

Protocol to set up the TELEMAC system on a new plant Deliverable 1.2 Final

*Deliverable type: Report
Number: D1.2*

Date of delivery: 29th Aug 2003

Final submission: 13th April 2004

Task: WP1.1

*Responsible:
M.Sc. M.Eng. ir. Usama Zaher
ir. Dirk J.W. De Pauw
Prof. Dr. ir. Vanrolleghem Peter*

Reviewers:

<i>Dr. Benoit Chachuat and Dr. Olivier Bernard</i>	<i>INRIA</i>
<i>Prof.Dr.ir. Denis Dochain</i>	<i>CESAME</i>
<i>Dr. Jean-Philippe Steyer</i>	<i>INRA</i>

Editors:

M.Sc. M.Eng. Ir. Usama Zaher

Prof. Dr. ir. Vanrolleghem Peter

e-mail: Usama.Zaher@ugent.be

e-mail: Peter.Vanrolleghem@ugent.be

Copyright © 2003, BIOMATH

Department of Applied Mathematics, Biometrics and Process Control

Ghent University

Coupure Links 653

B-9000 Gent, Belgium.

Phone: +32 (9) 264.59.32

Fax: +32 (9) 264.62.20

e-mail: Peter.Vanrolleghem@ugent.be

Abstract:

This is a draft version of D1.2 and it reports the protocol to set up the TELEMAC system on a new plant. The first step of the protocol consists of data collection and an initial model fit. Then it proceeds with optimal experimental design to determine which sensors, measurement intervals, ...etc. that should be used to obtain data with maximal information content for parameter estimations. Also, for the generality of the protocol and determination of the boundaries of the TELEMAC system, a back loop is considered to fine tune the plant, upgrade the model and collect additional data if needed to face unexpected problems. Therefore, a detailed description of the data collection guide, Optimal Experimental Design (OED) methods are provided. A virtual plant has been developed and a detailed example is given for the protocol implementation by applying it to the virtual plant. An application of optimal experimental design to an industrial case study is also provided. User-friendly software has been further developed to implement the OED methods in the proposed protocol. A brief description of the software interface is provided at the end of the report.

Factual summary D1.2

Concrete Results	1.1 Man-Months	
	BIOMATH*	Domecq UK
2 Proposed architecture – Global strategy, structure and framework for the protocol to set of the TELEMAT system in a new plant	0.5	
3 Developed concepts and methods – A guide to define the plant layout and the type of digester – A guide to determine a detailed influent characterisation for anaerobic digestion – A guide to analyse long term data recorded at an existing plant. – Development of a procedure to quantify the quality of local sensitivity analysis calculations (already charged in D3.2) – Development of an optimal experimental design procedure – Building a virtual plant to imitate a real digester treating vinasses – Provide an example of the protocol implementation and to investigate the added value of the TELEMAT system	0.1 0.2 0.1 0 0.75 0.5 1	
4 Developed software – Matlab code to interpolate simulated signals and add noise generated at predefined level. – MSL code for the implementation of the IWA ADM1 model in WEST (already partially charged to D2.2 and D3.1) – MSL code for TELEMAT models implementation in WEST platform – C++ code (4,700 lines): implementation of finite difference sensitivity (already charged in D3.2) – C++ code (500 lines): implementation of the optimal experimental design procedure using continuous optimisation algorithms	0.1 1 0.5 0 0.5	
5 Validations and tests Writing D1.2 Validation of D1.2	0.5 0.25	0.3
Sub TOTAL	6.00	0.3
TOTAL	6.30	

* Based on the accepted update of staff

TABLE OF CONTENTS

1 Introduction	7
2 Protocol description	8
2.1 Plant layout study	9
2.1.1 Type of Digester	9
2.1.2 Temperature Operation Range	16
2.1.3 Upstream Buffer Tank	16
2.1.4 pH Adjustment	16
2.1.5 Other Factors	17
2.2 Detailed influent characterisation	17
2.2.1 COD, TOC compared to theoretical COD and oxidation state of carbon	17
2.2.2 Theoretical gas composition	18
2.2.3 Substrate composition, hydrolysis and degradability	19
2.2.4 VFA levels	20
2.2.5 VFA fractions	20
2.2.6 Assessment of possible precipitation and adsorption	21
2.2.7 determination sulphates	21
2.2.8 determination of proteins and N compounds	21
2.2.9 Check nutrients and trace elements	22
2.2.10 Assessment of possible toxicity	22
2.2.10.1 Inorganic toxins	22
2.2.10.2 Toxicity of natural organic compounds	24
2.3 Analysing long term data	25
2.3.1 Determine seasonal changes	25
2.3.2 Which measurements are generally available	25
2.3.3 Gas composition	25
2.3.4 Fluctuation of the reactor pH	26
2.3.5 Assessment of the available buffering capacity	26

2.3.6 Assessment of COD hydrolysis and degradability	27
2.4 Telemac sensors installation	28
2.4.1 Installed Sensors	28
2.4.2 Interference and reference measurements	28
2.4.3 Installation of sensors (sample filtering, location)	29
2.5 Optimal experimental design	29
2.5.1 Introduction	29
2.5.2 Fisher information matrix	30
2.5.3 Local sensitivity analysis using finite differences	32
2.5.4 Iterative optimal experimental design procedure	35
3 Virtual case study	37
3.1 Description and modelling of the virtual plant	38
3.1.1 Introduction of ADM1 model will be used to simulate a real plant	38
3.1.2 Virtual plant description	41
3.2 Influent characteristics	43
3.3 Long term data	44
3.4 Telemac sensors	45
3.5 Optimal experimental design	46
3.5.1 Model calibration for long term data	46
3.5.2 Optimal experimental design: adding TELEMAC sensors	48
3.5.3 Optimal experimental design: designing an optimal influent flow profile	50
4 Agralco industrial case study	55
4.1 Model calibration based on long term data	55
4.2 Optimal experimental design	58
4.2.1 Defining the experimental degrees of freedom	58
4.2.2 Finding the optimal dilution rate profile	59
4.2.3 Finding the optimal experimental duration	60
4.2.4 Investigating the influence of different sensor arrangements	63
4.2.5 Determining the influence of a change in k _{la} and influent total alkalinity	64

5 Software developments	66
5.1 West [®] Sensitivity analysis module	66
5.2 West [®] Optimal Experimental Design module	70
6 Conclusion	73
References	74

1 Introduction

The TELEMAC project is focused on the control of the reactants of anaerobic treatment. This opens the applicability of the system to a wide range of treatment systems that implement the anaerobic digestion process. The system targets small enterprises as well as small sites of large companies and thus it is simple, robust and cost effective. These properties of the system most suit industrial applications whose wastewater characteristics are almost optimum for the anaerobic process, e.g. the treatment of vinasses which is planned for the first implementation and testing of the system. Treatment of vinasses is a big sector for the potential use of anaerobic digestion treatment. Since the system focuses on the core of the process, it is applicable to other anaerobic digestion applications provided that other process complexities are assessed and, if required, the proper extensions of the system are designed.

Therefore, this report illustrates a general protocol to process the information that can be collected at an anaerobic digestion plant, investigate possibilities to optimise the process, assess the applicability of the system and study possible issues. The data collection is the first important step in the protocol and, therefore, a detailed guide is provided to define/study the type of reactor, influent characterisation and long term data. This step is essential to properly install the TELEMAC system. For instance, the definition of the reactor type will influence how sensors will be installed and how the models will be applied. The assessment of the applicability provides the boundaries of use of the first version of the TELEMAC system. A study of possible issues will highlight possible future extensions to cope with particular applications.

After the initial assessment of the process with the available data at the plant, optimal experimental design is used to automate the improvement of practical identifiability of the TELEMAC system models. Also it will determine the optimum measurement set-up: measurement selection, measurement interval and influent dynamics to achieve the highest confidence in the parameter estimates and therefore the quality of the TELEMAC implementation at the plant

2 Protocol description

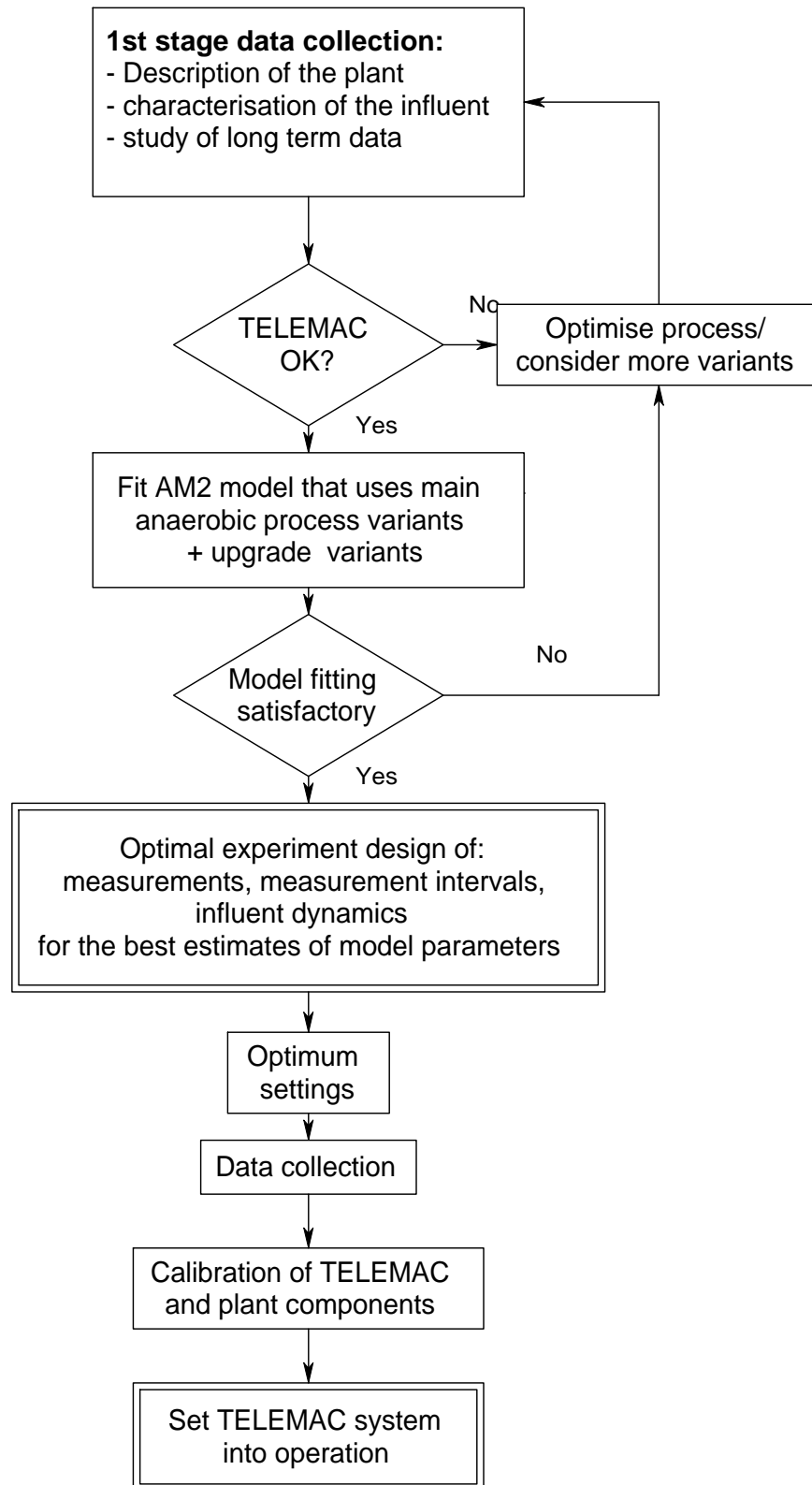


Figure 1: General protocol to optimise and set up the TELEMAC system

The general protocol to set up the TELEMAC system at a new treatment plant is shown in figure 1. It starts with data collection. Data collection, especially characterisation of the influent can be extended to a very detailed, laborious analysis but preferably it is not. In the application of the presented protocol it is advised to split the data collection into stages. Data is collected under three main categories: plant layout, influent characterisation and study of long term data collected at the treatment plant. A check list is made for each category highlighting common problems that can be found at the digestion plant and possible remedies. Those check lists are described in detail in section 2.1 to 2.3.

The intended first applications are devoted to the treatment of vinasses. For most of those plants it is expected that the protocol will proceed in the yes directions at the mentioned two check nodes (see figure 1). The simple models developed for TELEMAC are already validated on those applications. In case of unexpected problems at these plants or for application to other types of waste that is not strictly suitable for the TELEMAC system, process optimisation can be an option. Examples of problems that might be found are high solids contents, large variations in the influent, more particulate than expected and so on. Such problems can be solved by minor updates to the plant design or operation instead of investing in an extended monitoring system more than TELEMAC system. If the AM2 model is not able to describe the process dynamics, the process probably needs to be optimised and upgrade variants may have to be added to the model, e.g. toxicity that will be presented in the second set of models for abnormal conditions. These 2 back loops are not expected to cycle more than two to three times. Otherwise, an implementation of the present version of the TELEMAC system to a particular plant is not feasible.

To our opinion, the inclusion of optimal experimental design to set up a monitoring system for a biological process is innovative.

2.1 Plant layout study

2.1.1 Type of Digester

Most of the digester especially those installed at industrial plants are labelled by commercial names. Therefore, the following description will help in defining the type of digester and highlight the consequences of the type on the process optimisation, monitoring and control.

Practically, the digester can't be a perfect CSTR (Continuously Stirred Tank Reactor). Moreover, due to the slow growth of anaerobic bacteria, it is desired to maintain a long retention of the anaerobic biomass and keep them at high concentrations. Therefore, it is more common, for industrial applications in particular, to find what is called high rate anaerobic reactor systems. Contrary to aerobic processes, in an anaerobic or anoxic (denitrification) process the maximum permissible load is not governed by the maximum rate at which a necessary reactant can be supplied (e.g. oxygen during aerobic processes). Instead, to enable an anaerobic reactor system to accommodate high loading rates for treating a specific wastewater, the following conditions should be optimised:

- **High retention of viable sludge in the reactor under operational conditions.**
- **Sufficient contact between viable bacterial biomass and wastewater.**

- **High reaction rates and absence of serious transport limitations.**
- **The viable biomass should be sufficiently adapted and/or acclimatised.**
- **Prevalence of favourable environmental conditions for all required organisms in the reactor.**

Considerable efforts have been made in the last two decades to develop high rate anaerobic reactors that fulfill these conditions, first of all in order to enable the application of anaerobic treatment to cold and very low strength wastewaters. Additional reasons were:

1. to reduce the capital expenditure generally, because small reactors are lower in investment.
2. to reduce space requirements, because frequently space is the critical factor at industrial locations.
3. to improve the process stability.

Therefore, different configurations of high rate anaerobic reactors have been developed. The most common and widely applied configurations are classified and listed below:

Single stage anaerobic reactors.

The Contact Process

The contact process developed in the 1950's, is basically the anaerobic version of the aerobic activated sludge process, employing external settlers and sludge return. The various versions of the first generation of 'high rate' anaerobic treatment systems so far have not been very successful. The main difficulty appeared to be the separation of the sludge from the treated water. Later studies have shown that achieving the optimum mixing rate is the key point to have balanced micro-ecosystems for the rapid degradation of intermediates. Another benefit of achieving an optimum rate of mixing at a gentle and intermittent level is the excellent sludge sedimentation properties and better separation process. The maximum COD load to the contact (over the total reactor volume, including the settler) will not exceed 10 kg COD m³d⁻¹. A main disadvantage of the contact processes remains the need to return the separated sludge to the reactor by pumping. They are not widely used.

Anaerobic filters (AF)

The Upflow Anaerobic Filter (UAF) was developed in the USA by Young and McCarty (1969) in the late sixties. The performance of the UAF is mainly influenced by the sludge retention of the UAF and it is based on:

- The attachment of a biofilm to the solid (stationary) carrier material (e.g. synthetic packing, gravel, coke, bamboo segments). Size and weight are important properties of these materials. The most important are surface properties to increase bacterial attachment and the void ratio to prevent clogging. A preacidification step helps in clogging reduction
- The sedimentation and entrapment of sludge particles between the intersections of the packing material cause clogging.
- Formation of very well settling sludge aggregates.

The Downflow Stationary Fixed Film (DSFF) and Attached Anaerobic Fixed Film

(**AF**) is in many ways similar to the UAF except the direction of the flow. In a downflow packed bed, hardly any suspended biomass (or bacterial aggregates) can be retained (Kennedy et al., 1991). The system can also treat low strength wastewaters.

Anaerobic Sludge Bed (ASB) reactors

The ASB-reactor: is based on the formation of easily settling sludge aggregates (flocs or granules), and on the application of an internal gas/sludge/liquid separation system (GSS-device).

The Upflow Anaerobic Sludge Bed (UASB), was developed in the Netherlands in the early seventies (Lettinga et al., 1980). As in the AF-system the wastewater moves in an upward direction through the reactor. However, contrary to the AF system there is generally no packing material in the reactor. The sludge bed reactor concept is based on the following ideas:

1. Anaerobic sludge has or acquires good sedimentation properties provided that mechanical mixing in the reactor remains gentle and the process is operated correctly. For that reason, but also because it reduces the investment and maintenance costs, mechanical mixing usually is omitted in UASB reactors. Because of the excellent settling characteristics of the sludge, high superficial liquid velocities can be applied without any risk of considerable sludge wash-out.
2. The required good contact between sludge and wastewater in UASB-systems generally is accomplished:
 - by feeding the wastewater as uniformly as possible over the bottom of the reactor,
 - as a result of the agitation caused by the production of biogas.
 - In special cases, viz when the natural gas production doesn't suffice, some kind of gentle and intermittent mechanical mixing and/or gas recirculation could be applied. Such a situation may occur when applying the process to very low strength wastewaters at ambient temperatures.
3. The wash-out of sludge aggregates can largely be prevented by separating the biogas formed using a gas collection dome installed at the top of the reactor. In this way a zone with relatively little turbulence is created in the uppermost part of the reactor, consequently the reactor is equipped with a built-in clarifier. The gas collection dome acts like a three-phase separator (gas/sludge/liquidseparator: GSL separator). The GSL-device constitutes an essential part of a UASB reactor, it serves the following functions:
 - a) To collect, separate and discharge the biogas formed. For a satisfactory performance the gas-liquid area within the device should be sufficiently large, so that gas can evade easily. This is particularly important in case scum layers would develop. Sufficient mixing should prevail at the gas-liquid interface in order to combat this phenomenon. Since the formation of a scum layer is a very complex phenomenon, it is extremely difficult to give clear guidelines for the dimensions of the gas-liquid interface. In order to collect the biogas as completely as possible, one or more baffles should be installed beneath the aperture between the gas domes.

- b) To reduce liquid turbulence, resulting from the gas production, in the settler compartment. Also for this purpose the gas evolved in the sludge bed should be collected as completely as possible.
- c) To remove sludge particles by a mechanism of sedimentation, flocculation and/or entrapment in a sludge blanket (if present in the settler). The collected sludge can slide back into the digestion compartment, in case the sludge bed does not reach into the settler, or can be discharged occasionally together with excess sludge from the digester compartment.
- d) To limit the expansion of the sludge bed in the digester compartment. The system more or less acts as a barrier against excessive expansion of the lighter part of the sludge bed. In case the sludge bed expands into the settler, the sludge will tend to thicken (because the gas has been separated). This thickened and heavier sludge present in the settler lays on top of the more voluminous sludge blanket that tends to move into the settler.
- e) To reduce or prevent that buoying sludge particles washed out from the system. For this purpose a skim layer baffle should be installed in front of the effluent weir of the overflow. Such a baffle is particularly essential for treating very low strength wastewaters, because wash out of viable biomass then should be kept at very low levels.
- f) To accomplish some polishing of the wastewater with respect to suspended matter and - to some extent - also of remaining degradable soluble substances. For that purpose it is beneficial that the sludge bed expands into the settler.

Upflow Anaerobic Sludge Bed – Anaerobic filter (UASB-AF)

Some researchers suggest to replace the GSL -device by a packed bed in the upper part of the reactor. This recommendation is based on results obtained with some hybrid of the UASB-AF (AnaerobicFilter) system. Such combined reactors can maintain high treatment efficiencies at COD-loads considerably higher than those accommodated by upflow and downflow completely packed reactors. These hybrid reactor systems certainly offer quite interesting potentials for the treatment of low strength wastewaters containing a fraction of finely dispersed solids. However, in such a hybrid reactor it remains beneficial to also install a GSS-device. This certainly is the case when treating higher strength wastewaters. However, this combined system is still undergoing further research to improve its practical use.

Anaerobic Fluidized Bed (FB)

As a matter of fact, FB systems can be considered as a special variant of the sludge bed system. The FB-process is based on the occurrence of bacterial attachment to mobile carrier particles, e.g. consisting, of fine sand (0.1 - 0.3 mm), basalt, or plastic. The system relies on the formation of a more or less uniform (in thickness, density, strength) attached biofilm and/or particles. In order to maintain a stable situation with respect to the biomass film development, a high degree of pre-acidification was considered necessary (Heijnen, 1989) while dispersed matter should be absent in the feed (Ehlinger, 1994). It is virtually impossible to guarantee a stable process performance in practice. The complex dynamic process of film formation, viz. film attachment and release, can not be controlled sufficiently. Aggregates of

very different size and density will always develop in the system, and consequently - at a certain imposed specific superficial liquid velocity, segregation of sludge particles will occur. This means that at some time bare carrier particles will start to accumulate in the lower part of the reactor as a kind of stationary bed, while light fluffy aggregates (detached biofilms) will be present in the upper part, at least when they can be retained. Thus like UASB, gas-liquid separator may be needed in the top of the reactor. Moreover, when the biofilm thickness increases in size and the volume of the bed thus increases, a separate system (e.g. recycle pump) is needed to wash part of the media and return it to the reactor. In this way the density of the media is controlled and a more homogeneous reactor bed is created. Up to 30-90 kg VSS m⁻³ reactor can be retained in this way and because of the applied high liquid upflow velocities (i.e. up to 10 m/hr) an excellent liquid-biomass contact is accomplished. The system is applicable to wastewaters with a SS-concentration < 500 mg/l. The idea that a slight expansion of the sludge bed suffices for the required contact is the main difficulty of such system.

Anaerobic Attached Fixed Film Expanded Bed (AAFFEB) and Expanded Granular Sludge Bed (EGSB)

The **AAFFEB** and **EGSB** are further developed versions of the **FB**. The **AAFFEB** has a mobile carrier material, e.g. glass, plastic, an ion exchanger, or diatomaceous earth. These media are used for bacterial attachment. The **EGSB** relies on the formation of granular sludge due to the characteristics of the wastewater and the good settling properties of the formed sludge, which will act as a media by itself. The main difference in the design with a normal **FD** is that the specific external area of the media (or granules) is relatively small, ca. 100 m² m⁻³. Bed expansion is realised by recycling the effluent. To some degree, also settlement of sludge particles occurs in the bed. Because of the high settleability of the sludge superficial liquid velocities exceeding 6 m/hour, sometimes even higher than 10 m.h⁻¹, can be applied. These high liquid velocities, together with the lifting action of gas evolved in the bed, leads to a slight expansion of the sludge-bed. And as a result of that an excellent contact between sludge and wastewater prevails in the system, leading to significantly higher loading potentials compared to conventional UASB. These upgraded types of **FB** are less sensitive to some toxicities. In this type of reactor, the biogas formed is separated halfway up the reactor by means of partitions, and is then led upwards through a pipe. The lifting forces of the collected biogas are used to bring about a recirculation of granular sludge over the lower part of the reactor, which results in improved contact between sludge and wastewater. Thus, the control of liquid/sludge recirculation depends on the level of gas production.

The Baffled Anaerobic Reactor is a compartmentalised sludge bed reactor in which the compartments are operated in series. No special measures are taken to retain the sludge in the various compartments, except after the last one. Since the superficial liquid velocity in a baffled system is substantially higher than in a single step sludge bed reactor, most of the sludge will move with the liquid through the various compartments. The sludge is separated after the last compartment in a settler and then returned to the head of the reactor. The system has been tried on laboratory scale and pilot scale for sewage treatment. Uptill now there is no full scale implementation with a ratified performance. Upscaled installations of the concept are expected to resemble sequentially operated moduled UASB-reactors and or staged anaerobic reactors. Such system is promising since it was found that the sludge granulation process proceeds quite well in these systems.

The Anaerobic Sequencing Batch Reactor (ASBR): is another interesting configuration of anaerobic reactors. It consists of a set of anaerobic reactors operated in a batch mode using a 'fill and draw' method. A certain amount of the raw wastewater is supplied to the anaerobic reactor, after the supernatant liquid of a previously batch has been discharged. Then a 'gentle' type of mixing of the reactor contents is started in order to enable the settled viable sludge to contact the wastewater and to eliminate the biodegradable organics. After a sufficient period of reaction time, the sludge is allowed to settle and the supernatant solution is discharged. The next cycle is then started. Granulation proceeds well in an ASBR treating dilute wastewaters, also at lower ambient temperatures (Dague et al,1997).

Membrane Type Anaerobic Reactors of recent date are also systems using membrane technology for sludge retention. The reactor concepts may become attractive for treating very specific wastewaters with e.g. xenobiotics in case the adapted - specific - anaerobic cultures can not be retained in a simpler way. Particularly when the membrane permeate fluxes can be kept high and the overall treatment costs of these systems can be reduced, membrane type anaerobic reactors may offer prospects for anaerobic treatment. Moreover it can help in the separation of acidogenesis from the acetogenesis-methanogenesis step in the staged reactor, illustrated in the next section.

Staged reactor systems.

A very important improvement in anaerobic treatment undoubtedly concerns the **Staged Sludge Anaerobic Reactor (SSAR)** concept. It comprises a kind of anaerobic plug flow treatment system (Van Lier et al, 1994). Process staging can be accomplished by using a number of sequentially operated reactors (or compartments of a single reactor). In the various modules of the SSAR-system in principle still all phases of the anaerobic degradation process are allowed to proceed simultaneously. For a non- or partially acidified soluble substrate the acidogenic flora will develop mainly in the first module(s), but also acetogenic and methanogenic organisms are allowed to develop and to flourish (which in fact is contrary to ideas underlying the phase separation concept). Consequently the system is completely self-regulating in this respect. When treating a partially soluble wastewater, the first module mainly will serve for hydrolysis and acidification.

In the next modules of the system along with the evolution of the degradation process, gradually a sludge with a higher methanogenic activity will develop. It is clear that the sludge present in the separate compartments will differ depending on the remaining compounds/intermediates to be degraded and the specific environmental conditions prevailing in that compartment. Since *mixing of the sludge over the total reactor system is prevented*, indeed a very specific sludge can (and will) develop in each separate module. In case a complete phase separation (between acidogenesis and methanogenesis) would occur in a modulated reactor, this simply will be the result of a kind of self-selection process. A staged reactor system will provide a higher treatment efficiency, because also more difficult compounds like intermediates such as propionate will find a more optimal environment for degradation. Moreover the process stability of such a system will be higher. This is true for both mesophilic as well as thermophilic systems, but particularly for the thermophilic systems. Relative to mesophilic digestion, the thermophilic digestion suffers more seriously from a higher degree of product and/or substrate inhibition (Van Lier et al, 1994). A satisfactory treatment of concentrated wastewaters under high loading conditions therefore generally can not be accomplished in a single stage thermophilic reactor. The need for

staging was demonstrated by (Van Lier et al., 1994) in experiments conducted in a compartmentalized reactor using a sucrose-VFA mixture as feed. They found a high treatment efficiency (effluent VFA -COD < 300 mg/l) under loading conditions up to 120 kg COD m⁻³.day⁻¹. Sucrose was converted in the first compartment, followed by the conversion of butyrate and acetate in the next compartments, leaving the elimination of propionate to the last compartments.

The use of a SSAR-reactor certainly is also very profitable for mesophilic anaerobic treatment, particularly for the conversion of 'difficult' compounds, e.g. xenobiotics.

Acidifying reactors.

These reactors are not designed for complete anaerobic digestion. They are designed for the acidification step only when preacidification of the wastewater is required in certain situations e.g. treatment of cold wastewaters.

Liquefying (& acidifying) reactors.

In some cases, a high proportion of the COD load is in the particulate form. Some particulates are slowly hydrolysed. Therefore a pre-treatment step of hydrolysis can be performed in a separate reactor in the up-stream and can be combined with the acidification step.

The Hydrolysis Upflow Sludge Bed (HUSB), was investigated in China for domestic sewage pre-treatment. Later more detailed investigations were made in the Netherlands (Wang Kayun, 1994). The HUSB uses a flocculent type of sludge and the imposed upflow liquid velocities are approximately 1 m h⁻¹. Once again, all anaerobic degradation steps are allowed to proceed in the system. The potentials of the HUSB-system are:

- Removal of Suspended Solids from the wastewater
- Entrapment of precipitating substrates (e.g. proteins).
- Partial or (almost) complete liquefaction of entrapped insoluble or precipitated substrate ingredients.
- Partial or complete acidification of liquefied COD,
- Some conversion of acidified substrate into methane may occur.

The design and performance of the HUSB-system can be improved in the following ways:

- Imposing optimal environmental conditions, e.g. for substrate precipitation. Once sufficiently well understood, this process could be optimised so that a higher fraction of substrate can be brought to precipitation.
- Elevation of the temperature into the optimal range for hydrolysis. This can be done occasionally and obviously is of particular importance when the temperature of the waste water subjected to the treatment is relatively low. The rate of hydrolysis is low at temperatures beneath 15-20 °C.
- By applying intermittent mixing in the reactor in order to distribute organisms and enzymes more evenly over the organic matter.
- By supplying flocculants to improve the removal of SS.
- Combining the system with a sludge digester which is operated under optimal

environmental conditions. Compared to the amount of wastewater treated in the HUSB a relatively small amount of thick sludge will be conveyed to the digester. Therefore the heat requirements for such digester are small.

- Additions of specific chemicals which enhance the hydrolysis.

2.1.2 Temperature Operation Range

As in all biological processes, one can distinguish three temperature ranges in anaerobic biological treatment of waste and wastewater:

- a psychrophilic range, from 0-20 °C,
- a mesophilic range, from 20-42 °C,
- and a thermophilic range, from 42-75 °C.

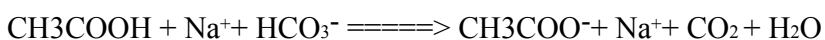
The upper limits of these ranges are defined by the temperature at which the decay rate of the bacteria starts to exceed the growth rate. This means that thermophilic bacteria are still quite active in the mesophilic temperature range, and that mesophilic bacteria are still active in the psychrophilic temperature range. Especially, this last feature is very important and one can treat wastewaters at very low temperatures, provided one doesn't have to grow the bacteria at this low temperature. Anaerobic sludge, grown at 35°C, can be used very well to handle wastewater at 10-15 °C. Then still 5-10 % of its activity remains. Also, at 20°C still volumetric loading rates of 4-5 kg COD m⁻³d⁻¹ can be effectively treated. In the mesophilic range, it is possible to say that the activities and growth of the bacteria decreases by one half for each 10°C decrease below 35°C.

2.1.3 Upstream Buffer Tank

Too high hydraulic and/or COD-loading leads to wash-out of sludge and decreased removal efficiencies. If such peaks are expected in the influent wastewater a buffer tank can be designed to accommodate a part of the wastewater during the hydraulic and/or COD peak time. The flows to the anaerobic reactor can be managed by a control. The control can be based on an on-line measurement of the process variables such as flow, TOC or flow depending on the situation. The use of the gas flow is based on the principle that at too high COD-loading the biogas load surpasses a certain critical value (around 0.5 m³m⁻³h⁻¹ for UASB). In a UASB, too high gas loading rates would lead to an unacceptable decrease in the performance of the gas collectors. Reduction of the flow to the reactor means that (temporarily) the flow to the buffer tank will increase. Then the flow from this buffer tank is added when the peak is gone (seasonal changes).

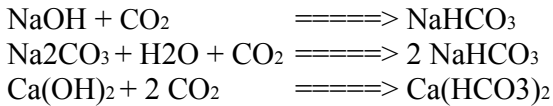
2.1.4 pH Adjustment

A very important aspect in the control of the pH in an anaerobic system, is the fact that VFA's are formed during the acidification phase and can accumulate in the reactor contents. When there is enough bicarbonate alkalinity available, the following reaction will occur, for instance with acetate (C2):



Since bicarbonate is the most crucial substance in the buffering of the reactor content, the most logical chemical for supplementation of bicarbonate alkalinity is sodium bicarbonate. It

is the only product which will gently shift the equilibrium to any desired value without disturbing the physical and chemical balance of the delicate biological community. Unfortunately bicarbonate is more expensive than caustic and lime. Alkaline reagents such as caustic soda, soda ash, and lime cannot increase bicarbonate alkalinity without reacting with soluble CO₂:



Addition of these components therefore will result in a lower CO₂-concentration of the biogas.

2.1.5 Other Factors

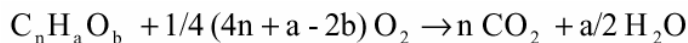
Other points that have to be added to the check list to describe the plant are:

- influent pump configurations
- possibility of influent dilution and/or recycle rates
- size of the reactor(s) for both liquid and gas compartments
- available sensors and off-line analysis
- available actuators and applied control systems

2.2 Detailed influent characterisation

2.2.1 COD, TOC compared to theoretical COD and oxidation state of carbon

The Chemical Oxygen Demand (COD) and Total Organic Carbon (TOC) are important parameters for the wastewater characterisation. Also, the determination of their fraction helps in the assessment of the corresponding process effects. The description of the analytical methods to determine COD and TOC are out of the scope of this report. For an organic compound C_nH_aO_b



$$\text{COD}_t : 8 * (4n + a - 2b) / (12n + a + 16b) \text{ g COD/g C}_n\text{H}_a\text{O}_b$$

$$\text{TOC}_t = 12n / (12n + a + 16b) \text{ g TOC/g C}_n\text{H}_a\text{O}_b$$

$$\text{COD/TOC} = 8(4n + a - 2b) / (12n) = 8/3 + 2(a - 2b) / (3n)$$

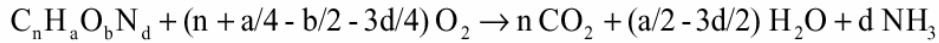
The COD/TOC ratio indicates the complexity of the substrates in the wastewater. Generally, if the ratio lies in the range 2 to 3, it is more likely that the substrates are less complex and are anaerobically degradable compounds. Thus if the ratio calculated from the analytical results is out of this range certainly further characterisation of the anaerobic degradability of the wastewater is required.

Including the organic nitrogen in the calculation of the theoretical COD, one considers that organic-N is converted into NH₃-N. This conversion is also what happens in the COD-test.

For organic compound $C_nH_aO_bN_d$ the average oxidation number of the C-atom :

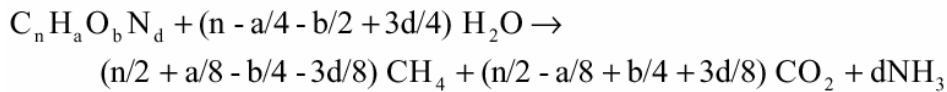
$$\text{OxSt C} = (2b - a + 3d)/n$$

and theoretical COD is calculated as the quantity of oxygen consumed in the following chemical oxidation



2.2.2 Theoretical gas composition

If the compound $C_nH_aO_bN_d$ is anaerobically degraded, the produced gases can be estimated from the Buswell equation (Buswell, et. al., 1930)



Thus if the influent substrate composition is determined (sections 2.2.3 and 2.2.4), it is possible to evaluate the produced gas composition. The following figure (2) shows the theoretical composition of the biogas produced from the anaerobic degradation of different compounds in relation to their carbon oxidation state.

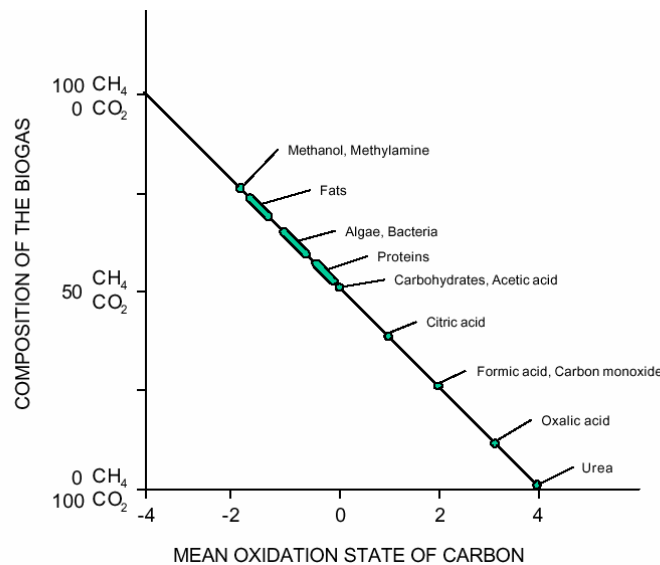


Figure (2): The theoretical composition of the biogas produced from the anaerobic degradation of different compounds in relation to their carbon oxidation state.

Similarly, the composition of the gas produced from a digester can provide information on the most probable substrate in the influent. This information can only be used as indication provided that the gas composition data are collected from the digester operating under normal operating conditions and with adequate retention time.

It should be noted that if other electron acceptors such as sulphate, sulphite, nitrates and nitrite will cause a drop in the methane production compared to the one estimated by the

Buswell equation. Also, presence of these electron acceptors will cause a drop in the COD measurement compared to the COD of the existing organic compounds.

For CO_2 estimated by the Buswell equation will be significantly larger than measured from the biogas for 2 reasons. First there is because the high solubility of CO_2 in water. The second reason is the transformation of the dissolved CO_2 to bicarbonate.

2.2.3 Substrate composition, hydrolysis and degradability

Carbohydrates:

Most of the carbohydrates are anaerobically degradable. The common polysaccharides and the sugar monomers are easily degraded by anaerobic sludge. Cellulose biodegradation is less rapid. However, the rate is sufficiently rapid that the biodegradation of pure cellulose is not a significant rate limitation. The cellulose of plants is never pure. Natural cellulose is present in lignocellulose. The percentage anaerobic biodegradation of the lignocellulose is less than pure cellulose due to the protective effect of lignin. Thus, if the quantity of particulate COD: COD particulate = CODt-CODs is high, quantification of lignocellulose might be needed. Particulate degradability can be assessed by comparing CODacid (defined 2.3.6) with CODt and CODs in the influent.

Proteins and amino acids:

Proteins are often easily hydrolysed into amino acids, however they are sometimes coagulated to insoluble forms when exposed to either heat, acids or tannins. Most types of proteins are hydrolysed and degraded anaerobically. In an anaerobic reactor protein COD is converted in anaerobic reactor to methane and protein organic N to NH_4^+ -N. The degradation of proteins is not rate limiting. For example, the rate of gelatine protein degradation in an acidifying reactor was found to be $12.0 \text{ g COD g}^{-1}\text{VSS d}^{-1}$.

Fats and long chain fatty acids:

Fats are polymers of long chain fatty acids (LCFA) linked to a glycerol molecule with ester bonds. The hydrolysis of fats by extracellular lipase enzymes is generally rapid if the fat is soluble. The fats are more soluble if the pH value is high (pH 8) compared to the pH of acidifying reactors (5.5 - 6.0) where the fat is mostly insoluble and the hydrolysis is slow.

The anaerobic degradation of the LCFA monomers of fat is actually not by fermentation but more similar to the anaerobic oxidation that is done by acetogenic bacteria. The most important products of the anaerobic oxidation of long chain fatty acids (LCFA) is C_2 and H_2 (67% and 33% of the LCFA acidification end product COD respectively). The degradation of the LCFA is inhibited by the presence of too much H_2 and C_2 . The LCFA acidification is therefore only possible if an active methanogenic population is present to remove the H_2 and C_2 from the media.

Phenolic Compounds:

The phenolic compounds present in wastewater are usually derived from the lignin and the tannin of plants. Lignin is apolar and is usually only soluble in alkaline conditions. However, some low MW forms are soluble. The tannins are water soluble compounds. Tannins have ester intermonomeric bonds that are easily biologically hydrolysed and acidified during anaerobic digestion

The phenolic compounds can be divided into two major groups: **Monomeric and Polymeric Phenol Compounds**. The anaerobic degradation of **Monomeric Phenolic Compounds** depends on the number of hydroxy (or methoxy) groups per ring of the phenol. If these compounds have three hydroxy (or methoxy) groups per ring, they are degraded anaerobically without any lag period by unadapted sludge. Their degradation does not depend on the methanogenesis. If these compounds only have one or two hydroxy (or methoxy) groups, they are not degraded rapidly and require a certain adaptation period before the anaerobic degradation occurs. Similar to LCFA their degradation is inhibited by acidification end products like C_2 and H_2 . The **Polymeric Phenol Compounds** (*Lignin and Tannins*) are generally less biodegradable than the monomeric phenols. The large native polymers of lignin in wood and straw are not biodegradable during anaerobic digestion. Smaller polymers (oligomers) of 2 to 7 monomeric units (300 to 1000 g mole⁻¹) and are partially degradable, indicating that the intermonomeric bonds of lignin can be hydrolysed by anaerobic bacteria. The poor biodegradability of native lignin is due to the bad access of hydrolytic enzymes to the intermonomeric bonds of these large polymers. The lignin in the wastewater is generally produced due to alkali treatment of the native lignin. Therefore, wastewater lignin is potentially partially biodegradable during anaerobic treatment. Similarly condensed tannins are partially degradable during anaerobic treatment. The natural tannins are usually oligomers indicating that partial degradation of the condensed tannins will occur. However, the natural colourless oligomeric condensed tannins can oxidise to dark coloured high MW tannins that are not degraded anaerobically.

Inert particulates:

This fraction can be determined approximately by determination of TSS and VSS where the difference is the estimate of the inerts. It is not that easy, however, TSS and VSS are usually quantified in mg/l and for many reasons they need to be quantified in other units, as COD for instance. Therefore, an approximate chemical composition of the inerts might be needed. This can be determined by further analysis or by investigating the process at the source of the wastewater.

2.2.4 VFA levels

VFAs are easily biodegradable substrates. In the anaerobic process they are the intermediate products between the main process kinetics acidogenesis, acetogenesis and methanogenesis. Some wastewaters are high in VFA concentration especially those originating from fermentation processes (e.g. vinasses). VFAs have pKa values between 4.7 and 4.9 and when produced they will cause a drop in the wastewater pH. Therefore, high VFA levels in wastewaters are favourable to anaerobic digestion provided that the pH is controlled to the optimum level (around 7).

2.2.5 VFA fractions

The acidification of monoscharides by anaerobic bacteria produces acetate C_2 , propionate C_3 and butyrate C_4 . Thermodynamically, the production of C_2 by anaerobic acidifying bacteria is more favourable than respectively C_3 and C_4 . Thus, for acidified wastewaters that mainly contain carbohydrates, it is expected that acetate will be the highest proportion of VFA. Under stress conditions to the process, propionate and butyrate concentrations are expected to increase. For example, at high hydrogen concentration more propionate is produced.

Wastewaters with high proteins (amino acids) content mainly yield acetate, butyrate, valerate (C_5) and propionate when acidified. Usually, the yields from amino acids acidification is

dependent on their type and lead mainly to:

- High acetate (C₂) production but less than the levels produced from monoscharides acidification,
- Butyrate (C₄) which is high compared to the proportion produced from monoscharides but less than acetate (C₂),
- Valerate (C₅) which is produced to a similar level of (C₄)
- A Low propionate (C₃) produced

Generally, acetate is the highest VFA fraction expected from acidification and higher carbon VFAs (but not propionate) are expected in high proportion from protein acidification.

2.2.6 Assessment of possible precipitation and adsorption

These two mechanisms could lead to some COD removal when removed with SS. However, they have adverse effect on high rate reactors (e.g. Upflow Anaerobic Sludge Bed (UASB) and Fixed Bed Reactors (FBR)). The entrapment of the solids in the high rate reactor may cause dilution of the methane bacteria population while the contact time with the wastewater is relatively short. Some mechanisms that might cause these effects are:

1. Lignin (lowering pH below 9 or adding calcium)
2. Fat (lowering pH below 8 or adding calcium)
3. Some proteins (lowering pH below 6)
4. Humic acids (lowering pH below 5 or adding calcium)
5. Pectin (adding calcium)
6. Protein plus tannins (form protein-tannin aggregates)

2.2.7 Determination of sulphates

Determination of sulfur and sulfates in the influent wastewater is important. It has been shown that sulphate reducers are able to outgrow the methanogens. This is due to the higher energy gained by sulphate reduction compared to methanogenesis. Therefore, the presence of high concentrations of sulphate will result in the determination of low methane potentials. More over, the reduction of sulphate to sulphides or H₂S is toxic to methanogens as will be illustrated latter.

2.2.8 Determination of proteins and N compounds

Similar to sulphates, nitrates and nitrites need to be measured since there reducers will compete with methanogens. Moreover other N-compounds should be checked. There are three reasons to check the N-concentrations of the influent and the effluent (the reactor):

- (1) NH₄⁺-N is an important nutrient for bacterial growth
- (2) NH₄⁺-N can cause toxicity to methanogenic bacteria
- (3) NH₄⁺-N can neutralise a pH drop, caused by hydrolysis and VFA formation

Also quantification of the total nitrogen helps to empirically determine the amount of proteins in the influent. The "crude protein" concentration can be estimated assuming that the org-N is mostly proteins and amino acids and can be calculated by the factors listed below:

$$\text{protein (g/l)} = \text{org-N (g/l)} * 6.25$$

$$\text{protein (g COD/l)} = \text{org-N (g/l)} * 7.81$$

2.2.9 Check nutrients and trace elements

Anaerobic biological wastewater treatment processes are performed by bacteria which have to grow during the wastewater treatment. Otherwise they will eventually wash out of the treatment system. For this reason the wastewater has to contain a number of compounds from which the bacteria can synthesise their cell constituents. The nutrient composition of methane bacteria is presented in Table 1. These include macronutrients like nitrogen, phosphate, and sulphur needed in moderate concentrations. All other elements (the micronutrients or trace elements) that the anaerobic bacteria need for growth must also be present.

Table 1. The elemental composition of methane bacteria (according to data published by Scherer, 1983).

MACRONUTRIENTS		MICRONUTRIENTS	
Element	Concentration mg/kg dry bacteria	Element	Concentration mg/kg dry bacteria
<i>N</i>	65000	<i>Fe</i>	1800
<i>P</i>	15000	<i>Ni</i>	100
<i>K</i>	10000	<i>Co</i>	75
<i>S</i>	10000	<i>Mo</i>	60
<i>Ca</i>	4000	<i>Zn</i>	60
<i>Mg</i>	3000	<i>Mn</i>	20
		<i>Cu</i>	10

The minimal concentration of macro- and micronutrients can be calculated based on the biodegradable COD (COD_{BD}) concentration of the wastewater and its degree of acidification (which determines the cell yield factor). The concentration of nutrients required can be calculated by the following equation:

$$\text{NUT}_{\text{required}} = \text{COD}_{\text{BD}} * Y * \text{NUT}_{\text{bacteria}} * 1.14$$

$$\text{NUT}_{\text{influent}} = \text{NUT}_{\text{required}} * 2$$

COD_{BD} = COD_{BD} concentration in the influent (mg COD L⁻¹); Y = cell yield (g VSS g⁻¹COD consumed); NUT_{Rbacteria} = concentration nutrient in bacterial cells (g nutrient g⁻¹ dry bacteria); 1.14 = assumed TSS:VSS ratio of bacterial cells; NUT_{Rrequired} = minimal nutrient concentration required in the influent (mg L⁻¹). The cell yield coefficients to use, depends on whether the wastewater is already acidified. For wastewaters that are not acidified, a Y value of 0.15 should be used. For wastewaters that are already acidified, a Y value of 0.03 should be used. The nutrient concentration in the influent should be adjusted to a value equal to twice the minimal nutrient concentration required to ensure that there is a small excess of the nutrients needed.

2.2.10 Assessment of possible toxicity

2.2.10.1 Inorganic toxins

Ammonium

Ammonium nitrogen is present in wastewaters that originally contain high concentrations of proteins or amino acids. The organic nitrogen is mineralised to ammonium during anaerobic

digestion. The toxicity of ammonium is due to the unionised form (free NH_3). The fraction of free NH_3 is low at a pH value of 7 (0.01) but it is about 10 times higher at pH 8. The concentration of free $\text{NH}_3\text{-N}$ responsible for a 50% inhibition (50% INHIB) of the methanogenic activity in unadapted granular sludge is 50 mg L^{-1} (Koster and Lettinga, 1983). The toxicity of NH_3 is reversible. Adaptation of the sludge to high levels of NH_3 is possible (Van Velsen, 1979).

Sulphur

Wastewater may contain inorganic forms of sulphur like sulphate (SO_4^{2-}) and sulphite (SO_3^{2-}). During the anaerobic digestion, these compounds are microbiologically reduced to hydrogen sulphide (H_2S). The toxicity of H_2S is due to the unionised form (free H_2S). The concentration of free H_2S which causes 50% inhibition of methanogenic activity in granular sludge is approximately 250 mg S L^{-1} (Koster et al., 1986). Sulphate is relatively non-toxic. Therefore, the biological reduction of SO_4^{2-} to H_2S during the anaerobic digestion increases the toxicity of sulphur. In contrast, sulphite is more toxic than H_2S . The biological reduction of SO_3^{2-} is very desirable because it will decrease the toxicity of sulphur.

Salt

Salt can cause toxicity problems if the concentration is very high. High concentrations of salts are sometimes present in the wastewaters of industrial processes where salt is added. For acidified wastewater, e.g from fermentation processes, high concentrations of salt are sometimes needed to neutralise high concentrations of VFA. In these cases, the possibility of salt toxicity should be considered. The methanogenic toxicity of various kinds of salts is listed in Table 2. The results of this table indicate that monovalent cations are less toxic than divalent cations like calcium (Ca^{2+}). However the low solubility of Ca^{2+} in the presence of bicarbonate at mild alkaline conditions may result in a low effective concentration in the reactor.

Table 2. The 50% Inhibitory concentration of salts to the methanogenic activity of digested domestic sludge (Kugelman and McCarty, 1965). pH= 7.0, T= 35°C.

Salt	50% inhibitory concentration (mg L^{-1})
Mg^{2+}	1930
Ca^{2+}	4700
K^+	6100
Na^+	7600

However, the toxicity of sodium (Na^+) is slightly greater at pH 8 compared to pH 7. The toxicity is twice as strong for the anaerobic degradation of C_3 and C_4 compared to the degradation of C_2 . This indicates the greater effect of Na^+ on the autotrophic compared to acetoclastic methanogenesis. The Na^+ toxicity is reversible but slow compared to adaption to other toxicities such as NH_3 toxicity

Heavy Metals

Heavy metals are sometimes present in wastewaters. Occasionally, it is necessary to add heavy metals as nutrients to wastewater for anaerobic treatment. However, care should be taken to avoid an over-dose which can have a toxic effect. The toxicity of heavy metals

depends on the soluble concentration. The soluble concentration of heavy metals would decrease as a result of precipitation reactions with carbonate (CO_3^{2-}) and sulphide (S^{2-}) which are generally present in anaerobic digesters. The precipitation of heavy metals is more effective at increasing pH due to the pH dependency of CO_3^{2-} and S^{2-} . The heavy metals 50% inhibition of methanogenesis is in the range of 30 to 300 mg L⁻¹.

2.2.10.2 Toxicity of natural organic compounds

The toxicity of natural organic compounds is generally due to one of two basic reasons:

1. Apolarity: compounds which are apolar damage the membrane systems of the bacteria.
2. Hydrogen (H) bonding: certain phenolic compounds called "tannins" are adsorbed by proteins due to hydrogen bonding. If the adsorption is strong, the bacterial proteins (enzymes) can be damaged.

Volatile Fatty Acids

The toxicity of VFA is dependent on the pH because only the unionised (free) acids are toxic. The 50% inhibition concentration of unionised C₂ and C₃ for the methanogenic activity was reported to be 16 and 6 mg COD L⁻¹ (Kroiss, 1985). If the reactor pH is low, the unionised fraction is sufficiently high that a high concentration of VFA cannot be tolerated. In contrast, the VFA are relatively non-toxic at pH 7 and greater. The toxicity of VFA solutions at pH 7.4 for adapted granular sludge is usually not evident at concentrations up to 15,000 mg L⁻¹.

Fortunately, the inhibition of methane bacteria by the combination of VFA and low pH is reversible. In general, VFA inhibited activity will be restored several days to several weeks after the low reactor pH is corrected. If the low pH conditions is only 12 hours or less, the methanogenic activity will recover at the time the pH is corrected (Anderson et al., 1982).

Long Chain Fatty Acids

The toxicity of long chain fatty acids (LCFA) is more severe than of the smaller VFA. The toxicity is largely affected by the anaerobic degradation of the LCFA. Generally, when the anaerobic degradation occurs an immediate recovery of the methanogenic activity will be evident. The LCFA causes severe inhibition of the methane production for a period of several weeks. After the sludge starts to degrade the LCFA, the methanogenic activity is completely recovered. Hanaki et al. (1981) have demonstrated that the presence of VFA in the media increases the toxic effect of LCFA. This is because the VFA inhibit the anaerobic degradation of LCFA by acetogenic bacteria.

The LCFA are not always completely soluble in the anaerobic digester. Low pH and Ca²⁺ can cause insolubilisation reactions. Additionally, LCFA are adsorbed onto the surfaces of anaerobic sludge (Keurentjes and Rinzema, 1986). LCFA toxicity to granular sludge (50% inhibition at concentrations > 1000 mg l⁻¹) is less than its toxicity to normal sludge (50% inhibition at concentrations 250 to 500 mg l⁻¹).

Apolar Phenolic Compounds

It is generally possible to estimate the toxicity of many monomeric phenols based on the degree to which the phenolic compounds have apolar characteristics. The phenolic compounds are more toxic for the degradation of C₃ compared to C₂. For example, Phenol

concentrations of 2250 and 875 mg l⁻¹ cause 50 % inhibition to the degradation of C₂ and C₃ VFA's respectively (Blum et al., 1986). Due to the apolar properties of lignin it is expected that low molecular weight phenolic compounds derived from lignin can cause methanogenic inhibition.

Tannic Organic Compounds

Polar phenolic compounds are able to form hydrogen bonds with bacterial proteins. The hydrogen bonding of polymeric tannins is very strong due to the presence of more than one bond with the protein per tannin molecule. The toxicity of tannins for methane bacteria are relative to the tannin molecular size. The oligomeric tannins are the most effective inhibitors because they bind strongly with enzymes and are small enough to penetrate into the cell. High molecular weight tannins can bind strongly with enzymes but are not able to penetrate into the cell and are therefore not very toxic to bacteria.

Other examples of organic toxicity that can be further investigated in particular cases are *Phenolic Amino Acids, Caramel Compounds* (e.g. *Furfurals*), *Xenobiotic Compounds, Chlorinated Hydrocarbons, Formaldehyde, Cyanide, Petrochemicals, Detergents and Antibiotics*.

2.3 Analysing long term data

2.3.1 Determine seasonal changes

Most of the wastewater treatment systems are subject to seasonal changes, especially in the influent, which introduces high levels of uncertainty to the system design and operation. For domestic wastewater treatment, some of the uncertainty is resolved by studying the daily and annual variation of water consumption, the main source of wastewater, as well as the annual variation of rain and infiltration. Similarly, in industrial applications uncertainty can be resolved by investigating the process ahead, the main source of wastewater. Also, other long term variation can be estimated from seasonal variation of the ore material, e.g. crop harvest seasons for industries based on agricultural products (grape for vinasses). Moreover, studying the long term records of the plant help in locating these variations and thus raise questions on these variations to the manufacturers.

2.3.2 Which measurements are generally available

Generally at an anaerobic digestion plant with a limited monitoring system relatively cheap measurement are expected to be available on-line. For example, pH, temperature, influent flow, gas flow and pressure. In addition, frequent off-line measurement will be done such that COD/ TOC, alkalinity data are available. Some detailed VFA records, but less frequent, records may be present. Indeed, this is the minimum level expected. If it is not available, an initial measuring campaign is mandatory before setting up the new monitoring system.

2.3.3 Gas composition

Three gases are normally produced in an anaerobic digester: methane, carbon dioxide and hydrogen. Additionally, hydrogen sulphide and/or ammonia are expected for digesters treating wastewaters with high sulphates and/or nitrogen compounds. Practically, the methane to carbon dioxide ratio is the most important information about the gas composition

when controlling the digester. The other gases are of particular importance depending on the composition of the waste or regarding their elimination before the methane/carbon dioxide sensor. In the latter case humidity also comes into the picture. Other factors that may cause more measurement of the gas components are the further processing of the gas (e.g. treatment before power generation), odour control and controlling corrosion /erosion of the transporting pipes.

2.3.4 Fluctuation of the reactor pH

The production of methane will perform optimally at pH values between 6.5-7.5. In some cases, however, the methane digestion will perform well beyond this range. Below a pH of 6.0 methane digestion on VFA will proceed very slowly if at all. The influence of the pH on the methane production is related with the concentration of the VFA. This is because C_2 , as well as the other VFA are toxic in their unionised form. This means that the lower the pH is, the more toxic VFA will be, as the fraction of free acid will increase with decreasing pH.

If it exists, data records collected during pH shocks to the reactor should be examined. The influence of a pH-shock for the anaerobic sludge will be depending on :

- the duration of the pH-shock
- situations of pH out of the optimum range (especially lower at values)
- the VFA concentration
- the VFA composition

2.3.5 Assessment of the available buffering capacity

It is very important to maintain the pH of an anaerobic system between the limits, in which the methane-forming bacteria are active. Otherwise, the acid-forming bacteria may be more active than the methane forming are, which may result in problems: the reactor may become acidified (i.e. too much VFA accumulate).

The buffer systems introduced with the influent will influence the buffering capacity in the reactor. The buffer capacity is the ability of the reactor contents to resist a change in the pH if acids (VFA for example) are accumulating. The reactor contents must have enough buffer capacity to neutralise an eventual VFA accumulation, and of course, mixing of the reactor contents must be adequate, to prevent build-up of localised acid zones in the reactor.

Buffering can be performed by a strong base combined with weak acids. The maximum buffer capacity is obtained at the smallest change in pH upon addition of a certain amount of strong base or strong acid. This usually happens when half of the weak acid is ionised at a pH = pKa value) and consequently neutralised by a strong base. It will be obvious that the extent of buffering mainly depends on the absolute amount of buffer present in the system. Buffer components exert their maximum buffer capacity at the pKa values. The pKa values of volatile fatty acids are around 4.8. Although these acids are relatively weak, it will be clear that they don't represent the proper buffer for optimal methanogenesis. A weaker acid is required for that purpose, i.e. carbonic acid derived from the dissolution of CO_2 from the biogas, which produces the bicarbonate buffering system.

The strong bases can come from three sources: naturally present in the wastewater, formed in

the degradation process in the form of NH_4^+ or added if the buffer-capacity (alkalinity) of the reactor content is too low. The alkalinity is equal to the strong bases minus the strong acids i.e the net sum of protons;

Alkalinity may also be defined as the excess of positive charges over the anions of strong acids:

$$[\text{Alk}] = [\text{NH}_4^+] + [\text{Na}^+] + [\text{K}^+] + 2[\text{Ca}^{++}] + 2[\text{Mg}^{++}] - [\text{R-COO}^-] - [\text{Cl}^-] - 2[\text{SO}_4^{--}]$$

In this example $[\text{R-COO}^-]$ represents the VFAs. although VFA are a weak acids, they can be considered as strong acids relative to carbonic acid. In most cases in anaerobic digestion processes, the VFA concentration is the most important variable effecting the bicarbonate alkalinity. This is due to the fact that the other factors, strong bases, inorganic acids of the wastewater and CO_2 composition of the biogas are generally stable parameters. An increase in VFA concentration causes a decrease in bicarbonate alkalinity. It is proven that the bicarbonate can be estimated (knowing the pH and the CO_2 composition of the gas and assuming the pH is less than 8.3):

$$[\text{HCO}_3^-] = a_1 K_h P\text{CO}_2 / a_0$$

where:

$$[\text{HCO}_3^-] = \text{bicarbonate concentration mole/L;}$$

$$a_0 = 1 / \{ 10^{(\text{pH} - \text{pK}_1)} + 1 \} = [\text{H}_2\text{CO}_3] / \{ [\text{H}_2\text{CO}_3] + [\text{HCO}_3^-] \}$$

$$a_1 = 1 - a_0 = [\text{HCO}_3^-] / \{ [\text{H}_2\text{CO}_3] + [\text{HCO}_3^-] \}$$

pK_1 for carbonic acid is 6.3

After most of this bicarbonate concentration is removed/replaced by the VFA formed due to the acidification in the reactor, the pH of the reactor starts to drop. Other buffer systems which buffer in the neutral pH-range are:

$\text{H}_2\text{S}/\text{HS}^-$ with a $\text{pK}_1 = 6.5$

$\text{H}_2\text{PO}_4^-/\text{HPO}_4^{2-}$ with a $\text{pK}_1 = 7.2$

However, these buffers generally are only present at relatively low concentrations so that in practice they in fact are of minor importance with respect to the pH stability. However, if the influent concentration of sulphates and phosphorous are high their effect on pH needs to be considered.

2.3.6 Assessment of COD hydrolysis and degradability

The amount of biodegradable COD is the amount available to the fermentative bacteria. Mainly, it consists of the COD turned into cell growth, COD converted to VFA and the COD of the produced methane gas. The last two components are defined as the acidifiable COD. For a high rate reactor the amount turned to cell growth will be difficult to be estimated and it is comparably small when compared with the other two. Thus, the percentage of acidifiable COD (COD_{acid}) is a good indicator of the degradability of the wastewater.

COD_{acid} is generally equal to 80% of degradable COD

Hydrolyzable COD

The percentage hydrolysis (%COD_{hyd}) of the insoluble COD is expected to be at least 90% and it is calculated as follows:

$$(\text{COD}_{\text{sol}} + \text{COD}_{\text{cells}} + \text{COD}_{\text{CH}_4}) / \text{COD}_t * 100$$

where:

COD_{sol} = soluble COD (including VFA)

$\text{COD}_{\text{cells}}$ = cells produced from acidification

COD_{CH_4} = methane produced

$\text{COD}_{\text{total}}$ = total COD substrate

2.4 Telemac sensors installation

2.4.1 Installed Sensors

figure 3 shows the possible arrangement of sensors and expected monitoring system expansions that can practically be found/provided in anaerobic digestion plants. Usually, at an anaerobic treatment plant, the pH, liquid flow and temperature are installed. Also, a gas flow sensor is typically installed especially when the gas is processed down-stream. For the gas composition, a gas sensor is developed within TELEMAC that measures % methane and thus %CO₂ as other interfering gases will be filtered. Also, a titrimetric sensor is developed within the TELEMAC project that will measure VFA and bicarbonate (alkalinity). Software sensors will be developed to estimate COD/TOC from other measurements using models developed in TELEMAC. Frequently, manual analysis is expected as well to be available, especially for COD.

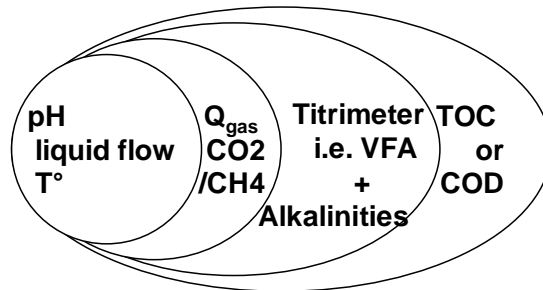


Figure 3: Sensor arrangements and expansion of monitoring

2.4.2 Interference and reference measurements

An advanced buffer capacity software sensor (BC sensor) will be implemented as a diagnosis tool. Frequently detailed and accurate titration curves will be produced by the titrimetric sensor and will be sent to the BC sensor. The BC sensor will detect and evaluate the interfering buffers such as phosphorous, hydrogen sulphide, ammonium, lactic acid, phenols...etc. An advantage of the diagnosis given by the BC sensor is the evaluation of the H₂S that interferes with the gas sensor. Despite the BC sensor evaluates the H₂S in the liquid phase, this indicates the possible presence of H₂S in the gas phase and thus alarms to filter H₂S before the gas sensor. In addition to the interference detection, the BC sensor will give accurate measurements of the VFA and bicarbonate alkalinity that can be used as a reference measurement to assess the quality of the titrimetric sensor results.

2.4.3 Installation of sensors (sample filtering, location)

Similar to the necessity to filter H₂S before the gas sensor, vapor and other interfering gases need to be filtered. If liquid samples are containing considerable amounts of solids, interference of suspended solids (SS) to the measurement device should be evaluated. In cases where the installed sensors or analysers are sensitive to SS an ultrafiltration loop should be added.

For most of the high rate reactors the distribution of the process variables is not uniform within the reactor, specially in the longitudinal/height direction. Thus, the location of sensors and sampling points within the reactor will be important. For that purpose models with extended dimensions (e.g. 2D) are developed within the TELEMAT context, task 2.5. the use of these models to locate sensors and sampling points will be described in detail in the deliverables D2.5 and D2.6.

2.5 Optimal experimental design

Sections 2.5.2 (Fisher Information Matrix) and 2.5.3 (local sensitivity analysis) have already been described and charged in Deliverable 3.2 (De Pauw et al., 2003). They have been repeated here for the sake of clarity and because the optimal experimental design procedure described in section 2.5.4 is based on these developments.

2.5.1 Introduction

In order to use models for diagnosis, prediction and control of an anaerobic digestion plant, model calibration is required. In short, model calibration is the exercise of minimising the weighted sum of squared errors between the model prediction and the available data by changing the model parameters (θ). The quality of the model calibration can be evaluated by analysing the parameter estimation covariance matrix $C(\theta)$ (Equation 1). This matrix is calculated by many minimisation programs and can also be estimated using different techniques (Dochain and Vanrolleghem, 2001). The diagonal elements are the variances of the parameter estimates and the off-diagonal elements are the covariances between the different parameters. The parameter estimation covariance matrix can then be used to calculate confidence intervals, confidence regions and parameter correlations. Small variances will result in small confidence regions and thus more accurate parameter estimations.

$$C = \begin{bmatrix} \sigma_{\theta_i}^2 & cov(\theta_i, \theta_j) & \cdots & cov(\theta_i, \theta_n) \\ cov(\theta_i, \theta_j) & \sigma_{\theta_j}^2 & & \\ \vdots & & \ddots & \\ cov(\theta_i, \theta_n) & & & \sigma_{\theta_n}^2 \end{bmatrix} \quad (1)$$

The quality of the model calibration greatly depends on the quality of the available data and

thus on the design of the experiment which was used to acquire the data. Poorly designed experiments result in poor data and will obviously result in poorly estimated parameters (large variances or strong correlations). Designing experiments is a difficult task and is often performed without a clear strategy. This is due to the large amount of choices that can be made and experimental constraints that have to be met when designing an experiment. These choices and constraints include:

- resources, including experimentation time and expenses, might be limited
- which measurements have to be performed
- which measuring frequencies can or have to be used
- where should the measurements be performed
- which experimental manipulations should be performed (e.g. reactor temperature, recycle flow rate) and what are the constraints for these manipulations
- ...

In order to design an experiment that will produce high quality data required for an accurate model calibration, optimal experimental design for parameter estimation (OED-PE) can be used. This is a mathematical technique which provides a solution to the complex problem of choices and constraints resulting in an optimal experiment. The basis of OED-PE is the Fisher Information Matrix (FIM) which under certain conditions (uncorrelated white measurement noise), gives the lower bound of the parameter estimation covariance matrix $C(\theta)$ according to the Cramer-Rao inequality (Ljung 1999; Walter and Pronzato, 1997):

$$C(\theta) \geq FIM(\theta)^{-1} \quad (2)$$

2.5.2 Fisher information matrix

The Fisher Information Matrix (Mehra, 1974) can be derived from the weighted least squares objective function $J(\theta)$ of the model calibration minimisation problem:

$$J(\theta) = \sum_{i=1}^N (y_i(\theta) - y_i)^T Q_i (y_i(\theta) - y_i) \quad (3)$$

in which y_i and $y_i(\theta)$ are vectors of N measured values and model predictions at times t_i ($i = 1$ to N) respectively. Q_i is a square matrix with user-supplied weighting coefficients, usually taken as the inverse measurement error covariance matrix.

The effect of a small parameter change $\partial\theta$ on the objective function (3) can be expressed by linearisation of the model along the trajectory:

$$E[J(\theta + \partial\theta)] = \sum_{i=1}^N (y_i(\theta + \partial\theta) - y_i)^T Q_i (y_i(\theta + \partial\theta) - y_i) \quad (4)$$

where:

$$y_i(\theta + \partial\theta) \approx y_i(\theta) + \left[\frac{\partial y(\theta)}{\partial \theta} \right]_{\theta} \partial\theta \approx y(\theta) + S_{\theta}(\theta) \partial\theta \quad (5)$$

In Equation 5, $S_{\theta}(\theta)$ is a vector of output sensitivity functions with respect to the parameters (for details, see Subsection 2.5.3). Using Equation 5, the expected value (4) can be rewritten as follows:

$$E[J(\theta + \partial\theta)] = J(\theta) + \partial\theta \left[\sum_{i=1}^N S_{\theta}(\theta)^T Q_i S_{\theta}(\theta) \right] \partial\theta \quad (6)$$

The term between brackets in Equation 6 is called the Fisher Information Matrix (FIM). The FIM expresses the information content of the experiment by combining the sensitivity functions $S_{\theta}(\theta)$ and the measurement error Q_i . As already discussed, the FIM is the inverse of the parameter estimation covariance matrix (Equation 2). This relationship is illustrated in figures 4 and 5 for a 2 parameter estimation problem. These figures represent the confidence regions of two parameters (θ_1 and θ_2). The size, shape and orientation of the confidence ellipse is determined by the eigenvalues and eigenvectors of the FIM. The largest axis is inverse proportional to the square root of the smallest eigenvalue ($\lambda_{min}(FIM)$), the smallest axis is inverse proportional to the square root of the largest eigenvalue ($\lambda_{max}(FIM)$). In this way properties of the FIM determine the properties of the confidence region and thus the accuracy of the parameter estimates. The information content of the experiment can be optimised by considering different measures of the FIM, table 3 lists these criteria. The D and A optimal design criteria aim at minimising the volume of the confidence ellipse. This is illustrated in figure 4. The modified E optimal design criterion on the other hand aims at reducing parameter correlations by getting the shape of the confidence region as close to a circle as possible which is illustrated in figure 5. The best value one can obtain for the modified E criterion is 1. However, one must be aware that such optimum only guarantees the confidence region is a circle, but it can be a very large circle (Dochain and Vanrolleghem, 2001).

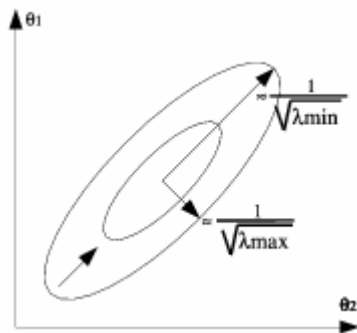


Figure 4. Effect of the D-criterion on the confidence ellipse of two parameters

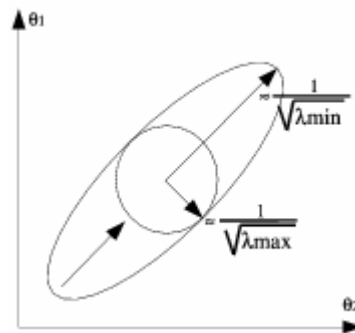


Figure 5. Effect of the Modified E-criterion on the confidence ellipse of two parameters

Table 3. Different optimal design criteria based on FIM properties

Name	Criterion
A-optimal design	$\min \left[\text{tr} \left(FIM^{-1} \right) \right]$
Modified A-optimal design	$\max \left[\text{tr} \left(FIM \right) \right]$
D-optimal design	$\max \left[\det \left(FIM \right) \right]$
E-optimal design	$\max \left[\lambda_{\min} \left(FIM \right) \right]$
Modified E-optimal design	$\min \left[\frac{\lambda_{\max} \left(FIM \right)}{\lambda_{\min} \left(FIM \right)} \right]$

tr(): sum of eigenvalues, det(): product of eigenvalues

2.5.3 Local sensitivity analysis using finite differences

In order to calculate the Fisher Information Matrix a sensitivity analysis has to be performed first. Sensitivity analysis studies the sensitivity of the outputs of a system to changes in the parameters, inputs or initial conditions which are often poorly known. Sensitivity analysis can be divided into two large categories: local and global sensitivity analysis. Local sensitivity analysis methods refer to small changes of parameters, while global methods refer to the effect of simultaneous, possibly orders-of-magnitude parameter changes. For a detailed review of existing sensitivity techniques reference is made to the reviews of Turanyi (1990) and Rabitz et al. (1983). The *FIM* is based on local sensitivity functions which will be discussed in detail in the next paragraphs.

The sensitivity of a state variable y to a parameter θ can be expressed as a sensitivity function (Equation 7). A state variable y is called sensitive to θ if small changes in θ produce significant changes in y .

$$S(t) = \frac{\partial y(t)}{\partial \theta} \tag{7}$$

This partial derivative can be analytically solved if the analytical solution of the model equations is known. Unfortunately, this is rarely the case and numerical methods have to be used in order to approximate the sensitivity function (Equation 7). Local sensitivity analysis techniques evaluate this partial derivative at one specific parameter value, also called the nominal parameter value.

Many techniques exist to calculate local sensitivity functions. Most of these techniques require complex manipulations of the model equations, often not practically feasible. The simplest way of calculating local sensitivities is to use the finite difference approximation. This technique is also called the brute force method or indirect method. It is very easy to implement because it requires no extra code beyond the original model solver. The partial derivative defined in Equation 7 can be mathematically formulated by the equation given below (forward difference).

$$\frac{\partial y_i}{\partial \theta_j} = \lim_{\Delta \theta_j \rightarrow 0} \frac{y_i(t, \theta_j + \Delta \theta_j) - y_i(t, \theta_j)}{\Delta \theta_j} \quad (8)$$

Quality control of sensitivity calculations

The practical implementation of Equation 8 has two major drawbacks. The resulting sensitivity function relates to the $(\theta_j + \Delta \theta_j / 2)$ parameter set and it does not provide any information on the quality of the sensitivity function. If sensitivities are required around the nominal values of the parameters then the central difference formula should be used (Equation 9).

$$\frac{\partial y_i}{\partial \theta_j} \approx \frac{y_i(t, \theta_j + \Delta \theta_j) - y_i(t, \theta_j - \Delta \theta_j)}{2\Delta \theta_j} \quad (9)$$

Although this method requires $2p$ model evaluations, it is innovative in the sense that it can be used to provide additional information concerning the quality of the sensitivity function. To implement the central difference method two sensitivity functions are calculated (equations 10 and 11). The first sensitivity function is calculated by increasing the nominal parameter value by $\xi \theta_j$, the second sensitivity function is calculated by decreasing the nominal parameter value by $\xi \theta_j$.

$$\frac{\partial y}{\partial \theta_{j+}} = \frac{y(t, \theta_j + \xi \theta_j) - y(t, \theta_j)}{\xi \theta_j} \quad (10)$$

$$\frac{\partial y}{\partial \theta_{j-}} = \frac{y(t, \theta_j - \xi \theta_j) - y(t, \theta_j)}{\xi \theta_j} \quad (11)$$

To calculate the centralised sensitivity function the average of both sensitivity functions is taken. To make the numerical error and the error introduced by the nonlinearity of the model as small as possible the difference between these two sensitivity functions should be minimal. Several criteria can be used to quantify this difference.

1. Sum of squared errors (SSE).

$$\frac{\sum \left(\frac{\partial y}{\partial \theta_{j+}} - \frac{\partial y}{\partial \theta_{j-}} \right)^2}{N} \quad (12)$$

For this criterion, the squared error between both sensitivity functions is calculated and summed over all times where the sensitivity is required (N).

2. Sum of absolute errors (SAE).

$$\frac{\sum \left| \frac{\partial y}{\partial \theta_{j+}} - \frac{\partial y}{\partial \theta_{j-}} \right|}{N} \quad (13)$$

For this criterion, the absolute error between both sensitivity functions is calculated and summed over all times where the sensitivity is required (N).

3. Maximum relative error (MRE).

$$\left| \frac{\frac{\partial y}{\partial \theta_+} - \frac{\partial y}{\partial \theta_-}}{\frac{\partial y}{\partial \theta_+}} \right|_{MAX} \quad (14)$$

This criterion returns the maximum value of the relative difference between both sensitivity functions. One should be careful with this criterion because $\partial y / \partial \theta_+$ or $\partial y / \partial \theta_-$ may become 0. In this special case the criterion returns 0.

4. Ratio between sensitivities (RATIO).

$$\left| 1 - \frac{\frac{\partial y}{\partial \theta_-}}{\frac{\partial y}{\partial \theta_+}} \right|_{MAX} \quad (15)$$

This criterion is based on the ratio of the sensitivity functions. The ideal case is when this ratio equals 1, because then both sensitivity functions are equal. The criterion returns the maximum deviation from this ideal situation. Like the MRE criterion one should be careful if $\partial y / \partial \theta_+$ becomes 0. In this special case the criterion returns 0.

figure 6 illustrates the use of these criteria for the sensitivity of organic substrate (S_1) to the maximum specific growth rate for acidogenic bacteria (μ_{m1}). This figure clearly illustrates the nonlinearity effects of the model for perturbation factors higher than 10^{-3} . An example is given in figure 7 for a perturbation factor 10^{-1} . In this figure the sensitivity function for the optimal perturbation factor 10^{-3} is also shown. It is clear that the $\partial y / \partial \theta_-$ differs significantly from $\partial y / \partial \theta_+$. figure 6 also illustrates the effect of numerical errors for perturbation factors lower than 10^{-3} . An example of this is given in figure 8 for a perturbation factor of 10^{-5} .

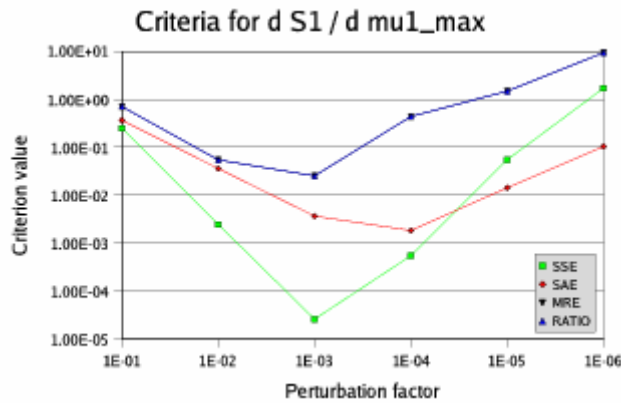


Figure 6. Different criteria calculated for the sensitivity of organic substrate (S_1) to the maximum specific growth rate (μ_{m1}) for acidogenic biomass

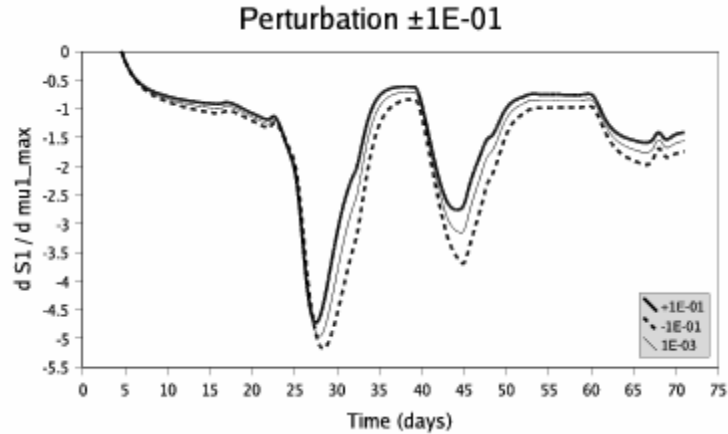


Figure 7: Effect of the nonlinearity of the model on the sensitivity calculations. Perturbation factor $1E^{\pm 1}$.

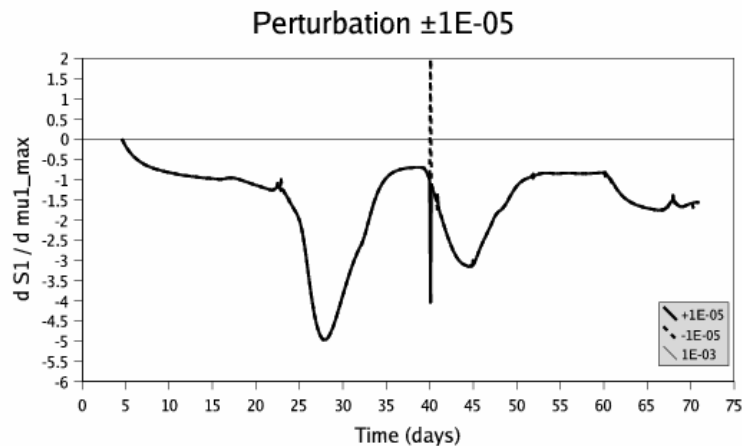


Figure 8: Effect of numerical error on the sensitivity calculations. Perturbation factor $1E^{\pm 5}$.

It can be concluded that perturbation factors are parameter dependent and to a lesser extent variable dependent. Using a fixed perturbation factor for all parameters is thus not advisable. In literature, however, this is frequently done Hoffmann et al. (2001); Marsili-Libelli et al. (2001); Reichert (1994); Saltelli et al. (2000).

Although the finite difference technique for sensitivity analysis is not new, the procedure to quantify the quality of the calculated sensitivity functions is innovative. In this way the finite difference technique can be used without numerical errors and errors introduced by nonlinearities of the model.

2.5.4 Iterative optimal experimental design procedure

Figure 9 shows the general procedure for optimal experimental design. Before starting experimental design a preliminary model should be available. In most cases the model is fitted to data acquired from initial experiments. For anaerobic digestion plants this is usually the long term data that is already available on the plant. If no data is available and initial

experiments can not be performed, default parameter values can be used. Once the model is identified, the iterative procedure is started. Based on the current model an experiment is proposed by choosing certain experimental degrees of freedom. This experiment is then simulated on the computer and an objective function is evaluated. Typically this objective function is a design criterion from the Fisher Information Matrix (table 3) which summarises the information content of the proposed experiment and is also a measure for the accuracy of the parameter estimates (in case this experiment would be performed in reality and the model would be fitted to the acquired data). A (non linear) optimisation algorithm can be used to propose different experiments and find an “optimal” experiment in the sense that it maximizes the parameter estimation accuracy and/or minimizes parameter correlations. Once the “optimal” experiment is found it can be performed in reality. Based on the data of this experiment the model can be refitted and the accuracy of the parameter estimates evaluated. If the required accuracy is not yet reached another iteration can be performed, leading to an even better “optimal” experiment. Finally a calibrated model is available for its intended use.

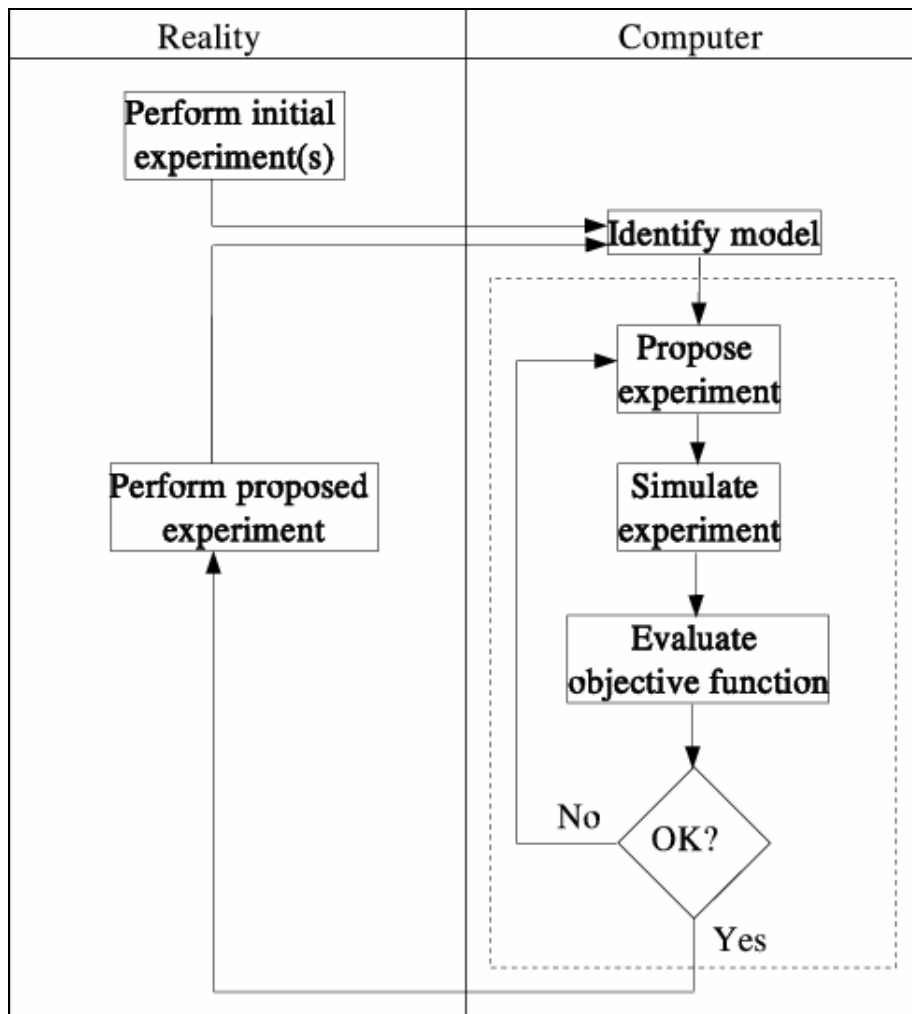


Figure 9: General procedure for optimal experimental design (from Dochain and Vanrolleghem, 2001)

3 Virtual case study

The idea of the virtual case study is to test the proposed protocol by building a virtual anaerobic digester and run the optimal experimental design tools to reliably set up the TELEMAC system on that plant. The virtual plant is created using a reference model (IWA ADM1 model) to simulate a real digester treating vinasses. Introduction of the ADM1 model and description of the plant are illustrated in section 3.1. The type of vinasses and the influent characterisation are described in section 3.2. The ADM1 implementation on the WEST[®] simulation platform is used to build this experiment and an extension is added to generate all possible practical measurements in the agreed TELEMAC standard units. Table (4) shows the generated list of the measurements

and the units as defined in deliverable 1.1. With this two data sets are generated. The first is presenting a long term data set that is expected as a record of data available at a digestion treatment plant (the virtual plant in this case study) as described in section 3.3. This first data set is then used to fit one of the main TELEMAC models (AM2) and get an initial estimate of its parameters. Then the optimal experiment design (OED) tool described in section (2.5 and 4) is used to design an optimum experiment based on available TELEMAC sensors. The proposed influent for the optimum experiment is then used to generate a second data set based on the TELEMAC sensors as illustrated in section 3.4. This second data set is then used to fit the AM2 model. The simulation results of AM2 for both data sets together with the OED analysis and prediction are illustrated in section 3.5.

To our knowledge this has been the first application that uses a reference model for a anaerobic digestion process as a data generator to evaluate methodologies and validate other simple models. Practical measurements are calculated in the most common units that are used at industrial applications and evaluated as functions of ADM1 state variables and parameters.

Table (4): Measured variables

Name	Unit	Meaning
Temp_dig	°C	Temperature inside the digester
GasFR	l/h	Gas flow rate
InfluentFR	l/h	Influent flow rate
pCO2		Percent CO2
pCH4		Percent CH4
H2	ppm	H2 ppm
PT	mbar	Total pressure
pH_dig		pH in the digester
Palk_dig	meq/l	Partial alkalinity in the digester
Talk_dig	meq/l	Total alkalinity in the digester
Ialk_dig	meq/l	Intermediate alkalinity in the digester
TOCt_dig	mg/l	Total organic carbon
TOCs_dig	mg/l	Total soluble organic carbon
CODt_dig	mg/l	Total chemical oxygen demand
CODs_dig	mg/l	Soluble chemical oxygen demand
VFA_dig	mg/l	Total volatile fatty acids
Bic_dig	meq/l	Bicarbonate in the digester
TSS_dig	mg/l	Total suspended solids
VSS_dig	mg/l	Volatile suspended solids
pH_inf		pH in the influent
Palk_inf	meq/l	Partial alkalinity in the influent
Talk_inf	meq/l	Total alkalinity in the influent
Ialk_inf	meq/l	Intermediate alkalinity in the influent
TOCt_inf	mg/l	Total organic carbon
TOCs_inf	mg/l	Total soluble organic carbon
CODt_inf	mg/l	Total chemical oxygen demand
CODs_inf	mg/l	Soluble chemical oxygen demand
VFA_inf	mg/l	Total volatile fatty acids
Bic_inf	meq/l	Bicarbonate in the influent
Acetate_dig	mg/l	Acetate
Propionate_dig	mg/l	Propionate
Butyrate_dig	mg/l	Butyrate

3.1 Description and modelling of the virtual plant

3.1.1 Introduction of ADM1 model

The International Water Association (IWA) Anaerobic Digestion Model no.1 (ADM1) is implemented as a reference model within the TELEMAT project. Implementation has been reported in deliverable 3.1a (Bernard, 2002). As shown in figure 4 the IWA ADM1 model (Batstone et. al, 2002) considers the conversion processes in the digester from the most complex (substrate) to the most simple. In other words, it starts from the disintegration of composite particulate till the gas.

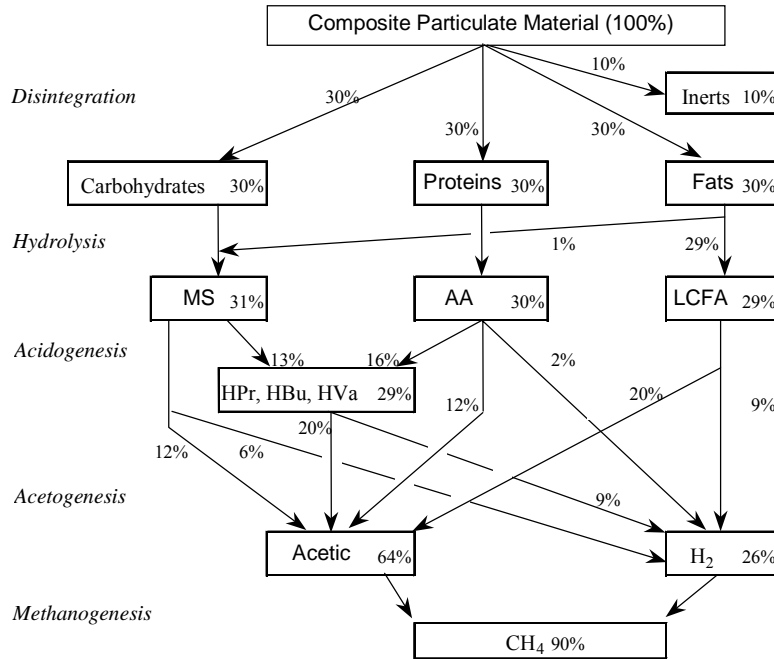


Figure 10: COD flux for a particulate composite comprised of 10% inerts, and 30% each proteins, carbohydrates and fats (in terms of COD). Propionic acid (10%), butyric acid (12%) and valeric acid (7%) are grouped in the diagram for simplicity (Batstone et. al, 2002).

In TELEMAC applications, however, the supply could be simpler. The process for supplied wastewater mainly takes the path of carbohydrates as shown in figure 10. Moreover, if the waste is already acidified, the process starts with the acetogenesis mainly degrading VFAs. However, implementation of the ADM1 in its extended version is still important for two reasons. First, it is possible to find considerable proportions of composite particulate originating from plant sources (e.g. grapes used in wineries, barley used in breweries and distilleries) in these wastewaters despite of the value of their separation and use as animal feed. Second, when biomass decay is modelled, the decaying species are lumped and added to the composite particulates, thus, paths described in figure 4 should be followed starting from the disintegration step.

Whereas, the biogas is a main product of the anaerobic process and the anaerobic reactor is a closed system, it is essential to also consider the gas - liquid transfer. Thus in the IWA ADM1, the gas transfer between two compartments (i.e gas compartment and liquid compartment), shown in figure 11, is modelled by considering three gas components and evaluating their concentrations in both phases using Henry's law. Also, the anaerobic process is very sensitive to pH changes and there are many buffering systems such as VFAs (acetate, propionate, butyrate and valerate), bicarbonate, ammonia affecting pH. Those buffers were considered in the ADM1 by the corresponding dissociation reactions that can be presented by either differential equations (DE implementation) or by algebraic equations of equilibrium (DEA implementation). From either implementations, the hydrogen ion concentration will be calculated from the equilibrium of the other ions and consequently other ion concentrations

are calculated. The results of these ion calculations are used to estimate inhibition factors that are also considered in the model. All these reactions/rates add more state variables that are in most cases not reasonable online and require a lot of effort for off-line analysis. It also implies a burden on the solver used for simulation. It should be able to handle this large number of equations with different reaction rates: slow (biological), intermediate (physical) and fast (chemical).

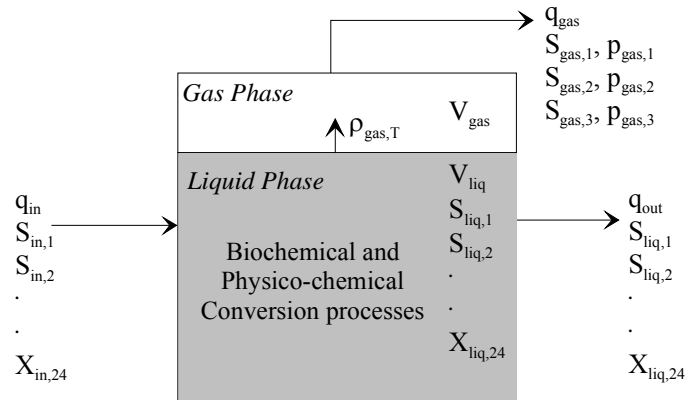


Figure 11: Schematic of a typical single-tank digester (q = flow, $m^3.d^{-1}$ V = volume, m^3 ; $S_{stream,i}$ = concentration of liquid components; $X_{stream,i}$ = concentration of particulate components; all in $kg\ COD.m^{-3}$; i is the component index), (Batstone et al, 2001)

Concluding, the ADM1, as considered in TELEMAC, is a detailed reference model that helps in understanding the process/process changes, validating simple models, generating balanced sets of data and experiment optimization. Of course, ADM1 is not to be used as a model for online control, but could be used to evaluate the proposed control systems.

3.1.2 Virtual plant description

The implemented Virtual Digester is a scaled up version of the USC lab CSTR system which is reported in deliverable 1.1 (Roca et al., 2002). A 180 days experiment was performed on the lab CSTR and results are reported in deliverable 1.3a (Roca and Lema, 2002). Validation of the experiment is reported in deliverable 2.2 part 2, (Zaher and Vanrolleghem, 2003). The ADM1 model that was applied to simulate this experiment and understand its dynamics (Zaher et. al., 2003). The simulation of the lab experiments was performed by the model implementations on three simulation platforms (WEST[®] SIMULINK[®] and AQUASIM[®]) where the results of the three implementation were in full agreement validating that the correct implementation has been achieved. Also in this work, the ADM1 model with a default parameter set suggested in (Rosen and Jeppsson, 2002) was used to determine an extended characterisation of the influent.

In the 180 days lab experiment, the reactor was a CSTR with a 2 litres liquid volume fed by an alcoholic distillery wastewater with characteristics mentioned in table 5. Due to its high concentration, the wastewater was diluted before it was introduced to the reactor to achieve the planned experimental protocol. The protocol Also included additional cation dosing to control the reactor pH and produce some shock loads. Further details about the experiment and the protocol are described in deliverable 1.1. A scale factor of 10^6 was used to size the virtual plant and approximate a real digester. Thus, the volume of the virtual plant is 2000 m^3 and the liquid flow is multiplied by the same scale factor.

Table 5: Average wastewater concentrations produced at the industrial plant

Characteristic	Unit	Average value
Total COD	(g/l)	75
Soluble COD	(g/l)	71
TOC	(g/l)	29.2
COD/TOC	(ratio)	2.6
Acetate	(g/l)	6.39
Propionate	(g/l)	0.07
NBuH	(g/l)	0.47
TSS	(g/l)	1.7
VSS	(g/l)	1.63
Protein	(g/l)	4.0
NH ₄ ⁺ -N	(g/l)	0.126
NO _x -	(g/l)	0
SO ₄ ²⁻	(g/l)	0.709
PO ₄ ³⁻	(g/l)	0.1142
Sugars	(g/l)	6.4

During the lab experiment the influent records were: flow and pH (on-line); TOC, soluble COD, total volatile fatty acids (VFA), total alkalinity (TA) and partial alkalinity (PA) were analysed off-line and recorded on a daily basis. The Digester/effluent records were: pH, gas

flow and gas composition (%CO₂ and %CH₄) (on-line); Specific VFA gas chromatography analysis, TA and PA, total organic carbon and COD were measured in the digester on daily basis. As already mentioned more detailed influent characteristics were estimated using ADM1 and the default parameter set. For the virtual treatment plant concentrations in the influent were maintained the same and simulation was performed for the 180 days of the experiment. The simulation results of the concentrations in the virtual plant were exactly the same as those simulated for the lab-scale reactor which validates the scaling assumption.

Since the lab experiment was designed to achieve an extreme overload towards the end of the experiment with some intermediate shock loads, it is not practical to consider the whole experiment period for the normal digester of the virtual plant. Such abnormal conditions wouldn't be allowed on a real plant. Thus, only the first 50 days are considered to avoid extreme loading and shocks. Also, the influent flow is filtered to avoid the rapid dynamics of the lab scale influent since these are not feasible for a normal digester. The result is a dynamic flow with a gradual increase in the influent through the designated 50 days. Accordingly, the hydraulic retention in the virtual plant is varying from 20 to 8 days which can be accepted for a real digester.

To better imitate the reality of seasonal changes in a real digester influent, the influent is extended steady for 10 more days and then the first 50 days are mirrored. This forms a 110 days period that can then be repeated to study the long term dynamics, as shown in the flow profile of figure (12).

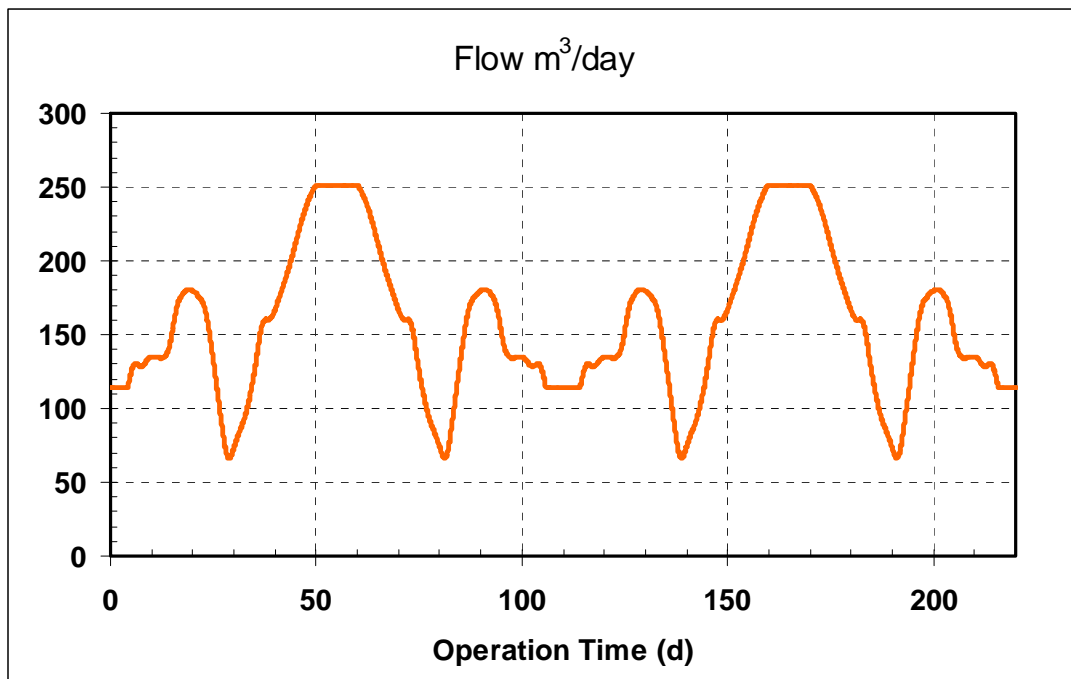


Figure (12): Designed long term flow profile to the virtual plant

3.2 Influent characteristics

The detailed influent characteristics previously estimated with the ADM1 were used to generate the influent characteristics for AM2 (detailed description of AM2 model is reported in deliverable 3.1a (Bernard, 2002)). Two influent characteristics were generated for AM2 simulation of two scenarios. The first scenario is for the long term data as would be typically collected from a digestion plant records. The second scenario is the one determined by the Optimal Experiment Design OED. To briefly illustrate how the AM2 influent is calculated from an ADM1 specified influent, let us recall the input variables for both ADM1 and AM2 (table (6) and table(7)).

Table (6): AM2 state variables

State Variable Description	Name	Units
Organic Substrate concentration	S_1	g/l
VFA concentration	S_2	mmol/l
Concentration of acidogenic bacteria	X_1	g/l
Concentration of methanogenic bacteria	X_2	g/l
Total alkalinity	Z	mmol/l
Inorganic carbon concentration	C	mmol/l

Table (7): State variables in ADM1.

State Variable Description	State Symbol	Units
Soluble inerts	S_i	kg COD m ⁻³
Monosaccharides	S_{su}	kg COD m ⁻³
Amino acids	S_{aa}	kg COD m ⁻³
Long chain fatty acids (LCFA)	S_{fa}	kg COD m ⁻³
Total valerate	S_{va}	kg COD m ⁻³
Total butyrate	S_{bu}	kg COD m ⁻³
Total propionate	S_{pro}	kg COD m ⁻³
Total acetate	S_{ac}	kg COD m ⁻³
Hydrogen gas	S_{h2}	kg COD m ⁻³
Methane gas	S_{ch4}	kg COD m ⁻³
Particulate inerts	X_i	kg COD m ⁻³
Composites	X_c	kg COD m ⁻³
Carbohydrates	X_{ch}	kg COD m ⁻³
Proteins	X_{pr}	kg COD m ⁻³
Lipids	X_{li}	kg COD m ⁻³
Sugar degraders	X_{su}	kg COD m ⁻³
Amino acid degraders	X_{aa}	kg COD m ⁻³
LCFA degraders	X_{fa}	kg COD m ⁻³
Valerate and butyrate degraders	X_{c4}	kg COD m ⁻³
Propionate degraders	X_{pro}	kg COD m ⁻³
Acetate degraders	X_{ac}	kg COD m ⁻³
Hydrogen degraders	X_{h2}	kg COD m ⁻³
Inorganic nitrogen	S_{in}	kmole N m ⁻³
Inorganic carbon	S_{ic}	kmole C m ⁻³
Anions	S_{an}	kmole m ⁻³
Cations	S_{cat}	kmole m ⁻³

Also, we need to recall two important assumptions of AM2. The first one is that total

alkalinity is equivalent to the net cation concentration. The second one is an assumption made when estimating the influent S_I from the practically measured total CODs and total VFA. The assumption is mainly that the VFA concentration is equivalent to a mixture of 70% of acetate and 30% of propionate. As a result 1g VFA (i.e S_2)=1.2 g COD and 1mol $S_2 = 64.2$ g. Despite the fact that it is straightforward to sum all ADM1 soluble COD components less VFA to calculate S_I , first CODs will be calculated in mg/l and VFA in mg/l as specified in the TELEMAT standard units (table 3). This is mainly done to take into account the effect of the second assumption as will happen in reality. Thus the AM2 influent characteristics are calculated as follows:

$$\begin{aligned} \text{CODs (mg/l)} &= (S_i + S_{su} + S_{aa} + S_{fa} + S_{va} + S_{bu} + S_{pro} + S_{ac}) \cdot 1000 \\ \text{VFA (mg/l)} &= (S_{va} \cdot 102 / 208 + S_{bu} \cdot 88 / 160 + S_{pro} \cdot 74 / 112 + S_{ac} \cdot 60 / 64) \cdot 1000 \\ S_I &= (\text{CODs} - \text{VFA} \cdot 1.2) / 1000 \\ S_2 &= \text{VFA} / 64.2 \\ X_I &= X_2 = 0 \\ Z &= S_{cat} - S_{an} \\ C &= S_{ic} \cdot 1000 \end{aligned}$$

Note that S_{ic} will be mainly equal to the bicarbonate concentration since it is expected to be measured after the addition of the cations and rising the pH.

The above estimated concentrations are projected to one-day intervals. In practice and at a digestion plant without on-line monitoring of the influent, these measurements are expected to be collected off-line on a daily basis. However, the flow was assumed to be recorded on-line and was thus collected every 30 min.

3.3 Long term data

The long term data was generated using the virtual plant. Since the long-term flow profile is assumed to be a repeated cycle of 110 days, the data was only generated for one period of 110 days. The data expected to be recorded at any digestion plant is assumed to be CODs, pH and total gas flow. It is expected that at a normal plant, the pH and the gas flow are collected on-line. Feasible measurement intervals are assumed to be 5min and 30min respectively. For CODs the measurement is assumed to be performed off-line on a daily basis. Thus, the long term measurements are interpolated from the virtual plant results at those intervals. Nevertheless, noise is expected in the measurements and thus signal is not expected to be smooth as simulated. Thus, a Matlab function was programmed to automatically interpolate individual records from the virtual plant and generate the noised signal. A white noise $N(0, E)$ is generated and introduced to the simulation signal. E is the corresponding error level listed in table 10. Figure 13 shows examples of the CODs and gas flow measured signal.

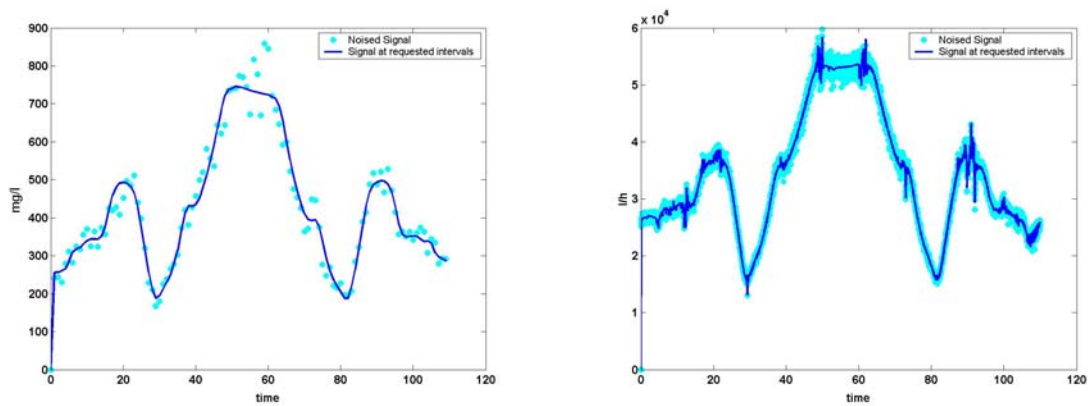


Figure 13: Example virtual plant simulated and noised signals (CODs to the left, Gas flow to the right)

Although the concept of white noise in the measurement is not new, the way it is used to investigate the effect of different noise levels (error in the measurement) on practical identifiability is innovative. Traditionally, the intention is to filter the measured signal from noise to clarify the signal trend. In the present application, noise is automatically generated on predefined error level and added to the signal that is generated by the virtual plant (simulated by the reference model).

3.4 Telemac sensors

Here we refer to figure 3 where the third circle presents the extension of the monitoring system with on-line measurements for VFA, %CO₂ and % CH₄. These additional on-line measurements will be provided by titrimetric and gas sensors developed within TELEMAC. Thus, to test the added value of these sensors using the Optimal Experimental Design (OED) procedure, the corresponding simulation signals were generated from the virtual plant. Similar to section 3.3, the simulated signals are interpolated and noised at the intervals and error levels listed in table 12.

After the check of the added value of the TELEMAC sensors, the OED tools are used to find an optimum experiment to set up the system. The proposed flow is then imposed on the virtual plant for a period of 28 days. The associated influent concentrations with the proposed flow are assumed the same as those of long term data in the period from 41 to 69. This mainly presents the middle period of the assumed seasonal.

Consequently, the measurements assumed available at the plant and the added measurement by the TELEMAC sensors are simulated using the intervals and errors listed in table 14, using the same procedure illustrated before. Section 3.5 describes the whole procedure of the experimental design and illustrates the approach indicating it is adequate for setting up the TELEMAC system on a plant.

3.5 Optimal experimental design

3.5.1 Model calibration for long term data

Using the available long term data, described in Section 3.3, model AM2 was calibrated. Initial conditions for state variables S_1 , S_2 , X_1 , X_2 , C and Z were obtained after running a long term simulation with a steady-state influent file. This influent file was constructed by averaging all influent components of the long term data over the period of 110 days. The initial conditions are listed in Table 8 and will be assumed fixed during the model calibration.

Table (8) Initial condition for AM2 state variables based on a steady state influent file

Name	Value
S_1 (g/l)	0.241
S_2 (mmol/l)	0.521
X_1 (g/l)	0.455
X_2 (g/l)	0.296
Z (mmol/l)	69.72
C (mmol/l)	71.89

Default AM2 parameters were assumed as initial parameter estimates. They are listed in Table 9.

Table (9) Default AM2 parameters used as initial parameter values

Name	Value
k_1 (-)	42.14
k_2 (mmol/g)	116.5
k_3 (mmol/g)	268
k_4 (mmol/g)	50.6
k_5 (mmol/g)	343.6
k_6 (mmol/g)	453
μ_{m1} (1/d)	1.2
μ_{m2} (1/d)	0.74
K_{S1} (g/l)	7.1
K_{S2} (mmol/l)	9.28
K_{I2} (mmol/l)	256
α (-)	0.5
$k_1 a$ (1/d)	19.8

Model AM2 was fitted to the three available measurement variables: soluble COD (COD_s),

pH and the gas flow rate (Q_{TOT}). Measurement errors displayed in Table 10 were used as weights in the parameter estimation objective function (weighted least squares) and the optimal experimental design. These measurement errors were assumed to be representative for a real case study.

Table 10: Measurement errors for the available long term data

Name	Error(%)
COD_s (g/l)	7
pH (-)	1
Q_{TOT} (l/d)	2

The model was calibrated using a constrained, derivative free optimisation algorithm, Simplex (Nelder and Mead, 1965). More information on the calibration of the TELEMAC models can be found in Deliverable 3.2 (De Pauw et al., 2003). The model calibration resulted in parameter values and estimation standard deviations listed in Table 11. Figures. 14, 15 and 16 show the fit of the model to the long term data. The model was able to describe the data adequately enough to be used as the starting point for the experimental design. The fit on the soluble COD measurements (COD_s) is less accurate than the fit on the other measurements because of the low quantity of data and the high measurement error. This causes the contribution of the COD_s to the optimisation objective function to be relatively small compared to the contributions of pH and total gas flow (Q_{TOT}) leading to a less accurate fit.

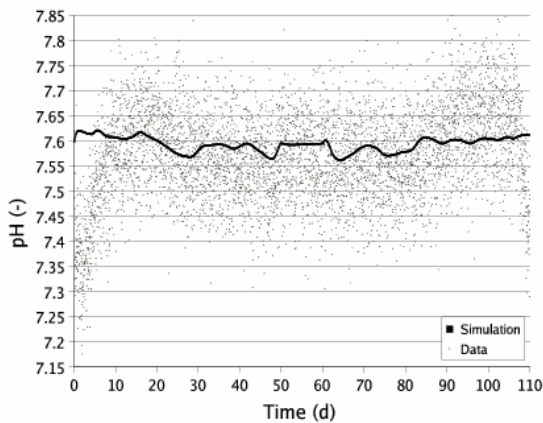


Figure 14: Optimal AM2 model fit to the long term pH measurements

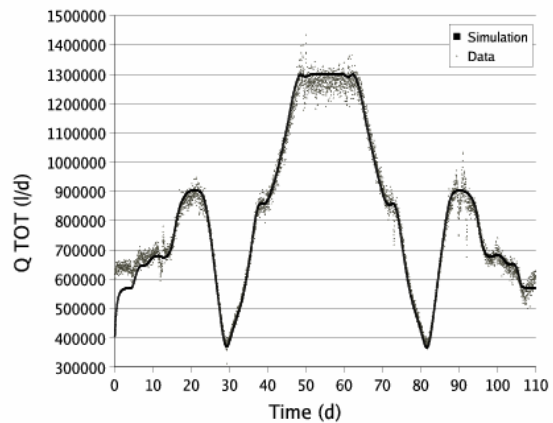


Figure 15: Optimal AM2 model fit to the long term gas flow rate measurements

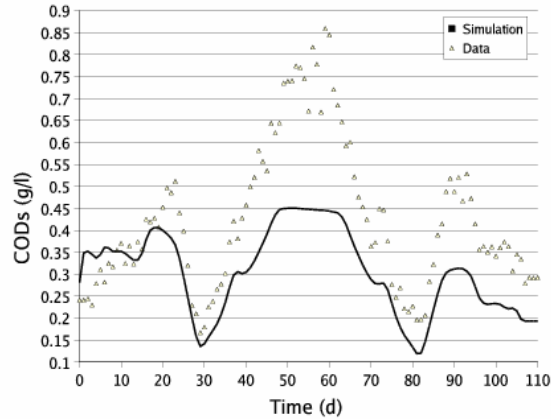


Figure 16: Optimal AM2 model fit to the long term COD_5 measurements

With the estimated parameters the Fisher Information Matrix (FIM) is calculated for the long term simulation and the existing data collection plan. The inverse of this FIM is used as an approximation of the parameter estimation covariance matrix (Equation 2). Based on this covariance matrix the estimation standard deviations are calculated as percentages of the parameter values. From Table 11 we can conclude that the confidence on the parameter estimates is very low (high estimation standard deviations). The information contained in the three available measurements for these parameters is very low. In fact, according to the parameter selection procedure described in Deliverable 3.2 (De Pauw et al., 2003) only parameters k_1 , k_2 and α could be estimated with adequate accuracy. To increase the confidence on the parameter estimates, optimal experimental design should be used to come up with experiments which yield data with a higher information content. This will be discussed in the next sections.

Table 11 Estimates of AM2 parameters and their standard deviations found by fitting the model to the long-term data

Name	Value	Standard deviation	Error(%)
k_1	20.63	1.53	7.43
k_2	243.17	19.03	7.83
k_3	344.55	56.18	16.31
k_4	30.42	7.83	25.74
k_5	184.44	31.01	16.81
k_6	466.25	75.86	16.27
μm_1	1.65	1.24	75.13
μm_2	0.70	1.31	186.77
K_{S1}	8.37	6.79	81.12
K_{S2}	7.17	15.02	209.61
K_{I2}	349.42	19.94E+03	57.07E+02
α	0.56	0.04	6.51
k_{la}	16.63	3.36	20.26

3.5.2 Optimal experimental design: adding TELEMAC sensors

To improve the information content of the experiment, the added value provided by

TELEMAC sensors will be investigated. *VFA* measurements (state variable S_2 in the AM2 model), CH_4 and CO_2 flow measurements will be considered. CH_4 and CO_2 flow measurements can be calculated from the total gas flow rate and the percentages of CH_4 and CO_2 in the gas provided by the sensor. With the availability *VFA* measurements, AM2 state variable S_1 (organic substrate) can also be calculated by subtracting the *CODs* measurement with the amount of *VFA* expressed as *COD*. In total six measurements will be considered in this proposed experiment: S_1 , S_2 , pH , Q_{TOT} , Q_{CH_4} and Q_{CO_2} . The total experiment time is again assumed to be 110 days. Table 12 lists the additional measurements together with the proposed measurement interval and measurement error.

Table 12: Measurement interval and error for the additional TELEMAC sensor measurements

Name	Measurement interval	Measurement error (%)
<i>VFA</i> (mmol/l)	2 hours	2
CH_4 flow rate (l/d)	30 minutes	2
CO_2 flow rate (l/d)	30 minutes	2

To illustrate the increase in information content that an experiment with these measurements would have, the FIM was calculated for the experimental conditions and parameter values used in the previous section with and without considering the additional measurements. Again, the information content of the newly proposed experiment is expressed in terms of the estimation standard deviations of the parameters. In Table 13 a comparison is made between the experiment with and without the TELEMAC sensor measurements. It can clearly be seen that the confidence on almost all parameters increases indicating that with the additional measurements the model parameters could be estimated with higher accuracy. Inhibition parameter K_{I2} however still shows very low confidence, probably because the reactor is not operated at inhibiting conditions. This parameter will not be considered further. However it is feasible to design an experiment for which this parameter does become identifiable, for example running the reactor at high S_2 concentrations.

Table 13: Added value of the TELEMAC sensors (*VFA*, CH_4 and CO_2) on the parameter estimation standard deviations

Name	Value	Without TELEMAC sensors		With TELEMAC sensors	
		Standard deviation	Error(%)	Standard deviation	Error(%)
k_1	20.63	1.53	7.43	0.23	1.14
k_2	243.17	19.03	7.83	2.91	1.20
k_3	344.55	56.18	16.31	2.73	0.79
k_4	30.42	7.83	25.74	1.44	4.73
k_5	184.44	31.01	16.81	2.21	1.2
k_6	466.25	75.86	16.27	3.08	0.66
μ_{m1}	1.65	1.24	75.13	0.53	32.21
μ_{m2}	0.70	1.31	186.77	0.06	8.42
K_{S1}	8.37	6.79	81.10	2.84	33.92
K_{S2}	7.17	15.02	209.60	0.63	8.83
K_{I2}	349.42	19.94E+03	57.07E+02	81.42E+02	23.30E+02
α	0.56	0.04	6.51	3.40E-03	0.61
k_{la}	16.63	3.36	20.26	0.72	4.36

By considering the TELEMAC sensors, errors on the parameter estimates could be reduced by a factor 5-20. This shows the importance of these sensors and the need for their development.

3.5.3 Optimal experimental design: designing an optimal influent flow profile

In order to further improve the accuracy of the AM2 model parameters, an experiment was designed using the optimal experimental design procedure described in Subsection 5.4. The measurements considered for this experiment included the already available measurements on the plant and the TELEMAC sensors. They are listed in Table 8 together with the measurement interval and error. The total experiment time is 1 month (28 days), rather than 110 days as before.

Table 14: Measurements considered for the design of the optimal experiment with their measurement interval and error.

Name	Model name	Interval	Error (%)
Organic substrate (g/l)	S_1	0.5 days	7
VFA (mmol/l)	S_2	2 hours	2
pH (-)	pH	30 minutes	1
Total gas flow rate (l/d)	Q_{TOT}	30 minutes	2
CH_4 gas flow rate (l/d)	Q_{CH_4}	30 minutes	2
CO_2 gas flow rate (l/d)	Q_{CO_2}	30 minutes	2

The experimental manipulation under investigation here is the influent flow profile. The profile consists of step changes in the influent flow rate at the end of every week, resulting in 3 steps during the entire experiment. Upper and lower bounds for the flow rate are 250000 l/d and 66000 l/d respectively. This corresponds with the maximum and minimum allowable hydraulic retention time of the plant. Figure 17 shows an example flow profile, including the upper and lower bounds. Using computer simulations this profile is optimised in order to maximise the D criterion of the FIM, which corresponds with a minimisation of the volume of the confidence region of the parameters (see Section 2.5.2). For the design of the experiment we will assume that the concentrations of the influent components (S_1 , S_2 , Z and C) are constant. They are taken as the average values of the influent components of the long-term data. In reality this will of course not be the case but taking the averages is a good approach because the real fluctuations cannot be known in advance. Moreover, additional dynamics in the real influent will only further increase the information content of the data.

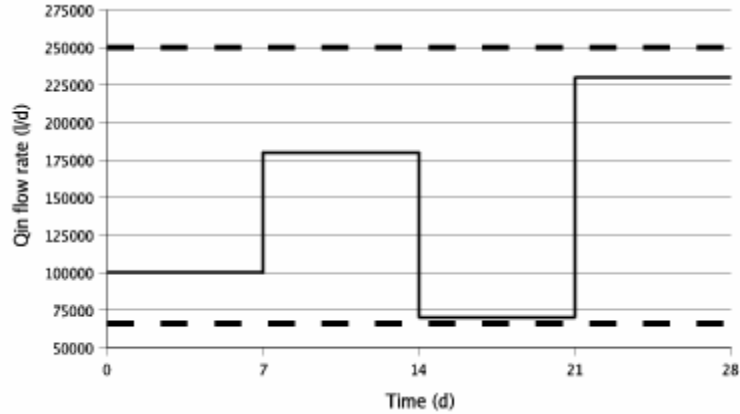


Figure 17: Example influent flow profile, including upper and lower bounds

Figure 18 represents the trajectory of the D-criterion during the optimisation of the influent flow rate. The optimisation algorithm used was Simplex (Nelder and Mead, 1965). For each experiment an influent flow rate is proposed by the optimiser. Using this profile, a sensitivity analysis of the measurements to the parameters is performed and the FIM is calculated. From this FIM, the D-criterion is calculated and new experiments are proposed by the optimiser based on previous D-criterion values. This process continues until the D-criterion reaches its maximal value, the “optimal” experiment has been found. In this case the optimum is reached after approximately 90 proposed and simulated experiments. Figure 19 shows the influent profile at the start of the optimisation (dotted line), experiment 1, and the optimal flow profile (full line), experiment 90. It can clearly be seen that the optimal profile is an alternation of the minimum and maximum allowable flow rate (bang-bang profile). This introduces the maximal amount of dynamics in the system and thus the highest possible information content of the data. The optimal experiment results in a FIM which can be used to predict what the parameter estimation standard deviations will be once the experiment is performed in reality and the model is fitted to the data. They are listed in Table 9.

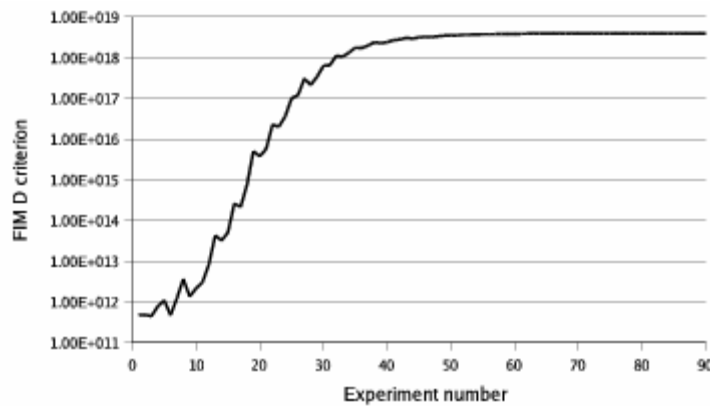


Figure 18. Evolution of the FIM D-criterion during the course of the experimental design optimisation

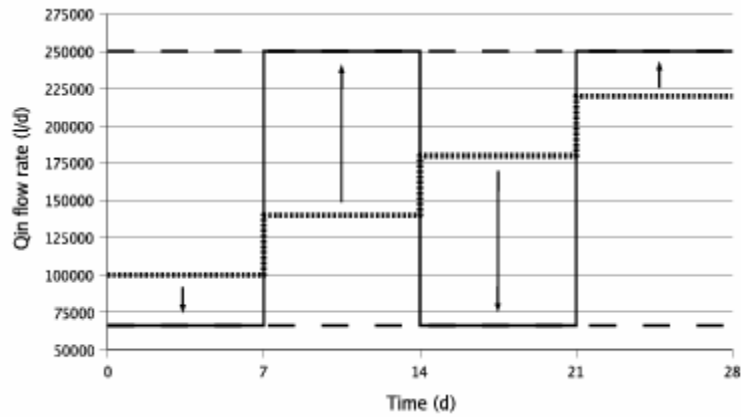


Figure 19: Initial (dotted line) and optimal (full line) influent flow profile. Upper and lower flow rate bounds are included as dashed lines. Arrows indicate the shift from initial to optimal influent flow profile

Table 15: Predicted parameter estimation standard deviations in case the model would be fitted to the data of the optimal experiment

Name	Value	Std. Dev.	Error (%)
k_1	20.63	0.46	2.21
k_2	243.17	9.16	3.77
k_3	344.55	11.84	3.44
k_4	30.42	14.57	47.93
k_5	184.44	19.18	10.4
k_6	466.25	9.46	2.03
μm_1	1.65	0.15	9.21
μm_2	0.70	0.01	2.02
K_{S1}	8.37	0.84	10.03
K_{S2}	7.17	0.13	1.8
α	349.42	0.02	2.69
k_{la}	0.56	1.34	8.07

Once the optimal experiment was determined it was performed in reality, in this case on the virtual plant. The complex ADM1 model was subjected to the optimal influent flow profile and the required variables (table 14) were sampled at the frequencies also listed in this table. Noise was added to these variables according to table 14 to produce the “virtual measurements”. The AM2 model was fitted again to the new data. Initial conditions for the state variables were obtained after long term simulations with a steady state influent file (average of the dynamic influent file). The Simplex optimisation algorithm (Nelder and Mead, 1965) was started from different initial parameter values to ensure the convergence to a global minimum. At the minimum the parameter estimation standard deviations were calculated based on the parameter estimation covariance matrix (inverse Fisher Information Matrix). Table 16 shows the parameter values at the optimum together with the predicted estimation standard deviations from the experimental design procedure and the calculated estimation standard deviations after the model fit. From the table we can conclude that all parameters were fitted with high confidence and that the experimental design procedure was able to predict the estimation standard deviations very well. As an illustration, model fits on

all 6 measurements are shown in Figures 20, 21, 22, 23, 24 and 25. The model is able to describe the data accurately, except for the pH where there is a slight offset but still within the noise band of the measurement.

Table 16: Optimal parameter values for the AM2 model fit to the optimal experiment data. Estimated and real parameter estimation standard deviations are also listed.

Name	Value	Standard Deviation	Error (%)	Predicted error (%)
k_1	11	0.4	3.64	2.21
k_2	175.87	6.61	3.76	3.77
k_3	262.94	10.07	3.83	3.44
k_4	14.11	1.98	14.04	47.93
k_5	134.81	5.72	4.24	10.4
k_6	251.36	9.43	3.75	2.03
μm_1	0.55	0.02	4.01	9.21
μm_2	0.42	0.04	8.36	2.02
K_{S1}	3.13	0.1	3.08	10.03
K_{S2}	15.44	1.24	8.06	1.8
α	0.59	0.03	4.6	2.69
k_{la}	18.83	1.56	8.29	8.07

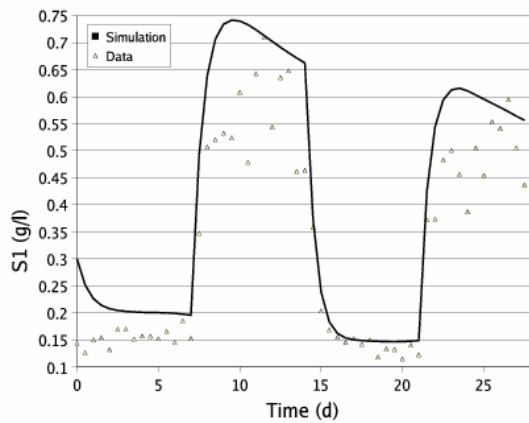


Figure 20: AM2 model fit on S_1 measurements from the optimal experiment

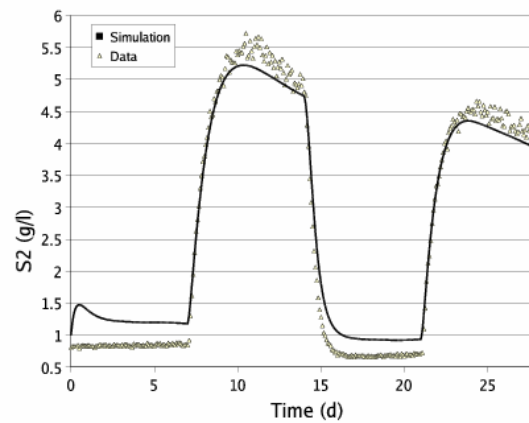


Figure 21: AM2 model fit on S_2 measurements from the optimal experiment

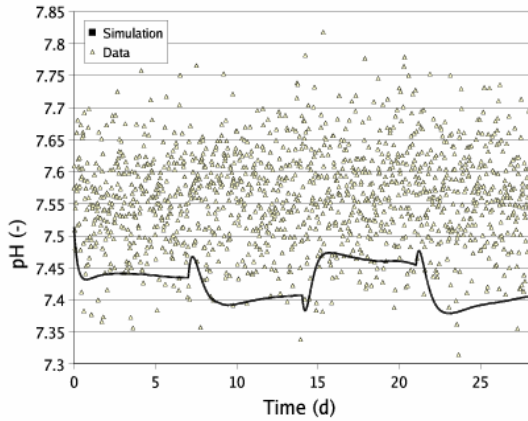


Figure 22: AM2 model fit on measurements from the optimal experiment

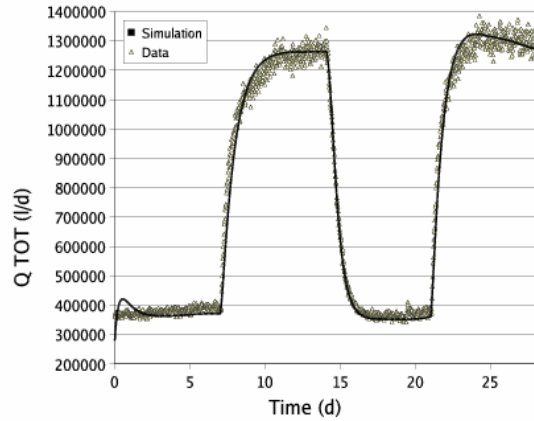


Figure 23: AM2 model fit on measurements from the optimal experiment

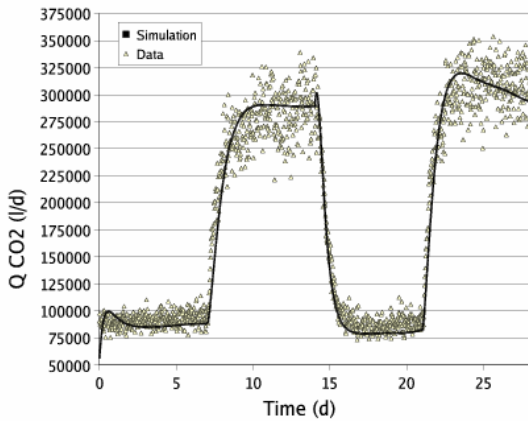


Figure 24: AM2 model fit on measurements from the optimal experiment

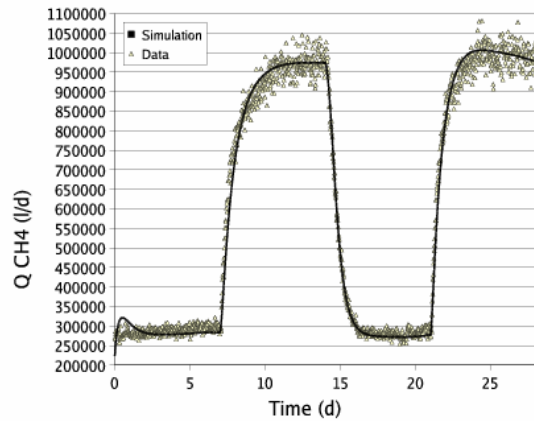


Figure 25: AM2 model fit on measurements from the optimal experiment

To our knowledge this has been the first application of optimal experimental design for anaerobic digestion. A wide range of experimental degrees of freedom has been considered: type of measurement, measurement frequencies and influent flow profile. This technique has proven itself to be highly valuable for the design of experiments in order to improve the model calibrations.

4 Agralco industrial case study

This section will deal with the application of optimal experimental design to the industrial scale CSTR of AGRALCO (2000 m³). A more detailed description of this system can be found in Roca *et al.* (2002).

4.1 Model calibration based on long term data

As a starting point for the experimental design procedure Telemac model AM2 (Bernard, 2002) was calibrated based on a dataset of 263 days (02/10/98 to 22/06/99).

For this period only flow and soluble COD measurements were available in the influent. Because other inputs were also required for the model (VFA, alkalinity and bicarbonate) some assumptions had to be made. From the influent wastewater characterization of the AGRALCO plant (Roca *et al.*, 2002) the ratio between the soluble COD and the VFA concentration was calculated to be 0.84 mmolVFA/gCODs. This ratio was used as a parameter during calibration and was found to be 0.8 mmolVFA/gCODs, which is very close to the calculated value. Another assumption that was made was that the influent total alkalinity concentration equaled that of VFA plus some fixed value. This value was also determined during the model calibration and was found to be 53,34 mmol/l. Influent bicarbonate was assumed to be very low and was taken 1 mmol/l. Figures 26-29 show the influent flowrate (Q_{in} , l/d), soluble COD concentration (CODs, g/l), volatile fatty acids concentration (VFA, mmol/l) and total alkalinity (Z , mmol/l) for the entire calibration period.

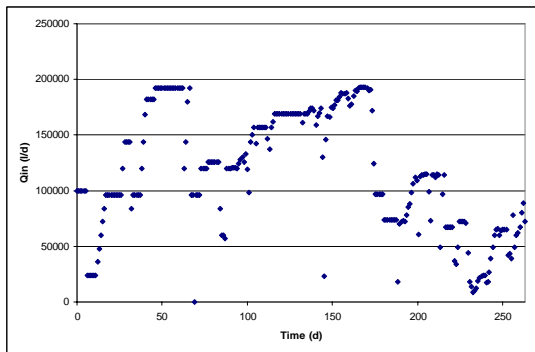


Figure 26: Influent flowrate (l/d) for the entire calibration period

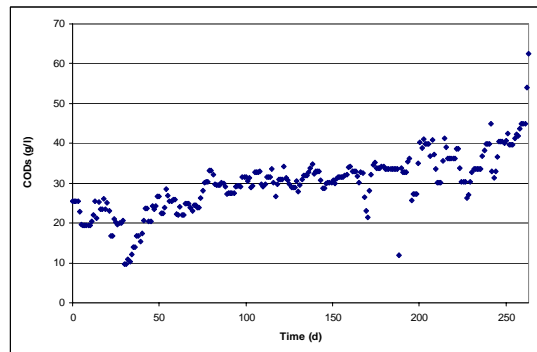


Figure 27: Influent soluble COD concentration (g/l) for the entire calibration period

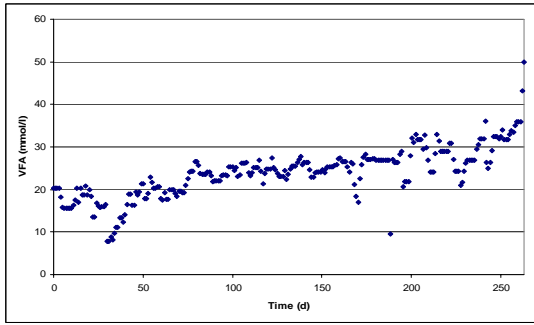


Figure 28: Influent VFA concentration (mmol/l) for the entire calibration period

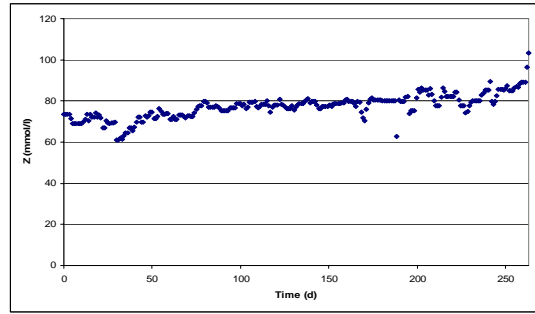


Figure 29: Influent total alkalinity concentration (mmol/l) for the entire calibration period

Daily measurements in the digester were available for soluble COD (CODs, g/l), volatile fatty acids (VFA, mmol/l), pH (-), total gasflow rate (l/d), percentage of CO₂ and CH₄ (%), total alkalinity (Z, mmol/l) and bicarbonate (C, mmol/l). The total gasflow rate was combined with the percentages of CO₂ and CH₄ to calculate the CO₂ and CH₄ flowrates.

Before proceeding with the model calibration the measurement errors (defined as standard deviations σ) were quantified. This was done based on expert knowledge. Table 17 lists the error type and magnitude for each measurement.

Table 17: Measurement error types and magnitudes for all available measurements

Measurement	Error type	%	Value
CH ₄ (l/d)	Absolute	4% of average (average = 1546083 l/d)	61843.32
CO ₂ (l/d)	Absolute	4% of average (average = 730668 l/d)	29226.72
pH (-)	Absolute		0.02
CODs (g/l)	Relative	10% of measurement value	
VFA (mmol/l)	Relative	5% of measurement value	
Z (mmol/l)	Relative	20% of measurement value	

The model was calibrated using a weighted sum of squares objective function and Simplex (Nelder and Mead, 1965) as the optimization algorithm. The weight for each measurement point was calculated based on the measurement errors listed in Table 17. The calibrated model parameters are listed in Table 18 together with their 95% confidence interval and error percentage. The error percentage was calculated by dividing the confidence interval value by the estimated parameter value. The initial conditions of the state variables were determined by running a long-term steady state simulation with the average influent composition. Figures 30-35 show the measurements together with the calibrated model output. From Table 18 and Figures 30-35 we can conclude that the model fit is adequate but certainly not perfect (large error %). The best estimate has a 25% error percentage and correlations among the parameters are very high. Looking at the model output we can conclude that the gas flow rates were fitted accurately while for the other measurements only the order of magnitude was predicted. This is probably caused by problems with the influent COD fractionation due to the abovementioned assumptions that had to be made and the relatively low information content of the data. The next step in the experimental design procedure is to find an optimal experiment that will result in more informative data.

Table 18: Estimated parameter values and initial conditions together with the 95% confidence interval and error percentage

Name	Value	95% confidence interval	Error %
k_1 (-)	189.567	± 51.098	26.95
k_2 (mmol/g)	115.772	± 69.068	59.65
k_3 (mmol/g)	56.348	± 19.832	35.19
k_4 (mmol/g)	51.229	± 609.520	1189.78
k_5 (mmol/g)	458.666	± 179.056	39.03
k_6 (mmol/g)	758.509	± 219.180	28.89
μ_{m1} (1/d)	0.201	± 0.200	99.36
μ_{m2} (1/d)	1.073	± 2.800	260.94
K_{S1} (g/l)	15.987	± 17.125	107.11
K_{S2} (mmol/l)	3.079	± 8.245	267.72
K_{I2} (mmol/l)	252.842	± 64208.546	25394.73
α (-)	0.614	± 0.159	25.93
$k_1 a$ (1/d)	19.385	± 6.662	34.37
S_1 (g/l)	0.694		
S_2 (mmol/l)	0.130		
X_1 (g/l)	1.420		
X_2 (g/l)	0.879		
Z (mmol/l)	132.458		
C (mmol/l)	137.576		

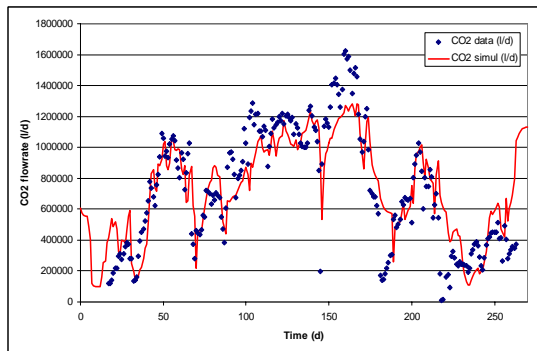


Figure 30: Measured CO₂ flowrate together with the calibrated model output

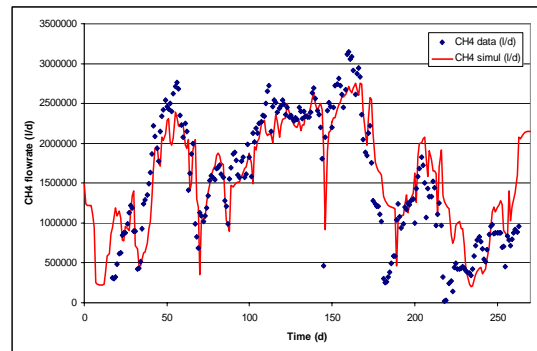


Figure 31: Measured CH₄ flowrate together with the calibrated model output

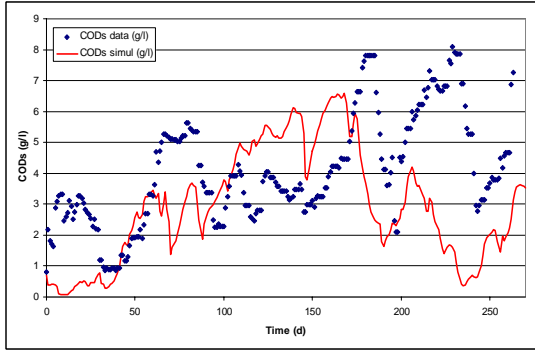


Figure 32: Measured soluble COD concentration together with the calibrated model output

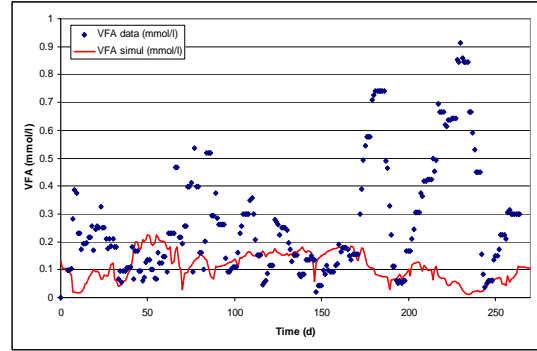


Figure 33: Measured VFA concentration together with the calibrated model output

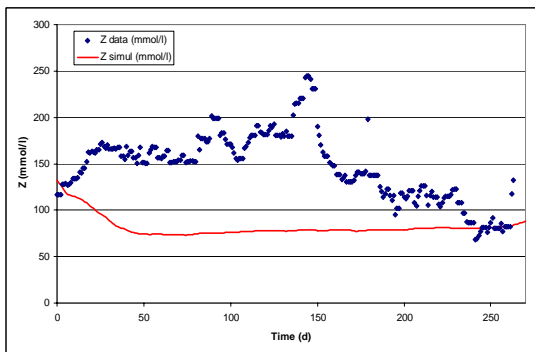


Figure 34: Measured total alkalinity together with the calibrated model output

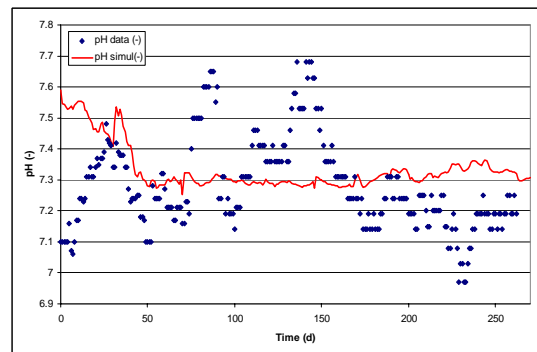


Figure 35: Measured pH together with the calibrated model output

4.2 Optimal experimental design

4.2.1 Defining the experimental degrees of freedom

The experimental degrees of freedom of a system are both experimental manipulations and measurements. For this case study the experimental manipulations are listed in Table 19 and the different available measurements in Table 20. It should be noted that the Telemac VFA sensor also measures total inorganic carbon and total alkalinity.

Table 19: Available experimental manipulations for the AGRALCO plant

Manipulation	Degrees of freedom
Dilution rate D (Q_{in}/V) (1/d)	Lower bound: 0.005 Upper bound: 0.2
$k_1 a$ (1/d)	$\pm 20\%$
Influent Z (mmol/l)	$\pm 20\%$

Table 20: Available measurements for the AGRALCO plant

Measurement	Measurement frequency
pH sensor	6 per hour
CODs measurement	1 per day
Classical VFA measurement	1 per day
Telemac VFA sensor	6 per hour
Total alkalinity measurement	1 per day
Classical gas sensor	1 per day
Telemac gas sensor	6 per hour

Another experimental degree of freedom that has to be considered is the model input, e.g. the influent wastewater composition. Since no information was available about the typical influent composition variations, the influent used during the experimental design was taken as the average influent composition of the calibration period.

4.2.2 Finding the optimal dilution rate profile

In order to find the dilution rate profile of the reactor that maximizes the information content of the measured data, a profile consisting of 14 periods was proposed (14 degrees of freedom). All other manipulations (k_1a and influent Z) remained fixed and the measurements were assumed to be: pH measurements, CODs measurements, Telemac VFA sensor measurements, total alkalinity measurements and Telemac gas sensor measurements. The experimental duration was fixed at 14 days.

The optimization of the 14 values of the dilution rate profile was performed using a real-value genetic algorithm (Herrera, 1998) because standard optimization algorithms (like e.g. Simplex) were unsuccessful in finding the global minimum. The objective of the optimization was to maximize the D-criterion of the Fisher Information Matrix (Section 2.5.2) in order to maximize the information content of the experiment and minimize the confidence region around the parameters. The profile which led to the highest D-criterion is illustrated in Figure 36. The D-criterion associated with this experiment is $1.84E+016$ and the predicted parameter estimation errors are given in Table 21.

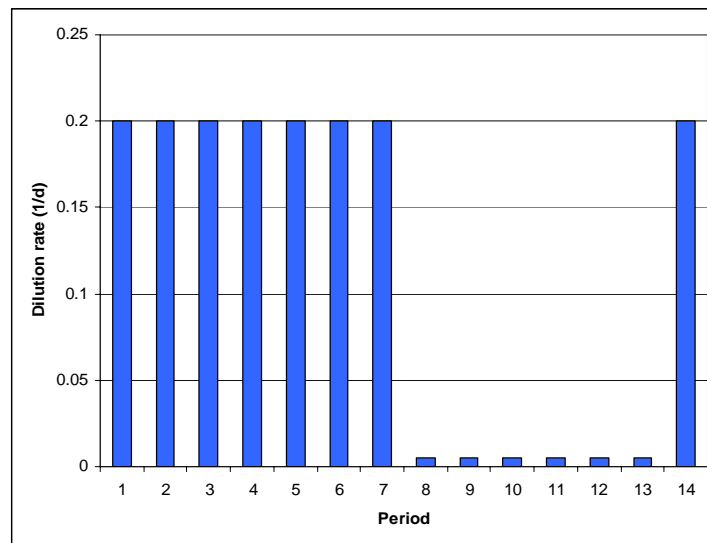


Figure 36: Optimal dilution rate profile

From Figure 36 it is clear that the optimal dilution profile consists of 3 phases. In the first phase, the reactor is operated at the highest dilution rate (0.2 d^{-1}), next the reactor is operated at the lowest dilution rate (0.005 d^{-1}) and finally at the end of the experiment, the dilution rate is switched back to the highest value. The characteristics of this optimal profile were also found to be valid for longer measurement campaigns.

4.2.3 Finding the optimal experimental duration

In this section it will be investigated what the influence of the experimental duration is on the predicted parameter estimation quality if the optimal dilution rate profile, found in the previous section, is used. Different experimental durations were tested ranging from 7 to 28 days. The estimated parameter error percentages together with the FIM D-criteria are listed in Table 21 for the proposed experimental durations and using the optimal dilution rate profile. Figure 37 also shows these results in a graphical way. From the table and the figure it is clear that longer experiments result in higher information content (higher D-criterion values) and in lower parameter estimation errors (lower error %). From Figure 37 it is also clear that the gain in parameter estimation accuracy is the highest from 7 to 14 days and that the added value for longer measurement campaigns of 14, 21 and 28 days is lower. Therefore, it is up to the experimenter to decide which parameter estimation accuracy is required. Based on this decision the experimental duration can be fixed. In the table the Modified-E criterion is also listed. This criterion is an indicator for correlations among parameters. It is found that increasing the experimental duration also increases the predicted correlations between the parameters.

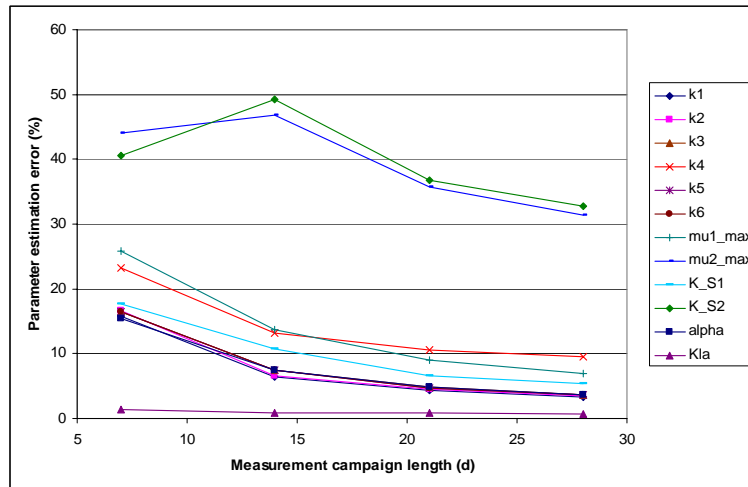


Figure 37: Optimal dilution rate profile

For the rest of this case study we will assume that the 28 day experiment is the optimal one. Also note that it is impossible to estimate K_{12} . This is due to the fact that this parameter is an inhibition parameter and that no inhibiting conditions occur for the proposed experiments. The system response for the optimal experiment (28 days, with optimal dilution rate profile and using all measurements) is shown in Figures 38-45. Notice that the pH for the proposed experiment hardly varies and that the CO_2 and CH_4 profiles are nearly identical in shape.

Table 21: Parameter estimation error % for different experimental durations

Parameter	Error % for different experimental durations			
	7 days	14 days	21 days	28 days
k_1 (-)	15.84	6.49	4.35	3.23
k_2 (mmol/g)	16.67	6.54	4.58	3.46
k_3 (mmol/g)	16.44	7.42	4.83	3.65
k_4 (mmol/g)	23.31	13.13	10.55	9.46
k_5 (mmol/g)	16.40	7.38	4.75	3.57
k_6 (mmol/g)	16.40	7.37	4.74	3.56
μ_{m1} (1/d)	25.83	13.77	8.97	6.93
μ_{m2} (1/d)	44.11	46.82	35.65	31.45
K_{S1} (g/l)	17.77	10.71	6.67	5.38
K_{S2} (mmol/l)	40.65	49.26	36.73	32.70
K_{I2} (mmol/l)	25540.75	32074.61	23413.05	20957.17
α (-)	15.42	7.49	4.78	3.66
$k_7 a$ (1/d)	1.30	0.90	0.78	0.68
D criterion	1.13E+10	1.84E+16	6.29E+19	8.65E+21
ModE criterion	9.24E+15	3.30E+16	4.75E+16	5.17E+16

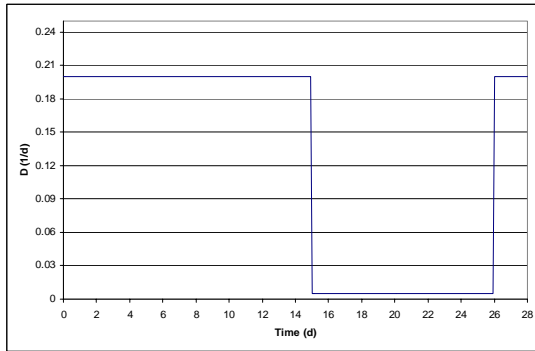


Figure 38: Optimal dilution rate profile for the optimal experiment

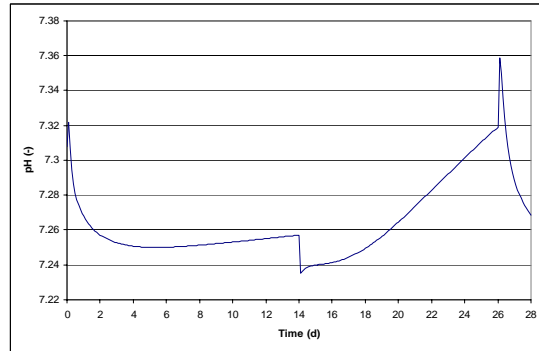


Figure 39: Expected pH response for the optimal experiment

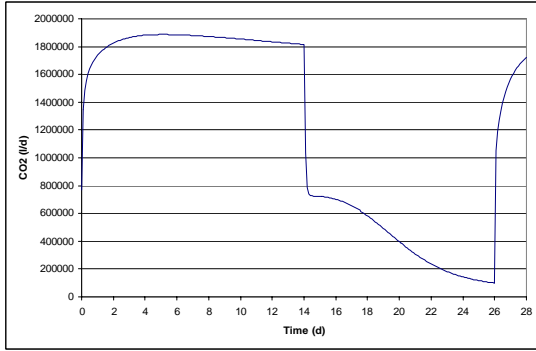


Figure 40: Expected CO₂ gas flow rate for the optimal experiment

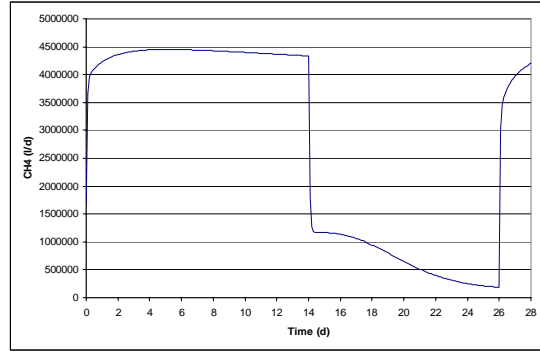


Figure 41: Expected CH₄ gas flowrate for the optimal experiment

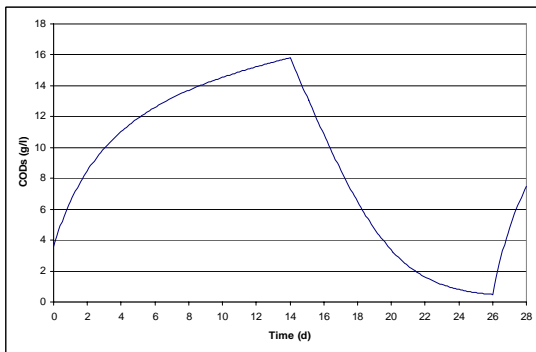


Figure 42: Expected soluble COD concentration for the optimal experiment

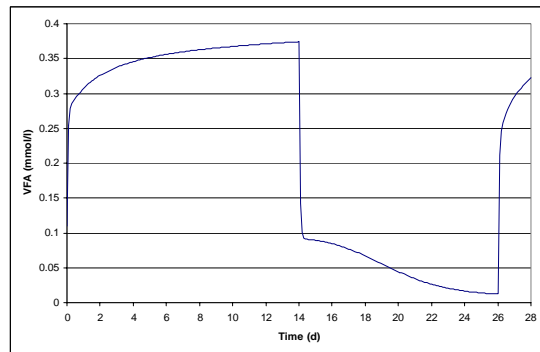


Figure 43: Expected VFA concentration for the optimal experiment

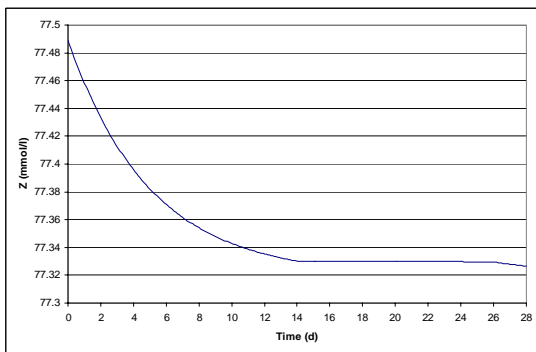


Figure 44: Expected total alkalinity concentration for the optimal experiment

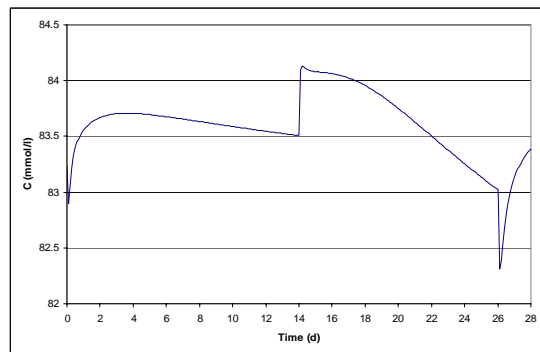


Figure 45: Expected bicarbonate concentration for the optimal experiment

4.2.4 Investigating the influence of different sensor arrangements

In this section it will be investigated what the added value for parameter estimation is of different possible sensor arrangements. Figure 46 shows some realistic sensor arrangements for anaerobic digestion plants with increasing complexity and cost. It will also be investigated what the added value of the Telemac VFA and gas sensor is compared to traditional VFA and gas measurements.

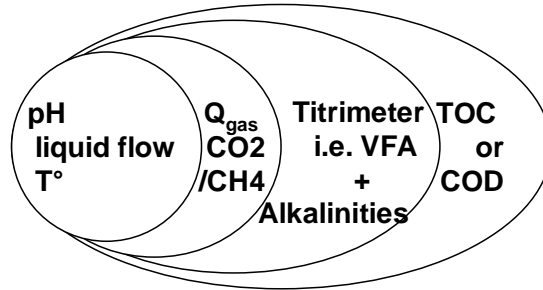


Figure 46: Possible sensor arrangements and expansion at anaerobic digestion plants

The basis for these investigations is the 28 day experiment using the optimal dilution rate profile. Table 22 lists the expected parameter estimation error percentages for different sensor arrangements. From the table it is clear that a pH measurement only is insufficient to calibrate the model. Adding a classical gas measurement (daily gas flow rate and CO₂-CH₄ ratio) improves the estimation accuracy for all parameters but the errors still remain too large, except for the k_{1a} which can now be estimated. Adding the Telemac gas sensor results in a large improvement in information content (D criterion going from 5.91E+05 to 1.98E+14) and in the parameter estimation quality. Adding classical VFA measurements (daily measurement) obviously decreases the expected estimation error for parameters related to VFA consumption: μ_{m2} and K_{S2} . The estimation errors for the other parameters only decrease slightly. Again, adding the Telemac VFA sensor increases the information content considerably (D criterion going from 2.82E+15 to 3.55E+21) and further improves the parameter estimation quality, especially for parameters related to VFA. Finally, adding daily soluble COD measurements only improves the parameter estimation quality slightly. The Modified-E criterion listed in the table indicates large correlations between the model parameters, which therefore remains a concern.

Table 22: Parameter error % for different sensor arrangements

Parameter	Sensor arrangement					
	pH	pH classical gas	pH Telemac gas	pH Telemac gas classical VFA alkalinity	pH Telemac gas Telemac VFA alkalinity	pH Telemac gas Telemac VFA alkalinity CODs
k_1 (-)	75.80	42.84	6.41	6.07	3.29	3.23
k_2 (mmol/g)	141.43	45.17	6.68	6.32	3.52	3.46
k_3 (mmol/g)	76.19	47.96	6.88	6.72	3.71	3.65
k_4 (mmol/g)	621.27	48.15	11.49	11.23	9.51	9.46
k_5 (mmol/g)	90.87	47.52	6.77	6.61	3.63	3.57
k_6 (mmol/g)	107.76	47.35	6.77	6.62	3.62	3.56
μ_{m1} (1/d)	188.83	94.85	14.50	13.70	7.09	6.93
μ_{m2} (1/d)	1025.67	336.44	112.94	77.27	31.83	31.45
K_{S1} (g/l)	160.89	72.27	11.39	10.74	5.52	5.38
K_{S2} (mmol/l)	1078.92	360.92	117.65	80.16	33.20	32.70
K_{I2} (mmol/l)	- *	- *	55773.73	44582.62	21300.14	20957.17
α (-)	93.64	49.50	7.44	7.05	3.73	3.66
$k_1 a$ (1/d)	110.44	2.56	0.69	0.68	0.68	0.68
D criterion	3.52E-05	5.91E+05	1.98E+14	2.82E+15	3.55E+21	8.65E+21
Mod-E criterion	3.88E11	1.02E+11	9.69E+17	2.42E+17	1.20E+17	5.17E+16

* could not be determined due to numerical problems

4.2.5 Determining the influence of a change in $k_1 a$ and influent total alkalinity

The effect of changing $k_1 a$ and the influent total alkalinity concentration by 20% was also investigated. Table 23 lists the results of these investigations. Similar to the previous section the basis for these investigations was the optimal experiment (28 days, with optimal dilution rate profile and using all measurements). Decreasing the influent Z concentration and decreasing the $k_1 a$ resulted in a slightly higher information content, the D-criterion going from 8.65E+21 to 1.02E+022 and 1.09E+022 respectively. Only the D-criterion values are listed in this table because these experimental manipulations did not result in significant improvements of the parameter estimation quality. Also note that the Modified-E criterion does not change significantly for the proposed experiments.

Table 23: Influence of changing k_{la} and the influent total alkalinity concentration on the information content

Experiment	D-criterion	Mod-E criterion
Optimal experiment	8.65E+21	5.17E+16
Increasing the influent Z concentration by 20%	2.50E+021	1.74E+17
Decreasing the influent Z concentration by 20%	1.02E+022	4.52E+16
Increasing the k_{la} by 20%	1.05E+021	2.29E+17
Decreasing the k_{la} by 20%	1.09E+022	8.50E+16

5 Software developments

The software developments done for this deliverable have partially been discussed and charged in Deliverable D3.2 (De Pauw et al., 2003). To implement the optimal experimental design procedure described in section 2.5.4 the software was extended (only these extensions are charged for this deliverable). For the sake of clarity the whole software package is described again, in order to give a clear view on the procedure a user has to follow to perform an optimal experimental design study.

The virtual case study presented in this deliverable was performed using West[®], a general purpose modelling and simulation environment (Vanhooren et al., 2003). In order to perform optimal experimental design, West[®] was extended with two modules: sensitivity analysis and optimal experimental design. Much attention was also given to the development of a Graphical User Interface (GUI) for these modules which allows non-expert TELEMAT users to use the optimal experimental design tool with a very short learning curve. West[®] was designed to be a software package with an open architecture for development, which allows other modules to be added very easily. Communication between the different West modules is done by using the West[®] API (Application Program Interface). This also has the benefit that virtually every software package can communicate with West (including MathWorks MatLab[®], Microsoft Excel[®], ...). This also provides easy communication with the TELEMAT system.

West[®] was programmed using C++, this choice was made based on different considerations:

- C++ is an object oriented programming language specifically designed to produce fast, compact and reusable code.
- C++ is generally available, including free versions for all existing platforms.
- Many Graphical User Interface (GUI) frameworks are available on all platforms to add GUI elements to C++ programs. Examples are Microsoft Foundation Classes (MFC), Trolltech Qt, Tcl/Tk.

To our knowledge this is the first software package to include the complete optimal experimental design procedure including a user friendly graphical user interface.

5.1 West[®] Sensitivity analysis module

The Sensitivity Analysis module added to West[®] implements local sensitivity analysis using the finite difference technique as described in this deliverable (Section 2.5.3).

The class architecture of this module is organised as follows:

SaGeneral. This class is the entry point of the module. It is responsible for (1) managing the parameters, variables, sensitivity functions and general options and (2) for the

communication with other West[®] modules (Simulation module, Model library module). It consists of a user interface (approximately 1350 code lines) and core functions (approximately 1300 code lines).

SaParameter. This class manages the properties of the selected parameters, including the perturbation factor. It consists of a user interface (approximately 430 code lines) and core functions (approximately 70 code lines).

SaVariable. This class manages the properties of the selected variables, including the instants where sensitivities are required. It consists of a user interface (approximately 620 code lines) and core functions (approximately 120 code lines).

SaSensitivityFunction. This class is used to define the sensitivity functions (combinations of variables and parameters). It also includes the calculation of the sensitivities and the quality indices. It consists of a user interface (approximately 240 code lines) and core functions (approximately 420 code lines).

The class organisation is also reflected in the user interface of this module. It can be divided in four sections, illustrated in figures 47, 48, 49 and 50. The first step in any sensitivity analysis is the selection of the parameters and variables for which one wants to know the sensitivity. Figure 47 shows the parameter selection tab where the user is able to select the required parameters, and to edit the relative parameter change (perturbation factor) that will be used during the sensitivity analysis. The variable selection window tab is shown in figure 48. In this window tab, variables can be selected and the time instants at which sensitivities need to be calculated can be entered. Selected parameters and variables can be combined to form sensitivity functions (figure 49). In short, a sensitivity function is the partial derivative of a variable to a parameter. When calculated, the sensitivity functions can be plotted and the quality criterion, used to assess the quality of the calculations (Section 2.5.3) will be displayed. A final window tab (figure 50) can be used to specify various calculation and output options.

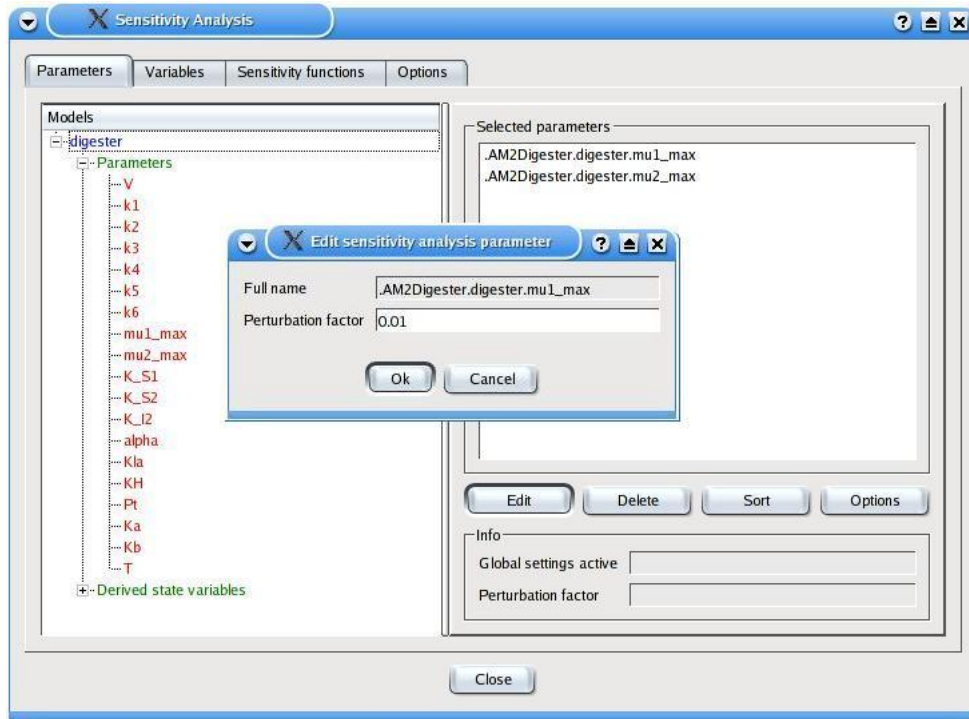


Figure 47: Sensitivity analysis parameter selection window tab

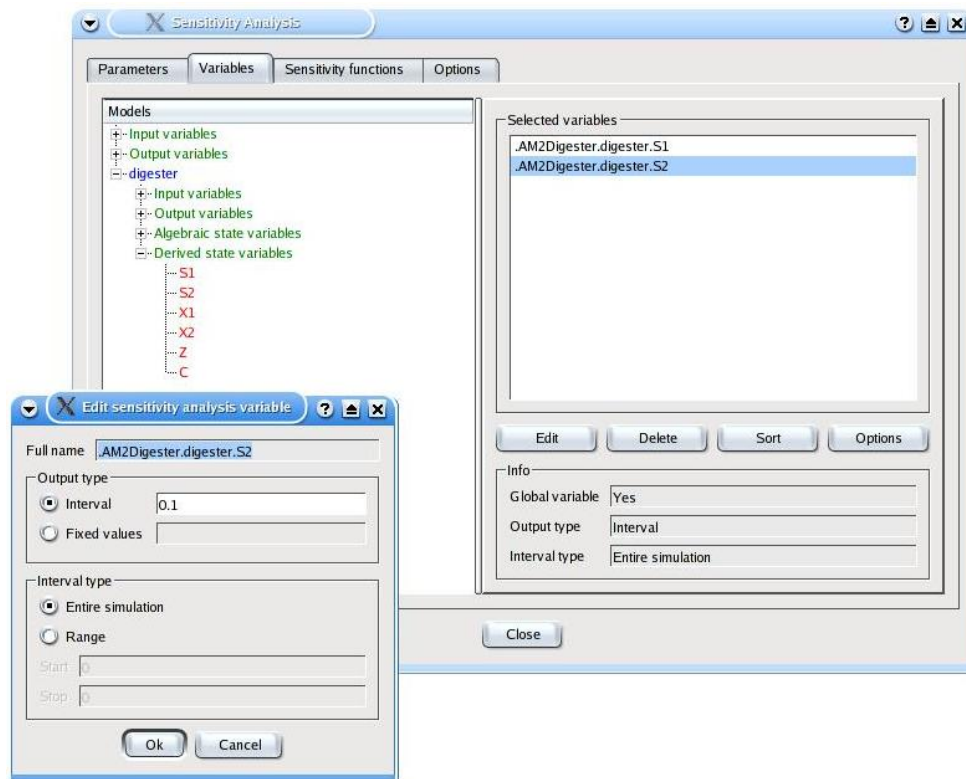


Figure 48: Sensitivity analysis variable selection window tab

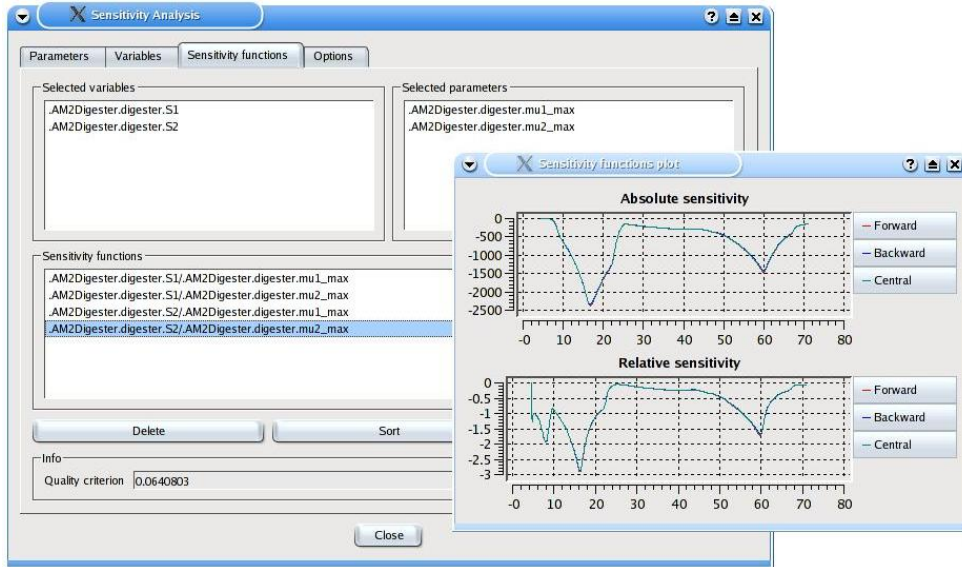


Figure 49: Sensitivity function window tab

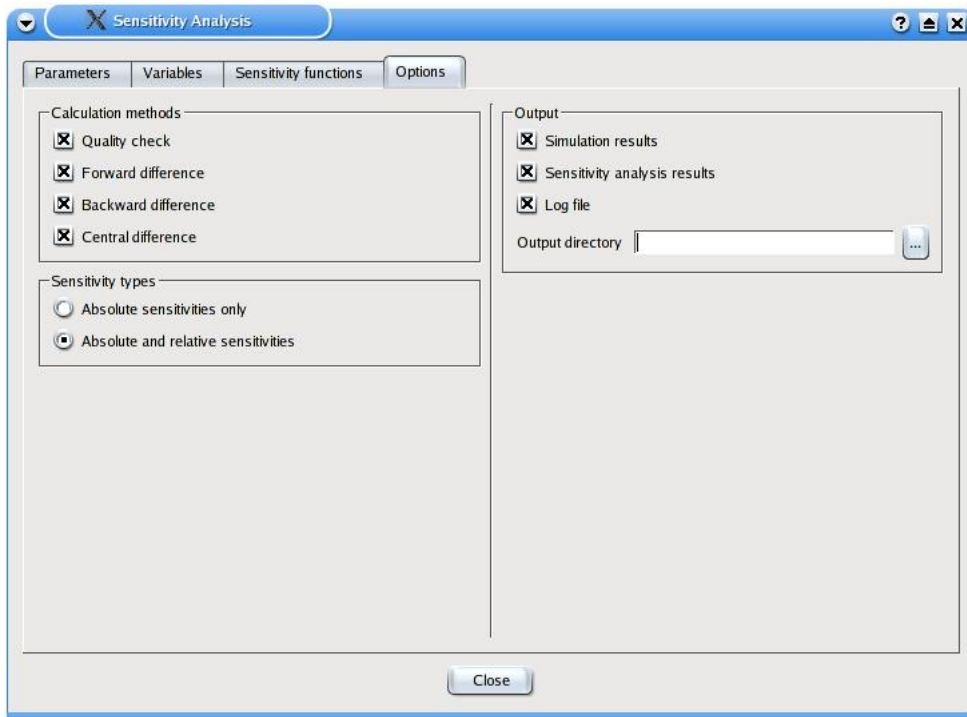


Figure 50: Sensitivity analysis options window tab

5.2 West[®] Optimal Experimental Design module

The Optimal Experimental Design module added to West[®] implements the general experimental design procedure described in Section 2.5.4.

The module is organised as follows:

OedpeGeneral. This class is the entry point of the module. It is responsible for (1) managing the parameters, measurements, manipulations, optimiser settings, objective settings and general options and (2) for the communication with other West[®] modules (Sensitivity analysis module, Optimisation module, Model library module). It consists of a user interface (approximately 1400 code lines) and core functions (approximately 1300 code lines).

OedpeParameter. This class manages the properties of the selected parameters, including the perturbation factors required for the sensitivity analysis. It consists of a user interface (approximately 430 code lines) and core functions (approximately 70 code lines).

OedpeMeasurement. This class manages the properties of the selected measurements. This includes the instants where measurements are available, measurement cost, optimiser specific settings and information about the correlation among the selected measurements. It consists of a user interface (approximately 1780 code lines) and core functions (approximately 720 code lines).

OedpeManipulation. This class is used to define the selected manipulations. For each manipulation, optimiser specific settings are stored including the type of manipulation, lower and upper bounds. It consists of a user interface (approximately 290 code lines) and core functions (approximately 120 code lines).

OedpeExperiment. This class is used to store the different experiments (number of measurements and experimental manipulations) proposed by the optimization algorithm. It is also responsible for the calculations of the Fisher Information Matrix and the different criteria associated with this matrix. The class consists of core functions only (approximately 880 code lines).

This class organization is reflected in the user interface of the module. It is organized in six parts. First of all, the parameters need to be selected (figure not shown). As a second step measurements (figure 51) and experimental manipulations (figure 52) need to be selected. For the measurements the choice can be made between variables that are measured with a measurement interval, a fixed number of measurement points or a set of measurement instants. Measurement costs can also be included in the optimization, in this way the costs of the measurement campaign can be weighted against the information content. The covariance button can be used to specify individual measurement errors and measurement correlations. The manipulations window tab (figure 52) can be used to specify experimental manipulations (degrees of freedom). Once the parameters, measurements and manipulations are selected, the objective function needs to be specified (figure 53) and an optimization algorithm should be chosen (figure 54). Different optimization algorithms are available for continuous and mixed continuous and discrete problems. Finally some calculation and output options can be set (figure not shown).

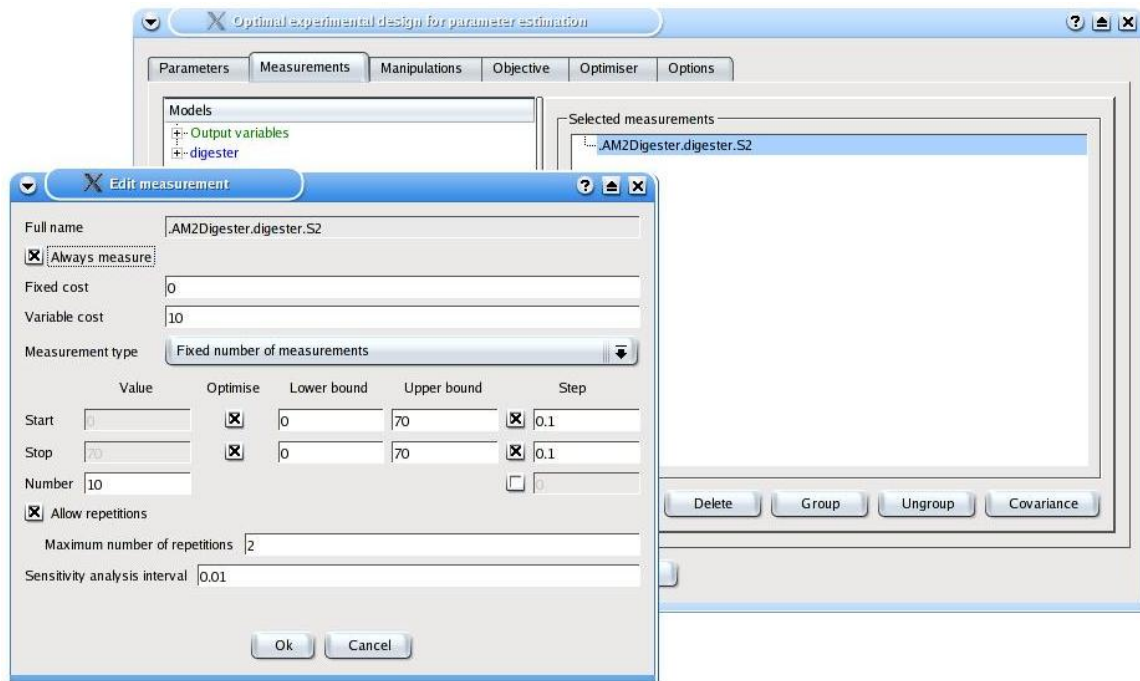


Figure 51: Optimal experimental design measurements window tab

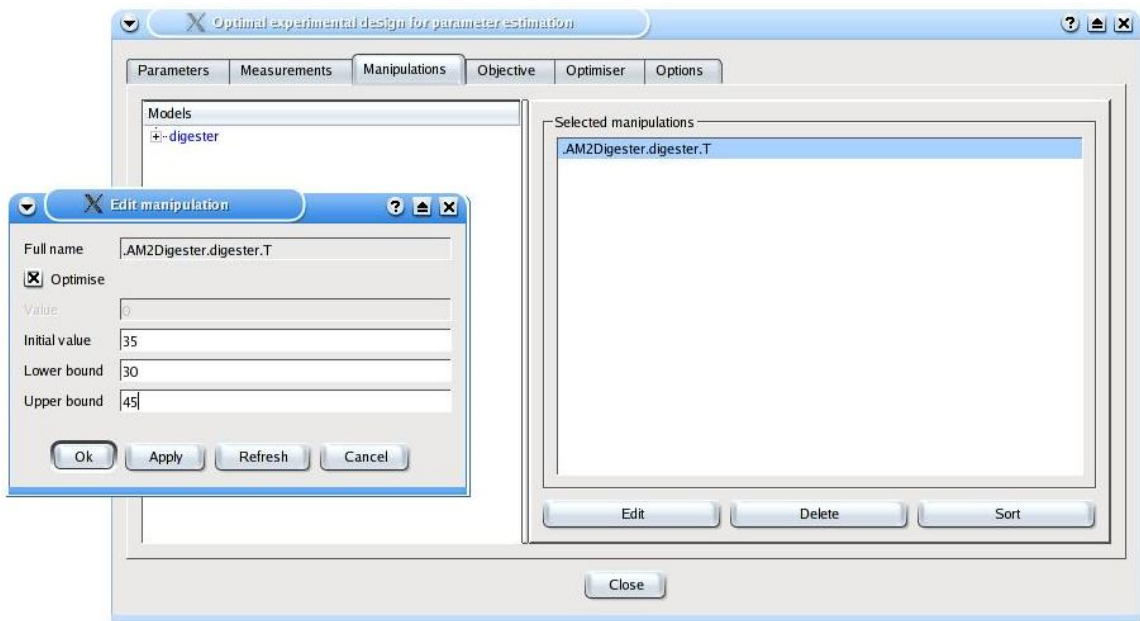


Figure 52: Optimal experimental design manipulation window tab

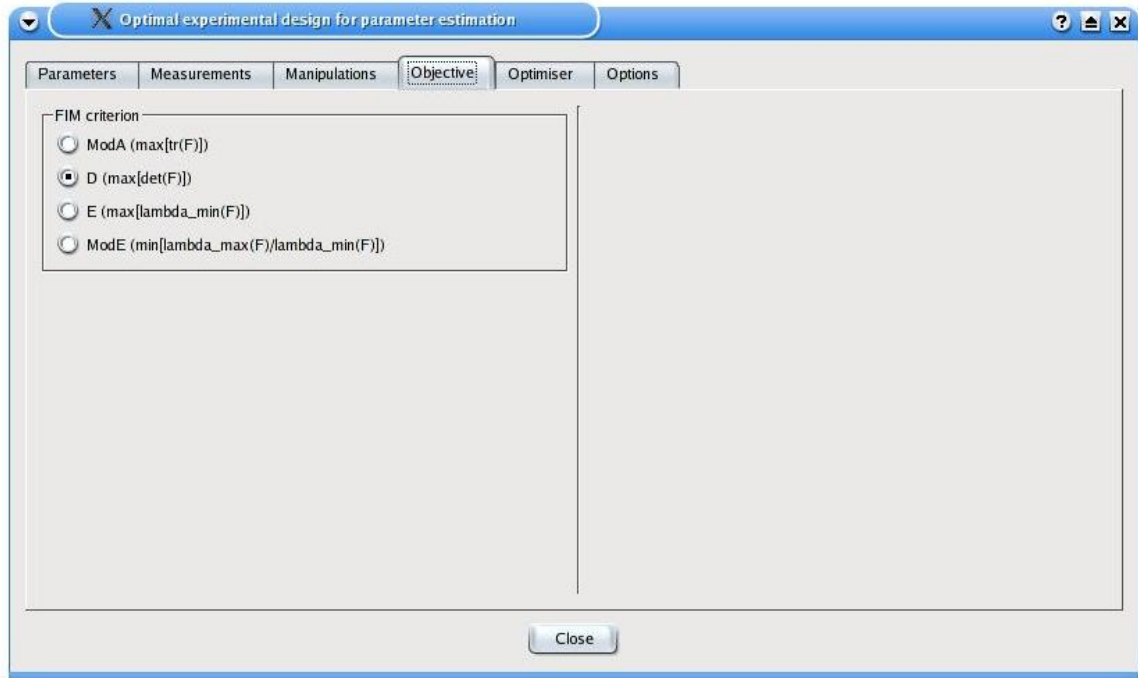


Figure 53: Optimal experimental design objective window tab

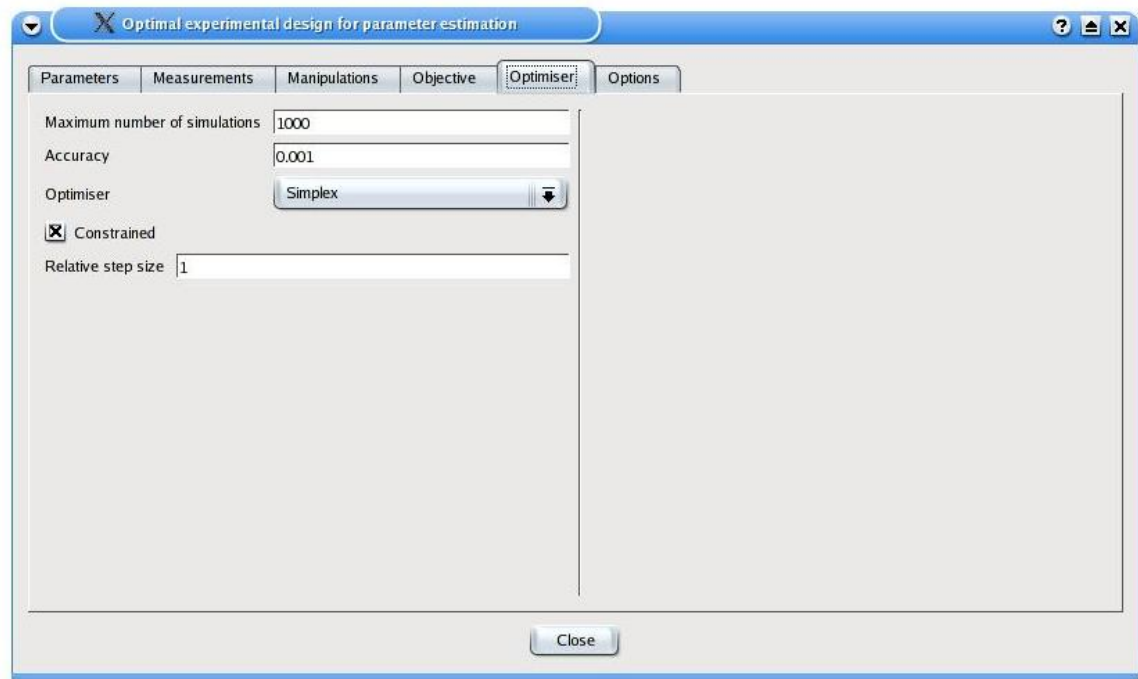


Figure 54: Optimal experimental design optimiser window tab

6 Conclusion

Through the report a protocol to set up the TELEMAT system has been developed. First, the protocol is presented schematically to show the main steps that will be followed to setup the system. The scheme consists of a straight sequence starting with data collection and fitting the main model of TELEMAT (AM2). Then it proceeds with the experimental design to determine the system settings. The determined settings for each plant will be used to calibrate the TELEMAT components (e.g. sensors) and then the TELEMAT system can be set into operation. These straightforward steps are suitable for the targeted applications e.g. (treating vinasses). Also, for the generality of the protocol and determination of the boundaries of the TELEMAT system, additional info is required. Fine tuning of the plant for process optimisation and possibility of model upgrade are considered in case of unexpected problems. Therefore, a detailed description of the data collection step is provided as know-how that can be exploited to guide and investigate the feasibility of the TELEMAT system installation at any anaerobic treatment plant.

The general BIOMATH know how in the field of Optimal Experimental Design is employed to develop a new, general and robust methodology to design the optimum set up experiment. The theory behind this methodology is briefly described in the report. Moreover, the methodology has been developed in a user-friendly software that will be easily used to support TELEMAT system installation.

A detailed example is given for the protocol implementation by applying it to a virtual plant and an industrial case.

A detailed reference model (IWA ADM1) is used to imitate a real digester and generate the necessary data for the experimental design. The virtual plant was designed by scaling up a lab experiment that was performed within the TELEMAT and that was earlier fitted by the ADM1 model. Then, through the user, the experimental design tools ran interactively with the virtual plant. The user here can be presented as the process expert who is involved in the installation of the system. The example does not only illustrate the implementation of the protocol but also highlights the added value of upgrading the plant with the TELEMAT sensors.

The industrial CSTR reactor of AGRALCO was used as an example for the implementation of optimal experimental design to a real case. The results gathered from this case are very similar to the virtual case study and clearly show the benefit of both Telemat sensors (VFA and gas).

References

- Anderson G. K., Donnelly T. and McKeown K. J. (1982) Identification and control of inhibition in the anaerobic treatment of industrial wastewaters, *Process Biochemistry*, 17(4), 28-32
- Batstone D.J., Keller J., Angelidaki R.I., Kalyuzhnyi S.V., Pavlostathis S.G., Rozzi A., Sanders W.T.M., Siegrist H. and Vavilin V.A. (2002), *Anaerobic Digestion Model no1*, ISBN: 1900222787, IWA publishing, London, UK.
- Bernard O. (2002) Design of models for normal working conditions, Technical report, TELEMAT IST 2000-28156, "Telemonitoring and Advanced Telecontrol of High-Yield Wastewater Treatment", Deliverable D3.1a.
- Blum J.W., Hergenroeder R., Parkin G. F. and Speece R.E. (1986) Anaerobic treatment of coal conversion wastewater constituents: biodegradability and toxicity. *Journal of the Water Pollution Control Federation*, 58(2), 122-131.
- Buswell E.G. and Neave S.L. (1930) Laboratory studies of sludge digestion, *Illinois Div of State Wat Survey*, 30.
- Dague R.R., Banik G.C., and Ellis T.G. (1997) ASBR treatment of dilute wastewater at psychrophilic temperatures, *Water Environment Research*, 70, 155-160.
- De Pauw D.J.W., Chachuat B., Bernard O. and Vanrolleghem P.A. (2003) A set of procedures to estimate the models parameters on-line and off-line, Technical report, TELEMAT IST 2000-28156, "Telemonitoring and Advanced Telecontrol of High-Yield Wastewater Treatment", Deliverable 3.2.
- Dochain D. and Vanrolleghem P.A. (2001), *Dynamical Modelling and Estimation in Wastewater Treatment Processes*, IWA Publishing, London, UK.
- Ehlinger F., Escoffier Y., Couderc J.P., Leyris J.P. and Moletta R. (1994) Development of an automatic control system for monitoring an anaerobic fluidized-bed, *Water Science and Technology*, 29(10-11), 289-295.
- Hanaki K., Matsuo T. and Nagase N. (1981), Mechanism of inhibition caused by long-chain fatty acids in anaerobic digestion process, *Biotechnology and Bioengineering*, 23, 1591-1610.
- Heijnen J.J., Mulder A., Enger W. and Hoeks F. (1989) Review on the application of anaerobic fluidized bed reactors in wastewater treatment, *The Chemical Engineering Journal*, 41(3), B37-B50.
- Herrera F., Lozano M. and Verdegay, J.L. (1998) Tackling real-coded genetic algorithms: Operators and tools for behavioural analysis, *Artificial Intelligence Review*, 12(4), 265-319.
- Hoffmann F., Posten C. and Rinas U. (2001) Kinetic model of in vivo folding and inclusion body formation in recombinant escherichia coli, *Biotechnology and Bioengineering*, 72(3), 315-322.
- Kennedy K.J. and Droste R.L., (1991) Anaerobic Treatment in Downflow Stationary Fixed Film Reactors, *Water Science and Technology*, 24(8), 157-177.
- Keurentjes, J. and Rinzema A. (1986) Adsorption of capric acid onto granular methanogenic sludge. *Anaerobic Treatment: A Grown-Up Technology, Industrial Presentations*, Schiedam, the Netherlands, Wastewater Treatment Conference, Amsterdam. 645-648.

- Koster I.W., Rinzema A., de Vegt A.L. and Lettinga G. (1986) Sulfide inhibition of the methanogenic activity of granular sludge at various pH-levels, *Water Research*, 20(12), 1561-1567.
- Koster I.W. and Lettinga G. (1988) Anaerobic digestion at extreme ammonia concentrations. *Biological Wastes*, 25, 51-59.
- Kroiss H. (1985) Toxicity problems during the anaerobic treatment of wastewater, Hannover *Industrieabwasser Tagung, Anaerobe Reinigung Industrieller Abwässer*, 12.
- Kugelman I.J. and McCarty P.L. (1965) Cation toxicity and stimulation in anaerobic waste treatment, *Water Pollution Control*, 37 (1), 97-116.
- Lettinga G., van Velsen A.F.M., Habma S.W., De Zeeuw W. and Klapwijk A. (1980) Use of the upflow blanket (UASB) reactor concept for biological wastewater treatment, specially for anaerobic treatment, *Biotechnology and Bioengineering* 22, 699-734.
- Ljung L. (1999), *System Identification: Theory for the User*, Prentice Hall, Englewood Cliffs, New Jersey, second edition.
- Marsili-Libelli S., Ratini P., Spagni A. and Bortone G. (2001), Implementation, study and calibration of a modified asm2d for the simulation of sbr processes, *Water Science and Technology*, 43(3), 69-76.
- Mehra R.K. (1974), Optimal input signal for parameter estimation in dynamic systems - survey and new results, *IEEE Transactions on Automatic Control*, 19(6), 753-768.
- Nelder J.A. and Mead R. (1965) A simplex method for function minimization, *Computer Journal*, 7, 308-313.
- Rabitz H., Kramer M. and Dacol D. (1983), Sensitivity analysis in chemical kinetics, *Annual Review of Physical Chemistry*, 34, 419-461.
- Reichert P. (1994) Aquasim - a tool for simulation and data analysis of aquatic systems, *Water Science and Technology*, 30(2), 21-30.
- Roca E., Steyer J.P., Bernard O. and Lema J.M.(2002) Experimental protocol specifications to reproduce normal and abnormal working conditions, Technical report, TELEMAT IST 2000-28156, "Telemonitoring and Advanced Telecontrol of High-Yield Wastewater Treatment", Deliverable 1.1.
- Roca E. and Lema J.M.(2002) Experimental data sets at laboratory, Technical report, TELEMAT IST 2000-28156, "Telemonitoring and Advanced Telecontrol of High-Yield Wastewater Treatment", Deliverable 1.3a.
- Rosen C. and Jeppsson U. (2002) Anaerobic COST benchmark model description, Technical Report, European Cooperation in the field of Scientific and Technical Research, COST-624 on Optimal Management of Wastewater Systems, pp. 14.
- Saltelli A., Chan K. and Scott E.M. (2000), *Sensitivity Analysis*, Wiley, New York.
- Turanyi T. (1990), Sensitivity analysis of complex kinetic systems, tools and applications, *Journal of Mathematical Chemistry*, 5, 203-248.
- Vanhooren H., Meirlaen, J., Amerlinck, Y., Claeys, F., Vangheluwe H. and Vanrolleghem P.A. (2003), West: modelling biological wastewater treatment, *Journal of Hydroinformatics*, 5(1), 27-50.

- Van Lier J. B., Boersma F., Debets M. M. W. H. and Lettinga G. (1994) High rate thermophilic anaerobic wastewater treatment in compartmentalized upflow reactors, *Water Science and Technology*, 30(12), 251–261.
- Van Velsen A. F. M. (1979) Anaerobic digestion of piggery waste, Department water pollution control, Agricultural University, Wageningen, Netherlands.
- Walter E. and Pronzato, L. (1997), *Identification of Parametric Models From Experimental Data*, Springer Verlag.
- WANG Kayun (1994) *Integrated Anaerobic and Aerobic Treatment of Sewage*, Ph.D. Thesis, Agric. Univ. of Wageningen, The Netherlands.
- Young J. C. and McCarty P. L. (1969) The anaerobic filter for waste treatment. *Journal of the Water Pollution Control Federation*, 41(5), 160-173.
- Zaher U., Rodríguez J., Franco A. and Vanrolleghem P. A. (2003) Application of the IWA ADM1 model to simulate anaerobic digester dynamics using a concise set of practical measurements, IWA Conference on Environmental Biotechnology, Advancement on Water and Wastewater Applications in the Tropics 9 – 10 December 2003, Kuala Lumpur, Malaysia.
- Zaher U., Vanrolleghem P.A. (2003) Data validation, Technical report, TELEMAC IST 2000-28156, "Telemonitoring and Advanced Telecontrol of High-Yield Wastewater Treatment", Deliverable 2.2(2).