in relation to the fraction of bacteria in the biofilm.

Process Performance

The results from one of the reactors, for two consecutive periods (including the 50 day start-up period, i.e., period I-III for reactor B as defined below), are summarized in Table 8.1 as average effluent values and standard deviations. Note that the transient values during the first week, as the system is changed from a non-inhibited to an inhibited one, are excluded when calculating the average values for the inhibited case. Although not shown, the results from the second reactor are very similar.

duration [days]	NH_4-N [mg N/l]	NO_2-N [mg N/l]	NO_3-N [mg N/l]	filt. COD [mg COD/l]	TSS [mg COD/l]	inhibition
126	21.4±2.5	0.6±0.2	11.5±2.0	47±10	76±8	no
109	10.7±1.8	0.5±0.1	21.2±2.2	40±16	120±15	yes

Table 8.1 Summary of the experimental results during two consecutive periods (measurements in the bulk water phase). The HRT is 3 hours and the pH is 7 during these periods.

As the microfauna was inhibited, the nitrification rate and concentration of TSS increased significantly, whereas the biofilm mass decreased. The total dry mass of the biofilm at the end of the experiment was found to be 1.67 g (i.e., 2.09 g/l) in the reactor currently operating without addition of inhi-

bitors and 1.11 g (i.e., 1.39 g/l) in the inhibited reactor (multiply by 1.5 to roughly transform the values into g COD). Both within the biofilm and the reactor effluent biomass, the VSS was approximately 95 % of the TSS throughout the experiment. The first addition of inhibitors resulted in a rapid increase in the nitrate production rate from 4 to 7 mg N/(l h) and in a corresponding increase in the ammonium removal rate. When the inhibition was stopped the nitrification slowly decreased to a minimum of 2 mg N/(l h) within a month. During the rest of the non-inhibited period the nitrate production recovered slowly, rising to approximately 4 mg N/(l h). Figures 8.2 and 8.3 show the nitrate production rates and the ammonium removal rates for the two reactors during the full duration of the experiment. The different periods are defined as follows:

- period I: start-up period after the decrease of the HRT to 3 h;
- period II: addition of inhibitors to reactor A, reactor B used as reference;
- period III: addition of inhibitors to reactor B, reactor A used as reference;
- period IV: addition of inhibitors to reactor A, reactor B used as reference;
- period V: same as period IV, except that the pH was increased from 7 to 8 in both reactors.

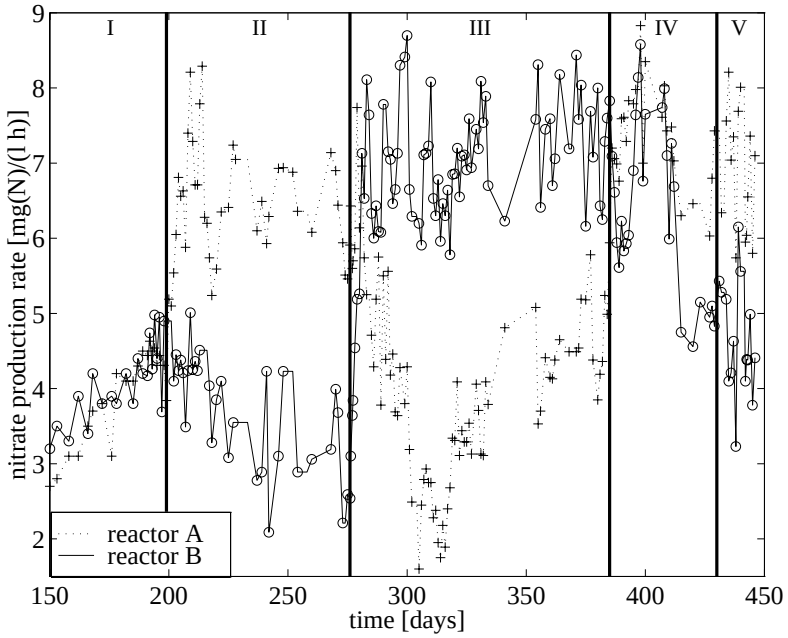


Figure 8.2 Measured nitrate production rates in the two reactors during the different test periods (I-V) of the experiment (Lee and Welander, 1994).

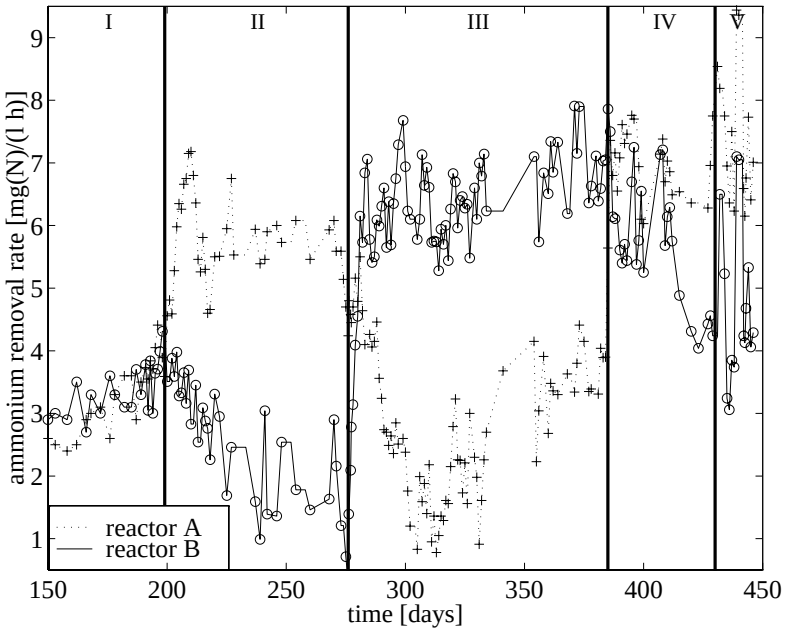


Figure 8.3 Measured ammonium removal rates in the two reactors during the different test periods (I-V) of the experiment (Lee and Welander, 1994).

The second switch of the addition of inhibitors between the two reactors (period IV) resulted in an increased nitrification in reactor A and a decreased nitrification in reactor B, although the response of reactor B was slower than the corresponding response of reactor A at the beginning of period III. The change of the pH from 7 to 8 at the end of the experiment did not influence the nitrification in either of the two reactors significantly. The nitrite production rate was low during the entire experiment and the removal of COD was not affected by the addition of inhibitors and remained fairly constant at approximately 85 % in both reactors throughout the experiment. The nitrification rate per carrier particle surface area, calculated from the nitrate production rate and the known surface area of the carriers, was on an average 0.46 g N/(m^2 d) and 0.85 g N/(m^2 d) for the non-inhibited and inhibited cases, respectively. These values are comparably low, which is mainly due to the low biomass concentration. The specific nitrification rate in the system operated with inhibition can be calculated to be 5.4 g N/(kg VSS h), which is higher than the specific nitrification rates generally found in trickling filters or activated sludge plants (Barnes and Bliss, 1983; Halling-Sørensen and Jørgensen, 1993). However, such a comparison is not relevant due to different characteristics of the wastewater, different operating conditions, etc.

Visual observations of the carrier particles in the two reactors gave the impression of the biofilm being significantly thicker in whichever reactor

was currently being operated without addition of inhibitors than in the current test reactor. This observation was in agreement with the measurements of total biofilm mass at the end of the experiment. Finally, as the inhibitors were added to a reactor, an initial temporary increase of up to 100 % could be detected in the effluent TSS.

The microscopical observations revealed a very diverse microflora/fauna in both reactors during the start-up period. The microbial population was composed of different kinds of adhering and polymer-forming bacteria, filamentous bacteria, dispersed bacteria, fungi, flagellates, ciliates, rotifers and nematodes. The most common higher order organisms were attached ciliates, rotifers and nematodes. The protozoa and metazoa together were roughly estimated to comprise approximately 30 % of the attached biomass. Addition of the specific eucaryotic inhibitors caused a significant change of the biofilm community. Certain species, such as rotifers and nematodes, appeared to be most strongly affected whereas the number of attached ciliates and flagellates were not reduced to the same extent. At the end of an inhibition period, the predators were estimated to make up about 5 % of the attached biomass. Moreover, the number of dense bacterial clusters in the biofilm, identified as agglomerates of nitrifying bacteria, increased considerably during the inhibition period. The colour of the biofilm also changed from dark brown to orange – a colour typical for many nitrifying bacteria (Watson *et al.*, 1989) – indicating an increase of the amount of nitrifiers in the biofilm. When the inhibition was stopped, a gradual recovery of the microfauna could be observed. The quantity of rotifers increased to their initial level, whereas the quantity of nematodes never reached the same level as prior to the addition of inhibitors.

The inhibitors could not be found to have any effect on the bacteria in the biofilm. It should also be noted that growth of biomass primarily took place on the inside walls of the carrier particles. Due to mechanical tear, no biofilm growth could be observed on the outer surface of the particles during the experiment.

Discussion

Protozoa and metazoa have generally been considered to influence the performance of aerobic treatment processes in a positive way. The studies described here indicate, however, that vigorous grazing of these organisms on the bacterial population can be harmful to bacterial transformations crucial to process performance. The positive response that nitrification

showed to the addition of inhibitors occurred rapidly and was accompanied by a rapid decrease in the quantity of predators, mainly the rotifers, whereas stopping the addition of inhibitors resulted in a slow recovery of the rotifer population, accompanied by a gradual decrease in nitrification capacity. This correlation suggests that the predators may affect the nitrification in biofilms in a negative way.

The nematodes, which are known to thrive within biofilms (e.g., Gray, 1989), never really recovered after the initial addition of inhibitors and were not abundant during the rest of the experiment. The flagellates and the attached ciliates, apparently less affected by the inhibitors, can probably only consume dispersed and loosely attached bacteria. Thus, the results indicate that the rotifers may play an important role in the biofilm turnover. This is of particular interest since the ecological niche, which the rotifers occupy in trickling filters, has not been adequately elucidated (Doohan, 1975; Gray, 1989). It should be emphasized that the difference in nitrification capacity between the reactors, cannot have been due to the occurrence of dispersed nitrifiers in the reactor to which inhibitors were added, since dispersed nitrifiers would not be able to grow fast enough to remain in the process at 3 hours HRT and 15 ˚C.

The dramatic increase in the effluent TSS during a transient period after the addition of inhibitors had been initiated was probably due to wash-out of predators, since a rapid decrease of the predator population in the biofilm was observed during the same period by means of microscopical observations. The difference in total biomass in the two reactors appeared comparable to the qualitatively estimated difference in predator biomass. Thus, the amount of attached bacterial biomass may have been approximately the same during the entire experiment. If this was the case then the increased nitrification capacity was not due to an increase of total biomass but to an increase in the fraction of nitrifiers in the biofilm. It should also be noted that the net biomass production on an average (TSS in the effluent) was 50 % higher for the system operated with addition of inhibitors than for the reference system (see Table 8.1). The lower sludge production demonstrates one positive effect of grazing predators in wastewater treatment processes.

The fact that predation limited nitrification but not the removal of readily biodegradable organic matter, indicates that the effects of grazing are more severe for slow-growing than for fast-growing bacteria. The latter are more readily able to keep up with the consumption of the active biomass and replace it with new. Theoretically, predation could lead to a complete loss

of nitrification in the biofilm process through a reduction in the actual mean cell residence time (MCRT) within the system to a value lower than that of the critical MCRT needed to maintain a population of nitrifying bacteria. The impact of selective grazing could further complicate this situation. The influence of predators on nitrification and indeed the entire biofilm behaviour can also be expected to be even more detrimental in, for example, trickling filters than in the type of biofilm process used in this study, since the trickling filter fauna includes larger metazoa, such as larvae, worms, flies and snails (e.g., Gray, 1989). Such organisms have a greater capacity for ingesting bacterial biofilm than the predators that dominated this study. In the discussed experiment, the establishment of larger metazoa in the biofilm was probably disfavoured by the use of submerged carrier particles in combination with the high turbulence in the reactors.

The ammonium removal rate was generally somewhat lower than the nitrate production rate (see Figures 8.2 and 8.3). This can be explained in terms of the conservation to nitrate not being the only ammonium transformation that occurred in the system. In addition to ammonium being consumed through nitrification, it was consumed through assimilation by heterotrophs and nitrifiers and was also produced by release from the organic nitrogen compounds of peptone and yeast extracts in the influent synthetic wastewater.

This study emphasized the importance of investigating the influence of biological – and not only abiotic – factors on nitrification. The two-fold increase in nitrification capacity achieved through inhibition of the predators is comparable to the increase that might be expected from an increase in temperature of 10˚C (Gujer and Boller, 1986). The results indicate that there may be a potential for increasing the nitrification capacity of trickling filters and other biofilm processes. However, efficient methods for selectively suppressing the microfauna under practical conditions remain to be developed. The full-scale use of the inhibitors employed in this study is unrealistic, both from an economical and an environmental point of view. Before applying selective suppression of predators on a full scale process, it is also important to consider the benefits of having these organisms in the system, for example, reduced sludge production and possibly improved transport of substrates within the biofilm.

8.2 Hypotheses on the Role of Microfauna

When discussing the effects of different species of microfauna in biofilm processes, it is necessary to consider their complex and different life strategies. Different influence patterns on the biofilm structure and activity may be possible depending on the type of protozoa or metazoa that dominates the biofilm and on the type of process that is used. The life strategies of different predators vary significantly and in terms of feeding behaviour, the microfauna may in a simplified manner be divided into three main groups.

- *Filter feeders*: organisms that feed on suspended particles by means of ciliary activity, either attached or in a free-swimming stage.
- *Burrowers (or crawlers)*: organisms that feed on bacteria within a biofilm or a sludge floc.
- *Carnivorous and cannibalistic organisms*: organisms that feed on other protozoa and metazoa, either inside or outside the biofilm.

Due to the complexity of the interactions between the microfauna and other organisms, such as bacteria, both negative and positive effects on the biofilm activity and structure as a whole may occur (Stal, 1989; WPCF, 1990). The hypotheses discussed below may therefore be limited to the specific experiment discussed in the previous section and the conditions applied there. In other experiments different results may be achieved. However, the hypotheses may be regarded as a first attempt to describe one aspect of microfauna influence, which may be used as a basis for future investigations.

The dominating species of the microfauna in this experimental study were, as earlier discussed, rotifers and stalked ciliates, in this case filter feeders that to a large part do not feed on particles larger than 10–20 μm (Fenchel, 1986; Arndt, 1993). However, since the majority of the nitrifying flocs were larger than 10 μm, questions about the influence of the filter feeders on the biofilm structure and activity in this particular experiment arose. It appeared unlikely that the nitrifiers could serve as the main food source for the predators in the non-inhibited case, thereby leading to a reduced nitrifying capacity.

Suspended particles constitute the primary food web for the filter feeders. By means of ciliary movements, detached particles from the biofilm may be captured by these organisms. This detachment may occur both at the main biofilm-bulk interface but also within the biofilm where the porous

structure form channels (see Figure 7.5). Based on visual inspections of the biofilm, the boundary between the biofilm and the bulk water phase was very unclear, especially for the non-inhibited case, where the outer part of the biofilm assumed a somewhat 'forest-like' appearance. However, the filter feeders naturally require oxygen for their respiration and, consequently, the activity of the filter feeders may affect the amount of oxygen available for the nitrifiers deeper down within the biofilm. Based on this possible oxygen limitation in the biofilm due to the large amount of microfauna, differences in the nitrification capacity could be postulated.

A second effect of the filter feeders is seen in Table 8.1 as a lower concentration of TSS in the bulk water phase when the system was not inhibited, that is, when the filter feeders were active. This is an expected effect of their feeding behaviour. A way to model this is to assume a higher attachment rate for the non-inhibited biofilm. Part of the attached suspended particles will certainly be consumed by the predators. However, it is also likely that the ciliary movements of a large number of filter feeders attached to the surface of the biofilm, will result in a higher turbulence in the water close to the biofilm surface and thereby increase the number of suspended particles that are incorporated into the biofilm, without being consumed. It has been shown that protozoa can cause micro-turbulence by the ciliary movements (Nisbet, 1984). This effect may be further accentuated by the 'forest-like' structure of the biofilm in the non-inhibited case, creating micro-eddies close to the surface of the biofilm.

Some other possible scenarios may also be considered. For example, the ciliary movements of the filter feeders could actually increase the oxygen flux into the biofilm. Depending on the amount of predators consuming oxygen, the oxygen transfer within the biofilm may be improved and, consequently, lead to an improved nitrification capacity. Another possibility is that the rotifers produce excretory and secretory substances, which may affect the structure and density of the biofilm (Doohan, 1975). Finally, some of the observed effects may occur as a consequence of changes in the hydraulic flow conditions over and within the film (the convective flow), due to a different biofilm structure when the microfauna is active compared to when it is not. However, such effects are complicated to verify experimentally.

Based on the above hypotheses with regard to the oxygen consumption of the microfauna and its effect on the attachment rate, attempts will be made to develop a mathematical model that may roughly explain and predict the behaviour of the experimental biofilm system discussed in Section 8.1.

8.3 Model Development

In order to investigate the above hypotheses and the possibility of modelling biofilm systems including effects due to microfauna, computer simulations of the suspended-carrier system will be performed. Three different model modifications will be presented and validated against experimental data. All models are developed and simulated with the AQUASIM software and the transport mechanisms are consequently based on the concepts described in Section 7.2.

Defining the Physical Properties of the System

The hydraulic retention time of the experimental system was 3 hours. The influent volumetric flow rate of the model was consequently set to $0.8/3 \approx 0.27$ l/h, as the volume of each reactor was 0.8 l. Of the influent COD concentration (300 mg COD/l), 96.5% was modelled as readily biodegradable substrate (S_S) and the remaining 3.5 % was considered to be soluble inert material (S_I). No particulate COD was present in the influent synthetic wastewater. All influent nitrogen (44 mg N/l) was modelled as ammonia (S_{NH}). No measurements of the biofilm density and the fraction of water in the biofilm were available from the experiment. However, the density of the different particulate components were set to 75 000 g/m^3 (as mass per unit of particle volume) and the water fraction of the biofilm was fixed at 80 %. These values were chosen for two reasons. Firstly, similar values were found in experimental investigations by Gujer and Boller (1990), when modelling a nitrifying biofilm process and, secondly, the values are reasonable in the sense that the mass of a modelled biofilm of one millimeter thickness, will be approximately equivalent to the amount of biomass measured at the end of the experiment (see Section 8.1). The thickness of the biofilm on the carrier particles measured during the experiment was close to the above value. Since no measurements of the water fraction of the biofilm were available, the possibility of modelling the water fraction of the biofilm as a varying dynamic variable provided in AQUASIM, was not used, that is, $r'_{\varepsilon_{l,F}}$ in equation (7.4) was set to zero.

The combination of attachment and detachment processes determines the steady-state thickness of the biofilm (Gujer and Boller, 1990). Although both the detachment coefficient in equation (7.16) and attachment coefficient in equation (7.17) may be modelled individually for each type of particulate material in AQUASIM, global detachment and attachment

rates were used, as the experimental measurements were not detailed enough to allow for individual rates to be modelled with any relevance. Instead the attachment coefficient was assumed to be constant, whereas the detachment coefficient was described as a function of the biofilm thickness according to

$$k_{de} = k_{\text{shear}} L_F \tag{8.1}$$

where k_{shear} was chosen to be constant. This implies that a thick biofilm is more sensitive to detachment than a thin film, which is a realistic assumption. In order to allow for simultaneous attachment and detachment in the model, an effective diffusion process, see equation (7.7), was assumed for the particulate material. The diffusion coefficients were considered to be constants and equal for all different particulate materials The effective diffusion coefficient was set to 10^{-10} m^2/d, that is, a considerably slower diffusion rate than the ones used for the soluble components.

The diffusion of soluble material was modelled according to equation (7.6). The used diffusion coefficients for the different soluble model components are given in Table 8.2.

Soluble COD (S_S)	$75 \cdot 10^{-6}$ m^2/d
Soluble inert COD (S_I)	$75 \cdot 10^{-6}$ m^2/d
Ammonia (S_{NH})	$170 \cdot 10^{-6}$ m^2/d
Nitrate (S_{NO})	$160 \cdot 10^{-6}$ m^2/d
Soluble organic nitrogen (S_{ND})	$75 \cdot 10^{-6}$ m^2/d
Dissolved oxygen (S_O)	$290 \cdot 10^{-6}$ m^2/d

Table 8.2 Diffusion coefficients used in the simulations.

The values for the diffusion coefficients are similar to what is commonly found in the literature (e.g., Henze *et al.*, 1992). Only the diffusion rate for the dissolved oxygen is somewhat high. However, the diffusion is highly dependent on the structure of the biofilm and no attempt was made to measure the actual diffusion rates during the experiment. Instead the oxygen diffusion coefficient was used for calibrating the model to the experimental data (within reasonable limits). Furthermore, a number of different processes may influence the molecular diffusion as previously discussed in Section 7.2. This implies that the diffusion coefficients should be regarded as effective diffusion rates including effects from other types

of processes and may consequently differ somewhat from values found in reference tables for molecular diffusion in pure water. It should also be noted that possible effects due to a liquid boundary layer (see Figure 7.2) was excluded from the model, in order to minimize the number of parameters and keep the model somewhat simpler.

In the experimental study, the growth of the biofilm took place on the inside walls of carrier particles with a special geometry (see Figure 8.1). The geometry leads to that the surface area of the biofilm will change as the film thickness varies. It is important to model a changing biofilm area correctly both for the calculation of fluxes into and out from the biofilm and for dynamically determining the biofilm volume and, consequently, the mass of attached biomass. The surface area of the biofilm was modelled as a function of the biofilm depth (z) according to the geometry of the carrier particles as

$$A(z) = 4n_c\left(2\pi l_c\left(r_c - z\right) + 2\pi\left(2r_c z - z^2\right)\right) \tag{8.2}$$

where n_c (= 400) is the number of carrier particles in each reactor and l_c (= 7 mm) and r_c (= 2.25 mm) are the length and radius of each carrier particle, respectively. Note that r_c is the radius of each cross-sectional quadrant of the carrier particle, approximated by a circle (cf. Figure 8.1). This means that the first term of (8.2) models the inside area of a cylinder and the second term describes the biofilm area of the edges of such a cylinder. The factor 4 is due to the fact that each carrier particle is made up of four such imaginary cylinders. The total surface area of the biofilm as a function of the biofilm depth (8.2) is illustrated in Figure 8.4. The graph shows that the modelled total surface area of the carrier particles without any biofilm (i.e., $z = 0$) in a reactor is 0.16 m^2, which is equivalent to the theoretical area of the particles (specified as 400 m^2/m^3, filling 50 % of a 0.8 l reactor). When $z = 2.25$ mm, the imaginary cylinders making up the carrier particles are completely clogged and the surface area of the biofilm is limited to the top and bottom areas of the cylinders, which equals an area of 0.05 m^2. The thickness of the modelled biofilm cannot exceed the radius of the cylinders making up the carrier particles, that is, 2.25 mm.

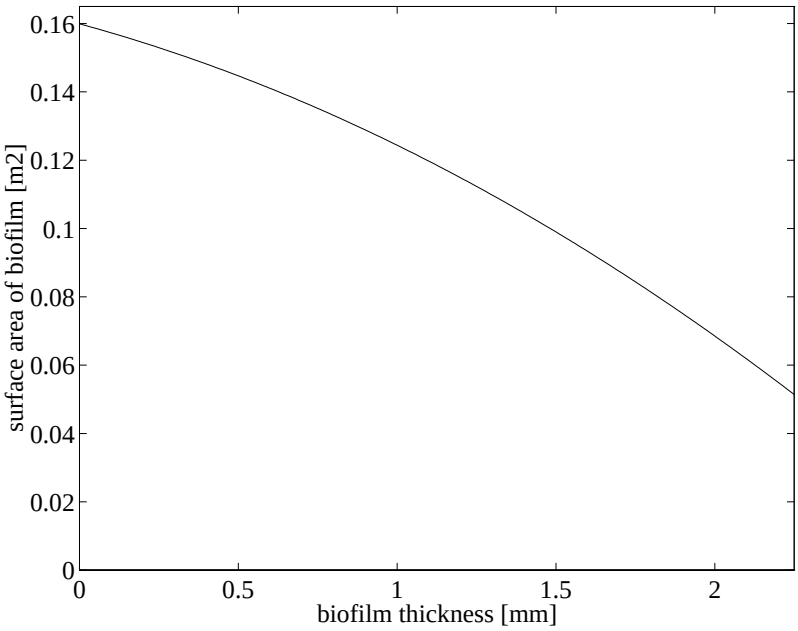


Figure 8.4 The biofilm surface area as a function of biofilm thickness according to equation (8.2).

Modelling the Biological Processes

When the physical description and the variables determining the fundamental mechanisms in the biofilm have been defined (e.g., attachment, detachment, diffusion, surface area variations), the biological mechanisms and processes must be described. AQUASIM only defines the physical structure and transport processes of the biofilm, whereas the biological processes must be included by the user.

As the hypotheses discussed in the previous section are mainly based on oxygen balances and oxygen consumption, it is natural to use a biological model that is based on maintaining an oxygen balance on a COD basis. The most widely used model of this type is the IAWQ AS Model No. 1 (Henze *et al.*, 1987), thoroughly described in Chapter 3. In this study, the processes both within the biofilm and in the bulk water phase was to be modelled. Due to the comparably high HRT (3 hours), it was decided that the reactions in the bulk phase could not be neglected. The IAWQ model has in numerous applications proven to be a reliable tool for modelling the mechanisms in suspended-growth systems and it would naturally be a great advantage if the same model could be used to model the biological processes within a biofilm as well. A difficulty when applying the IAWQ model to a biofilm process is the fact that it is not clear how the transport of slowly biodegradable particulate substrate (X_S) within the film takes

place (Gujer and Boller, 1990) and whether X_S is hydrolysed within the film or in the bulk phase, as discussed in Section 7.1. In this work, we assume that hydrolysis occurs within the biofilm and the IAWQ model may consequently be used in its original form. However, as there is no X_S present in the influent synthetic wastewater, the effects due to this assumption will be comparably small, although a certain amount of readily biodegradable substrate will be produced within the biofilm as a result of hydrolysis and reused for growth of biomass according to the death-regeneration principle (Dold *et al.*, 1980) as stipulated in the IAWQ model. Note that all internal processes of the model are equally active in the bulk phase and within the biofilm (using the same set of model parameters). The major difference is simply the concentration of biomass in the bulk phase and the biofilm.

In order to model the influence of microfauna on the behaviour of the biofilm, three different modelling approaches will be investigated. The main goal is to include the effects of the microfauna behaviour in a simplified way. In theory, it would be possible to introduce a detailed description of the growth and feeding behaviour of various higher organisms into a biological model. However, as a systematic knowledge of these organisms in biofilm systems is scarce (as discussed in Section 7.1), a simplified approach is chosen where certain observable *effects* – which may be influenced by the life strategy of the organisms and estimated by direct measurements in the bulk phase – are included in the model without any detailed description of the microfauna itself. The three different models are schematically outlined in Figure 8.5 using the notation earlier defined for the IAWQ model.

Model A

The first attempt is simply to use the original IAWQ model (not including alkalinity) and investigate if the influence of microfauna found in the experimental study may be described by a few simple changes of some model parameters according to the hypotheses in Section 8.3, that is, parameters influencing the oxygen consumption in the biofilm and rate of attachment. This model will be referred to as Model A.

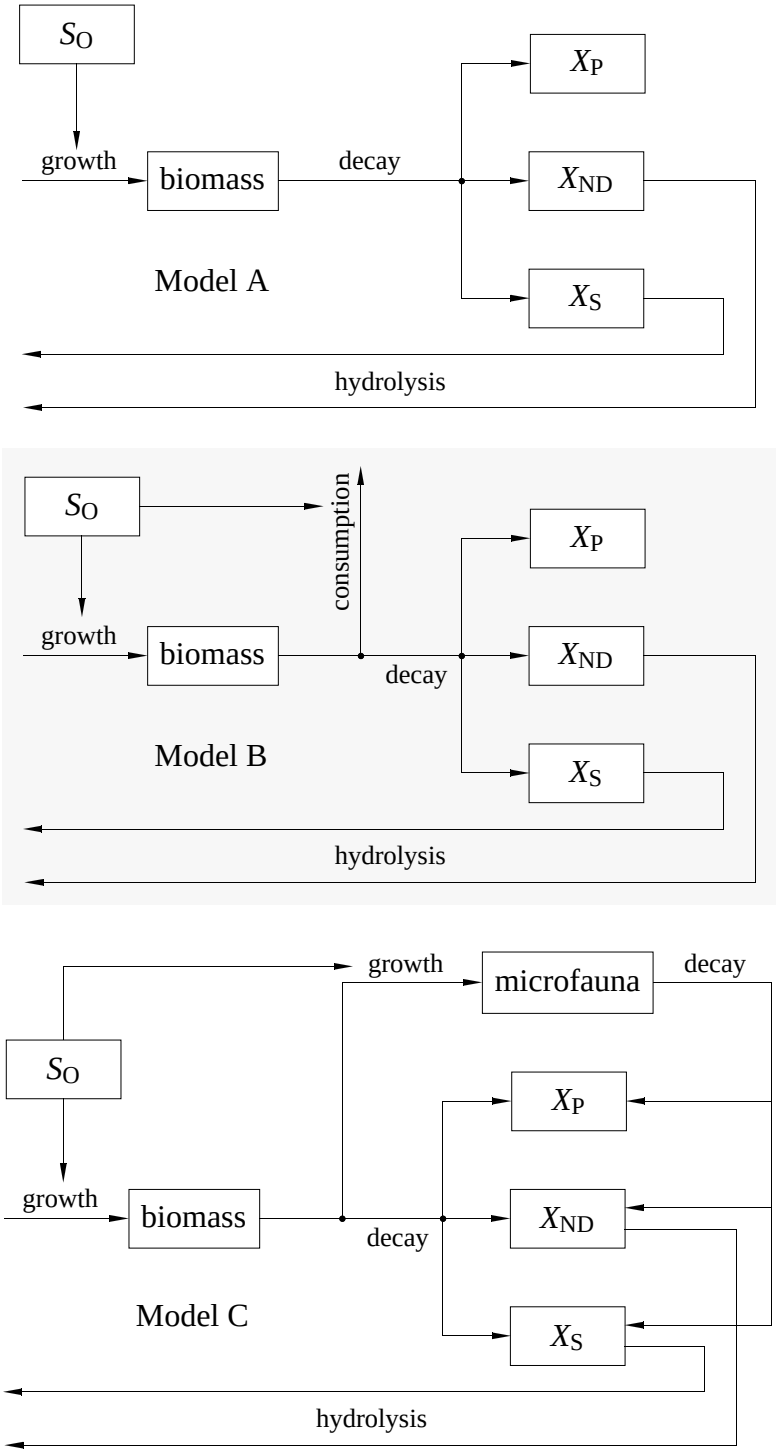


Figure 8.5 Schematic view of the differences of the three tested biological models (models A, B and C) used to describe the effects of microfauna behaviour in the investigated experimental biofilm system.

Model B

In order to more exactly be able to model the oxygen consumption in the biofilm due to microfauna activity, a slightly modified IAWQ model is proposed (referred to as Model B), see Figure 8.5. In this model, decay of bacteria is both due to traditional decay processes and to microfauna activity, where the consumption of bacteria by microfauna results in an immediate consumption of oxygen, that is, the process can only occur in an aerobic environment. However, no state variables are introduced to explicitly describe the microfauna. Instead a coefficient, γ, is used to describe the fraction of the biomass that is consumed by the microfauna. It is defined so that if $\gamma = 0$ then model B is identical to model A and if $\gamma = 1$ then the same amount of biomass that is lost through the traditional decay process is considered to be consumed by microfauna (under aerobic conditions), and so on. A switching function is used to avoid numerical problems with negative oxygen concentrations and the half-saturation coefficient $K_{O,Z}$ is set to 0.5 mg/l. Two new processes are included in the original model and the actual equations are described in Table 8.3 (cf. the IAWQ model in Appendix B). Note that the processes of traditional decay and consumption by microfauna may be written as a combined process but only the consumption part is given in Table 8.3 to simplify the description. The traditional decay processes are not affected by this extension due to microfauna activity but modelled exactly as in the IAWQ model, including death-regeneration.

Process	$X_{B,H}$	$X_{B,A}$	S_O	Process rate
Aerobic consumption of heterotrophs	–1		–1	$\gamma b_H \left(\frac{S_O}{K_{O,Z} + S_O} \right) X_{B,H}$
Aerobic consumption of autotrophs		–1	–1	$\gamma b_A \left(\frac{S_O}{K_{O,Z} + S_O} \right) X_{B,A}$

Table 8.3 Imposed extensions of the original IAWQ model, resulting in the proposed model B.

Model C

The third model proposed (Model C), is a further extension of model B, where the microfauna is introduced into the model as an explicit state variable and, consequently, growth and decay of microfauna can be

included in a traditional way, see Figure 8.5. In Table 8.4 the growth and decay processes are introduced. These equations are added to the original IAWQ model.

Process	X_S	X_Z	$X_{B,H}$	$X_{B,A}$	X_P	S_O	X_{ND}	Process rate
Aerobic growth of microfauna on heterotrophs		1	$-\frac{1}{Y_Z}$			$-\frac{1-Y_Z}{Y_Z}$		ν_1
Aerobic growth of microfauna on autotrophs		1		$-\frac{1}{Y_Z}$		$-\frac{1-Y_Z}{Y_Z}$		ν_2
Decay of microfauna	$1-f_P$	-1			f_P		$i_{XB}-f_P\,i_{XP}$	$b_Z X_Z$
$\nu_1 = \hat{\mu}_Z \left(\frac{X_{B,H}}{K_{B,H}+X_{B,H}} \right) \left(\frac{S_O}{K_{O,Z}+S_O} \right) X_Z$					$\nu_2 = \hat{\mu}_Z \left(\frac{X_{B,A}}{K_{B,A}+X_{B,A}} \right) \left(\frac{S_O}{K_{O,Z}+S_O} \right) X_Z$			

Table 8.4 The simplified growth and decay processes for higher order organisms (X_Z), included in model C.

The model described in Table 8.4 is still a very rough approximation of the true processes. Firstly, we are lumping all types of higher organisms into one variable, X_Z, although it is clear that different organisms have quite different life strategies. Secondly, the decay process is modelled equivalent to the decay processes of heterotrophs and autotrophs, that is, a constant decay rate coefficient, b_Z, is assumed and the microfauna is considered to decay into the same proportions of X_S, X_P and X_{ND} as the bacteria. The concept of death-regeneration is also applied.

The growth process is modelled as a Monod-type reaction (see Section 3.2) that only occurs when oxygen is available (although protozoa that thrive during anaerobic conditions exist, such organisms are neglected in the model). For simplicity we use switching functions to limit the growth rate if no heterotrophs or autotrophs are present. However, these functions are added mainly to avoid numerical problems (i.e., negative concentrations). We set the values of both the heterotrophic and autotrophic half-saturation coefficient ($K_{B,H}$ and $K_{B,A}$) to 100 mg COD/l, which means that the growth of microfauna will primarily be limited by the amount of oxygen available within the biofilm, as the biomass concentrations are generally much higher. The growth of microfauna is also accompanied by a direct consumption of oxygen. The oxygen half-saturation coefficient, $K_{O,Z}$, is set

to a value of 0.5 mg COD/l (same value as used for model B), that is, assuming the higher order organisms to be slightly more sensitive to oxygen limitation than the bacteria, since no information is available to determine the true value in the experimental study.

Some investigations have been performed to determine growth, yield and decay coefficients for various higher order organisms in suspended-growth systems (e.g., Sherr and Sherr, 1984; Ratsak *et al.*, 1994; Ratsak *et al.*, 1996). These results show significant spreading depending on the specific type of organism investigated and on the type of bacteria available as food source (e.g., Curds, 1977; Bloem *et al.*, 1988; Glaser, 1988). Generally the found values appear to be lower than for the heterotrophic organisms. Values for $\hat{\mu}_Z$, Y_Z and b_Z are not known for the microfauna present in the experimental study, however, it is a reasonable assumption that they are lower than for the bacteria. Therefore, Y_Z and b_Z are fixed to values that are 20% of the values used for the heterotrophs. The growth rate coefficient will be used to investigate the behaviour of the model when trying to mimic the effects of the microfauna in the experimental system. Note that the parameter values ($\hat{\mu}_Z$, Y_Z and b_Z) are identical whether the microfauna growth is based on consumption of heterotrophs or autotrophs, as no detailed knowledge of these processes is available.

However, initial simulations using model C indicated some severe problems. In a simulated dynamic biofilm system the different organisms are competing for space, as well as for substrate, oxygen, etc., and less favoured species will simply be outgrown and removed from the system when it has reached a steady state. A main reason why heterotrophs and autotrophs can coexist in a simulated biofilm process during steady-state conditions (see Figure 8.8) is that they are both competing for one resource (oxygen) but are also dependent on one more component independently of each other (heterotrophs need readily biodegradable substrate and autotrophs require ammonia). This means that the slow-growing autotrophs can compete with the heterotrophs in the deeper layers of the biofilm if the concentration of readily biodegradable substrate is sufficiently low (because it has been consumed in the upper layers) and there is still some oxygen available.

The situation is different for the microfauna, modelled according to Table 8.4, especially when the half-saturation coefficients influencing the microfauna are set to values comparable to those used for the bacteria. In simple terms this means that the microfauna is only competing for oxygen with the bacteria and, consequently, there exist three different steady-state

solutions. If the growth rate of the microfauna is set too low then no microfauna will exist within the biofilm, and if the growth rate is set somewhat higher then the slow-growing autotrophs will be completely consumed by the higher order organisms, while the heterotrophs are hardly affected. The third case occurs when the growth rate for the microfauna is high and the organisms can outgrow the heterotrophic bacteria as well. In this case the microfauna dominates the entire biofilm and the whole system collapses (the microfauna consumes everything). If the half-saturation coefficients affecting the microfauna are set to extremely large values, it is possible to find some other possible solutions where the microfauna may inhabit certain regions of the biofilm. For example, if $K_{O,Z}$ is set to a value of 100 mg O_2/l, this would imply that the growth of microfauna is very sensitive to the dissolved oxygen concentration available in the biofilm (< 10 mg O_2/l) and growth may be limited to the outmost region of the biofilm (due to the behaviour of the switching function). However, such high values are unreasonable if the coefficient is to be regarded as a half-saturation coefficient in the traditional sense. Using the more reasonable parameter values discussed above and setting the specific growth rate of the microfauna to 20% of the specific growth rate for heterotrophs leads to a situation where all autotrophic bacteria are consumed by the microfauna and no nitrification can occur.

However, the microscopical investigations of the experimental study had suggested that the higher order organisms present in the system primarily appeared to be dominating the outmost region of the biofilm. In order to achieve such a situation in the simulated system an empirical function was added to the model. This function may be regarded as a rough way of including information of the life strategy of the higher order organisms present in the experimental system. The predators present were to a large extent filter feeders which attach themselves to the surface of the biofilm, as earlier discussed in Section 8.1. The chosen empirical function is defined as

$$y\left(z/L_F\right) = \frac{e^{\alpha\left(\frac{z}{L_F}\right)} - 1}{e^{\alpha} - 1} \tag{8.3}$$

where z is the space coordinate of the biofilm perpendicular to the substratum ($z = 0$ at the substratum) and L_F is the thickness of the biofilm. Note that L_F is not a constant but a dynamic variable. The parameter α determines how rapidly the function is rising and was set to 7, which implies that the function equals 0.5 when the ratio of z and L_F equals 0.9.

By multiplying the two growth rate expressions in Table 8.4 with equation (8.3), we can describe a system in which the microfauna is very active at the outer regions of the biofilm (i.e., z/L_F close to 1) and inactive in the deeper regions of the film, that is, in principle the same behaviour as found in the experimental system. In Figure 8.6 the behaviour of equation (8.3) is demonstrated.

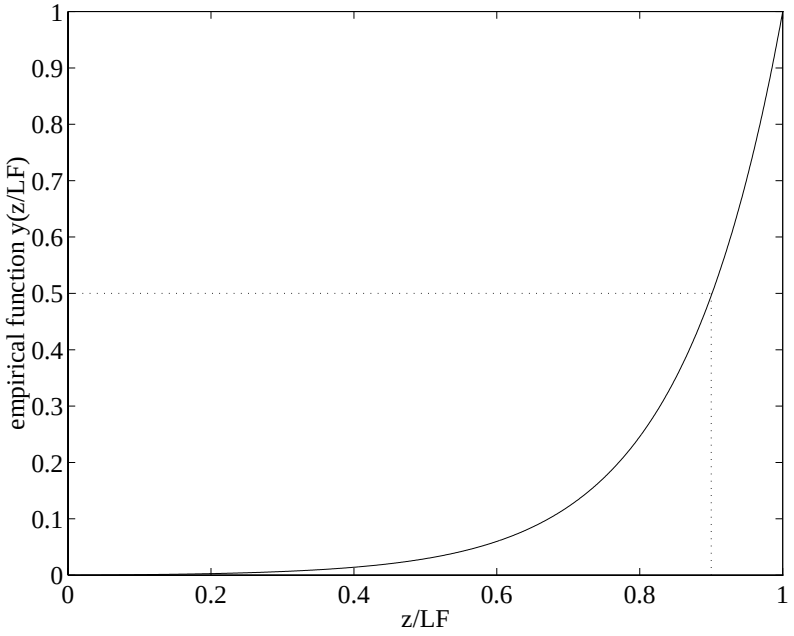


Figure 8.6 Behaviour of the empirical function (8.3) used to describe the life strategy of the microfauna in model C (α= 7).

If possible, empirical functions, such as equation (8.3), should be avoided in mathematical models that are based on an understanding (or at least a hypothesis) of the processes involved. However, there is no way to include information of the life strategies of higher order organisms in models available today. Therefore, an empirical function is used in model C, in order to obtain a description that can produce a situation similar to what was observed in the true process.

Another problem occurs when equation (8.3) is included in the growth rate expressions of model C. The specific growth rate of the microfauna must be set to a large value, which will allow the higher order organisms to dominate the outer region of the biofilm. Furthermore, the modelled biofilm is always influenced by a detachment process according to equations (7.16) and (8.1). The combination of bacteria consumption by the microfauna and detachment will lead to a shrinking biofilm because the growth of bacteria within the biofilm is not sufficiently high to compensate for this. Therefore, the attachment rate must also be increased in order to

sustain the biofilm when the microfauna is active. However, an increased attachment rate due to microfauna influence was discussed in Section 8.2 as one probable effect of filter feeders in a biofilm process. The model behaviour appears to strengthen this hypothesis.

An advantage of model C is the fact that the changing amount of microfauna can be followed during simulations, whereas its effect can only be observed as an increased oxygen consumption, when using model B. However, the higher model complexity and the necessary inclusion of an empirical function are significant drawbacks of model C. Note that all physical parameters affecting the microfauna in model C (for example, density, effective diffusion rate, attachment and detachment rate) are identical to the ones used for the other particulate components in the model.

The model complexity could be even further extended by, for example, including several types of higher order organisms, separating the growth process of microfauna depending on the food source by using different sets of parameters for the consumption of heterotrophs and autotrophs, and modelling the mobility of the microfauna. However, as there is no detailed knowledge of the microfauna influence and no experimental data to model the processes in any further details, such an approach would have no relevance even though it might improve the model fit to the experimental data from this specific experiment.

Model Calibration

The calibration of highly complex models as the ones described above is always an awkward task. Not only the parameters of the biological model must be adjusted but also the parameters controlling diffusion and other transport mechanisms, and the structure of the biofilm must be considered. Several parameter combinations may explain the experimental results equally well. The biofilm models are definitely not globally identifiable. In order to keep the calibration procedure as simple as possible, only a limited number of parameters were modified, whereas most parameters were kept fixed. The models were calibrated against the steady-state values for the inhibited case. The reason for this was that the majority of higher order organisms did not affect the process in this case and the mechanisms of the IAWQ model would be directly applicable. The *default* set of parameters for the IAWQ model (see Table 3.1), proposed by Henze *et al.* (1987), was applied independently of any calibration, only adjusted to the temperature of 15˚C by the traditional Arrhenius formula, i.e.,

$$k_{T^{\circ}\mathrm{C}} = k_{20^{\circ}\mathrm{C}} \mathrm{e}^{\zeta_k (T-20)} \tag{8.4}$$

where T is the temperature in ˚C, $k_{20^{\circ}\mathrm{C}}$ is the parameter value for 20˚C and ζ_k is a correction coefficient specific to every individual kinetic parameter k. The chosen values of ζ_k were also the default ones (Henze *et al.*, 1987) except that the coefficient for the heterotrophic growth rate was set equal to the coefficient for the autotrophic growth rate. The same value was also used for the correction factor of the autotrophic decay rate, since Henze *et al.* do not provide any suggestion for this coefficient. Only two of the default parameters had to be further modified. The heterotrophic yield, Y_{H}, was adjusted to 0.52 g COD/g COD (the default value is 0.67). This was an anticipated adjustment, firstly because some predators were still active in the biofilm in spite of the inhibition and, secondly, due to the type of organic substrate in the influent synthetic wastewater. Several investigations have revealed low yield values when a system is fed with a pure and very easily biodegradable substrate, such as acetate, methanol and ethanol (Nyberg, 1994). Moreover, the correction factor for anoxic growth of heterotrophs, η_{g}, was adjusted to 0.4 (default value 0.8). This modification was also an expected one, since a nitrogen mass balance based on the data from the experimental study indicated that very little denitrification occurred. Most of the physical parameters were also kept fixed during the calibration. As earlier discussed, the oxygen diffusion coefficient was used as a tuning parameter within reasonable limits. According to the hypotheses, the attachment rate was expected to be high when the system was not inhibited and lower for the inhibited system. For simplicity the attachment coefficient (k_{at}) was set to zero for the inhibited case, i.e., no attachment was assumed to occur. Instead the detachment coefficient (k_{shear}) was used to calibrate the model and a value of 0.37 d^{-1} was found to be appropriate. The above parameter modifications produced a model that well described the steady state of the experimental system during inhibition. This means that the entire model A (with approximately 50 parameters) could be calibrated by tuning only four parameters (the correction factor for anoxic growth of heterotrophs, the heterotrophic yield, the oxygen diffusion rate and the detachment rate coefficient) and using default values, and intelligent guesses based on various assumptions for the values of all other parameters.

In order to calibrate models B and C exactly the same set of parameters as for model A was used. This is possible by setting the parameter controlling the proportion of biomass (γ) that is consumed by the microfauna to zero in model B and the microfauna growth rate ($\hat{\mu}_{\mathrm{Z}}$) to zero in model C, that is, models A, B and C are identical for the inhibited case. The reason for

this is to simplify the comparison of the three models, although it is clear that some predators are active also during the inhibited period.

When the models have been calibrated for the inhibited case (i.e., limited microfauna influence), the idea is that the models will predict the behaviour of the non-inhibited system by modifying a few model parameters according to the proposed hypotheses. Firstly, the larger mass of the biofilm and lower TSS concentration in the bulk water phase for the non-inhibited case is modelled as an increased attachment rate. Secondly, the consumption of oxygen and bacteria by microfauna, both on the surface and within the biofilm, must be modelled. One possible way to achieve this in model A, would be to increase the decay rates (b_H and b_A) of the micro-organisms. However, as most of the decay material is transformed into X_S in the IAWQ model and only indirectly affects the oxygen consumption (unless the death-regeneration principle is rejected) and the fate of X_S within the biofilm is uncertain, this approach was rejected. Instead, lowering the yield coefficients (Y_H and Y_A) is considered to be a more direct modelling approach. Such a change will more directly affect the oxygen and substrate balance within the biofilm, that is, more oxygen and substrate is required per unit of formed biomass to mimic that a part of the biomass is consumed by microfauna.

For model B the increased oxygen consumption by the microfauna can be modelled directly by increasing the fraction of biomass (γ) that does not decay according to the death-regeneration principle but is simply removed from the system, while oxygen is consumed. The reduced amount of oxygen available within the biofilm will in turn affect the growth rate of new biomass.

For model C a similar approach is used, although here we define explicit growth and decay rate expressions to describe the amount of microfauna available in the biofilm and the transformation of higher order organisms into inert and degradable substrate by decay. Table 8.5 qualitatively describes how the most important model variables are affected by changing the yield coefficient values, the attachment rate and the fraction of biomass consumed by the microfauna. The parameter changes are imposed one at a time and are made in the vicinity of the values found during the initial calibration of the inhibited process.

	decrease Y_H and Y_A	increase k_{at}	increase γ
soluble COD conc.	–	+	o
TSS conc.	–	o	–
ammonia conc.	–	+	+
nitrate conc.	+	–	–
biofilm mass	–	+	–

Table 8.5 Effects of parameter changes on some important model variables (+: increase; –: decrease; o: insignificant effect).

It is noticeable that an increased attachment rate has almost no effect on the TSS concentration in the bulk phase. This is because the detachment rate is defined as a function of the biofilm thickness (8.1), that is, it will also increase and when new steady-state conditions are established the net detachment from the biofilm will be almost the same as before the attachment rate was increased. The effect of single parameter changes when using model C are not provided because in this model the microfauna growth rate and attachment rate must be modified simultaneously when the microfauna at the upper region of the biofilm becomes active. The effects will be demonstrated in the next section. Furthermore, some of the variations shown in Table 8.5 are also due to secondary effects of a changing biofilm surface area, for example the change in nitrification capacity when the attachment rate is increased. The surface area is modelled according to equation (8.2) and will consequently influence processes such as diffusion. Therefore, simulations with a different surface area model (e.g., a constant surface area) may exhibit somewhat different results. Note that the model extensions in models B and C are naturally only active within the actual biofilm, whereas the processes in the bulk water phase are in all cases described by the original IAWQ model.

8.4 Model Simulations and Validation

The simulations presented in this section are performed in AQUASIM using 25 grid points (2 for the bulk water phase and 23 for the biofilm itself) and a second-order approximation of the numerical flux terms (cf. Section 7.2). This means that the simulations require a significant amount of computer capacity. However, using 10 grid points and a first-order approximation of the flux terms will produce results that may differ 10 to

20% from the ones presented here. Increasing the accuracy even further (e.g., 50 grid points) will produce results that can hardly be distinguished from the ones achieved with 25 grid points. Therefore, 25 grid points and a second-order approximation of the flux terms were considered a good compromise between numerical accuracy and computational effort.

It should be noted that no optimization algorithms have been applied in order to minimize the differences between the measured variables and the simulated data. Such an approach would certainly improve the results in the sense that the simulated data would fit the measured data better. However, the purpose of the simulations presented below are more focused on demonstrating the principal behaviour of the models and their ability to describe some effects of microfauna influence on biofilm systems by fairly simple model modifications. The primary aim when modifying the models to describe a situation where the microfauna is fully active, have been to achieve models that predict the limited nitrification capacity of the experimental process. Not until the ammonia and nitrate concentrations could be accurately predicted by the model, the other measured variables were considered for calibration.

Inhibited Case

As discussed in the previous section, models A, B and C all produce identical solutions when no microfauna are active within the biofilm. The steady-state results for the inhibited case are presented in Table 8.6 and compared with the measured average values from the experimental study. Note that ‘steady state’ here means results obtained from dynamic simulations with constant operating conditions running for 400 days.

	measured	simulated
soluble COD [mg COD/l]	40	40
TSS [mg COD/l]	120	118
ammonia [mg N/l]	10.7	10.5
nitrate and nitrite [mg N/l]	21.7	21.4
biofilm mass [g]	1.11	1.10
biofilm surface area [m^2]	NA	0.13
biofilm thickness [mm]	NA	0.86

Table 8.6 Measured and simulated steady-state values for the calibrated model describing the inhibited case (NA: not available).

Table 8.6 clearly shows that it is possible to obtain a mathematical model, which is capable of mimicking the average experimental results with good accuracy. In Figures 8.7 and 8.8, the corresponding simulated steady-state concentration profiles in the biofilm are shown for the most important soluble and particulate variables, respectively.

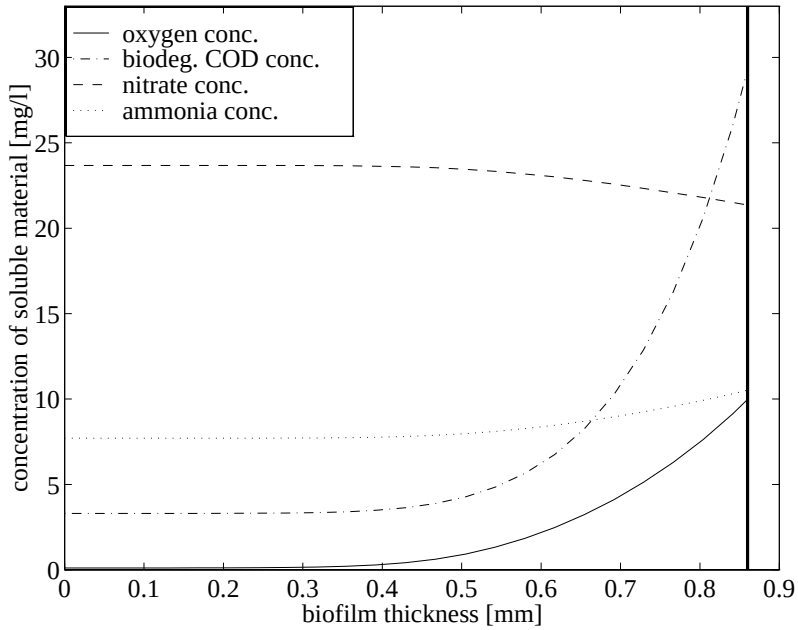


Figure 8.7 Steady-state concentration profiles of the soluble components within the biofilm for the inhibited case.

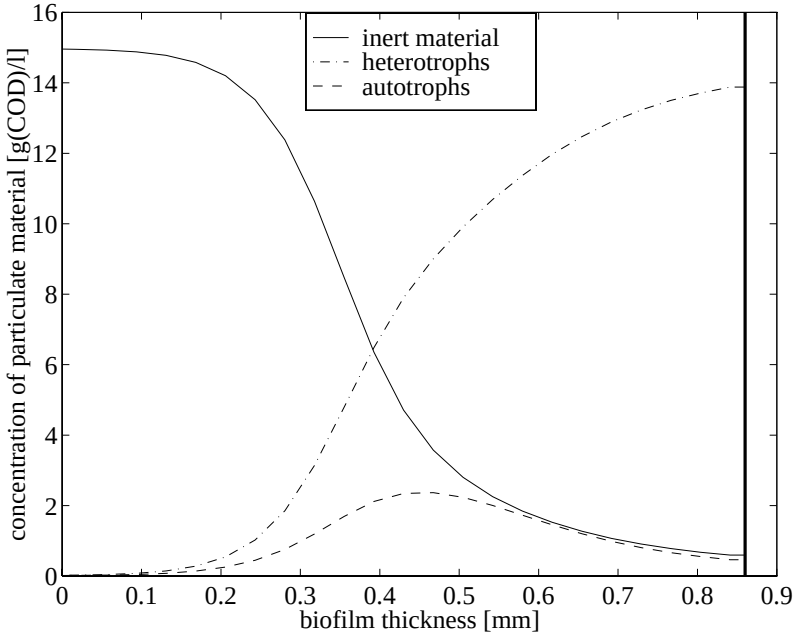


Figure 8.8 Steady-state concentration profiles of the particulate components within the biofilm for the inhibited case.

In both figures, the concentrations are presented as mass per total volume of biofilm, i.e., including the water fraction of the biofilm. The substratum-biofilm interface is positioned at $z = 0$ and the solid line at $z = 0.86$ indicates the interface between the biofilm and the bulk water phase (that is, the steady-state biofilm thickness).

Non-Inhibited Case

The first attempt to model the influence of the higher order organisms on the biofilm in general and on the nitrification capacity in particular is performed by using the original IAWQ model, describing the effects of the microfauna as an increased oxygen consumption (model A). By increasing the attachment rate, the behaviour of the filter feeding organisms on the surface of the biofilm can be described. The attachment rate parameter (k_{at}) is set to 0.055 m/d, identical for all different particulate matter (the value was zero for the inhibited case), see equation (7.17). Lowering the yield coefficients for both the heterotrophs and the autotrophs result in a situation where more oxygen and substrate are required to form the same amount of biomass. This effect will to some extent be similar to predators consuming bacteria and oxygen in the biofilm. For simplicity, the values of both yield coefficients are reduced in the same way although we have no information if both types of bacteria are affected by the microfauna to the same extent. The yield coefficients are set to values that are 30% lower than the calibrated values used for the inhibited case (Y_H changed from 0.52 to 0.37 and Y_A from 0.24 to 0.17). Apart from the parameter changes described above, the model is exactly the same as the one used for the inhibited case.

In Table 8.7, the steady-state results for the non-inhibited case are presented and compared with the measured average values from the experimental study. The reduced nitrification capacity of the non-inhibited system is reflected by the model simulations. This is due to the more severe oxygen limitation within the biofilm which in turn is a combined effect of the lower yield values and of the reduced biofilm surface area, which affects the diffusion process. The other measured variables are also close to the values predicted by the model.

	measured	simulated
soluble COD [mg COD/l]	47	45
TSS [mg COD/l]	76	78
ammonia [mg N/l]	21.4	22.2
nitrate and nitrite [mg N/l]	12.1	12.5
biofilm mass [g]	1.67	1.61
biofilm surface area [m^2]	NA	0.095
biofilm thickness [mm]	NA	1.56

Table 8.7 Measured and simulated steady-state values for model A describing the non-inhibited case (NA: not available).

In Figures 8.9 and 8.10, the corresponding simulated steady-state concentration profiles in the biofilm are presented for the main soluble and particulate components. The substratum-biofilm interface is positioned at $z = 0$ and the solid line at $z = 1.56$ indicates the interface between the biofilm and the bulk water phase.

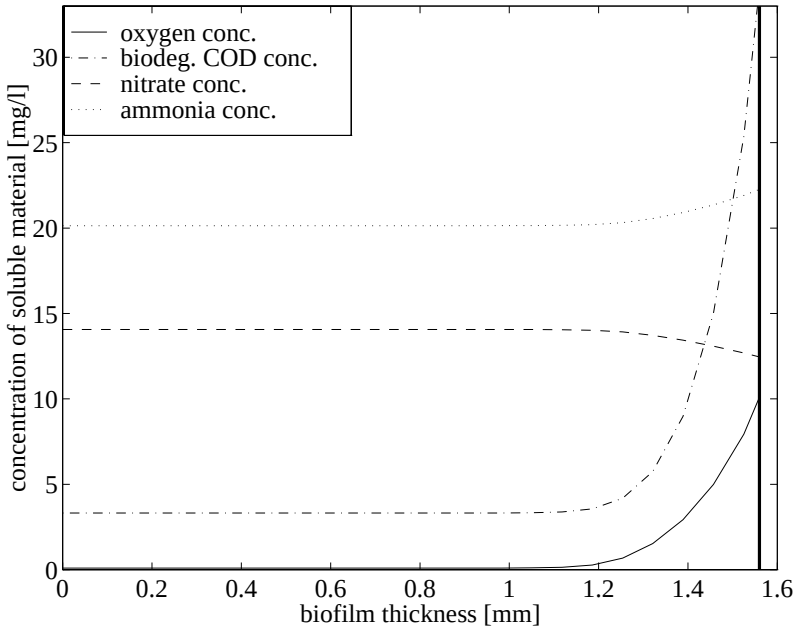


Figure 8.9 Steady-state concentration profiles of the soluble components within the biofilm for the non-inhibited case using model A.

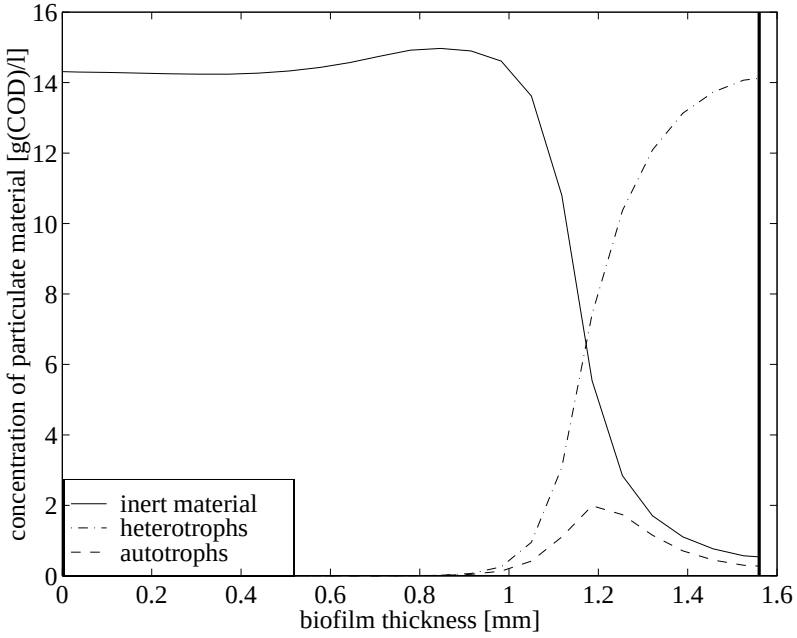


Figure 8.10 Steady-state concentration profiles of the particulate components within the biofilm for the non-inhibited case using model A.

One drawback of model A is that in order to describe the effects of microfauna influence, the yield coefficients for the bacteria are modified. There is no reason to assume that the yield values for the heterotrophs and autotrophs are actually affected by the higher order organisms. Moreover, the primary hypothesis discussed in Section 8.2 was that the oxygen consumption of the microfauna could have a major impact on the nitrification capacity of the biofilm. Although the change of the yield values will influence the oxygen consumption within the biofilm it will also affect the amount of substrate required for growth of the bacteria, which is not realistic. Especially, a lower autotrophic yield value will improve the nitrification capacity per mass unit of autotrophic biomass, which is contradictory to the purpose of decreasing the yield value in terms of an increased oxygen consumption. A better method is to apply the basic IAWQ model in its original form to model the behaviour of the bacteria and add on processes that will only occur due to the presence of microfauna in the system. Model B represents a simple approach to achieve this.

In model B two new processes are added to the IAWQ model to describe the consumption of bacteria by the predators, as shown in Table 8.3. The consumption of bacteria results in an immediate oxygen consumption. The process is modelled as a traditional decay process in order to keep the model simple and the factor γ describes the amount of bacteria that is

consumed by microfauna compared with the amount that is removed by the normal decay process. The rate of consumption is set directly proportional to the bacteria concentration, that is, the predators are more active in regions of the biofilm where there is more 'prey' available. To simulate the non-inhibited process we assume a value of $\gamma = 1$ and determine a suitable corresponding value of the attachment rate to be 0.032 m/d. The steady-state values for the non-inhibited system using model B are presented in Table 8.8.

	measured	simulated
soluble COD [mg COD/l]	47	46
TSS [mg COD/l]	76	84
ammonia [mg N/l]	21.4	20.2
nitrate and nitrite [mg N/l]	12.1	11.9
biofilm mass [g]	1.67	1.40
biofilm surface area [m^2]	NA	0.11
biofilm thickness [mm]	NA	1.22

Table 8.8 Measured and simulated steady-state values for model B describing the non-inhibited case (NA: not available).

Most of the simulated values correspond well to the measured ones, only the biofilm mass is somewhat different. However, this is easy to accept if we consider the fact that the consumption of bacteria by microfauna is modelled simply as an oxygen consumption and the bacteria are removed from the process. The normal decay process of bacteria results in production of inert material and substrate that may be used for new biomass growth. The sum of the steady-state ammonia and nitrate concentrations is also lower in model B than in model A. This is an effect of the removal of biomass when using model B, as the biomass contain nitrogen. The corresponding simulated steady-state concentration profiles of the soluble and particulate variables are shown in Figures 8.11 and 8.12.

The structure of model B allows us to describe the increased oxygen consumption in the biofilm due to the respiration of the microfauna – but only as a function of the bacteria concentration in the biofilm. Moreover, the mass of bacteria that is consumed by the predators are lost from the process.

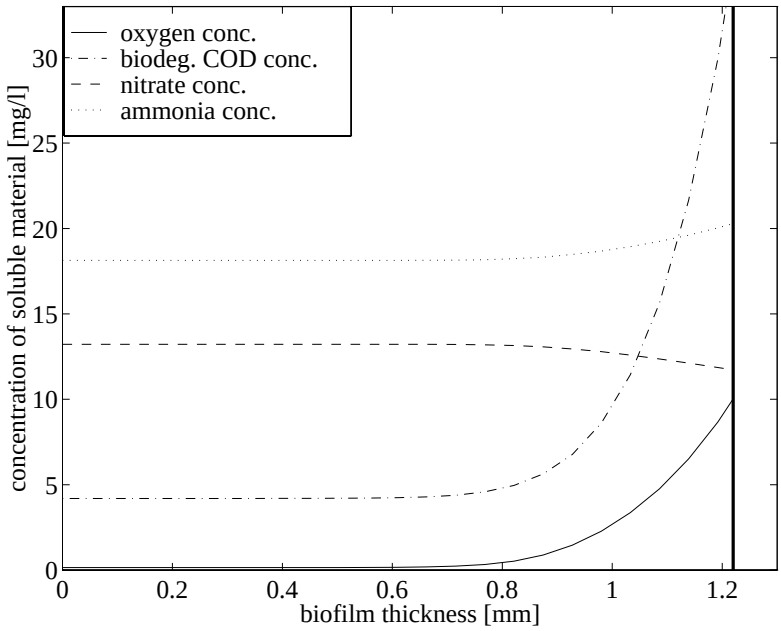


Figure 8.11 Steady-state concentration profiles of the soluble components within the biofilm for the non-inhibited case using model B.

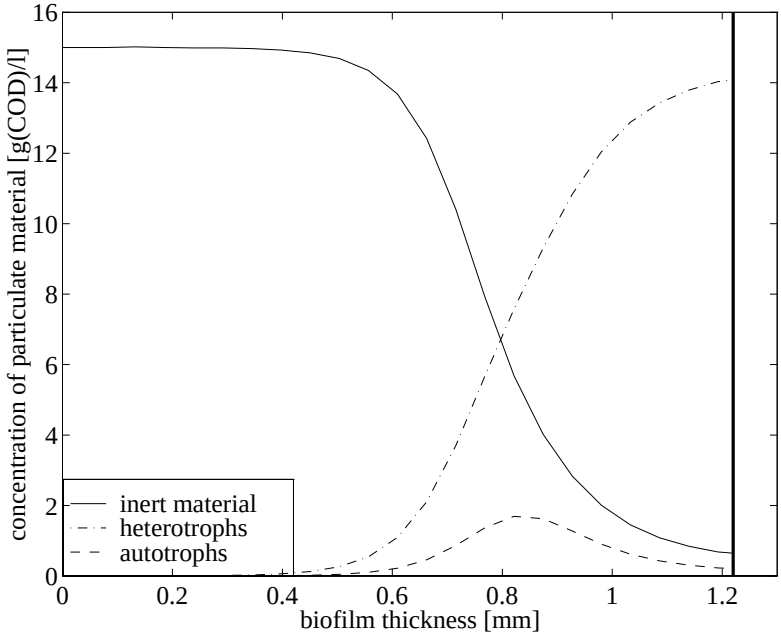


Figure 8.12 Steady-state concentration profiles of the particulate components within the biofilm for the non-inhibited case using model B.

In model C an attempt is made to directly include the growth and decay processes of microfauna into the model, see Table 8.4. The necessity to include an empirical function, i.e., equation (8.3), in the growth processes to describe some rudimentary information of the life strategy of the micro-

fauna dominating the experimental system, was discussed in the previous section. By setting the specific growth rate of the microfauna to a high value (3 d^{-1}), the higher order organisms will be able to dominate the outer region of the biofilm in steady state, which is in agreement with the microscopical examinations of the experimental system during the non-inhibited periods. The growth rate will rapidly decrease as a function of biofilm depth due to the exponential behaviour of the empirical function (see Figure 8.6).

However, if certain precautions are not taken then the microfauna in the biofilm will continue to consume bacteria within the biofilm while the outmost layer of the film is lost due to the detachment process. Over a short period, the simulated biofilm will get thinner and finally the system will collapse when all bacteria have been consumed. To balance this decrease of the biofilm thickness, the attachment rate coefficient is increased. A suitable value of k_{at} for the non-inhibited case was found to be 0.32 m/d, which is considerably higher than the values used for models A and B. This is a direct consequence of having a layer of active predators at the surface of the biofilm with a high growth rate (in order to outgrow the heterotrophs in this region). Any biomass in the bulk water phase that is attached to biofilm will be consumed immediately and only to a small part increase the thickness of the film because of the low yield value for the microfauna. Moreover, the growth of the most active bacteria (heterotrophs just below the region where the microfauna is dominating) will also be consumed by the microfauna and result in a limited increase of the biofilm thickness. In models A and B the attachment process is a more passive process, where attached biomass is simply incorporated into the biofilm, which in turn leads to a thicker biofilm. The active predators at the biofilm-bulk water interface dramatically affect this process and, consequently, the attachment rate has to be increased. The values of the other model parameters in model C are the ones that were discussed in Section 8.3.

In Table 8.9, the simulated steady-state results for the non-inhibited case using model C are presented and compared with the measured average values from the experimental study. The reduced nitrification capacity is correctly described by the model as an effect of the more severe oxygen limitations within the biofilm due to the growth of microfauna at the outer region of the biofilm. However, the predicted biofilm mass is more than 25 % lower than the measured biofilm mass at the end of the experimental study.

	measured	simulated
soluble COD [mg COD/l]	47	55
TSS [mg COD/l]	76	15
ammonia [mg N/l]	21.4	20.6
nitrate and nitrite [mg N/l]	12.1	11.6
biofilm mass [g]	1.67	1.25
biofilm surface area [m^2]	NA	0.12
biofilm thickness [mm]	NA	1.02

Table 8.9 Measured and simulated steady-state values for model C describing the non-inhibited case (NA: not available).

The simulated steady-state biofilm mass could be increased by increasing the attachment rate even further. However, this would reduce the surface area of the biofilm, see Figure 8.4, and, consequently, limit the diffusion of oxygen into the biofilm. This would in turn severely affect the nitrification process. The large difference between the measured and simulated TSS concentration in the bulk phase is due to a combination of the high attachment rate and the high consumption of biomass by the microfauna at the biofilm-bulk interface. Both these discrepancies indicate that the growth rate (and, consequently, the oxygen consumption) of the microfauna is too high in the model. However, in order to reach a steady-state solution where the microfauna is dominating the outer region of the biofilm, the growth rate of the microfauna must be sufficiently high so that the heterotrophs in this region cannot compete. For example, reducing the specific growth rate of the microfauna to 2 d^{-1} will result in a steady-state solution, where there is no microfauna at all in the biofilm. In reality, 'predator-prey' systems are probably never in steady state (cf. the famous Volterra cycle). In this early stage of the modelling work it is, however, necessary to focus on steady-state solutions. The dynamics of biofilm systems including higher order organisms are so complex that it would be impossible to start the investigation by looking at the dynamic behaviour. First the processes and mechanisms affecting the various predators in a biofilm need to be investigated and more knowledge gained.

The simulated steady-state concentration profiles of the soluble and particulate variables corresponding to Table 8.9 are shown in Figures 8.13 and 8.14.

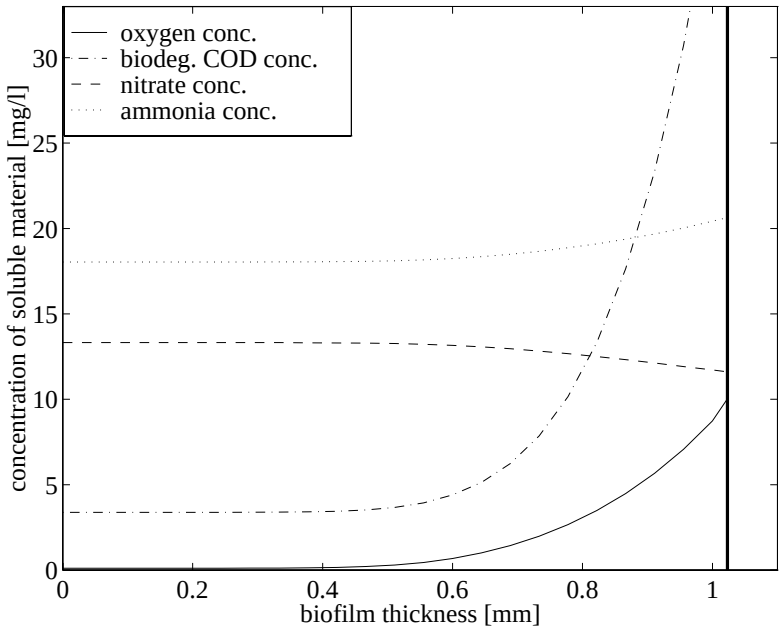


Figure 8.13 Steady-state concentration profiles of the soluble components within the biofilm for the non-inhibited case using model C.

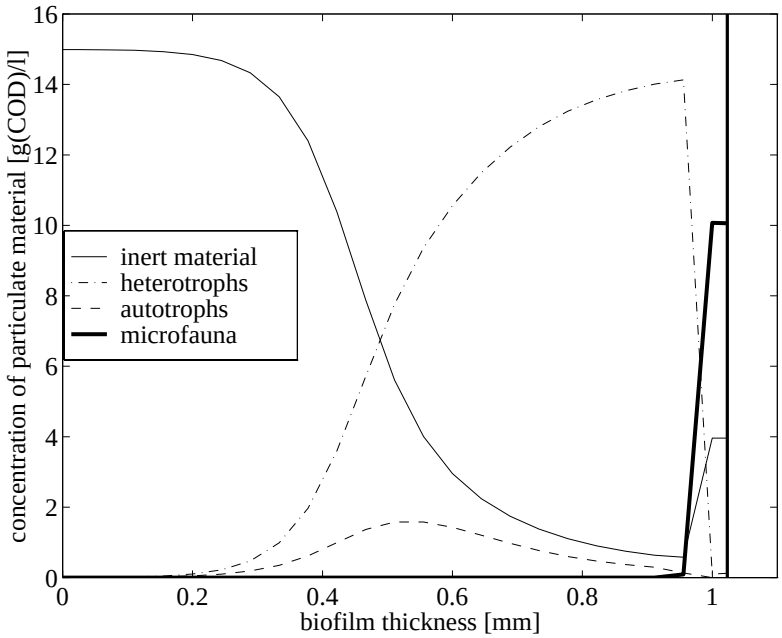


Figure 8.14 Steady-state concentration profiles of the particulate components within the biofilm for the non-inhibited case using model C.

Dynamic Behaviour

The discrepancies between the measured average values and the simulated steady-state results can partially be a consequence of the fact that the experimental process is not in steady state, as discussed above. There are long-term effects that influence the experimental study, which means that some of the measured average values used for the model validations are somewhat misleading. In Figures 8.15 and 8.16, the time series for the concentrations of nitrate and total suspended solids in the bulk phase are shown for the full experimental duration, i.e., periods I-V (cf. Section 8.1, reactor B). These graphs indicate that there are some slow dynamics involved in the process.

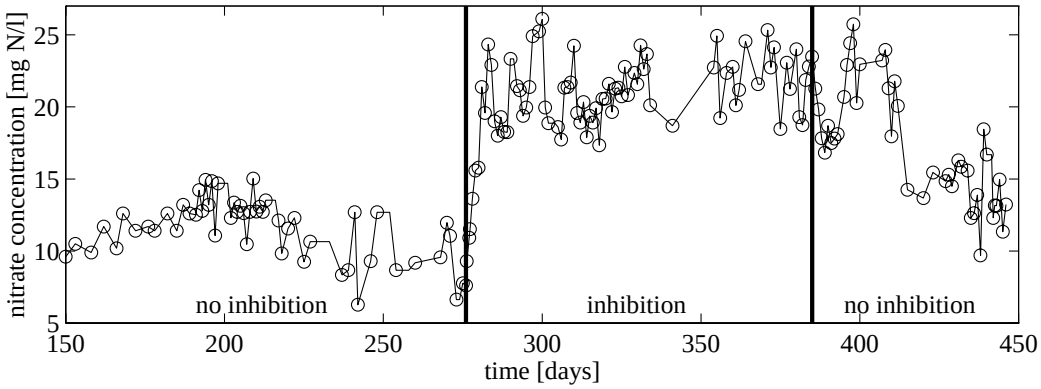


Figure 8.15 Measured values of the total suspended solids concentration in the bulk water phase during the full experimental duration.

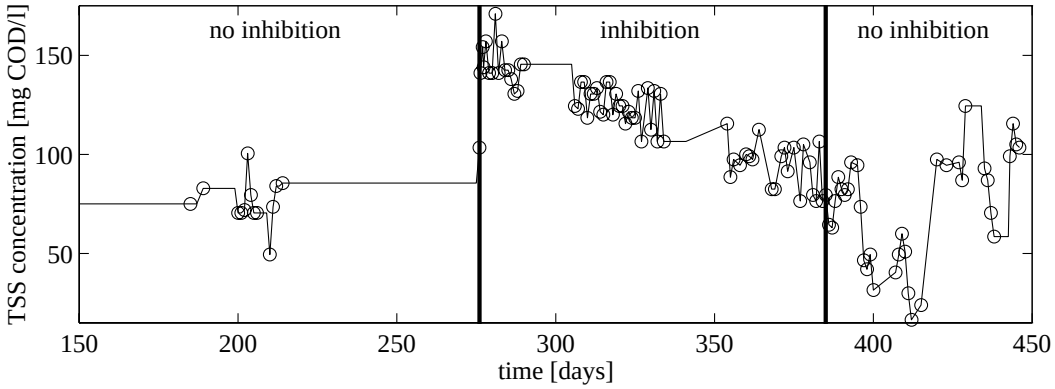


Figure 8.16 Measured values of the total suspended solids concentration in the bulk water phase during the full experimental duration.

The time series for the TSS concentration shows that the concentration seems to decrease with time during the inhibited period and that the concentration at day 385 (end of the inhibited period) is approximately the same as the one that was measured during the initial non-inhibited period.

The corresponding time series for the nitrate concentration remains unaffected by the strange behaviour of the TSS concentration. When the addition of inhibitors is stopped, the concentration of TSS continues to decrease whereas the nitrate concentration remains at a high level for almost 50 days. After this time, the concentration of nitrate decreases to a level which is slightly higher than what was measured during the initial non-inhibited period. Simultaneously with the decrease of the nitrate concentration, the concentration of TSS again returns to the level which was observed towards the end of the inhibited period.

No attempts have been made to model the dynamics of the experimental system in this work. Instead we have focused on explaining the observed effects based on steady-state reasoning. However, in order to demonstrate the dynamic behaviour of the models, a simulation is performed using model C, for the full experimental duration. The results of such a simulation are shown in Figures 8.17–8.20, and compared with the measured values from the real experiment. Note that no attempts have been made to introduce the parameter changes in a smooth manner, as the system changes from a non-inhibited system into an inhibited system and back again. The parameter variations of the specific growth rate of the microfauna and the attachment rate are simply imposed as step functions. The only special condition compared with the previously performed simulations is that during the inhibited period not all of the microfauna within the biofilm are considered to decay. Instead a small fraction of the predators (0.5 % of the biomass) is modelled to remain in a dormant state. This is a necessary modification, otherwise it is impossible to model the recovery of the microfauna when the inhibition is stopped. Moreover, the attachment rate is modelled to increase as a linear function of time during the first five days when the inhibition is stopped, as the increased attachment rate is assumed to coincide with the recovery of the filter feeding organisms at the outer region of the biofilm.

The simulations indicate that the model performs reasonably well when the inhibition is initiated, in spite of the fact that no special calibration has been carried out to model this dynamic situation. The slow dynamics of the TSS concentration during the inhibited case is not predicted by the model (see Figure 8.20). These dynamics may indicate that the detachment and attachment of biomass are not uniformly constant over time. The effect may be due to the slow forming of a more dense biofilm when the microfauna is not active, which would be less exposed to detachment processes than the non-inhibited biofilm.

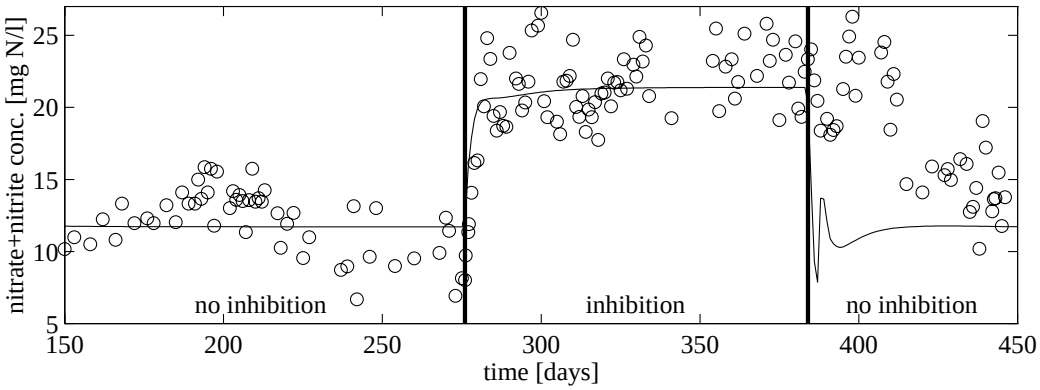


Figure 8.17 Measured (o) and simulated (solid) values of the nitrate+ nitrite concentration in the bulk water phase during the full length of the experimental study.

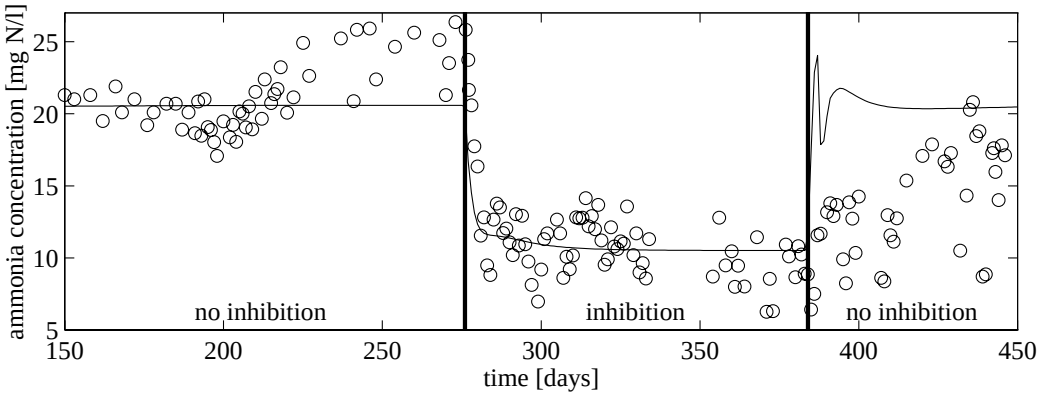


Figure 8.18 Measured (o) and simulated (solid) values of the ammonia concentration in the bulk water phase during the full length of the experimental study.

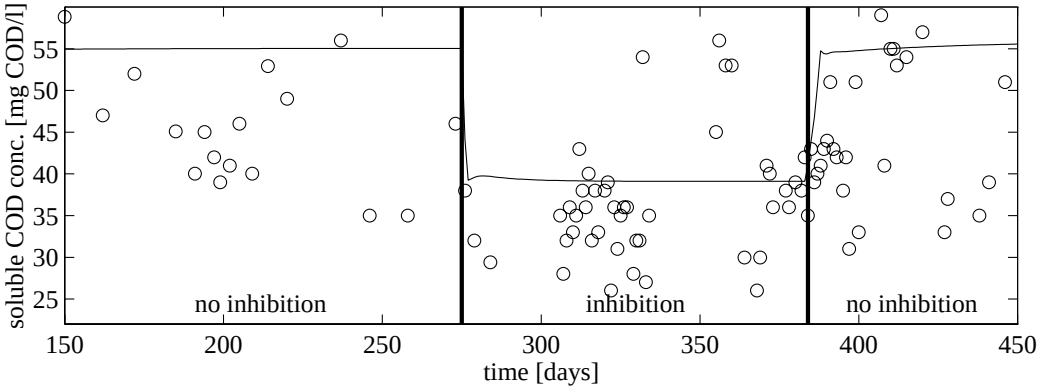


Figure 8.19 Measured (o) and simulated (solid) values of the soluble COD concentration in the bulk water phase during the full length of the experimental study.

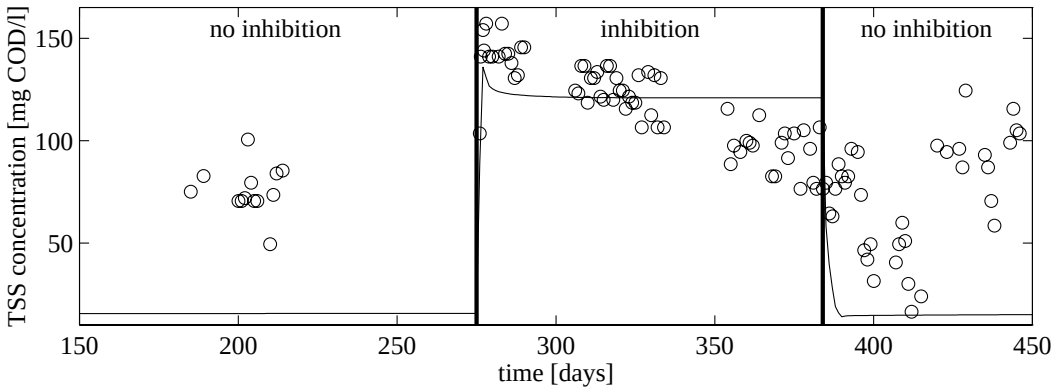


Figure 8.20 Measured (o) and simulated (solid) values of the TSS concentration in the bulk water phase during the full length of the experimental study.

It also seems reasonable to assume that the inhibited biofilm system and its bioscenosis will change over time due to effects from the long-term exposure to the inhibitory substances. Moreover, the inhibition of certain higher order organisms may encourage other species, occupying a similar ecological niche, to flourish.

The observed delayed reaction in the experimental system, both with regard to the nitrification process and the concentration of TSS when the inhibition is stopped, may be an indication of a gradual recovery of those higher order organisms that were present in the system prior to the inhibition. The model predicts this recovery to be significantly faster, primarily due to the necessary high growth rate of microfauna at the outer region of the biofilm, discussed earlier. A number of other processes are probably involved in the real system. In order to be able to evaluate and dynamically model these effects, a better understanding and a more elaborate experimental program – including future modelling aspects – will be required.

8.5 Future Model Development

Predators, such as rotifers and ciliates, can have a strong negative effect on nitrification in aerobic biofilm processes. However, in the literature the effects of microfauna are often neglected and there are no established mathematical models available for the prediction of the behaviour of biofilm processes exposed to predator influence. This is believed to be an important drawback of current biofilm models. As a first approximation,

the effects of microfauna can, during steady-state conditions, be modelled in a fairly simple way by adjusting the oxygen balance of the system, as demonstrated in this work. Model predictions show reasonable agreement with experimental results and appear to partially verify the postulated hypotheses on the effect of filter feeding organisms based on their oxygen consumption.

In order to describe the microfauna influence in a more mechanistic manner, a first attempt to incorporate the microfauna as an explicit state variable was investigated. This approach increases the model complexity and gives rise to a number of questions with regard to growth and decay of microfauna, for example,

- can the growth rate of microfauna be adequately described as a Monod-type expression;
- is the decay process of microfauna similar to the decay of bacteria;
- what are the appropriate parameter values for these processes;
- is the growth of microfauna limited by other factors than oxygen;
- how does the microfauna affect the structure and density of the biofilm;
- do microfauna and bacteria coexist in steady state or is the system inherently dynamic;
- how should information about the life strategies for different types of microfauna be included into a model;
- how should the mobility of higher order organisms be modelled;
- what type of defence mechanisms against predators do different bacteria exhibit and how should this be included in a model?

The above and many other questions need to be answered before the influence of microfauna in biofilm systems can be modelled accurately. There is also a great deal of uncertainty concerning the basic modelling of biofilm systems, neglecting the higher order organisms. These questions involve, for example, the existence of a hydrolysis process within the biofilm, various transport mechanisms within the film, hydraulic phenomena on the surface of the film, and the changing structure, porosity and density of the biofilm. Another important issue is the question if a one-dimensional biofilm model is at all capable of accurately describing the significant heterogeneity found within biofilm systems.

The future modelling work of biofilm systems will continue in the direction pointed out in this section. However, it is obvious that a large number of questions remain to be answered. There exist both a lack of fundamental knowledge of predator-bacteria interactions and how such knowledge should be incorporated into mathematical models. In order to model the details of microfauna behaviour, much more detailed experimental studies of such systems must be carried out (compared with the study described in this section), where molecular probes for bacteria are used to measure variables within the biofilm and not only in the bulk water phase. An important problem is due to the already extreme complexity of biofilm models and great effort must be made to avoid including a large number of new complex processes. Overly complex models can be fitted to almost any type of data and such models will not be operationally useful. Instead we must focus on keeping the models as simple as possible. It appears unrealistic to include different types of predators as different state variables and allowing them to be active simultaneously. Instead the dominating type of higher order organism in a specific process must be determined and a model should focus on describing the behaviour of only this organism (directly or indirectly), at least with regard to the limited knowledge that exists of predator-bacteria interactions today.

It is important to maintain an engineering approach with regard to modelling of biofilm systems. Researchers within the field of microbiology tend to describe processes at the individual cell level but this approach is completely unrealistic if the models are to be used for any practical operational purposes. In the work presented here, the model extensions suggested are reduced into one basic assumption, namely the possible oxygen limitation within the biofilm due to the respiration of higher order organisms. Other model modifications would also be possible, for example, modifying the diffusion rates, but somehow they will all be related to the amount of available oxygen within the biofilm. Based on the necessity for simple models and the limited available knowledge of microfauna influence on biofilms, we would recommend models where the microfauna is described in an indirect manner, that is, the use of model B.

Finally, the process of describing an experimental study and a mental model in an entity, encompassed within the framework of a mathematical model, is an efficient way of analysing experimental data, as areas that require further attention are highlighted. The modelling approach that was adapted in this work regarding the microfauna influence, should only be considered as a first attempt with serious limitations. However, it appears as if the basic hypotheses on the role of filter feeding organisms (i.e.,

increased oxygen consumption in the biofilm and a higher attachment rate) seem to provide a modelling direction that may be further explored to attain a better understanding of the effects of microfauna in different biofilm systems. However, it should be noted that biofilm systems dominated by other types of higher order organisms with different life strategies may require a different modelling approach. Some of the effects observed in this experimental study may also be influenced by the type of carrier material used, which may affect both the structure and the microbial population of the developed biofilm.

PART V

Conclusions and Perspectives

Chapter 9

Conclusions

Wastewater treatment processes can be considered the largest industry in terms of volumes of raw material treated. The large treatment plants that have been built to perform this task are generally operated with only elementary control systems that are often fed only with off-line data; this situation is rather irrational. It is believed that a more efficient use should and could be made of the large investments in construction and equipment for wastewater treatment by means of process control.

The biological behaviour of biotechnological processes occurring in a bioreactor has a complexity unparalleled in the chemical industry. As a consequence, to predict its behaviour from information about the environmental conditions is extremely difficult. The number of reactions and species that are involved in the system may be very large. An accurate description of such complex systems can therefore result in very involved models, which may not be useful from a control engineering viewpoint.

Modern control systems rely heavily on adequate process models. Design of advanced controllers is based on a mathematical description of the process. Since the involved biological processes are highly non-linear, time varying and subject to significant disturbances, the models require adjustment on-line, based on available data from various sensors. Partly due to the lack of available sensors and the complexity of the processes, a compromise must be made between the complexity and the accuracy of the used models.

9.1 Summary of Results

In this work the problem of over-parametrization in complex models describing the activated sludge process has been approached. Existing complex models of the AS system dynamics do not have a unique set of parameters which can explain a certain behaviour, that is, the models are not identifiable. An attempt has been made to develop reduced order models with a smaller number of states and parameters, which are capable of adequately describing the major dynamical behaviour of both the carbonaceous and nitrogenous activities of the AS process. Still, the mechanistic structure of the modelled reactions has been retained when possible. Furthermore, the lack of available on-line sensors emphasizes the need for a more realistic complexity of models for operational purposes.

A thorough investigation of the identifiability of the proposed reduced order models has been performed using both off-line and on-line algorithms. Results have been presented for a large number of different cases depending on which variables are assumed to be measurable. Under certain conditions the simplified models have been shown to be globally identifiable, even when the data are corrupted by a significant amount of noise. Comparisons between the IAWQ model and the reduced models based on numerical simulations have verified that the main features of the dynamics are retained.

Correlations between different model parameters under various conditions have also been investigated. One such correlation exists between the reaction rate factor and the decay rate factor. The difficulty of estimating these parameters separately in a global sense compared with estimating the combined net reaction rate was demonstrated. Finally, a sensitivity analysis of the reduced model to various parameter changes has been performed.

The primary objective of the reduced order models is to apply them as an on-line tool for supervision and control in a hierarchical control structure. Since the model parameters can be updated from on-line measurements, any deviation between the real plant and the model predictions may be used in an early warning system for process diagnosis purposes.

The major contribution related to the work on modelling the settling process was a detailed evaluation of the behaviour of a new robust settler model compared with other one-dimensional layer models available. It has also been demonstrated how the model is integrated with models of the

biological reactors. The settler model should be regarded as a 'first-order' model that captures the wave behaviour and the conservation of mass for any type of loading. The only calibration needed for the model is the establishment, via batch settling tests, of the batch settling flux function for the sludge under consideration. The robust settler model includes a new algorithm for describing the dynamic propagation of the individual biological components of the particulate material through the settler.

The evaluation, which was performed by numerical simulations, focused on the consistency of the model rather than its ability to predict a specific set of experimental data. A main advantage was that the numerical results of the robust model converged to the analytical solution as the number of layers increases. Traditional layer models are highly sensitive to the number of used layers because of the *ad hoc* flux terms commonly applied. Significant differences between the robust and the traditional settler models were observed both for the concentration profile in the settler and in the prediction of the effluent and underflow concentrations. It could also be concluded that a 10-layer model was too crude to resolve the detailed behaviour of the settler. At least 30 layers (preferably 50 layers) are recommended for producing more reliable results during normal operating conditions.

The new material propagation algorithm was compared with another commonly used procedure. Several advantages of the new method were demonstrated. This type of algorithm is of importance when simulating the settler in combination with the bioreactor if sludge is recirculated in the process, and it is even more crucial if biological reactions within the settler are to be modelled.

The robust settler model illustrates the importance of using a sound mathematical model structure and a consistent numerical algorithm. It does not suffer from the many numerical drawbacks commonly found in other settler models. Consequently, the robust model is a reliable platform for future model refinements.

Predators, such as rotifers and ciliates, can have a strong negative effect on the nitrification capacity in aerobic biofilm processes. In this work, it was demonstrated that the effects of the microfauna could, as a first approximation, be modelled in a fairly simple way by considering the oxygen balance of the system. By focusing on describing the *effects* due to microfauna influence rather than the detailed behaviour of the predators, it was possible to extend existing biofilm models in a way that hardly increased

the overall model complexity. Three different modelling approaches were discussed for describing the quasi steady-state behaviour of an experimental system where the influence of predators was studied in detail by alternately inhibiting the higher order organisms. Model predictions showed fair agreement with the experimental results and appeared to strengthen the postulated hypotheses on the effect of filter feeding organisms based on their oxygen consumption. These results were achieved by modifying a very limited number of model parameters, whereas the majority of the used parameter values were the default ones taken from the available literature.

However, over longer periods of time (months), the microfauna (or the effect of the inhibitory substances) appeared to influence the behaviour of the biofilm in a more complex way, which was not captured by the simplified modelling approach. These phenomena will require more detailed understanding if they are to be described by means of a mathematical model.

9.2 Topics for Future Research

In the course of the work presented in this thesis, several types of problems and questions that deserve future attention have been encountered. In relation to the results that have been presented, we can define a number of important issues and extensions that need to be focused upon. Some of them are summarized below.

An interesting approach would be to apply various mathematical methods for model reduction (e.g., perturbation methods) to the IAWQ model and compare the results with the reduced order models developed in this work (based on logical reasoning). If similar results are achieved, this would strengthen the possible applicability of the models. It may even be possible to simplify the models further. From an identifiability point of view, the simplest models should be based on net reaction rates, and the explicit description of active heterotrophic and autotrophic biomass as state variables ought to be eliminated.

A more detailed measurement strategy to be used in combination with the reduced order AS models must also be defined. Exactly what measurements – with what accuracy, how often, and carried out at which positions in a plant – are the minimum requirements to guarantee global identifiability of the reduced order models? How could other easily available measurements, such as redox potential or suspended solids, or more sophisticated measurements of the respiration rate, short-term BOD or sludge activity affect the structure of the models and improve the identifiability? In this context it is also important to investigate how the process (in full scale) should be perturbed in order to gain the most informative data to be used for on-line model identification without affecting the quality of the effluent water. The possibilities of temporarily changing flow rates and flow schemes, adding supernatant (from the sludge treatment) of high concentrations at specific locations, apply step-feed control for the largest natural variations of the influent wastewater, modify the control of the dissolved oxygen concentration, etc., in order to achieve more information from the transient behaviour of the process need to be further explored.

The reduced order AS models must naturally be evaluated for their true purpose, that is, model-based control. It is important to define rather direct links between the on-line measurements, the model predictions and the control actions. The possibility to develop control algorithms based on feed-forward and adaptive control principles should be investigated once the simple, identifiable AS models are available. Due to the large time constants of the AS process an optimal control scheme ought to be based on measurements of the influent wastewater, and control actions should be imposed before any problems actually occur (i.e., feed-forward control). On the other hand, adaptive control is the proper way of controlling processes that significantly change their behaviour over time, such as the AS system. Traditional feed-back control is not the best solution for wastewater treatment processes, especially as the process inputs can only be manipulated to a limited degree. In this context, a question regarding hierarchical control structures must also be addressed. The different low-level control modules need to be synchronized and the partly contradictory control criteria of an overall control strategy have to be formulated to allow for optimal plant performance.

With regard to the robust settler model evaluated in this work, there is a need to further refine the model. Natural extensions of the model are to take into account a decreasing cross-sectional area of the settler as a function of depth and also to include various biological processes occur-

ring in the settler. A varying cross-sectional area will uniquely define the height of the sludge blanket in the thickening zone (if it exists) based on the feed concentration and bulk flow rates in the settler. As for biological reactions in the settler, especially the denitrification process is of importance as extensive denitrification has been reported to occur in the settler. Another improvement to enhance the applicability of the model is to allow for larger time steps in the numerical algorithm. The time step is currently defined by equation (5.49) and calculated when a dynamic simulation is initiated and then fixed to this value, based on a worst case scenario. Instead the time step could be adjusted on-line by an adaptive algorithm based on the actual conditions in the settler. This way the computational effort would be greatly reduced, especially when the process is not exposed to large dynamic disturbances.

A fundamental question is whether one-dimensional settler models based on the solids flux theory are accurate enough to predict the behaviour of the settler. It is clear that hydrodynamic effects play an important role for the settling process in real wastewater treatment plants. It is theoretically possible to extend the robust settler model into a two-dimensional model although the computational effort to simulate such a model would be dramatically increased. Including hydrodynamic equations into the model would of course have a similar effect. It may be possible to include some of the observed phenomena in a simplified way. The ultimate purpose of the model will determine how accurate it needs to be. However, independently of which way we choose to proceed, it is believed that the robust settler model is a good platform for future model refinements due to its mathematically sound structure.

For both the reduced order AS models and the robust settler model there is a great need to evaluate the models using real data from pilot-scale or preferably full-scale WWT plants.

The work concerning modelling of biofilm processes including higher order organisms is still at such an early stage that it is difficult to define specific topics for future research. There is a great need for more fundamental knowledge of the processes and mechanisms occurring in such a system, especially regarding the dynamics. The need for simplified biofilm models is even more apparent than reduced order models for the AS process. A fundamental question with regard to biofilm modelling concerns the heterogeneity and the varying internal structure of biofilms. In this perspective it may be virtually impossible to describe biofilm systems accurately by one-dimensional models. With regard to the modelling work

presented in this thesis, a necessary extension is to carry out a number of experiments (using different influent wastewaters, other carrier materials, different operating conditions, etc.) in combination with more detailed analyses of the system performance (including measurements within the biofilm) in order to investigate the generality of the results as well as providing data for a more elaborate modelling study.

Some important areas for future research within the general field of modelling and control of wastewater treatment plants are suggested below. Many of these areas require an interdisciplinary approach, that is, joint efforts of many experts within different disciplines are needed to solve the fundamental problems.

- *Sewer system – WWT plant – receiving water interactions*: It is not sufficient to consider the operation of the WWT plants as a separate problem. The entire chain from the source of the wastewater to the receiving water must be taken into account in order to meet the demands of cost-effectiveness, good water quality and sustainable solutions in the future by utilizing all means of control and flexibility of the processes.
- *Performance indices*: An overall objective index would make it possible to evaluate the combined effects of both design and operation during the planning phase of new WWT plants, as well as allowing an objective comparison and evaluation of different operational strategies vs traditional expansion of plants already in operation. This could serve as an incentive for a higher degree of ICA at WWT plants and promote flexibility built into a plant.
- *New modelling approaches*: Grey-box models, neural networks and fuzzy logic are new approaches that need to be investigated.
- *Better overall control strategies*: The most important control aspect is not to optimize unit processes within a WWT plant but to optimize the performance of the entire plant.
- *Sensor development*: On-line sensors that are both robust and accurate, either in-situ (operating within the process) or in-line (operating in a side stream), are a necessity for any successful control of WWT processes.
- *Experimental design*: Any methods for identification and estimation require informative data to work well. Data quality can be improved dramatically by proper experimental design (equally important for both off-line and on-line methods).

- *Characterization problems*: Characterization of the active bacteria, flocs, biopolymers, phosphorus forms, dewatering properties, settling properties, floc-forming properties, etc., with respect to interactions with various wastewater components need to be improved in order to better understand the basic processes. Once we fully understand the processes, the performance of our wastewater treatment plants can be greatly improved.

Appendix A

Notation and Abbreviations

A	Cross-sectional area of a settler *or* surface area of a biofilm
AIC	The Akaike information criterion
AS	Activated sludge
BIC	The B information criterion
BOD	Biochemical oxygen demand
b	General decay rate coefficient
b_{A}	Decay rate coefficient for autotrophic organisms
b_{H}	Decay rate coefficient for heterotrophic organisms based on the death-regeneration hypothesis
b'_{H}	Traditional endogenous decay rate coefficient for heterotrophic organisms
b_{Z}	Decay rate coefficient for higher order organisms
$\boldsymbol{C}$	General matrix describing the model outputs
C	Carbon
COD	Chemical oxygen demand
D	Dispersion coefficient *or* height of settler above feed point *or* diffusion coefficient
DO	Dissolved oxygen
EBPR	Enhanced biological phosphorus removal
EKF	Extended Kalman filter
$\boldsymbol{F}, \boldsymbol{f}$	General non-linear vector functions
F	Godunov's flux term for the thickening zone
F/M	Ratio of substrate to biomass

$f(X)$	Total flux function in thickening zone
f_{ns}	Non-settleable fraction of the influent suspended solids concentration to the settler
f_{P}	Fraction of biomass yielding (inert) particulate products based on the death-regeneration hypothesis
f'_{P}	Fraction of biomass yielding (inert) particulate products based on the process of endogenous decay
$\boldsymbol{G}, \boldsymbol{g}$	General non-linear vector functions
G	Godunov's flux term for the clarification zone
GN	The Gauss-Newton optimization algorithm
$g(X)$	Total flux function in clarification zone
$\boldsymbol{H}, \boldsymbol{h}$	General non-linear vector functions
H	Hydrogen
H	Depth of settler below feed point
HRT	Hydraulic retention time
IAWQ	International Association on Water Quality (formerly IAWPRC)
ICA	Instrumentation, control and automation
i_{XB}	Mass N/mass COD in biomass
i_{XP}	Mass N/mass COD in products from biomass decay
$\hat{\boldsymbol{J}}$	General one-dimensional flux term
J	Loss function
J_{clar}	Special flux function for the clarification zone of the settler
J_{diff}	Flux due to an effective diffusion process
J_{dn}	Downward flux of SS due to downward bulk flow
J_{lim}	Limiting solids flux
J_s	Solids flux due to gravity settling
J_{up}	Upward flux of SS due to upward bulk flow
$\boldsymbol{K}$	General observer gain matrix *or* Kalman filter gain matrix
$K_{\mathrm{B,A}}$	Half-saturation coefficient for consumption of autotrophs by higher order organisms

$K_{B,H}$	Half-saturation coefficient for consumption of heterotrophs by higher order organisms
K_{NH}	Ammonia half-saturation coefficient for autotrophs
K_{NO}	Nitrate half-saturation coefficient for denitrifying heterotrophs
$K_{O,A}$	Oxygen half-saturation coefficient for autotrophs
$K_{O,H}$	Oxygen half-saturation coefficient for heterotrophs
$K_{O,Z}$	Oxygen half-saturation coefficient for higher organisms
K_S	Half-saturation coefficient for heterotrophic organisms
K_X	Half-saturation coefficient for hydrolysis of slowly biodegradable substrate
k, k_1, k_2	Parameters used to exemplify settling velocity functions
k_a	Ammonification rate
k_{at}	Attachment rate coefficient
k_{de}	Detachment rate coefficient
k_h	Maximum specific hydrolysis rate
k_{shear}	Coefficient for modelling of the biofilm detachment rate
$k_{T^\circ C}$	Any kinetic model parameter at temperature T ˚C
L	Likelihood function
L_F	Biofilm thickness
l_c	Length of carrier particles for biofilm growth
MCRT	Mean cell residence time
MLSS	Mixed-liquor suspended solids
$M(\cdot)$	Function that computes the local minimizer of $f(X)$
m	Number of model parameters
max(a, b)	Maximum value of a and b
min(a, b)	Minimum value of a and b
mumax	Maximum specific growth rate, identical to $\hat{\mu}$
N	Nitrogen
NM	The Nelder-Mead modified simplex optimization algorithm
n	Number of measured data points

n, n_1, n_2, n_3, n_4	Parameters used to exemplify settling velocity function
n_c	Number of carrier particles in a biofilm reactor
n_X	Number of modelled particulate substances
O	Oxygen
ODE	Ordinary differential equation
OUR	Oxygen uptake rate
PDE	Partial differential equation
$\boldsymbol{p}$	Percentage vector describing the proportions of different COD components of the suspended solids in the settler
Q	Volumetric flow rate
Q_{in}	Influent volumetric flow rate to the WWT plant
Q_{intr}	Internal recirculation volumetric flow rate from the last aerobic reactor to the first anoxic reactor
Q_r	Recycled volumetric flow rate from the settler underflow to the first anoxic reactor
Q_w	Wastage volumetric flow rate from the settler underflow
$\boldsymbol{P}$	Variance vector of the estimator error
$\overline{\boldsymbol{P}}$	Stationary variance vector of the estimator error
P	Phosphorus
PHA	Poly-hydroxyalkanoate
$p(\cdot)$	Probability density
$\boldsymbol{R}_1$, $\boldsymbol{R}_2$, $\boldsymbol{R}_{12}$	Covariance matrices
R	Reaction rate affecting the solids concentration
RAS	Return activated sludge (recycled from settler underflow)
$\boldsymbol{r}$	Percentage vector describing the proportions of the different soluble components in the settler
$\hat{\boldsymbol{r}}$	Net production rates per unit time and unit length
r	General reaction rate factor *or* net production rate per unit time and unit volume (in biofilms)
r_A	Reaction rate factor for autotrophic bacteria
r_c	Radius of carrier particles for biofilm growth

r_{H}	Reaction rate factor for heterotrophic bacteria
r_h	Settler model parameter for hindered settling
r_p	Settler model parameter for low solids concentrations
r_{ε_l}	Volume production of free water volume
round(a)	Closest integer value of a
S	Concentration of soluble material (sometimes mass per unit water phase volume)
SBR	Sequencing batch reactor
SRT	Sludge retention time
SS	Suspended solids
SVI	Sludge volume index
S_{ALK}	Molar concentration of alkalinity
S_{I}	Concentration of soluble inert organic matter
S_{ND}	Concentration of soluble biodegradable organic nitrogen
S_{NI}	Concentration of inert soluble nitrogen
S_{NH}	Concentration of ammonia nitrogen
S_{NO}	Concentration of nitrate and nitrite nitrogen
S_{O}	Concentration of dissolved oxygen
S_{S}	Concentration of readily biodegradable substrate
S_{sat}	Growth saturation concentration
s	Source function for settler feed inlet
T	Temperature
TCA	Tricarboxylic acid
TKN	Total Kjeldahl nitrogen
TOC	Total organic carbon
TSS	Total suspended solids
t	Time
UCT	University of Cape Town
$\boldsymbol{u}$	General vector of model or process inputs
V	Volume
VFA	Volatile fatty acids

VSS	Volatile suspended solids
v	Velocity
v_0	Maximum (theoretical) settling velocity
v_0'	Maximum (practical) settling velocity
v_{at}	Attachment velocity
v_{de}	Detachment velocity
v_{dn}	Downward bulk fluid velocity
v_F	Displacement velocity of a biofilm due to growth
$v_{\text{floc,dn}}$	Total particle velocity in the thickening zone
v_L	Interface velocity of a biofilm
v_s	Settling velocity of suspended solids
v_{up}	Upward bulk fluid velocity
$\boldsymbol{W}$	Weight matrix for defining the loss function J
WAS	Wastage activated sludge (removed from settler underflow)
WWT	Wastewater treatment
X	Concentration of particulate material
$X_{\text{B,A}}$	Concentration of active autotrophic biomass
$X_{\text{B,H}}$	Concentration of active heterotrophic biomass
X_{COD}	Concentration of biodegradable organic substrate
X_c	Steady-state constant concentration of suspended solids in settler thickening zone
X_D	Concentration of SS at the bottom of the settler
X_G	Local minimizer of $g(X)$
X_g	Concentration larger than X_G satisfying $g(X_g) = g(X_G)$
X_H	Concentration of SS at the top of the settler
X_{high}	Upper limit concentration for maximum settling velocity
X_{I}	Concentration of particulate inert organic matter
X_{infl}	Concentration of suspended solids where the batch settling curve has an inflection point
X_{low}	Lower limit concentration for maximum settling velocity
X_M	Local minimizer of $f(X)$

X_m	Concentration less than X_M satisfying $f(X_m) = f(X_M)$ *or* the concentration of SS in the feed layer of the settler
$X_{\max}$	Maximum packing concentration of particulate material
$X_{\min}$	Minimum concentration of suspended solids
X_{NB}	Concentration of active biomass nitrogen
X_{ND}	Concentration of particulate biodegradable organic nitrogen
X_{NI}	Concentration of nitrogen associated with inert organic particulate matter
X_{NP}	Concentration of nitrogen associated with inert organic particulate products
X_{P}	Concentration of particulate products from biomass decay
X_{pp}	Concentration of primary particles
$X_{pp,\min}$	Minimum concentration of primary particles
X_{S}	Concentration of slowly biodegradable substrate
X_t	Threshold concentration of suspended solids for onset of hindered settling behaviour
X_{tan}	Concentration of suspended solids where the settling behaviour changes from discontinuous to continuous ($= X_M$)
X_{Z}	Concentration of higher order organisms (microfauna)
$\boldsymbol{x}$	General vector of state variables
$\hat{\boldsymbol{x}}$	General vector of estimated state variables
$\tilde{\boldsymbol{x}}$	Vector for the reconstruction error ($\boldsymbol{x} - \hat{\boldsymbol{x}}$)
Y	General yield coefficient
Y_{A}	Yield coefficient for autotrophic organisms
Y_{H}	Yield coefficient for heterotrophic organisms
Y_{Z}	Yield coefficient for higher order organisms
$\boldsymbol{y}$	General vector of measurements or model outputs
$y(z/L_F)$	Empirical function for modelling growth of microfauna
$Z(\cdot)$	Function that computes the local minimizer of $g(X)$
z	Space coordinate in a settler or a biofilm *or* height of a layer within the settler

Greek letters

α	Arbitrary vector
α	Parameter to describe the life strategy of microfauna
	General matrix describing the model inputs
γ	Sludge compaction ratio *or* proportion of biomass consumed by microfauna compared to the biomass transformed by traditional decay
Δ	Distance between two grid points
δ	The delta function
$\delta_{a,a}$, $\delta_{a,r}$, $\delta_{r,a}$, $\delta_{r,r}$	Model sensitivity functions
ε	General vector describing measurement noise
ε	Small positive parameter to include an entropy condition
ε_l	Volume fraction of the water phase in a biofilm reactor
ζ_k	Arrhenius correction coefficient for kinetic parameters
η_{g}	Correction factor for anoxic growth of heterotrophs
η_{h}	Correction factor for anoxic hydrolysis
θ	General vector of model parameters
θ_X	Sludge retention time, i.e., sludge age
μ	General growth rate function
$\hat{\mu}$	General maximum specific growth rate
$\hat{\mu}_{\mathrm{A}}$	Maximum specific growth rate for autotrophic organisms
$\hat{\mu}_{\mathrm{H}}$	Maximum specific growth rate for heterotrophic organisms
$\hat{\mu}_{\mathrm{Z}}$	Maximum specific growth rate for higher organisms
ρ	Density
$\hat{\rho}$	General one-dimensional density
σ	Standard deviation
τ	Length of a time step
υ	General vector describing process noise
Φ	General matrix describing the model states
Ψ	Extended flux function including outlets

Ω	Short-circuiting factor
Ω-function	Corrective function for effects of compression

Subscripts

B	Refers to a variable in the bulk water phase
e	Refers to a variable describing the settler effluent
F	Refers to a variable within the biofilm
f	Refers to a variable describing the settler influent (feed)
i	Refers to the i^{th} layer (or grid point) of a settler model starting from the top
j	Refers to the biological components of the IAWQ AS Model No. 1 or other similar types of biological models
k	Refers to a specific step
L	Refers to a variable in the liquid boundary layer
m	Refers to the feed layer (or grid point) of a settler model
n	Refers to the bottom layer (or grid point) of a settler model
S	Refers to a dissolved variable
u	Refers to a variable describing the settler underflow
X	Refers to a particulate variable

superscripts

j	Refers to a specific time step
k	Refers to a specific iteration
T	Matrix transposition

Appendix B

The IAWQ AS Model No. 1

The most widely used model today, describing the biological processes in wastewater treatment systems, is the IAWQ Activated Sludge Model No.1. It was presented in 1987 (Henze *et al.*, 1987) as a result of the work by the 'Task Group on Mathematical Modelling for Design and Operation of Biological Wastewater Treatment Systems' formed by the International Association on Water Quality (IAWQ) in 1983. The main goal was to review existing models and reach a consensus concerning the simplest model having the capability of realistic predictions of the performance of single-sludge systems carrying out carbon oxidation, nitrification and denitrification. Most alternative biological models available today are to a large extent influenced by the IAWQ AS Model No. 1.

The model is a highly mechanistic model where the major components of relevance and the most important biological processes have been identified. It is based on a COD balance of the system (oxygen is expressed as negative COD) and is usually presented in the matrix format suggested by Peterson (1965) using the notation recommended by Grau *et al.* (1982). The matrix representation allows rapid and easy recognition of the fate of each component. By moving down a column for a specific component, the full differential equation with all the biological processes may immediately be formulated and by moving across the matrix, the continuity of the model can easily be checked by calculating the sum of the stoichiometric coefficients.

Table B.1 (Next two pages) Process kinetics and stoichiometry for carbon oxidation, nitrification and denitrification, according to the IAWQ AS Model No. 1 (Henze *et al.*, 1987).

Component → i		1	2	3	4	5	6	7	8	9
j	Process ↓	S_I	S_S	X_I	X_S	$X_{B,H}$	$X_{B,A}$	X_P	S_O	S_{NO}
1	Aerobic growth of heterotrophs		$-\frac{1}{Y_H}$			1			$-\frac{1-Y_H}{Y_H}$	
2	Anoxic growth of heterotrophs		$-\frac{1}{Y_H}$			1				$-\frac{1-Y_H}{2.86Y_H}$
3	Aerobic growth of autotrophs						1		$-\frac{4.57}{Y_A}+1$	$\frac{1}{Y_A}$
4	'Decay' of heterotrophs				$1-f_P$	–1		f_P		
5	'Decay' of autotrophs				$1-f_P$		–1	f_P		
6	Ammonification of soluble organic nitrogen									
7	'Hydrolysis' of entrapped organics		1		–1					
8	'Hydrolysis' of entrapped organic nitrogen									
Observed Conversion Rates [$ML^{-3}T^{-1}$]		$r_i = \sum_j \nu_{ij}\rho_j$								
Stoichiometric Parameters: Heterotrophic yield: Y_H Autotrophic yield: Y_A Fraction of biomass yielding particulate products: f_P Mass N/Mass COD in biomass: i_{XB} Mass N/Mass COD in products from biomass: i_{XP}		Soluble inert organic matter [$M(COD)L^{-3}$]	Readily biodegradable substrate [$M(COD)L^{-3}$]	Particulate inert organic matter [$M(COD)L^{-3}$]	Slowly biodegradable substrate [$M(COD)L^{-3}$]	Active heterotrophic biomass [$M(COD)L^{-3}$]	Active autotrophic biomass [$M(COD)L^{-3}$]	Particulate products arising from biomass decay [$M(COD)L^{-3}$]	Oxygen (negative COD) [$M(-COD)L^{-3}$]	Nitrate and nitrite nitrogen [$M(N)L^{-3}$]

10 S_{NH}	11 S_{ND}	12 X_{ND}	13 S_{ALK}	Process Rate, ρ_j [ML^{-3}T^{-1}]
$-i_{XB}$			$-\frac{i_{XB}}{14}$	$\hat{\mu}_H\left(\frac{S_S}{K_S+S_S}\right)\left(\frac{S_O}{K_{O,H}+S_O}\right)X_{B,H}$
$-i_{XB}$			$\frac{1-Y_H}{14\cdot 2.86Y_H}-\frac{i_{XB}}{14}$	$\hat{\mu}_H\left(\frac{S_S}{K_S+S_S}\right)\left(\frac{K_{O,H}}{K_{O,H}+S_O}\right)\cdot\left(\frac{S_{NO}}{K_{NO}+S_{NO}}\right)\eta_g X_{B,H}$
$-i_{XB}-\frac{1}{Y_A}$			$-\frac{i_{XB}}{14}-\frac{1}{7Y_A}$	$\hat{\mu}_A\left(\frac{S_{NH}}{K_{NH}+S_{NH}}\right)\left(\frac{S_O}{K_{O,A}+S_O}\right)X_{B,A}$
		$i_{XB}-f_P i_{XP}$		$b_H X_{B,H}$
		$i_{XB}-f_P i_{XP}$		$b_A X_{B,A}$
1	–1		$\frac{1}{14}$	$k_a S_{ND} X_{B,H}$
				$k_h\frac{X_S/X_{B,H}}{K_X+(X_S/X_{B,H})}\left[\left(\frac{S_O}{K_{O,H}+S_O}\right)+\eta_h\left(\frac{K_{O,H}}{K_{O,H}+S_O}\right)\left(\frac{S_{NO}}{K_{NO}+S_{NO}}\right)\right]X_{B,H}$
	1	–1		$\rho_7(X_{ND}/X_S)$
$r_i=\sum_j \nu_{ij}\rho_j$				
NH_4+NH_3 nitrogen [M(N)L^{-3}]	Soluble biodegradable organic nitrogen [M(N)L^{-3}]	Particulate biodegradable organic nitrogen [M(N)L^{-3}]	Alkalinity – Molar units	Kinetic Parameters: Heterotrophic growth and decay: $\hat{\mu}_H$, K_S, $K_{O,H}$, K_{NO}, b_H Autotrophic growth and decay: $\hat{\mu}_A$, K_{NH}, $K_{O,A}$, b_A Correction factor for anoxic growth of heterotrophs: η_g Ammonification: k_a Hydrolysis: k_h, K_X Correction factor for anoxic hydrolysis: η_h

Appendix C

Reduced Order AS Models

In this appendix, two reduced order models for the activated sludge process are presented. They are capable of roughly describing carbon oxidation, nitrification and denitrification in a single-sludge system. The models are presented in the same matrix format as used for the IAWQ AS Model No.1 (see Appendix B) and the same notation is applied, whenever possible.

The models completely separate the process behaviour during anoxic and aerobic conditions in order to avoid using a large number of switching functions. The number of state variables, parameters and processes have been greatly reduced when compared with the IAWQ model. Model A requires ten parameters to be estimated and model B is based on the estimation of seven model parameters. The models do not provide complete descriptions of the process mechanisms and should not be used for detailed analyses of an AS process. For example, the total amount of sludge in the system cannot be predicted because the models do not include inert material. The models are assumed to be used for predictions with a time horizon in the range of hours, which means that variations of the total sludge mass are not important. For the same reason, the hydrolysis process is not considered essential for the model behaviour. Moreover, dissolved oxygen concentration is not included in the models as this parameter is assumed to be controlled separately. The models have been developed for the purpose of control and, therefore, one important aspect is that the model parameters can be estimated from available on-line measurements. Consequently, the models should be considered and evaluated as possible tools for future control applications and not as models providing realistic predictions of all internal mechanisms of the AS process.

Model A		Anoxic environment						Aerobic environment					
Component →	*i*	1	2	3	4	5	Process rate	1	2	3	4	5	Process rate
j	Process ↓	X_{COD}	S_{NH}	S_{NO}	$X_{B,H}$	$X_{B,A}$	ρ_j [ML^{-3}T^{-1}]	X_{COD}	S_{NH}	S_{NO}	$X_{B,H}$	$X_{B,A}$	ρ_j [ML^{-3}T^{-1}]
1	Growth of heterotrophs	$-\frac{1}{\boldsymbol{Y_H}}$	$-i_{XB}$	$-\frac{1-\boldsymbol{Y_H}}{2.86\boldsymbol{Y_H}}$	1		$\boldsymbol{r_H} X_{COD} X_{B,H}$	$-\frac{1}{\boldsymbol{Y_H}}$	$-i_{XB}$		1		$\boldsymbol{r_H} X_{COD} X_{B,H}$
2	Growth of autotrophs								$-i_{XB} - \frac{1}{\boldsymbol{Y_A}}$	$\frac{1}{\boldsymbol{Y_A}}$		1	$\boldsymbol{r_A} S_{NH} X_{B,A}$
3	Decay of heterotrophs	1	$-i_{XB}$		-1		$\boldsymbol{b_H} X_{B,H}$	1	$-i_{XB}$		-1		$\boldsymbol{b_H} X_{B,H}$
4	Decay of autotrophs	1	$-i_{XB}$			-1	$\boldsymbol{b_A} X_{B,A}$	1	$-i_{XB}$			-1	$\boldsymbol{b_A} X_{B,A}$
Conversion rates [ML^{-3}T^{-1}]		$r_i = \Sigma \nu_{ij} \rho_j$											
Stoichiometric parameters: Heterotrophic yield: Y_H Autotrophic yield: Y_A Mass N/mass COD in biomass: i_{XB}		Biodegradable organic matter [M(COD)L^{-3}]	Ammonia nitrogen [M(N)L^{-3}]	Nitrate nitrogen [M(N)L^{-3}]	Active heterotrophic biomass [M(COD)L^{-3}]	Active autotrophic biomass [M(COD)L^{-3}]	Parameters to estimate: Anoxic: r_H, Y_H, b_H, b_A Aerobic: r_H, r_A,Y_H, Y_A, b_H, b_A	Biodegradable organic matter [M(COD)L^{-3}]	Ammonia nitrogen [M(N)L^{-3}]	Nitrate nitrogen [M(N)L^{-3}]	Active heterotrophic biomass [M(COD)L^{-3}]	Active autotrophic biomass [M(COD)L^{-3}]	Kinetic parameters: Heterotrophic reaction rate: r_H Autotrophic reaction rate: r_A Heterotrophic decay rate: b_H Autotrophic decay rate: b_A

Table C.1 Matrix formulation of the reduced order model A. The parameters written in bold are assumed to be variable and should be identified from available on-line measurements (separately for the anoxic and aerobic zone).

Model B	Anoxic environment							Aerobic environment						
Component → i	1	2	3	4	5	Process rate		1	2	3	4	5	Process rate	
j Process ↓	X_{COD}	S_{NH}	S_{NO}	$X_{B,H}$	$X_{B,A}$	ρ_j [ML^{-3}T^{-1}]		X_{COD}	S_{NH}	S_{NO}	$X_{B,H}$	$X_{B,A}$	ρ_j [ML^{-3}T^{-1}]	
1 Growth of heterotrophs	$-\frac{1}{\boldsymbol{Y_H}}$	$-i_{XB}$	$-\frac{1-\boldsymbol{Y_H}}{2.86\boldsymbol{Y_H}}$	1		$\boldsymbol{r_H} X_{COD} X_{B,H}$		$-\frac{1}{\boldsymbol{Y_H}}$	$-i_{XB}$		1		$\boldsymbol{r_H} X_{COD} X_{B,H}$	
2 Growth of autotrophs									$-i_{XB} - \frac{1}{\boldsymbol{Y_A}}$	$\frac{1}{\boldsymbol{Y_A}}$		1	$\boldsymbol{r_A} S_{NH} X_{B,A}$	
3 Decay of heterotrophs	1	$-i_{XB}$		-1		$\boldsymbol{b} X_{B,H}$		1	$-i_{XB}$		-1		$\boldsymbol{b} X_{B,H}$	
4 Decay of autotrophs	1	$-i_{XB}$			-1	$\boldsymbol{b} X_{B,A}$		1	$-i_{XB}$			-1	$\boldsymbol{b} X_{B,A}$	
Conversion rates [ML^{-3}T^{-1}]	$r_i = \Sigma v_{ij} \rho_j$													
Stoichiometric parameters: Heterotrophic yield: Y_H Autotrophic yield: Y_A Mass N/mass COD in biomass: i_{XB}	Biodegradable organic matter [M(COD)L^{-3}]	Ammonia nitrogen [M(N)L^{-3}]	Nitrate nitrogen [M(N)L^{-3}]	Active heterotrophic biomass [M(COD)L^{-3}]	Active autotrophic biomass [M(COD)L^{-3}]	Parameters to estimate: Anoxic: r_H, Y_H Aerobic: r_H, r_A, Y_H, Y_A Common: b		Biodegradable organic matter [M(COD)L^{-3}]	Ammonia nitrogen [M(N)L^{-3}]	Nitrate nitrogen [M(N)L^{-3}]	Active heterotrophic biomass [M(COD)L^{-3}]	Active autotrophic biomass [M(COD)L^{-3}]	Kinetic parameters: Heterotrophic reaction rate: r_H Autotrophic reaction rate: r_A Heterotrophic and autotrophic decay rate: b	

Table C.2 Matrix formulation of the reduced order model B. The parameters written in bold are assumed to be variable and should be identified from available on-line measurements (b common for both the anoxic and aerobic zone).

Appendix D

The Simplex Algorithm

Many early methods for optimization of algebraic functions are based on rough ideas without much theoretical background, i.e., *ad hoc* methods. One such possible method implies generating a number of points at random within a certain region of the function space and selecting the one which gives the best function value over a large number of trials. Unfortunately, this type of method suffers from the 'curse of dimensionality' since the amount of effort required to solve actual problems goes up rapidly (typically as 2^n) as the number of degrees of freedom (n) increases.

The most successful of the methods, which merely compare function values, is the *simplex method.* The algorithm is still widely used [Fletcher, 1987]. A regular simplex means a set of $n+1$ equidistant points in $\mathbb{R}^n$, such as the triangle for $n=2$ and tetrahedron for $n=3$. The current information kept in the method is the coordinates of the $n+1$ points and their corresponding function values, i.e., a very limited amount of data.

On the first iteration of the simplex method the vertex at which the function value is largest is determined. The vertex is then reflected in the centroid of the other n vertices, thus forming a new simplex. The function value at this new vertex is evaluated and the process repeated. On iterations after the first it might appear that the newest vertex still has the largest function value in the new simplex, and to reflect this vertex would cause oscillations. Hence, the largest function value other than that at the newest vertex is subsequently used to decide which vertex to reflect. Ultimately this iteration will fail to make further process, so an additional rule has to be introduced. When a certain vertex i has been in the current simplex for more than a fixed number of iterations (M), then the simplex should be contracted by replacing the other vertices by new ones, half way along the edge to the vertex i. A suitable value of M is normally determined by the dimension of the problem.

The typical progress of the iterations is illustrated in Figure D.1 using a two dimensional example. Vertices 1, 2 and 3 form the initial simplex and increasing numbers indicate the new vertices added at each iteration. Note that vertex 7 has the largest function value for the simplex (4, 6, 7) but is not reflected immediately since it is the newest vertex in that simplex. When simplex (6, 9, 10) is reached, vertex 6 has been in the current simplex for four iterations and if M is assumed to equal 3.5, the simplex is contracted at this stage to the new simplex (6, 11, 12) and the iteration continues from this simplex. The algorithm will continue to reflect and contract the simplex until the required tolerance has been achieved.

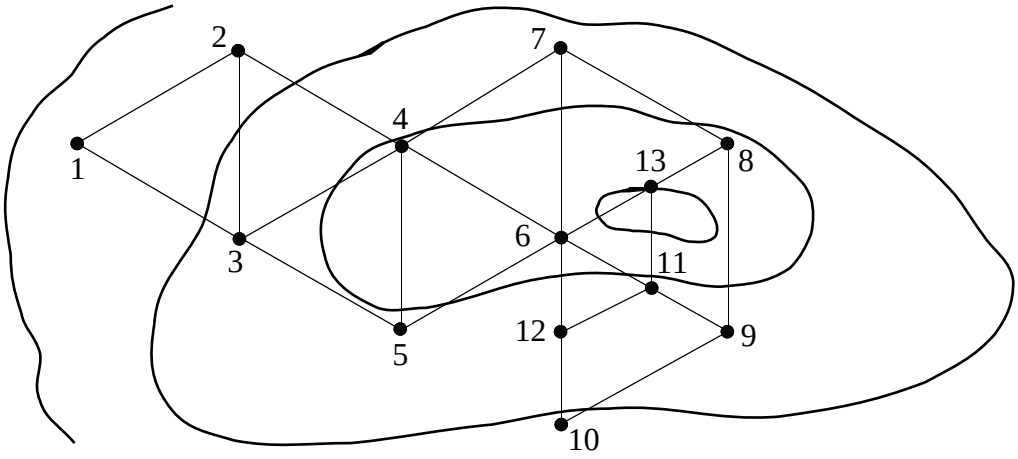


Figure D.1 The simplex algorithm for a two-dimensional problem.

The Nelder-Mead algorithm (Nelder and Mead, 1965) used in this work is a slightly modified simplex method, which allows irregular simplexes. Moreover, the distortions of the simplex are performed automatically in an attempt to take into account the local geometry of the function to optimize.

Due to the problems with the computational effort there is a practical limit to the size of systems which the method can be applied to. The convergence rate is slow but the algorithm is very robust and quite insensitive to noise. Often the method can be used in combination with more sophisticated ones. The simplex method is then applied in an early stage of the optimization in order to get the convergence going in the right direction and thereby producing suitable initial values for methods that converge faster. Such algorithms are usually less robust and likely to diverge if the initial estimates are far from the true ones (for example, the Gauss-Newton algorithm discussed in Section 2.6).

Appendix E

The Extended Kalman Filter

The technique of Kalman filters (Kalman, 1960) is a general filtering technique which can be applied to such problems as optimal estimation, prediction, noise filtering and stochastic control. Adaptive gain tuning capability is the characteristic of the Kalman filter. The method can also be applied to both stationary and non-stationary processes.

The following time-discrete linear system is assumed:

$$\begin{cases} \boldsymbol{x}(t_{k+1}) = \boldsymbol{\Phi x}(t_k) + \boldsymbol{\Gamma u}(t_k) + \upsilon(t_k) \\ \boldsymbol{y}(t_k) = \boldsymbol{C x}(t_k) + \varepsilon(t_k) \end{cases} \tag{E.1}$$

where υ and ε are Gaussian white noise processes with zero mean and the covariance matrices are given as

$$\begin{cases} E\left[\upsilon(t_k)\ \upsilon^T(t_k)\right] = \boldsymbol{R}_1 \\ E\left[\upsilon(t_k)\ \varepsilon^T(t_k)\right] = \boldsymbol{R}_{12} \\ E\left[\varepsilon(t_k)\ \varepsilon^T(t_k)\right] = \boldsymbol{R}_2 \end{cases} \tag{E.2}$$

Let the estimator have the form

$$\hat{\boldsymbol{x}}(t_{k+1}|t_k) = \boldsymbol{\Phi}\hat{\boldsymbol{x}}(t_k|t_{k-1}) + \boldsymbol{\Gamma u}(t_k) + \boldsymbol{K}(t_k)\left(\boldsymbol{y}(t_k) - \boldsymbol{C}\hat{\boldsymbol{x}}(t_k|t_{k-1})\right) \tag{E.3}$$

The reconstruction error, $\tilde{\boldsymbol{x}} = \boldsymbol{x} - \hat{\boldsymbol{x}}$, is governed by

$$\begin{aligned} \tilde{\boldsymbol{x}}(t_{k+1}|t_k) &= \boldsymbol{\Phi}\tilde{\boldsymbol{x}}(t_k|t_{k-1}) + \upsilon(t_k) - \boldsymbol{K}(t_k)\left(\boldsymbol{y}(t_k) - \boldsymbol{C}\hat{\boldsymbol{x}}(t_k|t_{k-1})\right) \\ &= \left(\boldsymbol{\Phi} - \boldsymbol{K}(t_k)\boldsymbol{C}\right)\tilde{\boldsymbol{x}}(t_k) + \upsilon(t_k) - \boldsymbol{K}(t_k)\varepsilon(t_k) \end{aligned} \tag{E.4}$$

The property of the noise is taken into account and the criterion is to minimize the variance of the estimator error, $\boldsymbol{P}(t_k)$, by determining the best gain matrix, $\boldsymbol{K}(t_k)$. $\boldsymbol{P}(t_k)$ is defined as

$$\boldsymbol{P}(t_k) = E\left[\left(\tilde{\boldsymbol{x}}(t_k) - E\left[\tilde{\boldsymbol{x}}(t_k)\right]\right)\left(\tilde{\boldsymbol{x}}(t_k) - E\left[\tilde{\boldsymbol{x}}(t_k)\right]\right)^T\right] \tag{E.5}$$

The mean value of $\tilde{\boldsymbol{x}}$ is obtained from equation (E.4) as

$$E\left[\tilde{\boldsymbol{x}}(t_{k+1})\right] = \left(\boldsymbol{\Phi} - \boldsymbol{K}(t_k)\boldsymbol{C}\right)E\left[\tilde{\boldsymbol{x}}(t_k)\right] \tag{E.6}$$

If $E[\boldsymbol{x}(0)] = m_0$ then the mean value of the reconstruction error is zero for times $t_k \geq 0$, independent of $\boldsymbol{K}$ if $E[\hat{\boldsymbol{x}}(0)] = m_0$. This is assumed to be true and equation (E.4) yields

$$\begin{aligned}\boldsymbol{P}(t_{k+1}) = E\left[\tilde{\boldsymbol{x}}(t_{k+1})\,\tilde{\boldsymbol{x}}^T(t_{k+1})\right] &= \left(\boldsymbol{\Phi} - \boldsymbol{K}(t_k)\boldsymbol{C}\right)\boldsymbol{P}(t_k)\left(\boldsymbol{\Phi} - \boldsymbol{K}(t_k)\boldsymbol{C}\right)^T \\ &+ \boldsymbol{R}_1 + \boldsymbol{K}(t_k)\boldsymbol{R}_2\boldsymbol{K}(t_k)^T - 2\boldsymbol{K}(t_k)\boldsymbol{R}_{12}\end{aligned} \tag{E.7}$$

The criterion is to minimize the scalar $\alpha^T\boldsymbol{P}(t_{k+1})\,\alpha$ where α is an arbitrary vector, by choosing the best possible $\boldsymbol{K}(t_k)$. If the criterion is developed using (E.7), two terms occur according to

$$\begin{aligned}&\alpha^T\boldsymbol{P}(t_{k+1})\alpha \\ &= \alpha^T\left\{\boldsymbol{\Phi}\boldsymbol{P}(t_k)\boldsymbol{\Phi}^T + \boldsymbol{R}_1 - \boldsymbol{\Phi}\boldsymbol{P}(t_k)\boldsymbol{C}^T\left(\boldsymbol{R}_2 + \boldsymbol{C}\boldsymbol{P}(t_k)\boldsymbol{C}^T\right)^{-1}\boldsymbol{C}\boldsymbol{P}(t_k)\boldsymbol{\Phi}^T\right\}\alpha \\ &+ \alpha^T\left\{\left[\boldsymbol{K}(t_k) - \boldsymbol{\Phi}\boldsymbol{P}(t_k)\boldsymbol{C}^T\left(\boldsymbol{R}_2 + \boldsymbol{C}\boldsymbol{P}(t_k)\boldsymbol{C}^T\right)^{-1}\right]\left[\boldsymbol{R}_2 + \boldsymbol{C}\boldsymbol{P}(t_k)\boldsymbol{C}^T\right]\right. \\ &\left.\left[\boldsymbol{K}(t_k) - \boldsymbol{\Phi}\boldsymbol{P}(t_k)\boldsymbol{C}^T\left(\boldsymbol{R}_2 + \boldsymbol{C}\boldsymbol{P}(t_k)\boldsymbol{C}^T\right)^{-1}\right]^T\right\}\alpha\end{aligned} \tag{E.8}$$

The first term of (E.8) is independent of $\boldsymbol{K}(t_k)$ whereas the second term is determined by $\boldsymbol{K}(t_k)$. If $\boldsymbol{K}(t_k)$ is chosen such that the second part of (E.8) is zero, a minimum is obtained. The following two equations result:

$$\boldsymbol{K}(t_k) = \left(\boldsymbol{\Phi}\boldsymbol{P}(t_k)\boldsymbol{C}^T + \boldsymbol{R}_{12}\right)\left(\boldsymbol{C}\boldsymbol{P}(t_k)\boldsymbol{C}^T + \boldsymbol{R}_2\right)^{-1} \tag{E.9}$$

$$\boldsymbol{P}(t_{k+1}) = \boldsymbol{\Phi}\boldsymbol{P}(t_k)\boldsymbol{\Phi}^T + \boldsymbol{R}_1 - \boldsymbol{K}(t_k)\left(\boldsymbol{C}\boldsymbol{P}(t_k)\boldsymbol{C}^T + \boldsymbol{R}_2\right)\boldsymbol{K}^T(t_k) \tag{E.10}$$

The reconstruction defined be equations (E.3), (E.9) and (E.10) is called the Kalman filter. The main difficulty is usually to determine the proper covariance matrices of (E.2) and to initially select a suitable variance matrix $\boldsymbol{P}(0)$. Note that $\boldsymbol{P}(t_k)$ does not depend on the observations. Thus, the gain can be precomputed in forward time and stored in the computer. Other methods with similar characteristics are the recursive instrumental variable method and the recursive prediction error method (Söderström and Stoica, 1989).

Extended Kalman filters (EKF) are a logical generalisation of linear Kalman filters for the case where the system dynamics vary with operating and control points in non-linear systems. The first step of the generalisation is to exchange the linear process model (E.1) for a non-linear one:

$$\begin{cases} \boldsymbol{x}(t_{k+1}) = \boldsymbol{f}(\boldsymbol{x}(t_k), \boldsymbol{u}(t_k)) + \boldsymbol{\varepsilon}(t_k) \\ \boldsymbol{y}(t_k) = \boldsymbol{h}(\boldsymbol{x}(t_k)) + \boldsymbol{\varepsilon}(t_k) \end{cases} \tag{E.11}$$

where $\boldsymbol{f}$ and $\boldsymbol{h}$ represent general non-linear vector functions.

The second step is to use a linearisation of the process dynamics in order to minimize the effect of process and measurement noise. This linearisation is performed around the current state estimates, $\hat{\boldsymbol{x}}(t_k)$, on-line. The main elements of an extended Kalman filter are thus a description of the process dynamics (and a linearized version of it) and a noise model.

There are three different kinds of extended Kalman filters; discrete EKF, continuous EKF and continuous-discrete EKF. The continuous-discrete EKF uses a continuous time update of the non-linear observer while it employs a discrete measurement update. Such a filter is often a good approach because the model can be kept in the traditional continuous form while the measurements are most conveniently digitized using a zero-order hold network. A continuous-discrete EKF was used for the on-line estimations in Chapter 4.

As was described in Section 4.2, the calculations are often divided in a *prediction* and a *correction* phase. If $\boldsymbol{R}_{12}$ for simplicity is assumed to equal zero, the EKF can be formulated in a straightforward way. The predictor phase includes the following calculations:

$$\hat{\boldsymbol{x}}(t_{k+1}|t_k) = \boldsymbol{f}(\hat{\boldsymbol{x}}(t_k|t_k), \boldsymbol{u}(t_k)) \tag{E.12}$$

$$\boldsymbol{P}\left(t_{k+1}|t_k\right)=\boldsymbol{F}\left(t_k\right)\boldsymbol{P}\left(t_k|t_k\right)\boldsymbol{F}^T\left(t_k\right)+\boldsymbol{R}_1 \tag{E.13}$$

and the corrector phase calculations include

$$\hat{\boldsymbol{x}}\left(t_{k+1}|t_{k+1}\right)=\hat{\boldsymbol{x}}\left(t_{k+1}|t_k\right)+\boldsymbol{K}\left(t_{k+1}\right)\left[\boldsymbol{y}\left(t_{k+1}\right)-\boldsymbol{h}\left(\hat{\boldsymbol{x}}\left(t_{k+1}|t_k\right)\right)\right] \tag{E.14}$$

$$\boldsymbol{P}\left(t_{k+1}|t_{k+1}\right)=\boldsymbol{P}\left(t_{k+1}|t_k\right)-\boldsymbol{K}\left(t_{k+1}\right)\boldsymbol{H}\left(t_{k+1}\right)\boldsymbol{P}\left(t_{k+1}|t_k\right) \tag{E.15}$$

$$\begin{aligned}&\boldsymbol{K}\left(t_{k+1}\right)\\&=\boldsymbol{P}\left(t_{k+1}|t_k\right)\boldsymbol{H}^T\left(t_{k+1}\right)\left[\boldsymbol{H}\left(t_{k+1}\right)\boldsymbol{P}\left(t_{k+1}|t_k\right)\boldsymbol{H}^T\left(t_{k+1}\right)+\boldsymbol{R}_2\right]^{-1}\end{aligned} \tag{E.16}$$

where $\boldsymbol{F}(t_k)$ and $\boldsymbol{H}(t_k)$ corresponds to the Jacobian matrices of $\boldsymbol{f}(\cdot)$ and $\boldsymbol{h}(\cdot)$, respectively. The Jacobians are defined as

$$\boldsymbol{F}\left(t_k\right)=\left.\frac{\partial \boldsymbol{f}\left(\boldsymbol{x}\left(t_k\right),\boldsymbol{u}\left(t_k\right)\right)}{\partial \boldsymbol{x}\left(t_k\right)}\right|_{\boldsymbol{x}\left(t_k\right)=\hat{\boldsymbol{x}}\left(t_k\right)} \tag{E.17}$$

$$\boldsymbol{H}\left(t_k\right)=\left.\frac{\partial \boldsymbol{h}\left(\boldsymbol{x}\left(t_k\right)\right)}{\partial \boldsymbol{x}\left(t_k\right)}\right|_{\boldsymbol{x}\left(t_k\right)=\hat{\boldsymbol{x}}\left(t_k\right)} \tag{E.18}$$

The significant real-time computational burden imposed by the use of extended Kalman filters have motivated the search for more simple estimators, which can retain the same robustness characteristics of the full EKF. The *constant* gain EKF is one such simplification. In this case a constant gain matrix is achieved for a selected operating point $\boldsymbol{x}_0$ of the system according to

$$\boldsymbol{K}=\left(\boldsymbol{\Phi}\bar{\boldsymbol{P}}\boldsymbol{C}^T+\boldsymbol{R}_{12}\right)\left[\boldsymbol{C}\bar{\boldsymbol{P}}\boldsymbol{C}^T+\boldsymbol{R}_2\right]^{-1} \tag{E.19}$$

where $\bar{\boldsymbol{P}}$ is obtained as the positive semi-definite solution of the stationary Riccati equation:

$$\bar{\boldsymbol{P}}=\boldsymbol{\Phi}\bar{\boldsymbol{P}}\boldsymbol{\Phi}^T+\boldsymbol{R}_1-\left(\boldsymbol{\Phi}\bar{\boldsymbol{P}}\boldsymbol{C}^T+\boldsymbol{R}_{12}\right)\left[\boldsymbol{C}\bar{\boldsymbol{P}}\boldsymbol{C}^T+\boldsymbol{R}_2\right]^{-1}\left(\boldsymbol{\Phi}\bar{\boldsymbol{P}}\boldsymbol{C}^T+\boldsymbol{R}_{12}\right)^T \tag{E.20}$$

Such a filter maintains its robust behaviour even when exposed to significantly varying signals (Hendricks, 1992). For practical reasons this approach was used in the study presented in Chapter 4.

The methods presented in this appendix not only hold for state estimation but also for simultaneously state and parameter estimation. The equations given are still valid although $\hat{\boldsymbol{x}}$ becomes a generalized state vector which includes both the unknown state variables and the uncertain model parameters. The parameter vector is modelled as a random walk or drift. The difficulty of determining the proper covariance matrices (E.2) is, however, more emphasized. $\boldsymbol{R}_1$ is used to describe how fast the different components of the parameter vector are expected to vary.

In order to avoid various numerical problems (due to accumulation of rounding errors in the covariance matrix) when applying the Kalman filter, several modifications have been proposed. Bierman (1977) suggests replacing the covariance matrix update with a stabilized one which will enhance the numerical stability. However, the method does not guarantee numerical stability and positive definiteness of $\boldsymbol{P}$. Thus, methods for updating the square root of the covariance matrix have been proposed in order to ensure that $\boldsymbol{P}$ is positive definite. These methods are often based on Cholesky factorization (e.g., Ljung, 1987). One such algorithm is presented in Thornton and Bierman (1980). The above modifications will, however, increase the computational requirements.

A more detailed and theoretical derivation of the Kalman filter and its variants is given, for example, by Ljung and Söderström (1983) and Ogata (1987).

Appendix F

Simulation Environment

When working with mathematical modelling, parameter identification and model validation it is important to have access to a good simulation environment – both hardware and software. All the computations included in this work have been performed using a single-processor Sun™ SPARCstation 10 with the simulation programs Simnon™, Matlab/Simulink™ and AQUASIM.

Simnon (SSPA Systems, 1991) is designed for solving ordinary differential and difference equations and for simulating dynamic systems. Numerical integration routines are used to simulate differential equations and difference equations are solved by iteration. No symbolic analysis of the systems is possible. The systems may be described as an interconnection of subsystems (promoting a hierarchical system description) which may be either in continuous or discrete time. The user interacts with the program by typing commands (a graphical interface is available as an add-on product). Parameters, initial conditions and system descriptions can be modified interactively and the results are graphically or numerically displayed on the screen. A built-in macro facility allows the user to create his own set of commands. The allowed complexity of the developed models is limited, although this is seldom a problem (the maximum number of state variables is 300). The program also has real-time capabilities, i.e., data can be transferred on-line from and to a real process connected to the I/O-devices of the computer. As an example, the straightforward text file for simulating the simple model (2.13) is illustrated in Figure F.1.

```
CONTINUOUS SYSTEM  model_2_13
″Model of the Monod growth equation in a single-
″substrate (S)/single-organism (X) batch reactor
″with no other growth limitations.

STATE  S X
DER    dS dX

TIME   t

″Values for the model parameters
mumax : 6          ″maximum specific growth rate
Ks    : 10         ″half-saturation coefficient
b     : 0.48       ″decay rate factor
Y     : 0.66       ″yield factor

″Initial values for the state variables
S : 100            ″substrate concentration
X : 2              ″organism concentration

″Dynamic equations
dS = -mu*X/Y
dX = (mu - b)*X
mu = mumax*S/(Ks + S)  ″Monod growth rate

END
```

Figure F.1 Simnon text file describing the simple model (2.13).

Simulink (MathWorks, 1995) is an interactive system for simulating dynamic systems. It is a graphical, mouse-driven program that allows the user to model a system by drawing a block diagram in a graphical editor and manipulating it dynamically. It handles linear, non-linear, continuous-time, discrete-time, multi-variable and multi-rate systems. A large number of predefined building blocks is included in the program and it is easy for the user to extend this library with blocks of his own. Hierarchical models are recommended since blocks may include other blocks and allows for graphical ‘information zooming’. Results are numerically and graphically available in numerous ways. The block diagram for describing the small model (2.13) is shown in Figure F.2.

The major advantage of Simulink is the fact that it is an integrated part (toolbox) of the complete Matlab™ (*mat*rix *lab*oratory) computing environment (MathWorks, 1992). Matlab is an interactive system whose basic data element is a matrix that does not require dimensioning. It includes a large library of predefined functions and a simple way for the

user to define functions of his own expressed as they are written mathematically – without any traditional programming. However, it is possible to include models written in C (using a special format) into the program, as well as creating stand-alone applications. Moreover, data can be transferred on-line in and out of the program from a real process connected to the I/O-devices of the computer, allowing for real-time analysis and on-line control.

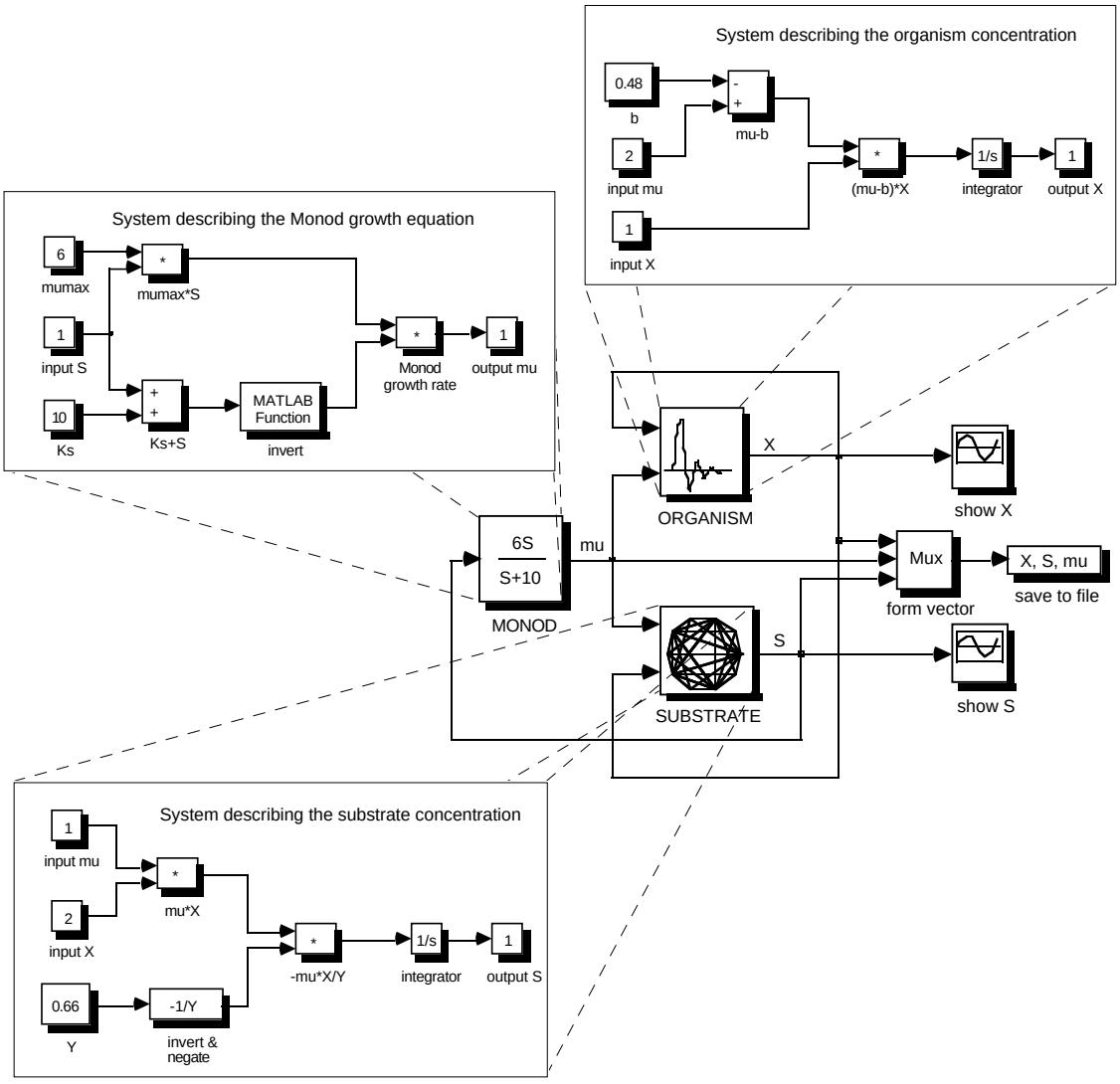


Figure F.2 Simulink graphical block diagram of the simple model (2.13).

A large number of application-specific toolboxes that extend the Matlab environment in order to solve particular classes of problems are available. These toolboxes include signal processing, control system design, system identification, optimization, neural networks, fuzzy logic, statistics, partial differential equations, symbolic math, etc. Altogether this means that the Simulink user not only has the possibility to perform simulations but an

enormous capability to manipulate and further validate the results. All this power is available at the user's fingertips in one integrated environment.

AQUASIM (Reichert and Ruchti, 1994) is an interactive system for simulating and analysing the dynamics of aquatic systems. It is a graphical, mouse-driven program that includes a number of predefined compartment models (at present: mixed reactors, biofilm reactors and river section). The introduction of compartment models limits the generality of the program but allows the selection of efficient numerical algorithms according to the type of partial differential equation used to describe the transport mechanisms. The internal dynamic processes within a compartment are formulated by the user without any restrictions. Different compartments can be combined by using different types of links. The program also contains built-in tools for identifiability analysis (by sensitivity analysis), parameter estimation and uncertainty analysis. Results are presented graphically (though not on-line and only as traditional two-dimensional plots) or saved as ASCII files. The program does not have any real-time capabilities but is only intended for off-line use.

All three programs are available for a large number of computer platforms (although Simnon is primarily aimed for systems running MS-Windows). The compatibility between the programs is also quite good. It is easy to exchange data files since all programs accept simple ASCII files in tabular form. Linear, time-invariant models can also be directly exported from Matlab into Simnon by a special translation script. Therefore, the user can combine the three programs and take advantage of their respective strongpoints and use them as a combined model building, simulation, data analysis and data manipulation software environment.

Bibliography

[1] Akaike, H. (1981), "Modern Development of Statistical Methods". *Trends and Progress in System Identification*, (P. Eykhoff ed.), Pergamon Press, Oxford, Great Britain, pp. 169-184.

[2] Albertson, O.E. (1991), "Bulking Sludge Control – Progress, Practice and Problems". *Wat. Sci. Tech.*, vol. 23, pp. 835-846.

[3] Anderson, N.E. (1945), "Design of Final Settling Tanks for Activated Sludge". *Sewage Works J.*, vol. 17, no. 1, pp. 50-63.

[4] Andersson, B. (1990), "Tentative Nitrogen Removal with Fixed Bed Processes in Malmö Sewage Treatment Plant". *Wat. Sci. Tech.*, vol. 22, no. 1-2, pp. 239-250.

[5] Andrews, J.F., ed. (1992), *Dynamics and Control of the Activated Sludge Process*, Technomic Publishing Co., Lancaster, Pennsylvania, USA.

[6] APHA, AWWA, WPCF (1985a), "Oxygen Demand (chemical) – Closed Reflux, Colorimetric Method". *Standard Methods for the Examination of Water and Wastewater* (16th ed.), American Public Health Association, Washington DC, USA, pp. 537-538.

[7] APHA, AWWA, WPCF (1985b), "Total Suspended Solids Dried at 103-105˚C/Fixed and Volatile Solids Ignited at 550˚C". *Standard Methods for the Examination of Water and Wastewater* (16th ed.), American Public Health Association, Washington DC, USA, pp. 537-538.

[8] Arden, E., Lockett, W.T. (1914), "Experiments on the Oxidation of Sewage without the Aid of Filters". *J. Soc. Chem. Ind.*, vol. 33, p. 523.

[9] Arndt, H. (1993), "Rotifers as Predators on Components of the Microbial Web (Bacteria, Heterotrophic Flagellates, Ciliates) – Review". *Hydrobiology*, no. 255-256, pp. 231-246.

[10] Arvin, E., Harremoës, P. (1990), "Concepts and Models for Biofilm Reactor Performance". *Wat. Sci. Tech.*, vol. 22, pp. 171-192.

[11] Aspegren, H., Andersson, B., Olsson, G., Jeppsson, U. (1990), "Practical Full Scale Experiences of the Dynamics of Biological Nitrogen Removal". *Instrumentation, Control and Automation of Water and Wastewater Treatment and Transport Systems,* (R. Briggs ed.), Pergamon Press, London, Great Britain, pp. 283-290.

[12] Aspegren, H. (1995), *Evaluation of a High Loaded Activated Sludge Process for Biological Phosphorus Removal.* Ph.D. dissertation, Dept of Water and Environmental Eng., Lund Inst. of Tech., Lund, Sweden.

[13] Åström, K.J., Wittenmark, B. (1990), *Computer-Controlled Systems, Theory and Design*, (2nd ed.), Prentice Hall, Inc., Englewood Cliffs, New Jersey, USA.

[14] Atkinson, B., Davies, I.J. (1974), "The Overall Rate of Substrate Uptake (Reaction) by Microbial Films, Part I. A Biological Rate Equation". *Trans. Inst. Chem. Engrs.*, vol. 52, pp. 248-259.

[15] Auzerais, F.M., Jackson, R., Russel, W.B. (1988), "The Resolution of Shocks and the Effects of Compressible Sediments in Transient Settling". *J. Fluid Mech.*, vol. 195, pp. 437-462.

[16] Ayesa, E., Florez, J., García-Heras, J.L., Larrea, L. (1991), "State and Coefficients Estimation for the Activated Sludge Process Using a Modified Kalman Filter Algorithm". *Wat. Sci. Tech.*, vol. 24, no. 6, pp. 235-247.

[17] Ayesa, E., Carstensen, J., Jeppsson, U., Vanrolleghem, P. (1996), "Identification of the Dynamic Processes in WWTP and Control of WWTP". *COST 682: Environment,* (D. Dochain, P. Vanrolleghem, M. Henze eds.), European Commission, Directorate-General XII: Science, Research and Development, Luxembourg, pp. 89-104.

[18] Banerji, S.K., Ewing, B.B., Engelbrecht, R.S., Speece, E.E. (1968), "Kinetics of Removal of Starch in Activated Sludge". *J. Water Pollution Control Fed.*, vol. 40, pp. 161-173.

[19] Barnard, J.L. (1973), "Biological Denitrification". *J. Water Pollution Control Fed.*, vol. 34, p. 920.

[20] Barnard, J.L. (1975), "Biological Nutrient Removal without the Addition of Chemicals". *Wat. Res.*, vol. 9, p. 485.

[21] Barnard, J.L. (1976), "A Review of Biological Phosphorus Removal in the Activated Sludge Process". *Water S.A.*, vol. 2, no. 3, p. 126.

[22] Barnard, J.L. (1983), "Background to Biological Phosphorus Removal". *Wat. Sci. Tech.*, vol. 15, p. 1.

[23] Barnes, D., Bliss, P.J. (1983), *Biological Control of Nitrogen in Wastewater Treatment,* University Press, Cambridge, Great Britain.

[24] Barton, N.G., Li, C.-H., Spencer, J. (1992), "Control of a Surface of Discontinuity in Continuous Thickeners". *J. Austral. Math. Soc. Ser.*, vol. B33, pp. 269-289.

[25] Bastin, G., Dochain, D. (1990), *On-Line Estimation and Adaptive Control of Bioreactors*, Elsevier Science Publishers B.V., Amsterdam, The Netherlands.

[26] Batchelor, B. (1982), "Kinetic Analysis of Alternative Configurations for Single-Sludge Nitrification-Denitrification". *J Water Pollution Control Fed.*, vol. 54, pp. 1493-1504.

[27] Beck, M.B. (1986), "Identification, Estimation and Control of Biological Wastewater Treatment Processes". *Proc. Inst. Elec. Eng.*, vol. 133, pp. 254-264.

[28] Beck, M.B. (1987), "Water Quality Modeling: A Review of the Analysis of Uncertainty". *Wat. Resourc. Res.*, vol. 23, no. 8, pp. 1393-1442.

[29] Beck, M.B. (1991), "Principles of Modelling". *Wat. Sci. Tech.*, vol. 24, no. 6, pp. 1-8.

[30] Benefield, L., Molz, F. (1984), "A Model for the Activated Sludge Process which Considers Wastewater Characteristics, Floc Behaviour, and Microbial Population". *Biotech. Bioeng.*, vol. 26, pp. 352-361.

[31] Benefield, L., Reed, R.B. (1985), "An Activated Sludge Model which Considers Toxicant Concentration: Simulation and Sensitivity Analysis". *Appl. Math. Modelling*, vol. 9, pp. 454-465.

[32] Berthouex, P.M., Fan, R. (1986), "Evaluation of Treatment Plant Performance: Causes, Frequency and Duration of Upsets". *J. Water Pollution Control Fed.*, vol. 58, pp. 368-375.

[33] Bhat, N., McAvoy, T. (1990), "Use of Neural Nets for Dynamic Modelling and Control of Chemical Process Systems". *Computer and Chem. Eng.*, vol. 14, no. 5, pp. 573-583.

[34] Bierman, G. (1977), *Factorization Methods for Discrete Sequential Estimation*, Academic Press, London, Great Britain.

[35] Billmeier, E. (1978), "Verbesserte Bemessungsvorschläge für horizontal durchströmte Nackklärbecken von Belebungsanlagen". *Berichte aus Wassergütewirtschaft und Gesundheitsingenieurwesen, vol. 21*, Institut für Bauingenieurwesen, Technische Universität, München, Germany.

[36] Bird, R.B., Stewart, W.E., Lightfoot, E.N. (1976), *Transport Phenomena*, John Wiley and Sons, Ltd., New York, New York, USA.

[37] Birkes, D., Dodge, Y (1993), *Alternative Methods of Regression*, John Wiley and Sons, Ltd., New York, New York, USA.

[38] Bishop, P.L., Kinner, N.E. (1986), "Aerobic Fixed-Film Processes". Chapter 3 in: *Biotechnology, vol. 8*, (H.J. Rehm, G. Reed eds.), VCH Verlagsgesellschaft, Weinheim, Germany.

[39] Bishop, P.L., Zhang, T.C., Fu, Y.-C. (1995), “Effects of Biofilm Structure, Microbial Distribution and Mass Transport on Biodegradation Processes”. *Wat. Sci. Tech.*, vol. 31, no. 1, pp. 143-152.

[40] Blackman, F. (1905), “Optima and Limiting Factors”. *Ann. Bot.*, vol. 19, pp. 281-295.

[41] Bloem, J., Starink, M., Bär-Gilissen, M.J.B., Cappenberg, T.E. (1988), “Protozoan Grazing, Bacterial Activity, and Mineralization in Two-Stage Continuous Culture”. *J. Appl. Environ. Microbiol.*, vol. 54, pp. 3113-3121.

[42] Bloodgood, D.E. (1938), “Studies of Activated Sludge Oxidation at Indianapolis”. *Sewage Works J.*, vol. 10, no. 1, p. 27.

[43] Boger, Z. (1992), “Application of Neural Networks to Water and Wastewater Treatment Plant Operation”. *ISA Transactions*, vol. 31, pp. 25-33.

[44] Bohlin, T. (1991), *Interactive System Identification: Prospects and Pitfalls*, Springer-Verlag, Berlin, Germany.

[45] Boller, M., Gujer, W. (1986), “Nitrification in Tertiary Trickling Filters Followed by Deep-Bed Filters”. *Wat. Res.*, vol. 20, no. 11, pp. 1363-1373.

[46] Bonhomme, M., Rogalla, F., Boisseau, G., Sibony, J. (1990), “Enhancing Nitrogen Removal in Activated Sludge with Fixed Biomass”. *Wat. Sci. Tech.*, vol. 22, no. 1-2, pp. 127-135.

[47] Bouwer, E.J., McCarty, P.L. (1985), “Utilization Rates of Trace Halogenated Organic Compounds in Acetate-Grown Biofilms”. *Biotech. Bioeng.*, vol. 27, pp. 1564-1571.

[48] Bouwer, E.J. (1987), “Theoretical Investigation of Particle Deposition in Biofilm Systems”. *Wat. Res.*, vol. 21, no. 12, pp. 1489-1498.

[49] Boyle, W.C., Berthouex, P.M. (1974), “Biological Wastewater Treatment Model Building Fits and Misfits”. *Biotech. Bioeng.*, vol. 16, pp. 1139-1159.

[50] Bretscher, U., Krebs, P., Hager, W.H. (1992), "Improvement of Flow in Final Settling Tanks". *J. Env. Eng.*, ASCE, vol. 118, no. 3, pp. 307-321.

[51] Bryant, J.O. (1972), *Continuous Time Simulation of the Conventional Activated Sludge Wastewater Renovation System*. Ph.D. dissertation, Clemson University, Clemson, South Carolina, USA.

[52] Busby, J.B., Andrews, J.F. (1975), "Dynamic Modeling and Control Strategies for the Activated Sludge Process". *J. Water Pollution Control Fed.*, vol. 47, no. 5, pp. 1055-1080.

[53] Buswell, A.M., Long, H.L. (1923), "Microbiology and Theory of Activated Sludge". *J. Am. Pollut. Wks. Assn.*, vol. 10, no.2, p. 309.

[54] Camp, T.R. (1936), "A Study of the Rational Design of Settling Tanks". *Sewage Works J.*, vol. 8, no. 5, pp. 742-758.

[55] Camp, T.R. (1946), "Sedimentation and the Design of Settling Tanks". *Transactions*, ASCE, vol. 111, paper no. 2258, pp. 895-958.

[56] Capodaglio, A., Novotny, V., Fortina, L. (1992), "Modelling Wastewater Treatment Plants through Time Series Analysis". *Environmetrics*, vol. 3, no. 1, pp. 99-120.

[57] Carlsson, B., Wigren, T. (1993), "On-Line Identification of the Dissolved Oxygen Dynamic in an Activated Sludge Process". *Proc. 12th World Congress of IFAC*, Sydney, Australia, 19-23 July.

[58] Carlsson, B., Lindberg, C.F., Hasselblad, S., Xu, S. (1994), "On-Line Estimation of the Respiration Rate and the Oxygen Transfer Rate at Kungsängens Wastewater Treatment Plant in Uppsala". *Wat. Sci. Tech.*, vol. 30, no. 4, pp. 255-263.

[59] Carstensen, J. (1994), *Identification of Wastewater Processes*. Ph.D. dissertation, Inst. of Mathematical Statistics and Operations Research (IMSOR), Technical University of Denmark, Lyngby, Denmark.

[60] Carstensen, J., Vanrolleghem, P., Ayesa, E., Jeppsson, U., Urrutikoetxea, A., Vanderhaegen, B. (1996), "Objective Functions for Wastewater Treatment Design and Operation". *COST 682: Environment*, (D. Dochain, P. Vanrolleghem, M. Henze eds.), European Commission, Directorate-General XII: Science, Research and Development, Luxembourg, pp. 105-108.

[61] Cech, J.S., Chudoba, J., Grau, P. (1985), "Determination of Kinetic Constants of Activated Sludge Microorganisms". *Wat. Sci. Tech.*, vol. 17, pp. 259-272.

[62] Chambers, B., Tomlinson, E.J. (1982), *Bulking of Activated Sludge: Preventive and Remedial Methods*, Ellis Horwood, Chichester, Great Britain.

[63] Chang, H.T., Rittmann, B.E. (1987a), "Mathematical Model of Biofilm on Activated Carbon". *Env. Sci. Tech.*, vol. 21, no. 3, pp. 273-280.

[64] Chang, H.T., Rittmann, B.E. (1987b), "Verification of the Model of Biofilm on Activated Carbon". *Env. Sci. Tech.*, vol. 21, no. 3, pp. 280-288.

[65] Chapman, D. (1984), *The Influence of Dynamic Loads and Process Variables on the Removal of Suspended Solids from an Activated Sludge Plant*. Ph.D. dissertation, University of Alberta, Edmonton, Alberta, Canada.

[66] Chen, B., Cunningham, A.B., Ewing, D., Peralta, R., Visser, E. (1994), "Two-Dimensional Modelling of Microscale Transport and Biotransformation in Porous Media". *Num. Methods for Part. Diff. Equations*, vol. 10, pp. 65-83.

[67] Cho, S.H., Colin, F., Sardin, M., Prost, C. (1993), "Settling Velocity Model of activated Sludge". *Wat. Res.*, vol. 27, no. 7, pp. 1237-1242.

[68] Christensen, M.H. (1975), "Denitrification of Sewage by Alternating Process Operation". *Prog. Wat. Tech.*, vol. 7, no. 2, pp. 339-347.

[69] Coe, H.S., Clevenger, G.H. (1916), "Methods for Determining the Capabilities of Slime Settling Tanks". *Transactions, Am. Inst. of Mining Eng.*, vol. 55, p. 356.

[70] Comeau, Y., Hall, K.J., Oldham, K.W. (1985), "A Biochemical Model for Enhanced Biological Phosphorus Removal". *Wat. Sci. Tech.*, vol. 17, no. 11-12, pp. 313-314.

[71] Concha, F., Bustos, M.C. (1983), "A Modification of the Kynch Theory of Sedimentation". *AIChE J.*, vol. 33, pp. 312-315.

[72] Cordoba-Molina, J., Hudgins, R.R., Silverston, P.L. (1978), "Settling in Continuous Sedimentation Tanks". ASCE, vol. 104, no. EE6, paper no. 14231, pp. 1263-1275.

[73] Costerton, J.W., Lewandowski, Z., De Beer, D., Caldwell, D., Korber, D., James, G. (1994), "Biofilms, the Customized Microniche". *J. Bacteriol.*, vol. 176, no. 8, pp. 2137-2142.

[74] Cunningham, A.B., Visser, E., Lewandowski, Z., Abrahamson, M. (1996), "Evaluation of a Coupled Mass Transport–Biofilm Process Model Using Dissolved Oxygen Microsensors". *Wat. Sci.Tech.*, vol. 32, no. 8, pp. 107-114.

[75] Curds, C.R. (1973), "A Theoretical Study of Factors Influencing the Microbial Population Dynamics of the Activated-Sludge Process – II. A Computer-Simulation Study to Compare two Methods of Plant Operation". *Wat. Res.*, vol. 7, pp. 1439-1452.

[76] Curds, C.R. (1975), "Protozoa. The Organisms and their Ecology". *Ecological Aspects of Used Water Treatment vol. 1*, (C.R. Curds, H.A. Hawkes eds.), Academic Press, London, Great Britain.

[77] Curds, C.R. (1977), "Microbial Interactions Involving Protozoa". *Aquatic Microbiology*, (F.A. Skinner, J.M. Shewan, eds.), Academic Press, London, Great Britain, pp. 67-97.

[78] Curds, C.R. (1992), *Protozoa in Water Industry*, Cambridge University Press, Cambridge, Great Britain.

[79] Daigger, G.T., Roper, R.E. (1985), "The Relationship between SVI and Activated Sludge Settling Characteristics". *J. Water Pollution Control Fed.*, vol. 57, no. 8, pp. 859-866.

[80] Davis, K.E., Russel, W.B. (1989), "An Asymptotic Description of Transient Settling and Ultrafiltration of Colloidal Dispersions". *Phys. Fluids*, vol. A1, no. 1, pp. 82-100.

[81] De Beer, D., Stoodley, P. (1996), "Relation between the Structure of an Aerobic Biofilm and Transport Phenomena". *Wat. Sci.Tech.*, vol. 32, no. 8, pp. 11-18.

[82] Derco, J., Kralik, M., Hutnan, M., Bodik, I. (1990a), "Dynamic Modelling of Activated Sludge Processes. (1) Nonlinear Kinetics Models". *Chemical Papers – Chemicke Zvesti*, vol. 44, no. 3, pp. 409-420.

[83] Derco, J., Kralik, M., Gubova, P., Bodik, I. (1990b), "Dynamic Modelling of Activated Sludge Process. (2) Linearized Kinetic Models". *Chemical Papers – Chemicke Zvesti*, vol. 44, no. 5, pp. 659-668.

[84] Dick, R.I., Ewing, B.B. (1967), "Evaluation of Activated Sludge Thickening Theories". ASCE, vol. 93, no. SA4, paper no. 5367, pp. 9-29.

[85] Dick, R.I. (1970), "Role of Activated Sludge Final Settling Tanks". *J. San. Div.*, ASCE, vol. 96, no. SA2, paper no. 7231, pp. 423-436.

[86] Dick, R.I., Young, K.W. (1972), "Analysis of Thickening Performance of Final Settling Tanks". *Proc. 27th Industrial Waste Conf.*, Purdue University, Lafayette, Indiana, USA.

[87] Diehl, S., Sparr, G., Olsson, G. (1990), "Analytical and Numerical Description of the Settling Process in the Activated Sludge Operation". *Instrumentation, Control and Automation of Water and Wastewater Treatment and Transport Systems*, (R. Briggs ed.), Pergamon Press, London, Great Britain, pp. 471-478.

[88] Diehl, S. (1995a), *Conservation Laws with Application to Continuous Sedimentation.* Ph.D. dissertation, Dept of Mathematics, Lund Inst. of Tech., Lund, Sweden.

[89] Diehl, S. (1995b), "On Scalar Conservation Laws with Point Source and Discontinuous Flux Function". *SIAM J. Math. Anal.*, vol. 26, no. 6, pp. 1425-1451.

[90] Diehl, S., Jeppsson, U. (1995), "A Simulation Model of the Reactor-Settler Interaction in Wastewater Treatment". Paper VI in *Conservation Laws with Application to Continuous Sedimentation*. Ph.D. dissertation by S. Diehl, Dept of Mathematics, Lund Inst. of Tech., Lund, Sweden.

[91] Diehl, S. (1996a), "Scalar Conservation Laws with Discontinuous Flux Function: I. The Viscous Profile Condition". *Comm. Math. Phys.*, vol. 176, pp. 23-44.

[92] Diehl, S. (1996b), "A Conservation Law with Point Source and Discontinuous Flux Function Modelling Continuous Sedimentation". To appear in *SIAM J. Appl. Math.*, vol. 56, no. 2.

[93] Diehl, S. (1996c), "Dynamic and Steady-State Behaviour of Continuous sedimentation". To appear in *SIAM J. Appl. Math*.

[94] Diehl, S. (1996d), "Continuous Sedimentation of Multi-Component Particles". Submitted to *Math. Methods Appl. Sci.*

[95] Diehl, S., Jeppsson, U. (1996), "A Model of the Settler Coupled to the Biological Reactor". Submitted to *Wat. Res.*

[96] Diehl, S., Wallin, N.-O. (1996), "Scalar Conservation Laws with Discontinuous Flux Function: II. On the Stability of the Viscous Profiles". *Comm. Math. Phys.*, vol. 176, pp. 45-71.

[97] DIN (1993), "Water Quality – Determination of Ammonia, Nitrite and Nitrate (38406, E5-1; 38405, D10; 38405, D9-2)". *German Standard and Technical Rules*, Deutsches Informationszentrum für technische Regeln, Institut für Normung, e.V. Beuth Verlag, GmbH, Berlin, Germany.

[98] Dixon, D.C., Souter, P., Buchanan, J.E. (1976), "A Study of Inertial Effects in Sedimentation". *Chem. Eng. Sci.*, vol. 31, pp. 737-740.

[99] Dobbins, W.E. (1944), "Effects of Turbulence on Sedimentation". *Transactions*, ASCE, vol. 109, paper no. 2218, pp. 629-678.

[100] Dochain, D., Bastin, G. (1984), "Adaptive Identification and Control Algorithms for Nonlinear Bacterial Growth Systems". *Automatica*, vol. 20, pp. 621-634.

[101] Dold, P.L., Ekama, G.A., Marais, G.v.R. (1980), "A General Model for the Activated Sludge Process". *Prog. Wat. Tech.*, vol. 12, pp. 47-77.

[102] Dold, P.L., Marais, G.v.R. (1986), "Evaluation of the General Activated Sludge Model Proposed by the IAWPRC Task Group". *Wat. Sci. Tech.*, vol. 18, no. 6, pp. 63-89.

[103] Dold, P.L., Wentzel, M.C., Billing, A.E., Ekama, G.A., Marais, G.v.R. (1991), *Activated Sludge System Simulation Programs*, Water Research Commission, Pretoria, Republic of South Africa.

[104] Dold, P.L. (1992), "Activated Sludge System Model Incorporating Biological Nutrient (N & P) Removal". *Tech. Report.*, Dept of Civil Eng. & Eng. Mechanics, McMaster University, Hamilton, Ontario, Canada.

[105] Doohan, M. (1975), "Rotifera. The Organisms and their Ecology". *Ecological Aspects of Used Water Treatment, vol. 1*, (C.R. Curds, H.A. Hawkes eds.), Academic Press, London, Great Britain, pp. 289-304.

[106] Downing, A.L., Painter, H.A., Knowles, G. (1964), "Nitrification in the Activated Sludge Process". *J. Proc. Inst. Sewage Purification*, vol. 64, pp. 130-258.

[107] Dupont, R., Henze, M. (1989), "Mathematical Modelling of Nitrogen and Phosphorus Removal". *Trans. Seminar on Nutrient Removal from Municipal Waste Water*, 4-6 Sep., Tampere, Finland.

[108] Dupont, R., Henze, M. (1992), "Modelling of the Secondary Clarifier Combined with the Activated Sludge Model No. 1". *Wat. Sci. Tech.*, vol. 25, no. 6., pp. 285-300.

[109] Dupont, R., Dahl, C. (1995), "A One-Dimensional Model for a Secondary Settling Tank Including Density Current and Short-circuiting". *Wat. Sci. Tech.*, vol. 31, no. 2, pp. 215-224.

[110] Eckenfelder, W.W., O'Connor, D.J. (1955), "The Aerobic Treatment of Organic Wastes". *Proc. 9th Industrial Waste Conf.*, Purdue University, Lafayette, Indiana, USA.

[111] Eckenfelder, W.W., Melbinger, N. (1957), "Settling and Compaction Characteristics of Biological Sludges". *Sewage and Industrial Wastes*, vol. 29, no. 10, pp. 1114-1122.

[112] Eckenfelder, W.W. (1966), *Industrial Water Pollution Control*, McGraw-Hill Book Co., New York, New York, USA.

[113] Einfeldt, J. (1992), "The Implementation of Biological Phosphorus and Nitrogen Removal with the Bio-Denipho Process on a 265,000 PE Treatment Plant". *Wat. Sci. Tech.*, vol. 25, pp. 161-168.

[114] Ekama, G.A., Marais, G.v.R. (1979), "Dynamic Behaviour of the Activated Sludge Process". *J. Water Pollution Control Fed.*, vol. 51, pp. 534-556.

[115] Ekama, G.A., van Haandel, A.C., Marais, G.v.R. (1979), *The Present Status of Research on Nitrogen Removal: a Model for the Modified Activated Sludge Process*, Water Research Commission, Pretoria, Republic of South Africa.

[116] Ekama, G.A., Marais, G.v.R. (1984), "Nature of Municipal Wastewaters". *Theory, Design and Operation of Nutrient Removal Activated Sludge Process*, Water Research Commission, Pretoria, Republic of South Africa.

[117] Ekama, G.A., Dold, P.L., Marais, G.v.R. (1986), "Procedures for Determining COD Fractions and the Maximum Specific Growth Rate of Heterotrophs in Activated Sludge Systems". *Wat. Sci. Tech.*, vol. 18, no. 6, pp. 91-114.

[118] Fenchel, T. (1986), "Protozoan filter feeding". *Progress in Protistology, vol. 1*, Biopress, Ltd., Bristol, Great Britain, pp. 65-114.

[119] Finnson, A. (1994), *Computer Simulations of Full-Scale Activated Sludge Processes*. Tech. Lic. dissertation, Dept of Water Resources Eng., Royal Inst. of Tech., Stockholm, Sweden.

[120] Fitch, B. (1983), "Kynch Theory and Compression Zones". *AIChE J.*, vol. 29, pp. 940-947.

[121] Fitch, B. (1993), "Thickening Theories – an Analysis". *AIChE J.*, vol. 39, pp. 27-36.

[122] Fletcher, R. (1987), *Practical Methods of Optimization*, John Wiley and Sons, Ltd., Chichester, Great Britain.

[123] Font, R. (1988), "Compression Zone Effect in Batch Sedimentation". *AIChE J.*, vol. 34, pp. 229-238.

[124] Fu, Y.-C., Zhang, T.C., Bishop, P.L. (1994), "Determination of Effective Oxygen Diffusivity in Biofilms Grown in a Completely Mixed Biodrum Reactor". *Wat. Sci. Tech.*, vol. 29, no. 10-11, pp. 455-462.

[125] Fuhrman, R.E. (1984), "History of Water Pollution Control". *J. Water Pollution Control Fed.*, vol. 56, no. 4, pp. 306-313.

[126] Fujie, K.T., Tsubone, T., Moriya, H., Kubota, H. (1988), "A Simplified Kinetic Model to Simulate Soluble Organic Substances Removal in an Activated Sludge Aeration Tank". *Wat. Res.*, vol. 22, no. 1, pp. 29-36.

[127] Gabb, D.M.D., Still, D.A., Ekama, G.A., Jenkins, D., Marais, G.v.R. (1991), "The Selector Effect on Filamentous Bulking in Long Sludge Age Activated Sludge Systems". *Wat. Sci. Tech.*, vol. 23, pp. 867-877.

[128] Garrett, M.T., Sawyer, C.N. (1952), "Kinetics of Removal of Soluble BOD by Activated Sludge". *Proc. 7th Industrial Waste Conf.*, Purdue University, Lafayette, Indiana, USA, pp. 51-77.

[129] Gibbs, J.T., Bishop, P.L. (1996), "A Method for Describing Biofilm Surface Roughness Using Geostatistical Techniques". *Wat. Sci. Tech.*, vol. 32, no. 8, pp. 91-98.

[130] Glaser, D. (1988), "Simultaneous Consumption of Bacteria and Dissolved Organic Matter by *Tetrahymena pyriformis*". *Microb. Ecol.*, vol. 15, pp. 189-201.

[131] Godfrey, K.R., DiStefano, J.J. (1985), "Identifiability of Model Parameters". *Identification and System Parameter Estimation 1985*, (H.A. Barker, P.C. Young eds.), Proc. 7th IFAC/IFORS Symposium, vol. 1, pp. 89-114.

[132] Godfrey, K.R., DiStefano, J.J. (1987), "Identifiability of Model Parameters". *Identifiability of Parametric Models*, (E. Walter ed.), Pergamon Press, Oxford, Great Britain, pp. 1-20.

[133] Goodman, B.L., Englande, A.J. (1974), "A Unified Model of the Activated Sludge Process". *J. Water Pollution Control Fed.*, vol. 46, pp. 312-332.

[134] Godunov, S.K. (1959), "A Finite Difference Method for the Numerical Computations of Discontinuous Solutions of the Equations of Fluid Dynamics". *Mat. Sb.*, pp. 271-290.

[135] Gould, R.H. (1942), "Operating Experiences in New York City". *Sewage Works J.*, vol. 14, no. 1, p. 70.

[136] Grady, Jr., C.P.L., Lim, H.C. (1980), *Biological Wastewater Treatment, Theory and Applications*, Marcel Dekker, Inc., New York, New York, USA.

[137] Grady, Jr., C.P.L., Dang, J.S., Harvey, D.M., Jobbagy, A., Wang, X.-L. (1989), "Determination of Biodegradation Kinetics through Use of Electrolytic Respirometry". *Wat. Sci. Tech.*, vol. 21, pp. 957-968.

[138] Grant, S., Hurwitz, E., Mohlman, R.W. (1930), "The Oxygen Requirement of the Activated Sludge Process". *Sewage Works J.*, vol. 2, p. 228.

[139] Grau, P., Sutton, P.M., Henze, M., Elmaleh, S., Grady, Jr., C.P.L., Gujer, W., Koller, J. (1982), "Recommended Notation for Use in the Description of Biological Wastewater Treatment Processes". *Wat. Res.*, vol. 16, pp. 1501-1505.

[140] Gray, N.F. (1989), "Biological Processes – Fixed-Film Reactors". *Biology of Wastewater Treatment*, Oxford University Press, Oxford, Great Britain, pp. 263-334.

[141] Grijspeerdt, K., Vanrolleghem, P., Verstraete, W. (1995), "Selection of One-Dimensional Sedimentation: Models for On-Line Use". *Wat. Sci. Tech.*, vol. 31, no. 2, pp. 193-204.

[142] Gujer, W., Boller, M. (1986), "Design of a Nitrifying Tertiary Trickling Filter Based on Theoretical Concepts". *Wat. Res.*, vol. 20, no. 11, pp. 1353-1362.

[143] Gujer, W., Wanner, O. (1989), "Modeling Mixed Population Biofilms". *Biofilms*, (W.G. Characklis, K.C. Marshall eds.), John Wiley and Sons, Ltd., New York, New York, USA, pp. 397-443.

[144] Gujer, W., Boller, M. (1990), "A Mathematical Model for Rotating Biological Contactors". *Wat. Sci. Tech.*, vol. 22, no. 1-2, pp. 53-73.

[145] Gujer, W., Henze, M. (1991), "Activated Sludge Modelling and Simulation". *Wat. Sci. Tech.*, vol. 23, pp. 1011-1023.

[146] Gujer, W., Kappeler, J. (1992), "Modelling Population Dynamics in Activated Sludge Systems". *Wat. Sci. Tech.*, vol. 25, no. 6, pp. 93-104.

[147] Halling-Sørensen, B., Jørgensen, S.E. (1993), *The Removal of Nitrogen Compounds from Wastewater*, Elsevier Science Publishers B.V., Amsterdam, The Netherlands.

[148] Hamilton, J., Jain, R., Antoniou, P., Svoronos, S.A., Koopman, B., Lyberatos, G. (1992), "Modelling and Pilot-Scale Experimental Verification for Predenitrification Process". *J. Env. Eng.*, vol. 118, pp. 38-55.

[149] Hendricks, E. (1992), *Identification and Estimation of Nonlinear Systems Using Physical Modelling.* Ph.D. dissertation, Inst. of Mathematical Statistics and Operations Research (IMSOR), Technical University of Denmark, Lyngby, Denmark.

[150] Harremoës, P. (1976), "The Significance of Pore Diffusion to Filter Denitrification". *J. Water Pollution Control Fed.*, vol. 48, no. 2, pp. 377-388.

[151] Harremoës, P. (1978), "Biofilm Kinetics". *Water Pollution Microbiology, vol. 2,* (R. Mitchell ed.), John Wiley and Sons, Ltd., New York, New York, USA, pp. 71-109.

[152] Harremoës, P., Jansen, J. la. C., Kristensen, G.H. (1980), "Practical Problems Related to Nitrogen Bubble Formation in Fixed Film Reactors". *Prog. Wat. Tech.,* vol. 12, no. 6, pp. 253-269.

[153] Härtel, L. (1990), *Modellansätze zur Dynamischen Simulation des Belebtschlammverfahrens.* Ph.D. dissertation, Technische Hochschule Darmstadt, Darmstadt, Germany, WAR-Schriftenreihe, vol. 47.

[154] Härtel, L., Pöpel, H.J. (1992), "A Dynamic Secondary Clarifier Model Including Processes of Sludge Thickening". *Wat. Sci. Tech.*, vol. 25, no. 6, pp. 267-284.

[155] Hazen, A. (1904), "On Sedimentation". *Transactions,* ASCE, vol. 53, paper no. 980, pp. 45-88.

[156] Henze, M., Grady Jr., C.P.L., Gujer, W., Marais, G.v.R., Matsuo, T. (1987), "Activated Sludge Model No. 1". *IAWQ Scientific and Technical Report No. 1*, IAWQ, London, Great Britain.

[157] Henze, M. (1988), "Constants in Mathematical Models". *IAWPRC Technology Transfer Seminar on Mathematical Modelling of Biological Wastewater Treatment Processes*, 4-6 May, Rome, Italy, pp. 1-13.

[158] Henze, M. (1992), "Characterization of Wastewater for Modelling of Activated Sludge Process". *Wat. Sci. Tech.*, vol. 25, no. 6, pp. 1-15.

[159] Henze, M., Harremoës, P., Jansen, J. la C., Arvin, E. (1992), *Wastewater Treatment, Biological and Chemical* (2nd ed.), (in Danish), Polyteknisk Forlag, Lyngby, Denmark.

[160] Henze, M., Dupont, R., Grau, P., de la Sota, A. (1993), "Rising Sludge in Secondary Settlers due to Denitrification". *Wat. Res.*, vol. 27, no. 2, pp. 231-236.

[161] Henze, M., Gujer, W., Mino, T., Matsuo, T., Wentzel, M.C., Marais, G.v.R. (1995), "Activated Sludge Model No. 2". *IAWQ Scientific and Technical Report No. 3,* IAWQ, London, Great Britain.

[162] Henze, M., van Loosdrecht, M., Wanner, J., Krebs, P., Lafuente, J., Ayesa, E., Vanrolleghem, P. (1996), "Research Projects in Modelling and Control of Wastewater Treatment". *COST 682: Environment,* (D. Dochain, P. Vanrolleghem, M. Henze eds.), European Commission, Directorate-General XII: Science, Research and Development, Luxembourg, pp. 19-36.

[163] Hermanowicz, S.W., Schindler, U., Wilderer, P. (1996), "Fractal Structure of Biofilms: New Tools for Investigation of Morphology". *Wat. Sci. Tech.*, vol. 32, no. 8, pp. 110-117.

[164] Hill R.D. (1985), *Dynamics and Control of Solids-Liquid Separation in the Activated Sludge Process*. Ph.D. dissertation, Rice University, Houston, Texas, USA.

[165] Hiraoka, M., Tsumura, K., Fujitsa, I., Kanaya, T. (1990), "System Identification and Control of Activated Sludge Process by Use of Autoregressive Model". *Instrumentation, Control and Automation of Water and Wastewater Treatment and Transport Systems,* (R. Briggs ed.), Pergamon Press, London, Great Britain, pp. 121-128.

[166] Holmberg, A. (1981), *A Systems Engineering Approach to Biotechnical Processes – Experiences of Modelling, Estimation and Control Methods.* Ph.D. dissertation, Acta Polytechnica Scandinavica, Helsinki, Finland.

[167] Holmberg, A. (1982), "On the Practical Identifiability of Microbial Growth Models Incorporating Michaelis-Menten Type Nonlinearities". *Math. Biosciences*, vol. 62, pp. 23-43.

[168] Holmberg, A., Ranta, J. (1982), "Procedures for Parameter and State Estimation of Microbial Growth Process Models". *Automatica*, vol. 18, pp. 181-193.

[169] Holmberg, U., Olsson, G. (1985), "Simultaneous On-Line Estimation of the Oxygen Transfer Rate and Respiration Rate". *Modelling and Control of Biotechnological Processes*, (A. Johnson ed.), Pergamon Press, Oxford, Great Britain, pp. 185-189.

[170] Holmberg, U., Olsson, G., Andersson, B. (1989), "Simultaneous DO Control and Respiration Estimation". *Wat. Sci. Tech.*, vol. 21, pp. 1185-1195.

[171] Holmberg, U. (1991), *Relay Feedback of Simple Systems.* Ph.D. dissertation, Dept of Automatic Control, Lund Inst. of Tech., Lund, Sweden.

[172] Hultman, B., Hultgren, J. (1980), "Deviations from the Kynch Theory in Thickening Sludges from Wastewater Treatment Plants". *Trib. Cebedeau*, vol. 33, no. 441-442, pp. 375-389.

[173] Hultman, B. (1992), "The Baltic Sea Environment". *Proc. Water and Wastewater Management in the Baltic Region,* Royal Inst. of Tech., Stockholm, Sweden.

[174] Hydromantis (1992), *GPS-X, Technical Reference,* Hydromantis, Inc., Hamilton, Canada.

[175] Imam, E., McCorquodale, J.A., Bewtra, J.K. (1983), "Numerical Modelling of Sedimentation Tanks". *J. Hydr. Eng.,* ASCE, vol. 109, pp. 1740-1754.

[176] Ingersoll, A.C., McKee, J.E., Brooks, N.H. (1955), "Fundamental Concepts of Rectangular Settling Tanks". ASCE, vol. 81, p. 590.

[177] Jamshidi, M. (1983), *Large-Scale Systems Modeling and Control*, North-Holland Publishing Company, Amsterdam, The Netherlands.

[178] Jenkins, D., Richard, M.G., Daigger, G.T. (1993), *Manual on the Causes and Control of Activated Sludge Bulking and Foaming* (2nd ed.), Lewis Publishers, Inc., Chelsea, Michigan, USA.

[179] Jeppsson, U. (1990), “On-Line Estimation of the Nitrosomonas and Nitrobacter Concentrations in a Nitrification Process”, (in Swedish). *Technical Report, CODEN:LUTEDX/(TEIE-7025)/1-93/(1990)*, IEA, Lund, Sweden.

[180] Jeppsson, U. (1990), “Expert System for Diagnosis of the Activated Sludge Process”, (in Swedish). *Technical Report, CODEN: LUTEDX/(TEIE-7026)/1-83/(1990)*, IEA, Lund, Sweden.

[181] Jeppsson, U. (1993), *On the Verifiability of the Activated Sludge System Dynamics*. Licentiate’s thesis, CODEN:LUTEDX/(TEIE-1004)/1-177/(1993), IEA, Lund, Sweden.

[182] Jeppsson, U., Olsson, G. (1993), “Reduced Order Models for On-Line Parameter Identification of the Activated Sludge Process”. *Wat. Sci. Tech.*, vol. 28, no. 11-12, pp. 173-183.

[183] Jeppsson, U. (1994), “Simulation and Control of the Activated Sludge Process – a Comparison of Model Complexity”. *Proc. IMACS Symposium on Mathematical Modelling*, (I. Troch, F. Breitenecker eds.), Technical University Vienna, Austria, vol. 3, pp. 444-451.

[184] Jeppsson, U. (1994), “A Comparison of Simulation Software for Wastewater Treatment Processes – a COST 682 Program Perspective”. *Technical Report, CODEN:LUTEDX/(TEIE-7078)/1-12 /(1994)*, IEA, Lund, Sweden.

[185] Jeppsson, U. (1995), “A Simplified Control-Oriented Model of the Activated Sludge Process”. *Mathematical Modelling of Systems*, vol. 1, no. 1, pp. 3-16.

[186] Jeppsson, U., Diehl, S. (1995), “Validation of a Robust Dynamic Model of Continuous Sedimentation”. *Proc. 9th Forum Applied Biotechnology*, Med. Fac. Landbouww., University of Gent, Belgium, vol. 60, pp. 2403-2415.

[187] Jeppsson, U., Lee, N., Aspegren, H. (1995), “Modelling Microfauna Influence on Nitrification in Aerobic Biofilm Processes”. *Proc. Int. IAWQ Conf. Biofilm Structure, Growth and Dynamics*, Delft University of Technology, The Netherlands, pp. 77-85.

[188] Jeppsson, U., Diehl, S. (1996a), "An Evaluation of a Dynamic Model of the Secondary Clarifier". Accepted for presentation at *Water Quality International '96, IAWQ 18th Biennial International Conference*, Singapore, June 23-28.

[189] Jeppsson, U., Diehl, S. (1996b), "On the Modelling of the Dynamic Propagation of Biological Components in the Secondary Clarifier". Accepted for presentation at *Water Quality International '96, IAWQ 18th Biennial International Conference*, Singapore, June 23-28.

[190] Johansson, P. (1994), *SIPHOR a Kinetic Model for Simulation of Biological Phosphate Removal*. Ph.D. dissertation, Dept of Water and Environmental Eng., Lund University, Lund, Sweden.

[191] Jørgensen, S.E. (1992), *Integration of Ecosystem Theories: A Pattern*, Kluwer Academic Publishers, Dordrecht, The Netherlands.

[192] Kabouris, J., Georgakakos, A. (1991), "Stochastic Control of the Activated Sludge Process". *Wat. Sci. Tech.*, vol. 24, no. 6, pp. 249-255.

[193] Kahaner, D., Moler, C.B., Nash, S. (1989), *Numerical Methods and Software*, Prentice Hall, Inc., Englewood Cliffs, New Jersey, USA.

[194] Kalman, R.E. (1960), "A New Approach to Linear Filtering and Prediction Theory". *Trans. ASME, J. Basic Eng.*, vol. 82, pp. 35-45.

[195] Kappeler, J., Gujer, W. (1992), "Estimation of Kinetic Parameters of Heterotrophic Biomass under Aerobic Conditions and Characterization of Wastewater for Activated Sludge Modelling". *Wat. Sci. Tech.*, vol. 25, no. 6, pp. 125-139.

[196] Keinath, T.M., Ryckman, M.D., Dana, C.H., Hofer, D.A. (1977), "Activated Sludge – Unified System Design and Operation". *J. Env. Eng. Div.*, ASCE,. vol. 103, no. EE5, pp. 829-849.

[197] Kessler, L.H., Nichols, M.S. (1935), "Oxygen Utilization by Activated Sludge". *Sewage Works J.*, vol. 7, p. 810.

[198] Kevorkian, J., Cole, J.D. (1980), "Perturbation Methods in Applied Mathematics". *Applied Mathematical Sciences*, Springer-Verlag, Berlin, Germany.

[199] Kinner, N. (1983), *A Study of the Microorganisms Inhabiting Rotating Biological Contactor Biofilms during various Operating Conditions.* Ph.D. dissertation, University of New Hampshire, Durham, New Hampshire, USA.

[200] Kissel, J.C., McCarty, P.L., Street, R.L (1984), "Numerical Simulation of Mixed-Culture Biofilm". *J. Env. Eng.*, ASCE, vol. 110, no. EE2, pp. 393-411.

[201] Kokotovic, P.V., O'Malley, R.E., Sannuti, P. (1976), "Singular Perturbations and Order Reduction in Control Theory – an Overview". *Automatica*, vol. 12, pp. 123-132.

[202] Koo, K.Y., McInnis, B.C., Goodwin, G.C. (1982), "Adaptive Control and Identification of the Dissolved Oxygen Process". *Automatica*, vol. 18, pp. 727-730.

[203] Krebs, P. (1991a), "The Hydraulics of Final Settling Tanks". *Wat. Sci. Tech.*, vol. 23, no. 4-6, pp. 1037-1046.

[204] Krebs, P. (1991b), "Density Currents in Final Settling Tanks". *Proc. Interactions of Wastewater, Biomass and Reactor Configurations in Biological Treatment Plants*, 21-23 Aug., Copenhagen, Denmark, paper no. 27.

[205] Krebs, P. (1995), "Success and Shortcomings of Clarifier Modelling". *Wat. Sci. Tech.* , vol. 31, no. 2, pp. 181-191.

[206] Kuo, B.C. (1991), *Automatic Control Systems,* Prentice Hall, Inc., Englewood Cliffs, New Jersey, USA.

[207] Kynch, G.J. (1952), "A Theory of Sedimentation". *Trans. Faraday Soc.*, vol. 48, pp. 166-176.

[208] Laikari, H. (1989), "Simulation of the Sludge Blanket of a Vertical Clarifier in an Activated Sludge Process". *Wat. Sci. Tech.*, vol. 21, no. 6-7, pp. 621-629.

[209] La Motta, E.J. (1976), "Internal Diffusion and Reaction in Biological Films". *Env. Sci. Tech.*, vol. 10, pp. 765-769.

[210] Larrea, L., García-Heras, J.L., Ayesa, E., Florez, J. (1992), "Designing Experiments to Determine the Coefficients of Activated Sludge Models by Identification Algorithms". *Wat. Sci. Tech.*, vol. 25, no. 6, pp. 149-165.

[211] Larsen, P., Gotthardsson, S. (1976), "On the Hydraulics of Sedimentation Basins", (in Swedish). *Bulletin series A*, no. 51, Dept of Water and Environmental Eng., Lund Inst. of Tech., Lund, Sweden.

[212] Larsen, P. (1977), *On the Hydraulics of Rectangular Settling Basins, Experimental and Theoretical Studies.* Ph.D. dissertation, Dept of Water Resources Eng., Lund, Sweden.

[213] Larsen, T.A., Harremoës, P. (1994), "Modelling of Experiments with Colloidal Organic Matter in Biofilm Reactors". *Wat. Sci. Tech.*, vol. 29, no. 10-11, pp. 479-486.

[214] Lawrence, A.W., McCarty, P.L. (1970), "Unified Basis for Biological Treatment Design and Operation". *J. San. Eng. Div.*, ASCE, vol. 96, pp. 757-778.

[215] Lee, N. M., Welander, T. (1994), "Influence of Predators on Nitrification in Aerobic Biofilm Processes". *Wat. Sci. Tech.*, vol. 29, no. 7, pp. 355-363.

[216] Lens, P.N., Verstraete, W.H. (1992), "Aerobic and Anaerobic Treatment of Municipal Wastewater". *Profiles on Biotechnology*, (T.G. Villa, J. Abalde eds.), Servicio de Publicacions, Universidade de Santiago, Santiago de Compostella, Spain, pp. 333-356.

[217] Le Roux, A.Y. (1976), "On the Convergence of the Godunov's Scheme for First Order Quasi Linear Equations". *Proc. Japan Acad. Ser. A Math. Sci.*, vol. 52, pp. 488-491.

[218] LeVeque, R.J. (1990), *Numerical Solution of Conservation Laws*, Birkhäuser, Basel, Switzerland.

[219] Lewandowski, Z., Stoodley, P., Altobelli, S. (1995), "Experimental and Conceptual Studies on Mass Transport in Biofilms". *Wat. Sci. Tech.*, vol. 31, no. 1, pp. 153-162.

[220] Lewandowski, Z., Stoodley, P. (1996), "Flow Induced Vibrations, Drag Force, and Pressure Drop in Conduits Covered with Biofilm". *Wat. Sci. Tech.*, vol. 32, no. 8, pp. 19-26.

[221] Li, D., Ganczarczyk, J.J. (1987), "Stroboscopic Determination of Settling Velocity, Size and Porosity of Activated Sludge Flocs". *Wat. Res.*, vol. 21, no. 3, pp. 257-262.

[222] Little, J.N., Shure, L. (1988), *Signal Processing Toolbox*, The MathWorks, Inc., Natick, Massachusetts, USA.

[223] Ljung, L., Söderström, T. (1983), *Theory and Practice of Recursive Identification*, MIT Press, Cambridge, Massachusetts, USA.

[224] Ljung, L. (1987), *System Identification: Theory for the User*, Prentice Hall, Inc., Englewood Cliffs, New Jersey, USA.

[225] Ljung, L., Glad, T. (1991), *Model Development and Simulation*, (in Swedish), Studentlitteratur, Lund, Sweden.

[226] Ludzack, F.J., Ettinger, M.B. (1962), "Controlling Operation to Minimize Activated Sludge Effluent Nitrogen". *J. Water Pollution Control Fed.*, vol. 34, pp. 920-931.

[227] Lumley, D.J. (1985), *Settling of Activated Sludge – a Study of Limiting Factors and Dynamic Response*. Licentiate's thesis, Dept of Sanitary Eng., Chalmers University of Technology, Göteborg, Sweden.

[228] Lyn, D.A., Stamou, A.I., Rodi, W. (1992), "Density Currents and Shear-Induced Flocculation in Sedimentation Tanks". *J. Hydr. Eng.*, ASCE, vol. 118, no. 6, pp. 849-867.

[229] Marais, G.v.R., Ekama, G.A. (1976), "The Activated Sludge Process: Part 1 – Steady State Behaviour". *Water S.A.*, vol. 2, pp. 163-200.

[230] Marsili-Libelli, S. (1993), "Dynamic Modelling of Sedimentation in the Activated Sludge Process". *Civil Eng. Sys.*, vol. 10, pp. 207-224.

[231] Martinez, E.C., Drozdowicz, B. (1989), "Multitime-Scale Approach to Real-Time Simulation of Stiff Dynamic Systems". *Comput. Chem. Eng.*, vol. 13, no. 7, pp. 767-778.

[232] MathWorks (1992), *Matlab® – User's Guide,* The MathWorks, Inc., Natick, Massachusetts, USA.

[233] MathWorks (1995), *Simulink® – Dynamic System Simulation Software, User's Guide,* The MathWorks, Inc., Natick, Massachusetts, USA.

[234] McKee, J.E., Fair, G.M. (1942), "Load Distribution in the Activated Sludge Process". *Sewage Works J.*, vol. 14, no. 1, p. 121.

[235] McKinney, R.E. (1962), "Mathematics of Complete Mixing Activated Sludge". *J. San. Eng. Div.*, ASCE, vol. 88, no. 3, pp. 87-113.

[236] McKinney, R.E., Ooten, R.J. (1969), "Concepts of Complete Mixing Activated Sludge". *Trans. 19th San. Eng. Conf.*, University of Kansas, Kansas, USA, pp. 32-59.

[237] Mehra, R.K. (1980), "Nonlinear System Identification: Selected Survey and Recent Trends". *Identification and System Parameter Estimation,* (R. Isermann ed.), Proc. 5th IFAC Symposium, Darmstadt, Germany, 1979, pp. 77-83.

[238] Michaelis, L., Menten, M.L. (1913), "Die Kinetik der Invertinwirkung". *Biochem. Z.*, vol. 49, pp. 333-369.

[239] Middeldorf, J.M. (1989), "Biologische Aspekte zum Einsatz submerser Festbettkörper in Belebungsanlagen". *Korrespondenz Abwasser*, vol. 36, no. 10, pp. 1165-1169.

[240] Mino, T., Arun, V., Tsuzuki, Y., Matsuo, T. (1987), "Effect of Phosphorus Accumulation on Acetate Metabolism in the Biological Phosphorus Removal Process". *Biological Phosphate Removal from Wastewater*, (R. Ramadori ed.), Pergamon Press, London, Great Britain, pp. 27-38.

[241] Monod, J. (1942), "Recherche sur la Croissance des Cultures Bacteriennes". *Herman et Cie*, Paris, France.

[242] Monod, J. (1949), "The Growth of Bacterial Cultures". *Ann. Rev. Microbiol.*, vol. 3, pp. 371-394.

[243] Murthy, D.N.P., Page, N.W., Rodin, E.Y. (1990), *Mathematical Modelling*, Pergamon Press, New York, New York, USA.

[244] Nam, S.W., Myung, N.J., Lee, K.S. (1996), "On-Line Integrated Control System for an Industrial Activated Sludge Process". *Wat. Environ. Res.*, vol. 68, no. 1, pp. 70-75.

[245] Nelder, J.A., Mead, R. (1965), "A simplex method for function minimization". *Computer J.*, vol. 7, pp. 308-313.

[246] Nicholls, H.A. (1975), "Full Scale Experimentation of the New Johannesburg Aeration Plants". *Water S.A.*, vol. 1, no. 3, p. 121.

[247] Nisbet, B. (1984), *Nutrition and Feeding Strategies in Protozoa*, Croom Helm, London, Great Britain.

[248] Novotny, V., Jones, H., Feng, X., Capodaglio, A. (1991), "Time Series Analysis Models of Activated Sludge Plants". *Wat. Sci. Tech.*, vol. 23, no. 4-6, pp. 1107-1116.

[249] Nyberg, U. (1994), *Systems for Nutrient Removal at the Klagshamn Wastewater Treatment Plant*, (in Swedish). Tech. Lic. dissertation, Dept of Env. Eng., Lund Inst. of Tech., Lund, Sweden.

[250] Ogata, K. (1987), *Discrete-Time Control Systems*, Prentice Hall, Inc., Englewood Cliffs, New Jersey, USA.

[251] Oleinik, O.A. (1959), "Uniqueness and Stability of the Generalized Solution of the Cauchy problem for a Quasi-Linear Equation". *Uspekhi Mat. Nauk*, vol. 14, pp. 165-170, *Amer. Math. Soc. Transl. Ser.*, vol. 33, pp. 285-290, (1964).

[252] Olsson, G., Andrews, J.F. (1978), "The Dissolved Oxygen Profile – a Valuable Tool for Control of the Activated Sludge Process". *Wat. Res.*, vol. 12, pp. 985-1004.

[253] Olsson, G., Chapman, D. (1985), "Modelling the Dynamics of Clarifier Behaviour in Activated Sludge Systems". *Proc. Instrumentation and control of water and wastewater treatment and transport systems,* (R.A.R. Drake ed.), Pergamon Press, Oxford, Great Britain, pp. 405-412.

[254] Olsson, G., Stephenson, J., Chapman, D. (1986), "Computer Detection of the Impact of Hydraulic Shocks on Plant Performance". *J. Water Pollution Control Fed.*, vol. 58, no. 10, pp. 954-959.

[255] Olsson, G., Piani, G. (1992), *Computer Systems for Automation and Control*, Prentice Hall, Inc., Englewood Cliffs, New Jersey, USA.

[256] Olsson, G. (1993), "Advancing ICA Technology by Eliminating the Constraints". *Wat. Sci. Tech.*, vol. 28, no. 11-12, pp. 1-7.

[257] Olsson, G., Jeppsson, U. (1994a), "Establishing Cause-Effect Relationships in Activated Sludge Plants - what Can Be Controlled". *Proc. 8th Forum Applied Biotechnology,* Med. Fac. Landbouww., University of Gent, Belgium, vol. 59, pp. 2057-2070.

[258] Olsson, G., Jeppsson, U. (1994b), "Modelling, Simulation, and Identification Technologies". *Invited paper, Water Environment Federation (WEF), 67th Annual Conference,* Chicago, Illinois, USA, Oct. 15-19.

[259] Orhon, D., Artan, N. (1994), *Modelling of Activated Sludge Systems,* Technomic Publishing Company, Inc., Lancaster, Pennsylvania, USA.

[260] Ossenbruggen, P.J., Constantine K., Collins, M.R., Bishop, P.L. (1987), "Toward Optimum Control of the Activated Sludge Process with Reliability Analysis". *Civ. Eng. Sys.*, vol. 4, pp. 77-86.

[261] Ossenbruggen, P.J., McIntire, S. (1990), "Using Shock Wave Theory to Predict Secondary Clarifier Performance". *Instrumentation, Control and Automation of Water and Wastewater Treatment and Transport Systems*, (R. Briggs ed.), Pergamon Press, London, Great Britain, pp. 479-486.

[262] Otterpohl, R., Freund, M. (1992), "Dynamic Models for Clarifiers of Activated Sludge Plants with Dry and Wet Weather Flows". *Wat. Sci. Tech.*, vol. 26, no. 5-6, pp. 1391-1400.

[263] Padukone, N., Andrews, G.F. (1989), "A Simple, Conceptual Mathematical Model for the Activated Sludge Process and Its Variants". *Wat. Res.*, vol. 23, no. 12, pp. 1535-1543.

[264] Parker, D.S. (1983), "Assessment of Secondary Clarification Design Concepts". *J. Water Pollution Control Fed.*, vol. 55, No. 4, pp. 349-359.

[265] Parker, D., Lutz, M., Dahl, R., Bernkopf, S. (1989), "Enhancing Reaction Rates in Nitrifying Trickling Filters through Biofilm Control". *J. Water Pollution Control Fed.*, vol. 61, pp. 618-631.

[266] Patry, G.G., Chapman, D., eds. (1989), *Dynamic Modelling and Expert Systems in Wastewater Engineering*, Lewis Publishers, Inc., Chelsea, Michigan, USA.

[267] Patry, G.G., Takács, I. (1992), "Settling of Flocculent Suspensions in Secondary Clarifiers". *Wat. Res.*, vol. 26, no. 4, pp. 473-479.

[268] Petersen, E.E. (1965), *Chemical Reaction Analysis*, Prentice Hall, Inc., Englewood Cliffs, New Jersey, USA.

[269] Petzold, L. (1983), "A Description of DASSL: A Differential/ Algebraic System Solver". *Scientific Computing*, (R. Stepleman *et al.*, eds.), North-Holland Publishing Company, Amsterdam, The Netherlands, pp. 65-68.

[270] Pflanz, P. (1966), *Über das Absetzen des Belebten Schlammes in Horizontal Durchströmten Nachklärbecken*, Veröffentlichungen des Instituts für Siedlungswasserwirtschaft, Technische Hochschule Hannover, Hannover, Germany, vol. 25.

[271] Piela, P.C., Epperly, T.G., Westerberg, K.M., Westerberg, A.W. (1991), "ASCEND: An Object-Oriented Computer Environment for Modeling and Analysis: The Modeling Language". *Comput. Chem. Eng.*, vol. 15, no. 1, p. 53.

[272] Pitman, A.R. (1985), "Settling of Nutrient Removal Activated Sludges". *Wat. Sci. Tech.*, vol. 17, no. 4-5, pp. 493-504.

[273] Pohjanpalo, H. (1978), "System Identifiability Based on Power Series Expansion of the Solution". *Math. Biosciences*, vol. 41, pp. 21-33.

[274] Powell, O.E. (1967), *Microbial Physiology and Continuous Culture*, (E.O. Powell, C.G.T. Evans, R.E. Strange, D.W. Tempest eds.), Her Majesty's Stationery Office, Liverpool, Great Britain, pp. 34-55.

[275] Ratkowsky, D.A. (1986), "A Suitable Parametrization of the Michaelis-Menten Enzyme Reaction". *Biochem J.*, vol. 240, pp. 357-360.

[276] Ratsak, C.H., Kooijman, S.A.L.M., Kooi, B.W. (1993), "Modelling the Growth of an Oligochaete on Activated Sludge". *Wat. Res.*, vol. 27, no. 5, pp. 739-747.

[277] Ratsak, C.H., Kooi, B.W., van Verseveld, H.W. (1994), "Biomass Reduction and Mineralization Increase due to the Ciliate *Tetrahymena pyriformis* Grazing on the Bacterium *Pseudomonas fluorescens*". *Wat. Sci. Tech.*, vol. 29, no. 7, pp. 119-128.

[278] Ratsak, C.H., Maarsen, K.A., Kooijman, S.A.L.M. (1996), "Effects of Protozoa on Carbon Mineralization In Activated Sludge". *Wat. Res.*, vol. 30, no. 1, pp. 1-12.

[279] Reichert, P. (1994a), "AQUASIM - A Tool for Simulation and Data Analysis of Aquatic Systems". *Wat. Sci. Tech.*, vol. 30, no. 2, pp. 21-30.

[280] Reichert, P. (1994b), *Concepts Underlying a Computer Program for the Identification and Simulation of Aquatic Systems*, Swiss Federal Institute for Environmental Science and Technology (EAWAG), Dübendorf, Switzerland.

[281] Reichert, P., Ruchti, J. (1994), *AQUASIM: Computer Program for the Identification and Simulation of Aquatic Systems - User Manual*, Swiss Federal Institute for Environmental Science and Technology (EAWAG), Dübendorf, Switzerland.

[282] Reichert, P.,von Shulthess, R., Wild, D. (1995), "The Use of AQUASIM for Estimating Parameters of Activated Sludge Models". *Wat. Sci. Tech.*, vol. 31, no. 2, pp. 135-147.

[283] Renshaw, E. (1993), *Modelling Biological Populations in Space and Time*, Cambridge University Press, Cambridge, Great Britain.

[284] Richardson, J.F., Zaki, W.N. (1954), "Sedimentation of Fluidisation: Part I". *Trans. Inst. Chem. Eng.*, vol. 32, pp. 35-53.

[285] Rittmann, B.E., McCarty, P.L (1980), "Models of Steady-State-Biofilm Kinetics". *Biotech. Bioeng.*, vol. 22, pp. 2343-2357.

[286] Rittmann, B.E. (1982), "The Effect of Shear Stress on Biofilm Loss Rates". *Biotech. Bioeng.*, vol. 24, pp. 501-506.

[287] Rittmann, B.E., Brunner, C.W. (1984), "The non Steady-State Biofilm Process for Advanced Organics Removal". *J. Water Pollution Control Fed.*, vol. 56, no. 7, pp. 874-880.

[288] Rittmann, B.E. (1989a), "Mathematical Modelling of Fixed-Film Growth". *Dynamic Modelling and Expert Systems in Wastewater Engineering*, (G.G. Patry, D. Chapman eds.), Lewis Publishers, Inc., Chelsea, Michigan, USA, pp. 39-57.

[289] Rittmann, B.E. (1989b), "Detachment from Biofilms". *Structure and Function of Biofilms*, (W.G. Characklis, P.A. Wilderer eds.), John Wiley and Sons, Ltd., New York, New York, USA, pp. 49-58.

[290] Robertson, G.A. (1992), *Mathematical Modelling of Startup and Shutdown Operations of Process Plants*. Ph.D. dissertation, University of Canterbury, New Zealand.

[291] Rousseeuw, P.J., Leroy, A.M. (1987), *Robust Regression and Outlier Detection*, John Wiley and Sons, Ltd., New York, New York, USA.

[292] Sabin, G.C., Summers, D. (1993), "Chaos in a Periodically Forced Predator-Prey Ecosystem Model". *Math. Biosciences*, vol. 113, pp. 91-113.

[293] Samstag, R.W., McCorquodale, J., Zhou, S.P. (1992a), "Prospects for Transport Modeling of Process Tanks". *Wat. Sci. Tech.*, vol. 26, no. 5-6, pp. 1401-1410.

[294] Samstag, R.W., Dittmar, D.F., Vitasovic, Z., McCorquodale, J. (1992b), "Underflow Geometry in Secondary Sedimentation". *Wat. Environ. Res.*, vol. 64, no. 3, pp. 204-212.

[295] Sargent, R.G. (1982), "Verification and Validation of Simulation Models". *Progress in Modelling and Simulation*, Academic Press, New Jersey, USA.

[296] Schiesser, W.E. (1991), *The Numerical Method of Lines Integration of Partial Differential Equations*, Academic Press, San Diego, California, USA.

[297] Schindler, U., Potzel, U., Flemming, H.-C., Wilderer, P.A. (1995), "Laser Doppler Anemometry Studies on Mass Transport between Bulk Water Phase and Surface Biofilm". *Proc. Int. IAWQ Conf. Biofilm Structure, Growth and Dynamics*, Delft University of Technology, The Netherlands, pp. 45-52.

[298] Schlegel, S. (1988), "The Use of Submerged Biological Filters for Nitrification". *Wat. Sci. Tech.*, vol. 20, no. 4-5, pp. 177-187.

[299] Schwarz, G. (1978), "Estimating the Dimension of a Model". *The Annals of Statistics*, vol. 6, no. 2, pp. 461-464.

[300] Scott, K.J. (1966), "Mathematical Models of Mechanisms of Thickening". *I. & E. C. Fundamentals*, vol. 5, pp. 109-113.

[301] Scott, K.J. (1968), "Theory of Thickening: Factors Affecting Settling rate of Solids in Flocculated Pulps". *Trans. Inst. Min. Met.*, vol. 77, sec. C, pp. 85-97.

[302] Seborg, D.E., Edgar, T.F., Mellichamp, D.A. (1989), *Process Dynamics and Control*, John Wiley and Sons, Ltd., New York, New York, USA.

[303] Sekine, T., Tsugura, H., Urushibara, S., Furuya, N., Fujimoto, E., Matsui, S. (1989), "Evaluation of Settleability of Activated Sludge Using a Sludge Settling Analyzer". *Wat. Res.*, vol. 23, no. 3, pp. 361-367.

[304] Setter, L.R., Carpenter, W.T., Winslow, G.C. (1945), "Principle Application of Principles of Modified Sewage Aeration". *Sewage Works J.*, vol. 17, no. 4, p. 669.

[305] Shannon, P.T., Stroupe, E., Tory, E.M. (1963), "Batch and Continuous Thickening, Basic Theory, Solids Flux for Rigid Spheres". *I. & E. C. Fundamentals*, vol. 2, pp. 203-211.

[306] Shapiro, R., Hogan, J.W.T. (1945), "Discussion – Practical Application of Principles of Modified Sewage Aeration". *Sewage Works J.*, vol. 17, no. 4, p. 689.

[307] Sheffer, M.S., Hiraoka, M., Tsumura, K. (1984), "Flexible Modelling of the Activated Sludge System – Theoretical and Practical Aspects". *Wat. Sci. Tech.*, vol. 17, pp. 247-258.

[308] Sherr, B.F., Sherr, E.B. (1984), "Role of Heterotrophic Protozoa in Carbon and Energy Flow in Aquatic Ecosystems". *Current Perspectives in Microbial Ecology*, (M.J. Klug, C.A. Reddy eds.), American Society for Microbiology, Washington DC, USA, pp. 412-423.

[309] Shin, B.S., Dick, R.I. (1980), "Applicability of Kynch Theory to Flocculent Suspensions". *J. Env. Eng. Div.*, ASCE, vol. 106, no. EE3, pp. 505-526.

[310] Siegriest, H., Tschui, M. (1992), "Interpretation of Experimental Data with Regard to the Activated Sludge Model No. 1 and Calibration of the Model for Municipal Wastewater Treatment Plants". *Wat. Sci. Tech.*, vol. 25, no. 6, p. 167.

[311] Siegrist, H., Krebs, P., Bühler, R., Purtschert, I., Röck, C., Rufer, R. (1995), "Denitrification in Secondary Clarifiers". *Wat. Sci. Tech.*, vol. 31, no. 2, pp. 205-214.

[312] Smolders, G.J.F., van Loosdrecht, M.C.M., Heijnen, J.J. (1995), "A Metabolic Model for the Biological Phosphorus Removal Process". *Wat. Sci. Tech.*, vol. 31, no. 2, pp. 79-93.

[313] Smollen, M, Ekama, G.A. (1984), "Comparison of Empirical Settling Velocity Equations in Flux Theory for Secondary Settling Tanks". *Water S.A.*, vol. 10, pp. 175-184.

[314] Söderström, T., Stoica, P. (1989), *System Identification*, Prentice Hall, Inc., Englewood Cliffs, New Jersey, USA.

[315] Sollfrank, U., Gujer, W. (1991), "Characterization of Domestic Wastewater for Mathematical Modelling of the Activated Sludge Process". *Wat. Sci. Tech.*, vol. 23, no. 4-6, pp. 1057-1066.

[316] Spriet, J.A. (1985), "Structure Characterization – An Overview". *Identification and System Parameter Estimation*, (H.A. Barker, P.C. Young eds.), Proc. 7th IFAC/IFORS Symposium, vol. 1, pp. 749-756.

[317] Sprouse, G., Rittmann, B.E. (1990), "Colloid Removal in Fluidized-Bed Biofilm Reactors". *J. Env. Eng.*, vol. 116, pp. 314-329.

[318] SSPA Systems (1991), *Simnon™, User's Guide for UNIX Systems*, SSPA Systems, Göteborg, Sweden.

[319] Stahl, L.J. (rapporteur), Bock, E., Bouwer, E.J., Douglas, L.J., Gutnick, D.L., Heckmann, K.D., Hirsch, P., Kölbel-Boelke, J.M., Marshall, K.C., Prosser, J.I., Schütt, C., Watanabe, Y. (1989), "Group Report: Cellular Physiology and Interactions of Biofilm Organisms". *Structure and Function of Biofilms*, (W.G. Characklis, P.A. Wilderer eds.), John Wiley and Sons, Ltd., New York, New York, USA, pp. 269-286.

[320] Stehfest, H. (1984), "An Operational Dynamic Model of the Final Clarifier". *Trans. Inst. M. C.*, vol. 6, no. 3, pp. 160-164.

[321] Steinour, H.H. (1944), "Rate of Sedimentation, Nonflocculated Suspensions of Uniform Spheres". *I. & E.C.*, vol. 36, pp. 618-624.

[322] Stenstrom, M.K. (1975), *A Dynamic Model and Computer Compatible Control Strategies for Wastewater Treatment Plants*. Ph.D. dissertation, Clemson University, Clemson, South Carolina, USA.

[323] Stephanopolous, G., Henning, G., Leone, H. (1990), "MODEL.-LA: A Modeling Language for Process Engineering - 1: The Formal Framework". *Comput. Chem. Eng.*, vol. 14, no. 8, p. 813.

[324] Stull, R.B. (1988), *An Introduction to Boundary Layer Meteorology*, Kluwer Academic Publishers, Dordrecht, The Netherlands.

[325] Szalai, L., Krebs, P., Rodi, W. (1994), "Simulation of Flow in Circular Clarifiers with and without Swirl". *J. Hydr. Eng.*, ASCE, vol. 120, no. 1, pp. 4-21.

[326] Takács, I., Patry, G.G., Nolasco, D. (1991), "A Dynamic Model of the Clarification-Thickening Process". *Wat. Res.*, vol. 25, no. 10, pp. 1263-1271.

[327] Takahashi, S., Fujita, T., Kato, M., Saiki, T., Maeda, M. (1969), "Metabolism of Suspended Matter in Activated Sludge Treatment". *Proc. 4th IAWPRC Int. Conf. on Advances in Water Pollution Research*, Prague, Czech Republic, pp. 341-359.

[328] Tang, C.-C., Brill, Jr., E.D., Pfeffer, J.T. (1987), "Comprehensive Model of Activated Sludge Wastewater Treatment System". *J. Env. Eng.*, ASCE, vol. 113, no. 5, pp. 952-969.

[329] Te Braake, H., Babuska, R., Van Can, E., Hellinga, C. (1994), "Predictive Control in Biotechnology Using Fuzzy and Neural Models". *Proc. Advanced Instrumentation, Data Interpretation, and Control of Biotechnological Processes*, 24-27 Oct., Gent, Belgium.

[330] Thensen, A. (1974), "Some Notes on Systems, Models and Modelling". *Int. J. Sys. Sci.*, vol. 5, pp. 145-152.

[331] Thomann, R.V. (1982), "Verification of Water Quality Models". *J. Environ. Eng. Div.*, ASCE, vol. 108, no. 5, pp. 923-940.

[332] Thomas, D.G. (1963), "Transport Characteristics of Suspensions. Relation of Hindered Settling Floc Characteristics to Rheological Parameters". *AIChE J.*, vol. 9, pp. 310-316.

[333] Thornton, C., Bierman, G. (1980), "UDU^T Covariance Factorization for Kalman Filtering". *Control and Dynamic Systems vol. 16*, (C. Leondes ed.), Academic Press, London, Great Britain.

[334] Tiessier, C. (1936), *Ann. Physiol. Physicochem Biol.*, vol. 12, p. 527.

[335] Tracy, K.D., Keinath, T.M. (1973), "Dynamic Model for Thickening of Activated Sludge". *AIChE Symposium Series*, vol. 70, no. 136, pp. 291-308.

[336] Trulear, M.G., Characklis, W.G. (1982), "Dynamics of Biofilm Processes". *J. Water Pollution Control Fed.*, vol. 54, pp. 1288-1301.

[337] Ullrich, A.H., Smith, M.W. (1951), "The Biosorption Process of Sewage and Waste Treatment". *Sewage and Industrial Wastes*, vol. 23, no. 10, p. 1248.

[338] Vaccari, D.A, Christodoulatos, C. (1990), "Estimation of the Monod Model Coefficients for Dynamic Systems using Actual Activated Sludge Plant Data". *Wat. Sci. Tech.*, vol. 10, pp. 449-454.

[339] Vaerenbergh, E.V. (1980), "Numerical Computation of Secondary Settler Area Using Batch Settling Data". *Trib. Cebedeau*, vol. 33, pp. 369-374.

[340] Van Haandel, A.C., Ekama, G.A., Marais, G.v.R. (1981), "The Activated Sludge Process: Part 3 – Single Sludge Denitrification". *Wat. Res.*, vol. 15, pp. 1135-1152.

[341] Van Impe, J. (1994), *Modeling and Optimal Adaptive Control of Biotechnological Processes*. Ph.D. dissertation, Dept of Elektrotechniek (Afdeling ESAT), Katholieke Universiteit Leuven, Heverlee, Belgium.

[342] Vanrolleghem, P.A., Verstraete, W. (1993), "On-Line Monitoring Equipment for Wastewater Treatment Processes: State of the Art". *Proc. TI-KVIV Studiedag Optimalisatie van Waterzuiveringsinstallaties door Proceskontrole en -Sturing*, Gent, Belgium, pp. 1-22.

[343] Vanrolleghem, P.A. (1994), *On-Line Modelling of Activated Sludge Processes: Development of an Adaptive Sensor*. Ph.D. dissertation, Laboratory of Microbial Ecology, University of Gent, Gent, Belgium.

[344] Vanrolleghem, P.A., Van der Schueren, D., Krikilion, G., Grijspeerdt, K., Willems, P., Verstraete, W. (1995), "On-Line Quantification of Settling Properties with In-Sensor-Experiments in an Automated Settlometer". *IAWQ Spec. Conf. on Sensors in Wastewater Tech.*, Copenhagen, Denmark, Oct. 25-27.

[345] Vanrolleghem, P.A., Jeppsson, U. (1996), "Simulators for Modelling of WWTP". *COST 682: Environment*, (D. Dochain, P. Vanrolleghem, M. Henze eds.), European Commission, Directorate-General XII: Science, Research and Development, Luxembourg, pp. 67-78.

[346] Vanrolleghem, P., Jeppsson, U., Carstensen, J., Carlsson, B., Olsson, G. (1996), "Integration of WWT Plant Design and Operation – a Systematic Approach Using Cost Functions". Accepted for presentation at *Water Quality International '96, IAWQ 18th Biennial International Conference*, Singapore, June 23-28.

[347] Vansteenkiste, G.C., Spriet, J.A. (1982), "Modelling Ill-Defined Systems". *Progress in Modelling and Simulation*, (F.E. Cellier ed.), Academic Press, London, Great Britain, pp. 11-38.

[348] Vaughan, G.M., Scott, P.H., Holder, G.A. (1973), "Simulation of the Trickling Filtration Process". *J. Chem. Eng. of Japan*, vol. 6, no. 6, pp. 532-540.

[349] Vesilind, P.A. (1968a), "Theoretical Considerations: Design of Prototype Thickeners from Batch Settling Tests". *Water and Sewage Works*, vol. 115, no. 7, pp. 302-307.

[350] Vesilind, P.A. (1968b), "Discussion of 'Evaluation of Activated Sludge Thickening Theories', by R.I. Dick and B.B. Ewing", *J. Sanitary Eng. Div.*, ASCE, vol. 94, no. SA1, pp. 185-191.

[351] Vesilind, P.A. (1979), *Treatment and Disposal of Wastewater Sludges*, Ann Arbor Science, Ann Arbour, Michigan, USA.

[352] Vialas, C., Cheruy, A., Gentil, S. (1985), "An Experimental Approach to Improve the Monod Model Identification". *Proc. Modelling and Control of Biotechnological Processes, 1st IFAC Symposium*, Noordwijkerhout, The Netherlands, pp. 175.179.

[353] Vitasovic, Z. Z. (1985), *An Integrated Control System for the Activated Sludge Process.* Ph.D. dissertation, Rice University, Houston, Texas, USA.

[354] Vitasovic, Z.Z. (1989), "Continuous Settler Operation: A Dynamic Model". *Dynamic Modelling and Expert Systems in Wastewater Engineering,* (G.G. Patry, D. Chapman eds.), Lewis Publishers, Inc., Chelsea, Michigan, USA, pp. 59-81.

[355] Vitasovic, Z., Ji, Z., McCorquodale, J.A., Zhou, S.P. (1994), "A Dynamic Solids Inventory Model for Activated Sludge Systems". *Proc. WEFTEC'94*, Water Environment Federation, Chicago, Illinois, USA, Oct. 15-19, pp. 129-140.

[356] Wahlberg, E.J., Keinath, T.M. (1988), "Development of Settling Flux Curves Using SVI". *J. Water Pollution Control Fed.,* vol. 60, no. 12, pp. 2095-2100.

[357] Wanner, O., Gujer, W. (1984), "Competition in Biofilms". *Wat. Sci. Tech.*, vol. 17, no. 2-3, pp. 27-44.

[358] Wanner, O., Gujer, W. (1986), "A Multispecies Biofilm Model". *Biotech. Bioeng.*, vol. 28, pp. 314-328.

[359] Wanner, O. (1996), "New Experimental Findings and Biofilm Modelling Concepts". *Wat. Sci. Tech.*, vol. 32, no. 8, pp. 133-140.

[360] Wanner, O., Reichert, P. (1996), "Mathematical Modelling of Mixed-Culture Biofilms". To appear in *Biotech. Bioeng.*

[361] Watson, S.W., Bock, E., Harms, H., Koops, H.-P., Hooper, A.B. (1989), "Nitrifying Bacteria". *Bergey's Manual of Systematic Bacteriology,* (J.T. Stanley, M.P. Bryant, N. Pfennig, J.G. Holt eds.), Williams & Wilkins, Baltimore, Maryland, USA, pp. 1808-1833.

[362] Wentzel, M.C., Lötter, L.H., Ekama, G.A., Marais, G.v.R. (1986), "Metabolic Behaviour of Acinetobacter spp. in Enhanced Biological Phosphorus Removal – Biochemical Model". *Water S.A.*, vol. 12, no. 4, pp. 209-224.

[363] Wentzel, M.C., Dold, P.L., Ekama, G.A., Marais, G.v.R. (1989), "Enhanced Polyphosphate Organism Cultures in Activated Sludge Systems. Part III: Kinetic Model". *Water S.A.*, vol. 15, no. 2, pp. 89-102.

[364] Wentzel, M.C., Lötter, L.H., Ekama, G.A., Loewenthal, R.E., Marais, G.v.R. (1991), "Evaluation of Biochemical Models for Biological Excess Phosphorus Removal". *Wat. Sci. Tech.*, vol. 23, no. 4-6, pp. 567-576.

[365] Wentzel, M.C., Ekama, G.A., Marais, G.v.R. (1992), "Processes and Modelling of Nitrification Denitrification Biological Excess Phosphorus Removal Systems – A Review". *Wat. Sci. Tech.*, vol. 25, no. 6, pp. 59-82.

[366] Williamson, K., McCarty, P.L. (1976), "A Model of Substrate Utilization by Bacterial Films". *J. Water Pollution Control Fed.,* vol. 48, no. 1, pp. 9-24.

[367] Willoughby, R. (1974), *Stiff Differential Systems*, Plenum Press, New York, New York, USA.

[368] WPCF (1990), "Wastewater Biology: The Microlife". Special publication prepared by the *Task Force on Wastewater Biology, Water Pollution Control Fed.*, Virginia, USA.

[369] Wuhrmann, K. (1964), "Grundlagen für die Dimensionierung der Beluftung bei Belebt-Schlammanlagen". *Schweiz. Z. Hydrol.*, vol. 26, p. 310.

[370] Yang, Y.Y., Linkens, D.A. (1993), "Modelling of Continuous Bioreactors via Neural Networks". *Trans. Inst. M. C.*, vol. 15, no. 4, pp. 158-169.

[371] Yoshioka, N., Hotta, Y., Tanaka, S., Naito, S., Tsugami, S. (1957), "Continuous Thickening of Homogeneous Flocculated Slurries". *Kagaky Kogaku* (Chem. Eng.), vol. 21, no. 2, pp. 66-74.

[372] Young, P. (1984), *Recursive Estimation and Time-Series Analysis*, Springer-Verlag, Berlin, Germany.

[373] Zhang, T.C., Bishop, P.L. (1994a), "Density, Porosity, and Pore Structure of Biofilms". *Wat. Res.*, vol. 28, no. 11, pp. 2267-2277.

[374] Zhang, T.C., Bishop, P.L. (1994b), "Evaluation of Tortuosity Factors and Effective Diffusivities in Biofilms". *Wat. Res.*, vol. 28, no. 11, pp. 2279-2287.

[375] Zhou, S., McCorquodale, J.A. (1992a), "Influence of Skirt Radius on Performance of Circular Clarifier with Density Stratification". *Int. J. for Numerical Methods in Fluids*, vol. 14, pp. 919-934.

[376] Zhou, S., McCorquodale, J.A. (1992b), "Modelling of Rectangular Settling Tanks". *J. Hydr. Eng.*, ASCE, vol. 118, no. 10, pp. 1391-1405.

[377] Zhou, S., McCorquodale, J.A., Vitasovic, Z. (1992), "Influences of Density on Circular Clarifier with Baffles". *J. Env. Eng.*, ASCE, vol. 118, no. 6, pp. 829-847.