

Table des matières

| | |
|--|-----------|
| Résumé | iii |
| Abstract | iv |
| Table des matières | v |
| Liste des tableaux | vii |
| Table des figures | viii |
| Remerciements | xi |
| Introduction | 1 |
| 1 Respirometry in literature | 4 |
| 1.1 Basic background | 4 |
| 1.2 Usage of different water quality parameters | 6 |
| 1.3 Respirometer designs and their use | 11 |
| 1.4 Overview of the Activated Sludge Models | 14 |
| 1.5 Structure of ASM1 | 17 |
| 1.6 Wastewater and biomass characterization for ASM1 | 19 |
| 2 Objectives | 31 |
| 3 Methodology | 33 |
| 3.1 Study site | 33 |
| 3.2 Anatomy of the RODTOX | 36 |
| 3.3 Fast loop | 38 |
| 3.4 Operation of the RODTOX | 39 |
| 3.5 Data processing | 43 |
| 3.6 Modelling tools | 45 |
| 4 Implementation of the RODTOX sensor | 46 |
| 4.1 Installation of the <i>fast loop</i> system | 46 |
| 4.2 Start-up tests of the RODTOX sensor | 50 |
| 4.3 Default output of the RODTOX sensor | 58 |
| 4.4 Python-based data analysis | 63 |
| 4.5 Building a mathematical model of the RODTOX sensor | 81 |

| | | |
|----------|--|------------|
| 5 | Conclusions and perspectives | 105 |
| 5.1 | Conclusions | 105 |
| 5.2 | Paths for improvement | 107 |
| | Bibliographie | 110 |
| A | RODТОX python functions | 116 |
| A.1 | Function to import RODТОX data into a python environment | 116 |
| A.2 | Function to smooth the RODТОX signal and to obtain the first and second derivative | 119 |
| A.3 | Function to analyze every respirogram to determine K_La and $stBOD$. . . | 121 |
| A.4 | Function to plot respirograms and their associated parameters | 129 |
| A.5 | Function to create WEST-compatible input files | 133 |
| A.6 | Function to create WEST-compatible objective DO and OUR time series for calibration | 135 |
| A.7 | Function to automatically create WEST-compatible objective DO and OUR time series from wastewater respirograms | 136 |

Liste des tableaux

| | | |
|------|---|-----|
| 1.1 | List of respirometer design configurations (ROD TOX in bold) | 13 |
| 1.2 | Oxygen mass balance components for different types of respirometers (ROD TOX in bold) (Spanjers et al., 1998) | 13 |
| 1.3 | Organic matter fractionation for ASM1 according to STOWA (1996) | 20 |
| 1.4 | Nitrogen fractionation for ASM1 according to Vanrolleghem et al. (2003) | 21 |
| 1.5 | ASM1 coefficients and the availability of ROD TOX -compatible methods for their determination (adapted from Petersen (2000) and Henze et al. (1987)) | 28 |
| 3.1 | Sensor array of the pil EAU te primary effluent | 35 |
| 4.1 | Evolution VSS concentration within the ROD TOX | 55 |
| 4.2 | Examples of raw probe data coming from the ROD TOX | 63 |
| 4.3 | Comparison of on-line sensor outputs and their associated laboratory measurement | 77 |
| 4.4 | Analytical measurements and ROD TOX on-line output | 78 |
| 4.5 | Example of WEST input file representing the injection of two doses of calibration solution. | 84 |
| 4.6 | Example of WEST objective file based on processed ROD TOX data. | 84 |
| 4.7 | Typical fractions of particulate components in activated sludge (adapted from Henze et al. (2000)) | 86 |
| 4.8 | Values of ROD TOX model relevant to endogenous respiration. | 88 |
| 4.9 | Result of the parameter estimation experiment for the calibration substrate biodegradation | 93 |
| 4.10 | Comparison of <i>stBOD</i> evaluations of Python-based data analysis and WEST simulations | 102 |
| 4.11 | Evaluation of the toxicity check <i>stBOD</i> using different techniques. | 103 |

Table des figures

| | | |
|------|---|----|
| 1.1 | Typical RODTOX respirogram | 6 |
| 1.2 | View of the consumption of different <i>BOD</i> fractions over time | 8 |
| 1.3 | Basic representation of a respirometer (adapted from Spanjers et al. (1998)) . . | 12 |
| 1.4 | Structure of the ASM1 model (adapted from Gujer et al. (1999)) | 15 |
| 1.5 | Structure of the ASM3 model (adapted from Gujer et al. (1999)) | 17 |
| 1.6 | COD fractions of the ASM1 model. | 18 |
| 1.7 | Nitrogen fractions of the ASM1 model. | 19 |
| 1.8 | Step-wise decrease in OUR due to S_S depletion (Ekama et al., 1986) | 22 |
| 1.9 | Characterization of RODTOX respirogram curves | 23 |
| 3.1 | Localization of the pilEAUte treatment plant and its catchment (Google Inc., 2017) | 34 |
| 3.2 | Overview of pilEAUte treatment plant | 34 |
| 3.3 | Overview of the RODTOX | 36 |
| 3.4 | Devices inside the RODTOX reactor | 37 |
| 3.5 | Flow diagram of the RODTOX reactor | 38 |
| 3.6 | Sketch of the essential components of a fast loop system | 39 |
| 3.7 | RODTOX operation flowchart | 43 |
| 4.1 | Design of the <i>fast loop</i> near the measurement tank | 47 |
| 4.2 | Position of the <i>fast loop</i> drainage valve. | 48 |
| 4.3 | Layout of pilEAUte primary treatment installations. | 49 |
| 4.4 | Mass transfer coefficient of the RODTOX reactor in clean water at different air flow rates. | 50 |
| 4.5 | Probe dynamics of the OxyGuard 420 when plunged alternately in oxygen-free and oxygen-saturated solutions (τ in seconds). | 52 |
| 4.6 | Respirograms of two « optimal experiments » created by the calibration solution. | 53 |
| 4.7 | Respirogram of a « non-optimal » calibration solution (left) | 53 |
| 4.8 | Evolution of pH within the RODTOX reactor with and without sodium bicarbonate addition. | 56 |
| 4.9 | Typical RODTOX automatically-generated respirometry report. | 58 |
| 4.10 | DO profile associated with Figure 4.9's measurement report. | 59 |
| 4.11 | Typical automatically-generated RODTOX toxicity report. | 60 |
| 4.12 | DO profile associated with Figure 4.11's toxicity report. | 60 |
| 4.13 | pilEAUte plant data acquisition network | 61 |
| 4.14 | Example of the output of the <i>Deriv</i> function ($n = 21$). | 66 |
| 4.15 | Example of the output of the <i>Deriv</i> function ($n = 7$). | 66 |

| | | |
|------|--|-----|
| 4.16 | Localization of the beginning of the re-aeration phase of respirograms. | 70 |
| 4.17 | Identification of the important points of a series of respirograms. | 70 |
| 4.18 | Respirograms from Figure 4.17 and their modelled DO concentration in the re-aeration phase and C_e | 72 |
| 4.19 | Proposed RODTOX PLC flowchart for the inclusion of OUR_{end} estimation. . . | 73 |
| 4.20 | Respirograms from Figure 4.17 and their associated $stBOD$ measurements. . . | 75 |
| 4.21 | $stBOD$ measurements and their associated OUR_{ex} profiles. | 76 |
| 4.22 | On-line sensor output in the primary effluent measurement basin on August 10, 2017. | 77 |
| 4.23 | Comparison of different water quality measurements relevant to oxygen demand. . | 78 |
| 4.24 | Comparison of different sludges' exogenous $SOUR$ | 79 |
| 4.25 | Respirograms used in the following modelling exercise | 82 |
| 4.26 | Layout of the WEST RODTOX model. | 83 |
| 4.27 | RODTOX simulation before any biokinetic calibration | 87 |
| 4.28 | RODTOX simulation with calibrated saturation DO concentration (C_S). | 89 |
| 4.29 | Time series from a typical parameter estimation experiment in which μ_H , K_S , μ_A and K_{NH} were estimated with DO as the calibration variable. (a) shows the predicted OUR , while (b) shows the poor quality of the parameter estimation. | 91 |
| 4.30 | Output of parameter estimation experiment for μ_H , K_S , μ_A and K_{NH} using OUR as the calibration variable. (a) shows the predicted OUR , while (b) shows the good quality of the parameter estimation. | 92 |
| 4.31 | S_{NH} concentration in the bioreactor during a calibration respirogram. | 95 |
| 4.32 | Effect of the time-delay constant for the modelling of the beginning of respirograms | 97 |
| 4.33 | Modelled and measured OUR of sequential respirograms. | 100 |
| 4.34 | Modelled and measured DO of sequential respirograms. | 101 |
| 4.35 | Modelled and measured OUR and DO during toxicity check (2017-08-09 00 :53 :00) | 103 |

The pleasure we derive from
journeys is perhaps dependent
more on the mindset with which
we travel than on the destination
we travel to.

Alain de Botton

Introduction

Good effluent quality is always the ultimate goal in wastewater treatment. However, a slew of factors conspires to make reaching good effluent quality ever more difficult. Among those factors figure the population explosion of the last century and its accompanying intensification of industry, in addition to the increasingly high standard to which effluents are held. For treatment plants to keep up with these standards, there has been widespread adoption of new sensors and process control strategies in recent years. These have made the operation of wastewater treatment plants ever more efficient and effective at removing pollutants. Beyond just instrumentation-based control strategies, the widespread use of mathematical models of wastewater treatment processes has also led to a better understanding of the processes underlying treatment, and thus to the development of more robust, efficient and versatile treatment strategies.

The activated sludge process figures among the oldest and most studied wastewater treatment processes. This process consists in harnessing the organisms that thrive in wastewater and use their ability to degrade the pollutants it contains for our own purposes. In this process, the growth of these organisms is carefully fostered within biological reactors by feeding them fresh wastewater. The organisms then use the pollutants present in wastewater to grow, thus forming themselves into a concentrated sludge, and leaving the water surrounding them relatively free of soluble contaminants. After sedimentation of the sludge, this water can be discharged to the receiving waters.

One of the methods used to characterize activated sludge and the pollutants it consumes is respirometry, which consists in the measurement and interpretation of the sludge's respiration rate (Spanjers et al., 1998). The usefulness of respirometry rests upon the fact that the respiration rate of organisms present in activated sludge is different for different organisms populations, and for the biodegradation of different types of substrate. Specifically, organisms capable of transforming nitrogen compounds into nitrate — autotrophs — exhibit a different respirometric « signature » than organisms responsible for organic carbon removal — heterotrophs. Similarly, compounds which need to be transformed before being metabolized by cells create a different respirometric signal than compounds which are directly available to organisms. Thus, information about the nature of the pollutants present in water can be extracted from

respirometric experiments, in addition to information on the organism populations present in the sludge.

Several respirometers have been designed over the years to extract this information out of respirometric experiments. Most of those designs are meant to be exploited within a laboratory setting. However, others, like the RODTOX (short for ***R**apid **O**xygen **D**emand and **TOX**icity *tester**), are meant to be installed within treatment plants in an on-line fashion, meaning that wastewater can be sampled at will from the treatment and be analyzed in real-time. The RODTOX performs its measurements by exposing a sample of the plant’s bioreactor sludge to wastewater samples, and by then tracking the changes in the sludge’s dissolved oxygen concentration as it biodegrades the substances contained in the wastewater.

In 2016, Université Laval’s pilEAUte treatment plant was equipped with a RODTOX sensor to analyze its wastewater. The pilEAUte plant treats 24 m^3 of domestic wastewater per day. It consists of a primary treatment, which removes particulate matter from wastewater, followed by an activated sludge-based secondary treatment capable of carbon removal, nitrification and denitrification. The RODTOX was purchased to monitor the water quality at the effluent of the primary treatment, directly upstream of the plant’s biological reactors. It was hoped that this new piece of equipment could complement the information already provided for research and operation by the current array of on-line probes and sensors installed in this location.

In its default configuration, the RODTOX can produce two measurements : a toxicity measurement which tracks the activity of activated sludge, and a quantification of the short-term biochemical oxygen demand (*stBOD*), which measures the amount of readily biodegradable organic matter and nitrifiable nitrogen inside a wastewater sample. Though those measurements are interesting, they represent a small fraction of the information which can potentially be extracted from respirometric experiments (Spanjers and Vanrolleghem, 2016).

The goal of this study was, therefore, to install the RODTOX within the plant so that it could sample water from the effluent of the primary treatment, and then try to unlock as much of the information contained in its respirometric experiments as possible. The latter was done by extracting the data produced by the RODTOX’s built-in probes and then developing software and modelling tools to process this data.

The decoding process was done in two main stages. First, the raw RODTOX data was imported into a Python-based computing environment, where a series of custom functions were developed to extract relevant information from different parts of the probe signals. Then, this processed data was used in conjunction with a mathematical model of the RODTOX reactor and its biomass to decode the signal from respirometric experiments further. This process eventually made it possible to estimate the sludge’s biokinetic characteristics. Using this sludge characterization, it was then possible to separately estimate the oxygen demand coming from

the biodegradation of nitrogenous and carbonaceous substrates — the wastewater samples' $stBOD_N$ and $stBOD_C$, respectively.

The latter two results are particularly interesting, as the two components of $stBOD$ they describe serve very different functions within the activated sludge process. Therefore, the success of this separation opens a path towards their use for monitoring or control of the pilEAUte plant based on the nitrifiable nitrogen loading or the availability of readily biodegradable organic matter. Indeed having access to these parameters could help to better monitor and operate the nitrification and denitrification processes. These values could also be used to determine the required amounts of additional external carbon or alkalinity needed to optimize these treatments' performance.

After the methods used to implement the RODTOX sensor have been outlined, the tools developed to extract information from it are described in detail. Then, a short discussion is made where future improvements are suggested, and eventual uses of the fully-implemented RODTOX sensor are proposed.

Chapter 1

Respirometry in literature

1.1 Basic background

According to [Spanjers et al. \(1998\)](#), respirometry is the measurement and interpretation of the respiration rate of activated sludge. In turn, the respiration rate is defined as the rate at which oxygen (O_2), or any other electron acceptor for that matter, is consumed by activated sludge to support its metabolic processes. Though wastewater and activated sludge hold a complex ecosystem with several trophic levels, the bulk of its biomass is composed of bacteria and archaea. Taken as a whole, this biomass can be separated into two categories. The heterotrophic organisms use external sources of organic carbon as an electron donor. On the other hand, autotrophic organisms use inorganic carbon dioxide as their carbon source, and other substrates such as ammonia, nitrite, iron III or sulfur as the electron donor for their metabolic processes ([Metcalf & Eddy, 2013](#)). In wastewater, however, the most prevalent of these autotrophic substrates is ammonia.

The primary electron acceptor used by microorganisms, regardless of whether they are heterotrophic or autotrophic, is dissolved oxygen. However, some organisms, such as facultative or anoxic heterotrophs, use nitrate (NO_3^-) as an electron acceptor. These organisms are instrumental to wastewater treatment, as they are responsible for denitrification. Finally, anaerobic bacteria and archaea, which are responsible for the fermentation and methanogenesis occurring in anaerobic processes, use inorganic molecules such as CO_2 as their final electron acceptor. Though they do play a key role in sludge treatment, anaerobic organisms are not involved in the activated sludge process ([Spanjers and Vanrolleghem, 2016](#)).

Given the role of O_2 as the primary electron acceptor in activated sludge, its importance in activated sludge treatment cannot be overstated. Respirometry can, therefore, provide relevant information about the processes occurring within activated sludge by allowing one to assess these processes' oxygen consumption.

In broad terms, the DO concentration within activated sludge (S_O) can be understood as resulting from the transport of dissolved oxygen in and out of the system and the uptake of DO by the organisms, as is described in Equation 1.1 for a generic control volume.

$$\frac{dS_O}{dt} = \frac{J_{O, in}}{V} - \frac{J_{O, out}}{V} - r_O \quad (1.1)$$

Here, the net change in DO concentration in the studied system $\frac{dS_O}{dt}$ is a function of the mass flows of oxygen J_O both in and out of the control volume, the reactor volume V and the net respiration rate r_O . The respiration rate itself can be further broken down into two main components. The endogenous oxygen uptake rate (OUR_{end}) is caused by cells using oxygen to maintain their metabolic functions and to break down the waste and lysis products of other dying cells. The exogenous oxygen uptake rate (OUR_{ex}), on the other hand, accounts for the oxygen required for the consumption of fresh substrate by the cells (Spanjers and Vanrolleghem, 2016). The contribution of both components to the net respiration rate is shown in Equation 1.2.

$$r_O = OUR = OUR_{end} + OUR_{ex} \quad (1.2)$$

Several devices and clever experimental methods have been developed over the years to track both the endogenous and exogenous oxygen uptake rates and extract from them different characteristics of either the sludge being used or the substrate being fed to it (Spanjers et al., 1998).

In the context of this study, the device being used is a RODTOX (*Rapid Oxygen Demand and TOXicity Tester*) sensor (Vanrolleghem, 1994). The makeup of this device is thoroughly explained below but for now, suffice it to say that the RODTOX consists of a constantly-aerated, sludge-filled reactor equipped with a DO sensor. Wastewater samples can be added to this reactor, which increases the sludge's exogenous respiration rate. Doing so disturbs the DO equilibrium within the reactor, resulting in a DO decrease during the degradation of the wastewater sample. After the sample has been depleted of readily oxidizable substrates, OUR_{ex} decreases, and the DO concentration goes back to its equilibrium state. This decrease in DO, followed by its return to the baseline concentration, can be plotted over time, and the result is called a 'respirogram'. A typical RODTOX respirogram is shown in Figure 1.1.

It is through the study of those respirograms that RODTOX-aided respirometry allows one to learn about the composition of the wastewater samples, as well as about the sludge within the RODTOX itself. To efficiently communicate about those wastewater and sludge characteristics, however, some water quality parameters must be defined.

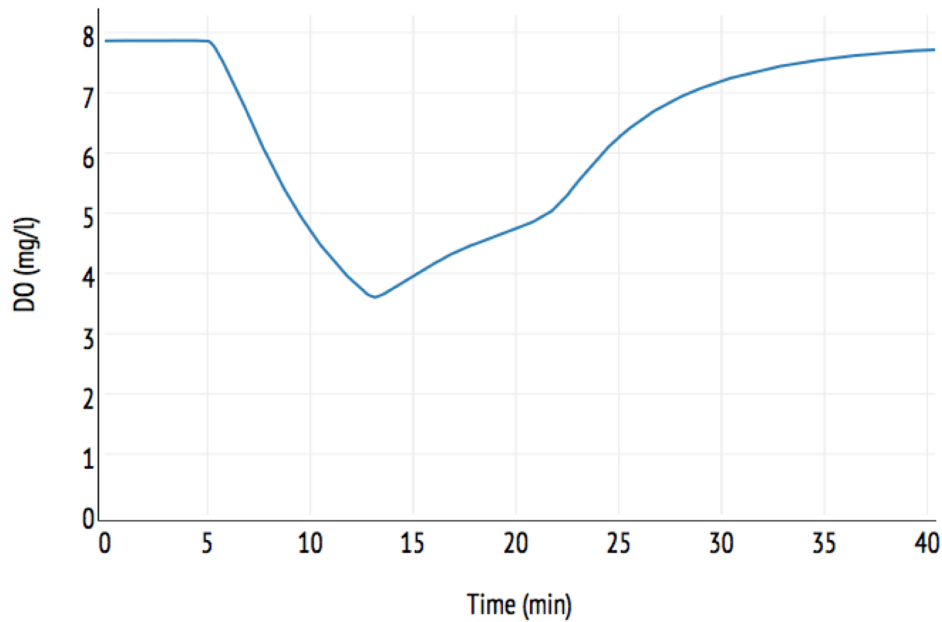


FIGURE 1.1 – Typical RODTOX respirogram

1.2 Usage of different water quality parameters

Besides its being important for respiration measurements, oxygen measurements can also be used to construct useful water quality parameters. Below are examples of such parameters and their uses.

1.2.1 Biochemical Oxygen Demand (BOD)

Given that oxygen is poorly soluble in water, it can often become scarce if biodegradable substances are added to the water since bacteria then increase their respiration to consume it. Thus, measuring a water sample's oxygen demand makes it easy to determine at-a-glance whether that water is likely to develop anoxic conditions that are detrimental to the ecosystem. The first use of this method of wastewater characterization in a regulatory setting is attributed to the Royal Commission on Sewage Disposal ([House of Commons of the Parliament of Great Britain, 1912](#)). This groundbreaking regulation established the often-used 30:20 effluent standard, meaning that a maximum concentration of 30 mg/l TSS and 20 mg/l of biological oxygen demand being consumed over a period of five days (BOD_5 for short) were allowed in wastewater treatment plant effluents.

Assessing BOD_5 is done by measuring the amount of dissolved oxygen consumed by a sample of biomass while it biodegrades the organic matter present in the water sample. Given its importance for regulatory purposes, the procedure for measuring BOD_5 is tightly standardized

(Spanjers and Vanrolleghem, 2016). In its typical incarnation, the test consists of measuring the change in dissolved oxygen concentration inside airtight, dark glass bottles containing a known volume of sample, which is kept at a constant temperature. Using Spanjers et al. (1998)’s classification, *BOD* bottles can thus be thought of as rudimentary ‘LSS’ respirometers (see section 1.3). The organisms used to consume the BOD_5 during the test come either from the sample itself or a small amount of ‘seed’ culture containing microorganism populations consistent with the water where the sample being tested is meant to be released. The final DO measurement, taken after five days, enables the calculation of the BOD_5 parameter following Equation 1.3.

$$BOD_5 = \frac{D_1 - D_2}{P} \quad (1.3)$$

Here, D_1 and D_2 represent the DO measurements taken at time $t = 0$ and $t = 5$ days respectively, whereas P represents the decimal fraction of the sample in the overall dilution found in the *BOD* bottle.

Depending on the DO probe being used, one may track the change in DO inside the bottles, and thus the *BOD* being consumed, continuously. Using this continuous tracking, one observes that for typical domestic wastewater, the *BOD* curve can be broken down into three distinct parts, as can be seen in Figure 1.2. The first part is the lag phase which occurs while the organisms inside the bottle acclimate to their new substrate and environmental conditions. Then, the first increase in *BOD* consumption is caused by the organisms starting to digest organic matter. This is followed by the second uptick in *BOD* consumption, which is caused by autotrophic organisms beginning to oxidize nitrogen compounds. This last increase in *BOD* occurs several days after the beginning of the test since the growth rate of autotrophs is small in comparison with heterotrophs. To avoid any consumption of the nitrogen compounds during the test, and thus only measure carbonaceous *BOD*, one can add a nitrification inhibitor to the *BOD* bottle, e.g., allylthiourea (ATU) (Spanjers and Vanrolleghem, 2016). After several weeks of digestion, nearly all the biodegradable matter has been consumed, and *BOD* asymptotically approaches a maximum value. ‘Ultimate *BOD*’ or BOD_u , is measured 28 days into a *BOD* incubation, at which time *BOD* consumption nears its maximum value. This parameter is meant to represent the total amount of biodegradable matter found in a water sample.

When only the carbonaceous portion of the *BOD* is considered, the consumption rate can often be modelled using a first-order kinetic equation, such as the one described by Equation 1.4.

$$BOD(t) = BOD_u(1 - e^{-k(t-t_0)}) \quad (1.4)$$

Here, k is the degradation constant and t_0 is the time at which the lag phase ends. The value of k indicates the speed at which the sample is degraded. A high value suggests that the compounds in the water sample are mostly readily biodegradable, whereas a lower constant

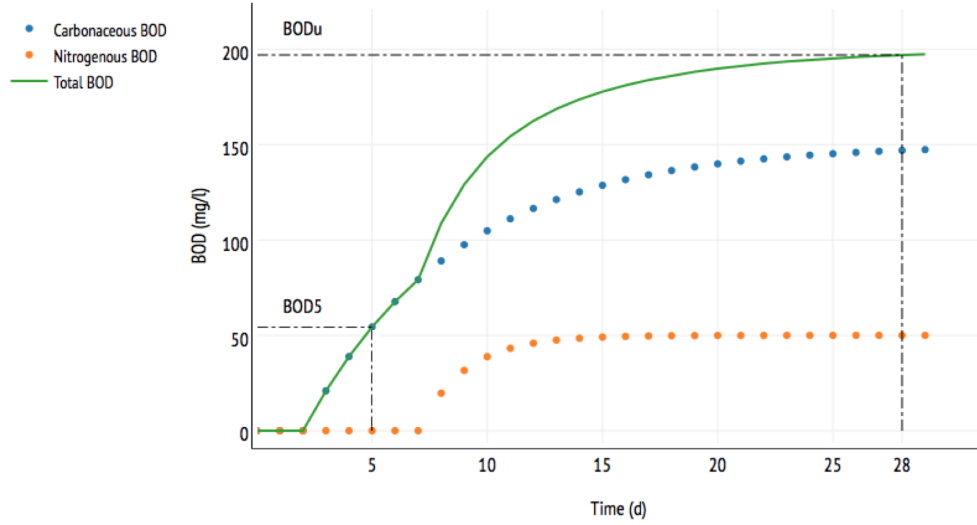


FIGURE 1.2 – View of the consumption of different *BOD* fractions over time

indicates that the substrate is relatively complex and therefore organisms need more time to break it down and consume it.

Thus, *BOD* measurements help to quantify how biodegradable a sample is, and how fast this biodegradation takes place. However, they fail to provide information on the proportion of readily biodegradable and slowly biodegradable substrates present in the sample. Additionally, the amount of microorganisms involved in the degradation of the sample is quite low compared to how much sample there is — this is called the substrate-to-biomass (S_0/X_0) ratio). It, therefore, may be assumed that a sizeable amount of substrate will not be used for respiration, but will instead be used for biomass growth. To account for this growth, one must know the organic yield Y (the ratio of the mass of new organisms produced relative to the total mass of organic matter consumed by the organisms that produced them. However, this yield may change across different samples, and its value cannot be extracted from the *BOD* test itself. This means that it is difficult to account for the biomass growth in calculations. Furthermore, as was discussed above, BOD_5 and BOD_u measurements are made over a period of several days or weeks. This presents a significant disadvantage for any task related to control or monitoring of wastewater treatment since the value of any measurement is known much too late for any corrective action to be made to the system if a problem is ever detected. However, their focus on biodegradable matter make *BOD* tests excellent environmental quality indicators, which is why they have been widely used in effluent regulations for wastewater treatment plants all over the world (Gouvernement du Québec (2003), US-EPA (2010) and van Rijswick and Havekes (2012)).

1.2.2 Chemical Oxygen Demand (*COD*)

Another — much quicker — way to measure water quality is by using chemical oxygen demand (*COD*). This test uses a strong oxidizer which is meant to react with the entirety of the oxidizable carbon inside the sample. The equivalent amount of oxygen needed to oxidize the sample is then obtained through the colour change of the reagent. In addition to this colourimetric method, sensors also have been developed to quantify *COD* in real time using lamps operating in the visible-to-UV spectrum. By shining different wavelengths of light into a water sample and detecting the reflected spectrum, these sensors can calculate the concentration of organic molecules, since these molecules contain double-carbon covalent bonds, which absorb specific wavelengths of light (Mesquita et al., 2017).

Regardless of the method being used, however, *COD* measurements are much quicker to take than *BOD* measurements, with a digestion time of only 2 hours with the colourimetric method, and a quasi-instantaneous reading with spectrometry. However, despite their expediency, *COD* measurements have some drawbacks, e.g., they do not include nitrogenous oxygen demand, while they do include demand from molecules which are not biodegradable and which will therefore not be broken down by biological treatment. However, by carefully choosing and refining water samples, *COD* can be used to differentiate different organic matter ‘fractions’. This feature of the *COD* test, coupled with its speed, make it an attractive indicator for influent characterization in wastewater treatment plants (Roeleveld and van Loosdrecht, 2002).

1.2.3 Short-term Biochemical Oxygen Demand (*stBOD*)

In the context of this study, another important oxygen-based water quality indicator is ‘short-term *BOD*’, or *stBOD*. This parameter shares some features with *BOD*₅ and *BOD*_u, in that all three measure only biodegradable substances in water samples. All three can also measure oxygen demand from nitrogenous compounds, though this demand can be suppressed by using a nitrification inhibitor. However, the time frame required for the measurement of these three parameters, and the methods used, are quite different. Instead of using a high S_0/X_0 ratio, *stBOD* is measured using a low ratio (usually under $\frac{1}{20}$). This means that in addition to the sample being oxidized very quickly (typically, approximately 30 minutes are needed), the expected biomass growth caused by the sample digestion is sure to be negligible, which means that almost the entire sample is used for respiration (Vanrolleghem et al., 1994).

stBOD is measured as follows: a sample is added to a respirometer (any respirometer configuration may be used, but the most common for this measurement is the ‘LFS’ type (see section 1.3)). The respirometer must contain a sample of active biomass in a state of endogenous respiration. When the water sample is added, the biomass begins to consume the readily biodegradable matter inside the sample almost immediately, which causes a disruption in the equilibrium oxygen concentration within the reactor as exogenous respiration increases. In the

short timeframe of the test, most of the slowly biodegradable molecules cannot be hydrolyzed or broken down quickly enough, meaning that only the readily biodegradable molecules are consumed. Once those molecules have been depleted, the oxygen concentration in the reactor climbs back up to the equilibrium point thanks to the reactor's continuous supply of oxygen (see Figure 1.1). One then obtains the *stBOD* value by calculating the amount of oxygen consumed between the two equilibria.

1.2.4 Uses for *stBOD*

The main advantage of *stBOD* in comparison with other biological parameters is the speed at which it is measured, which is counted in minutes rather than hours, days or even weeks for $BOD_{5/u}$. The fact that it can either include or exclude nitrogenous demand also gives the test flexibility. Most importantly, however, is the fact that by automatically filtering out the demand coming from slowly biodegradable compounds, the test shows how much substrate inside the sample is directly available to the biomass. This information is of great importance for some of the processes being used in wastewater treatment plants.

For instance, the denitrification process, which relies on heterotrophic bacteria, requires easily biodegradable substrate for rapid nitrate removal while using a small reactor (Metcalf & Eddy, 2013). Depending on the plant design being used, the carbon source used can differ. In post-anoxic configurations (in which the anoxic tank, which is where denitrification occurs, is placed downstream of the aerobic tank and their associated secondary clarifiers), external carbon sources are often employed, as most of the wastewater's organic matter is destroyed inside the aeration tank upstream of the denitrifiers. Pre-anoxic configurations, on the other hand, place the anoxic tank upstream of the aeration tank, and heavily recycle aerobic sludge to the former. In this configuration, wastewater still contains biodegradable compounds when it comes into contact with denitrifying organisms, thus reducing the need for external carbon. However, incoming readily biodegradable carbon may still not be sufficient to achieve complete denitrification. *stBOD* measurement can, therefore, help to dose appropriate amounts of external carbon for optimal denitrification in pre-anoxic plants by showing how much readily biodegradable carbon is being fed to the denitrifiers by the influent (Copp et al., 2002).

Phosphorus accumulating organisms (PAOs) also need readily biodegradable organic matter to function effectively. These organisms, as their name suggests, store phosphorus inside themselves as poly-phosphates, and the energy required to fuel this process comes from cell internal storage products. In turn, these internal storage products are created when cells assimilate fermentation products, which are a type of readily biodegradable organic matter. Therefore, the ability of PAOs to accumulate phosphorus hinges on the availability of readily biodegradable matter in the wastewater (Gujer et al., 1995). Thus, *stBOD* can help determine whether conditions are favourable for dephosphatation even though it is not a selective enough indicator

to specifically quantify fermentescible organic matter or to distinguish between carbon-related and nitrifiable nitrogen-related oxygen demand, for that matter.

The RODTOX does not differentiate between readily fermentescible organic matter and other types of readily biodegradable matter. Similarly, it also has no way to distinguish between ammonia nitrogen and other forms of nitrifiable nitrogen present in wastewater, e.g., organic nitrogen coming from dead cells' lysis products or organic molecules present in the water (e.g., proteins). These other types of nitrifiable nitrogen are hydrolyzed, ammonified and finally oxidized by the autotrophic organisms in the same way as the ammonia nitrogen present as such in the influent. Therefore, telling apart those sources of nitrogen is simply impossible using only the respirometer's DO or *OUR* measurements. However, this catch-all approach to nitrifiable nitrogen detection may prove to be an advantage for the RODTOX versus other types of nitrogen sensors (i.e., ion-sensitive electrodes), as it allows it to estimate the total nitrification load of a wastewater sample, instead of only that coming from ammonia nitrogen specifically.

1.2.5 Retrospective

It can, therefore, be said that a large gamut of parameters is available for organic matter characterization. These, as was said above, can be used either for environmental quality assays or wastewater and sludge characterization. Additionally, given their different advantages and drawbacks, some parameters are better suited than others to characterize different wastewater characteristics.

Thus, the RODTOX, like other respirometers, can help evaluate water quality inside a wastewater treatment plant by providing influent fractionation details. Additionally, since some of these respirometers contain the same sludge as the plant's activated sludge units, they may also be used to describe the activated sludge itself (Vanrolleghem and Coen, 1995). A thorough understanding of what each respirometer is capable of doing, and of the models needed to gain insights into a given phenomena, is, therefore, necessary if one is to take full advantage of respirometry within a plant setting.

1.3 Respirometer designs and their use

Over the years, a plethora of different respirometers have been designed to fulfill the needs of specific applications investigating different aspects of wastewater treatment. However, all of them share certain characteristics. For one thing, they all consist of at least one vessel in which an active biomass sample consumes oxygen and where the oxygen concentration can be tracked over time.

According to the classification proposed by Spanjers et al. (1998), respirometers can be abstracted as having two figurative ‘compartments’: the liquid phase (the biomass and wastewater samples) and the gas phase (the air or oxygen being fed to the liquid phase). This basic representation can be seen in Figure 1.3. The DO probe(s) can be placed anywhere within the system, including in the flow upstream or downstream of the gas and liquid compartments.

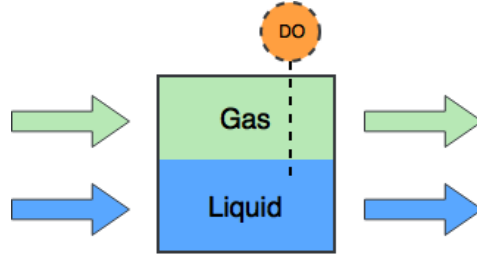


FIGURE 1.3 – Basic representation of a respirometer (adapted from Spanjers et al. (1998))

Respirometers can be broken down into eight categories using this principle. This classification is made according to:

1. Whether the oxygen concentration is measured in the gas phase (G) or the liquid phase (L).
2. Whether the gas phase is flowing (F) through the respirometer or is static inside of it (S).
3. Whether the liquid phase is flowing (F) or static (S).

The eight possible combinations of those factors are listed in Table 1.1 and examples of the implementation of the most common designs are given. The configuration used by the RODTOX is emphasized.

Depending on the design being used and the resulting position of the DO probe(s) within the respirometer, the equations needed to calculate the oxygen mass balance within the respirometer changes. These changes are done to account for liquid flow Q , gaseous flow F , changes in concentration in the liquid phase (S) and gaseous phase (C), as well as the exchange of DO between the compartments. The components of the overall DO mass balance for all respirometer types with the appropriate coefficients can be found in Table 1.2. By multiplying the mathematical equation next to every process with its corresponding stoichiometric coefficient, and then summing all those products, one obtains the complete oxygen mass balance for any of the eight types of respirometers.

Though they all have their particularities, generally speaking, any respirometer design can be used to track the oxygen uptake rate of a sludge sample. Among the different respirometer types, the ‘LSS’ and ‘LFS’ designs have proven to be the most popular. This may be explained by the fact that their mass balance equations are amongst the most concise (and thus easiest

TABLE 1.1 – List of respirometer design configurations (ROD TOX in bold)

| DO measurement site | Gas phase | Liquid phase | Examples of implementation |
|---------------------|-----------|--------------|---|
| G | F | S | Ekama et al. (1986) |
| G | F | F | — |
| G | S | S | Surmacz-Gorska et al. (1996) |
| G | S | F | — |
| L | F | S | Vanrolleghem (1994) Volskay and Grady (1990) Kong et al. (1996) Spanjers and Vanrolleghem (1995) |
| L | F | F | — |
| L | S | S | Kappeler and Gujer (1992) Gapes et al. (2003) Pratt et al. (2003) |
| L | S | F | — |

TABLE 1.2 – Oxygen mass balance components for different types of respirometers (ROD TOX in bold) (Spanjers et al., 1998)

| Process [$M T^{-1}$] | Rate equation | LSS | LFS | LSF | LFF | GSS | GFS | GSF | GFF |
|-----------------------------|---------------------------------------|-----|------------|-----|-----|-----|-----|-----|-----|
| Respiration | $V_L r_{O_2}$ | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 |
| DO accumulation | $\frac{d}{dt} (V_L S_{O_2})$ | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 |
| Liquid flow | $Q_{in} S_{O_2,in} - Q_{out} S_{O_2}$ | | | 1 | 1 | | | 1 | 1 |
| Gas exchange | $V_L K_{La} (S_{O_2}^* - S_{O_2})$ | | 1 | | 1 | 1 | 1 | 1 | 1 |
| Gaseous oxygen accumulation | $\frac{d}{dt} (V_G C_{O_2})$ | | | | | -1 | -1 | -1 | -1 |
| Gas flow | $F_{in} C_{O_2,in} - F_{out} C_{O_2}$ | | | | | | 1 | | 1 |
| Gas exchange | $V_L K_{La} (S_{O_2}^* - S_{O_2})$ | | | | | -1 | -1 | -1 | -1 |

to measure and/or calculate). Additionally, their use of a static liquid phase means that only one oxygen probe is needed to be able to calculate the oxygen mass balance, which removes all eventual difficulties involved with signals from different probes — which may be subject to drift — drifting away from one another.

In the context of a wastewater treatment plant, respirometers can be used as either *on-line* or *off-site* sensors. In the first case, they are connected directly to the system being measured, meaning that they can sample at-will from a particular point inside the system. In an *off-site* implementation, however, the respirometer is not connected to the plant and is instead operated in a lab setting, where researchers may use it to analyze wastewater from any part of the plant, or any other plant for that matter (Plana, 2015).

Regardless of the respirometer being used and its implementation style, however, the interpretation of a respirometer signal greatly depends on the theoretical framework used to understand the phenomena being studied. In a plant setting, choosing a framework with rigorous mathematical basis enables one to simulate the processes involved in substrate degradation, biomass growth and biomass decay in activated sludge. And since most of those processes rely on oxygen to function, respirometers can be a great help in describing those processes.

1.4 Overview of the Activated Sludge Models

Several models have been developed over the years to model the activated sludge process. The most widely known ones surely are those created by the International Water Association Task Group on Mathematical Modelling for Design and Operation of Biological Wastewater Treatment (Henze et al., 2000). From the task group’s efforts have arisen models which have been used in countless publications over the last 30 years. Moreover, many researchers have used ASM models as a base model to which extensions and modules could be added to either increase accuracy (e.g., accounting for the effects of pH (Sin and Vanrolleghem, 2007)), or to model additional processes. For example, ASM models can be extended to account for the biodegradation of xenobiotic molecules (Plosz et al., 2012), the production of the notorious greenhouse gas nitrous oxide (Guo and Vanrolleghem, 2014) or any other phenomena relevant to the activated sludge process.

The Task Group has released four major versions of the Activated Sludge Model, the first of which is called — appropriately — ASM1 (Henze et al., 1987). This model is meant to model carbon oxidation, nitrification and denitrification, which it does by modelling the behaviour of two organism populations inside the sludge. One of those is the heterotrophic biomass, and it takes care of carbon oxidation and denitrification, while the other is autotrophic and takes care of nitrification. A schematic view of the processes and components included in ASM1 is included in Figure 1.4.

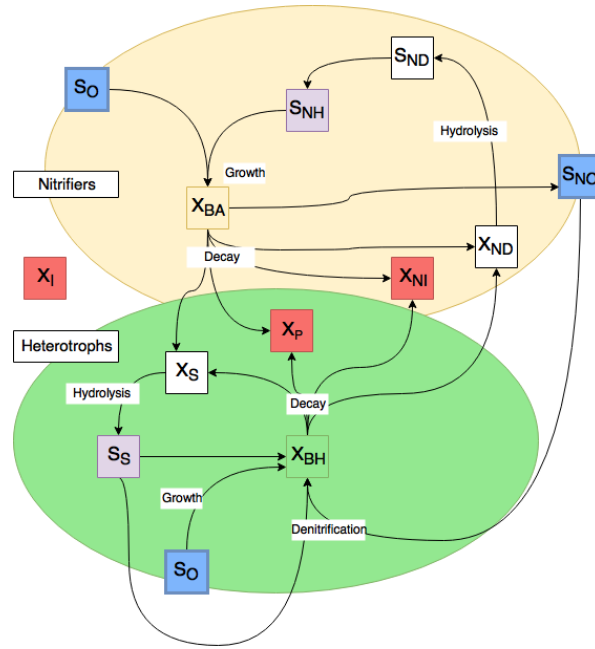


FIGURE 1.4 – Structure of the ASM1 model (adapted from Gujer et al. (1999))

In Figure 1.4, the large circles represent the autotrophic (yellow) and heterotrophic (green) biomasses. The electron acceptors are shown in blue, the readily biodegradable substrates are in purple, and the substrates which need to be broken down further before being ready for consumption by the biomass are in white. Finally, the end products of the metabolism of the organisms are shown in red.

Eight years after the release of ASM1 came ASM2 (Gujer et al., 1995), which was meant to add modelling of the removal of excess phosphorus in wastewater to the Activated Sludge Model, though the biology behind that process was not yet thoroughly understood at the time of ASM2's publication. As scientific knowledge grew in the late 1990's however, a new release of the model called ASM2d (Henze et al., 1999) was needed to reflect newfound understanding. Crucially, this version included into the model the insight that denitrifying Phosphorus Accumulating Organisms (PAOs) are partially responsible for denitrification in addition to phosphorus removal.

The ASM2d model included 21 rate equations, 19 wastewater components and an intimidating 45 kinetic and stoichiometric coefficients. The incremental nature of the changes made to ASM1 to produce ASM2 and ASM2d meant that the basic structure of those models is largely similar, e.g., all three posit the existence of a 'death-regeneration' loop in which some of the decaying biomass is turned into biodegradable substrate which is useable by sludge organisms. However, since each newer model accounts for more biological processes, more equations become necessary to describe the system fully. This increased complexity in comparison with ASM1 is, of course, beneficial in the sense that it enables the latest models to more accurately

describe sludge behaviours. However, this also leaves a lot of room for problems to arise in the calibration process, since a larger number of components to calibrate inevitably makes the calibration process longer and more experimentally tricky. A large number of components also means that the model may become unidentifiable in practice, meaning that not enough information can be extracted from the system being modelled to be able to assign a unique value to each parameter. Instead, several sets of values for the parameters may end up yielding the same modelling results, which is not ideal as it defeats the purpose of having a mechanistic model in the first place (Vanrolleghem et al., 1995).

Also, by keeping the same fundamental structure, each new model also carried forward some of the drawbacks of ASM1. Among those drawbacks, Gujer et al. (1999) list the following:

1. A shortage of nitrogen or alkalinity does not affect the heterotrophic organisms in ASM1, which is known to be the case in reality.
2. ASM1 differentiates between nitrogen fractions which are difficult to tell apart analytically, such as S_{ND} and X_{ND} .
3. The rate equations of heterotrophs and autotrophs are intertwined in ASM1 (e.g., the decay of both autotrophic and heterotrophic biomass create X_S , which is then hydrolyzed and consumed by heterotrophs). Neatly separating the different biomasses' rate equations would make it easier to characterize both accurately.
4. By having only one entry point for oxygen for each biomass, the rate of each aerobic process becomes strongly dependent on the first rate in the substrate degradation process.

In 1999, the Task Group, therefore, decided to address the main shortcomings of ASM1 — and, by the same token, those of ASM2(2d) — by completely revising the model structure, thus yielding ASM3, the most recent activated sludge model produced by the Task Group (Gujer et al., 1999). In this iteration of the model, the death-regeneration concept implemented in ASM1 was completely abandoned. Instead, ASM3 views biological activity within each biomass as a daisy-chain of reactions. In the heterotrophs, this chain begins with the hydrolysis of slowly biodegradable substrate into readily biodegradable substrate. Then, this substrate is used to create storage products within the cells of the biomass. Then these products are turned into new biomass, and finally, this biomass decays into non-biodegradable decay products. A shorter but similar chain of reactions describes the autotrophs. A schematic view of this new model structure can be seen in Figure 1.5.

This new model structure allows ASM3 to account for the oxygen consumption (or nitrate consumption, for that matter) inherent to each step of the process separately. Additionally, this change in structure reduces the importance of the rate of hydrolysis (a rate which is difficult to measure in practice) on the overall *COD* transformation rate. Instead of relying so heavily on hydrolysis to account for the transformation of slowly biodegradable compounds into readily biodegradable ones, ASM3 adds an extra step between the production of readily biodegradable

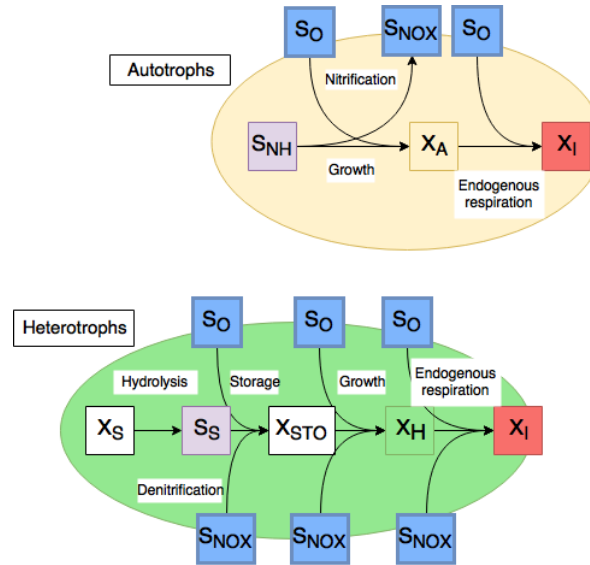


FIGURE 1.5 – Structure of the ASM3 model (adapted from Gujer et al. (1999))

substrate and its use by the biomass, which takes the form of a storage component. This helps better reflect the known process of substrate storage which occurs within PAOs and other heterotrophic organisms (Gujer et al., 1999).

Though each version of the Activated Sludge Model was meant to improve upon the preceding iteration, almost 20 years after its release, ASM3 still has not entirely replaced older ASMs in recent literature. Instead, each of those models is still being used to this day (with the possible exception of ASM2, which has been widely supplanted by ASM2d). This is explained by the assumptions and simplifications made by each model which all boast clear advantages compared to the others in any given situation. By considering those factors, one may, therefore, confidently select the model which best suits one's modelling needs.

1.5 Structure of ASM1

The goal of the Task Group that developed ASM1 was to develop a solid mathematical basis for the modelling of carbon oxidation, nitrification and denitrification. This model could then be shared among modellers worldwide, and act as a basis from which groups could either add or remove modules according to their needs. The main practical goals the Task Group set for itself was to develop a model which could accurately predict the following quantities:

1. The *COD* and *N* removal.
2. The sludge production
3. The oxygen consumption.

These goals were chosen as modelling goals because other — perhaps more intuitive — goals, e.g., predicting the effluent concentration of pollutants, tend to create modelling targets which

lie in a much narrower range of values, thus making calibration of the model more difficult and error-prone (Henze et al., 2000).

ASM1 uses COD as its basic measurement unit for organic carbon, as it has the advantage of encompassing both inert and biodegradable carbon. Additionally, COD is easy to measure both in soluble and in raw form, and it is a conserved quantity within a given system. Indeed, since COD can be considered a measurement of available electron donors in the wastewater, the COD used for biodegradation processes can be calculated according to Equation 1.5 (Copp and Dold, 1998), which makes it possible to perform a COD mass balance across the system.

$$COD_{consumed} = COD_{biomassgenerated} + \text{oxygen equivalence of electron acceptor reduced} \quad (1.5)$$

The COD fractions considered in ASM1 are laid out in Figure 1.6.

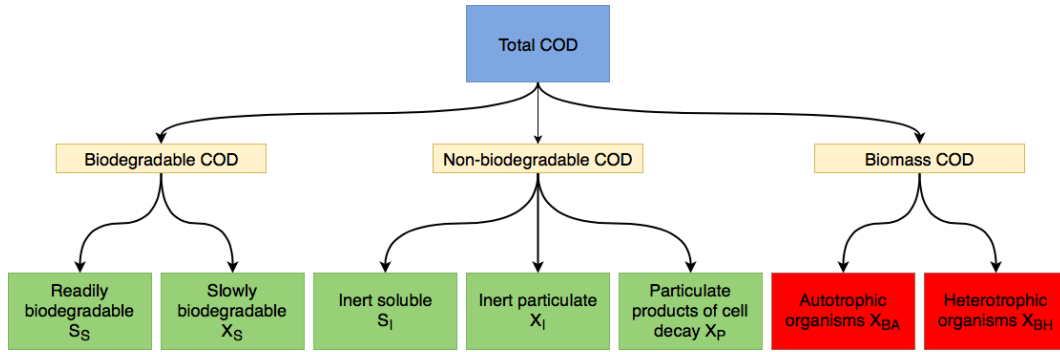


FIGURE 1.6 – COD fractions of the ASM1 model.

In wastewater treatment plant influents, the amount of biomass is usually negligible (Henze et al., 1987), which leaves only the biodegradable and non-biodegradable COD fractions. The non-biodegradable portion is split into soluble and non-soluble fractions, while on the other hand, the biodegradable portion is split by the rate at which it biodegrades. Slowly biodegradable COD consists mostly of particulate COD , and the latter takes more time to biodegrade because it must be enzymatically and physically broken down before it can be absorbed through the cell wall.

In the mixed liquor, however, biomass accounts for a significant proportion of the COD . This biomass is split into heterotrophic and autotrophic portions. Since these two organism populations participate in different processes, i.e. carbon oxidation and denitrification for heterotrophs, and nitrification for autotrophs, both are accounted for with different COD fractions, which are called X_{BH} and X_{BA} , respectively.

ASM1 also deals with the nitrogen portion of wastewater. Much like COD , nitrogen can be split into several fractions, which are separated according to their biodegradability. The

biodegradable fraction is itself made up of the ammonium and ammonia fraction (S_{NH}), the readily biodegradable organic nitrogen fraction (S_{ND}), and the particulate organic nitrogen fraction (X_{ND}). The latter fraction represents the nitrogen incorporated into the particulate *COD* fraction which is not directly usable by the nitrifiers. After hydrolysis, that fraction becomes available as readily biodegradable organic nitrogen S_{ND} . This organic nitrogen is then rapidly ammonified into S_{NH} which enables its use by the autotrophic bacteria. These then release NO_2^- and NO_3^- as a result (which correspond to S_{NO}). The portion of X_{ND} which is not turned into S_{ND} corresponds to the inert nitrogen. The inert particulate fraction X_{NI} is included in ASM1, whereas the soluble fraction S_{NI} is deemed negligibly small and thus isn't modelled at all. A breakdown of the nitrogen fractions used in ASM1 is shown in Figure 1.7.

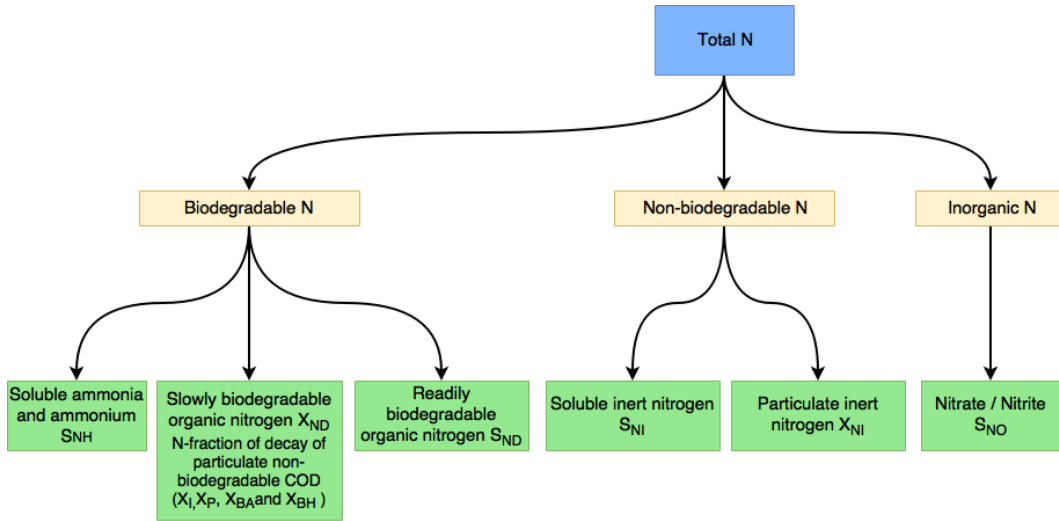


FIGURE 1.7 – Nitrogen fractions of the ASM1 model.

In all, carbon oxidation, nitrification and denitrification are described by eight differential equations and 13 stoichiometric and kinetic parameters in the ASM1 model. To understand this complex web of interactions, one can turn to graphical aids such as Figure 1.4. Another, more mathematically rigorous way to make the model legible is through the use of a ‘Gujer matrix’, which brings together the rate equations of every process, along with their coefficients (see Henze et al. (1987)).

1.6 Wastewater and biomass characterization for ASM1

1.6.1 Influent fractionation

To use a model, one must describe the system being studied in terms that the model can understand. The process of measuring influent wastewater characteristics and assigning the results to different model influent fractions is called influent fractionation. In the case of ASM1,

influent *COD* is split amongst seven recognized *COD* fractions (see Figure 1.6). Though fractionation is integral to successful modelling, no guideline for this process has been put forward or endorsed by the publishers of the ASM models themselves.

Several research groups have released characterization guidelines of their own over the years to fill this gap. One of these groups, the Dutch Foundation for Applied Water Research (STOWA in the original Dutch), has released characterization guidelines which are specifically tailored to ASM1 (STOWA, 1996; Rooleveld and van Loosdrecht, 2002). In those guidelines, STOWA only selects physicochemical methods for influent organic matter characterization, deeming respirometric methods too complicated to implement in most wastewater treatment plants where modelling would be taking place. Like Henze et al. (1987) before them, STOWA assumes that the amount of biomass in the influent is negligible. The characterization of the remaining organic matter fractions is represented in Table 1.3:

TABLE 1.3 – Organic matter fractionation for ASM1 according to STOWA (1996)

| | | |
|---------------|---|--|
| $COD_{t,inf}$ | $= S_S + X_S + S_I + X_I$ | |
| S_I | $= sCOD_{eff}$ | |
| S_S | $= sCOD_{inf} - S_I$ | |
| X_S | $= \frac{BOD_{u,inf}}{1-f_{BOD}} - S_S$ | |
| X_I | $= COD_{t,inf} - (S_S + X_S + S_I)$ | |

The brackets on the right of the figure in Table 1.3 show which *COD* fractions are included in each lab measurement, while the equations on the left show how to calculate the correct fraction concentrations using those lab measurements. Most of those measurements are already in terms of *COD*, except for *BCOD*, which is calculated from an influent *BOD_u* measurement, and then converted back into *COD* using Equation 1.6, where *f_{BOD}* is a correction factor accounting for biomass growth and decay during the *BOD_u* test (Rooleveld and van Loosdrecht, 2002).

$$BCOD = \frac{BOD_u}{1 - f_{BOD}} \quad (1.6)$$

The characterization technique put forward by STOWA for nitrogen is less exhaustive than for organic matter. Among the physicochemical tests available to researchers to characterize nitrogen, one finds some for the measurement of ammonia ($N-NH_4^+$, soluble), nitrate ($NO_3^- - N$, soluble), nitrite ($NO_2^- - N$, soluble), total nitrogen (N_t , soluble and particulate), and Kjeldahl nitrogen (TKN , soluble and particulate). One may also measure the Kjeldahl nitrogen of the soluble portion of the water (SKN , soluble) by filtering the sample before digestion.

However, some of the nitrogen fractions included in ASM1, such as S_{NI} and X_{NI} , are not analytically differentiable through physicochemical methods. Instead, they are determined by applying a factor i to the analogous organic matter fraction concentrations of the water being studied (Vanrolleghem et al., 2003). The resulting fractionation procedure can be seen in Table 1.4.

TABLE 1.4 – Nitrogen fractionation for ASM1 according to Vanrolleghem et al. (2003)

| | | |
|-------------|---|--|
| N_{total} | $= S_{NH} + S_{NI} + S_{ND} + X_{ND} + X_{NI} + S_{NO}$ | |
| TKN | $= S_{NH} + S_{NI} + S_{ND} + X_{ND} + X_{NI}$ | |
| SKN | $= sCOD/COD_t TKN$, or TKN of soluble fraction | |
| S_{NO} | $= NO_2^- - N + NO_3^- - N$ | |
| S_{NH} | $= NH_4^+ - N$ | |
| S_{NI} | $= i_{NSI} S_I$ | |
| X_{NI} | $= i_{NXI} X_I$ | |
| S_{ND} | $= SKN - S_{NI} - S_{NH}$ | |
| X_{ND} | $= TKN - X_{NI} - SKN$ | |

1.6.2 Respirometric tests for influent fractionation

Several protocols have been developed for influent characterization. Some of those protocols, such as the one proposed by Hulsbeek et al. (2002)), opt out entirely of using respirometry — respirometers are, after all, not available in most wastewater treatment plants to this day — while respirometric techniques are at the heart of some other protocols (see Vanrolleghem et al. (2003)).

For example, a respirometric method for the determination of the COD fraction S_S developed by Ekama et al. (1986) consists of feeding wastewater to a pilot-scale activated sludge reactor, complete with a return line and secondary clarifier, in a square-wave on/off pattern while monitoring the oxygen uptake rate OUR of the sludge. At the end of each feeding period, the OUR decreases almost immediately in a step-like manner (see example in Figure 1.8). This decrease in OUR is due to the depletion of rapidly biodegradable substrate S_S , while the remaining OUR is due to the slowly biodegradable substrates still being broken down (albeit at a slower pace than the rapidly biodegradable substrate), and to the endogenous respiration occurring in the reactor. The concentration of S_S can thus be calculated with Equation 1.7:

$$S_S = \frac{\Delta OUR \cdot V_p}{Q \cdot (1 - f_{cv} Y_H)} \quad (1.7)$$

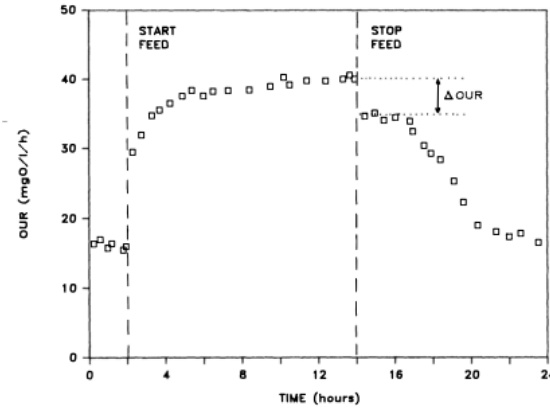


FIGURE 1.8 – Step-wise decrease in OUR due to S_S depletion (Ekama et al., 1986)

Where V_p is the reactor volume, Q is the influent flow rate, Y_H is the heterotrophic biomass yield and finally f_{cv} is the $^{COD}/_{VSS}$ ratio of the activated sludge. This method is advantageous because it doesn't require prior knowledge of $\hat{\mu}_H$. Besides, the typical value of Y_H is well-known and doesn't change much across different systems (Henze et al., 1987). However, the apparatus needed to carry out this test is relatively complex (pilot-scale activated sludge reactor, return line and secondary clarifier), and the required experiment duration is somewhat long (square feed waves with a 12–24 h period, over several days). Several other ways of quantifying S_S , as well as X_S , via respirometric measurements have been reviewed by Henze (1992) (see Table 1).

Spanjers and Vanrolleghem (1995) have also developed methods for influent fractionation involving respirometers which determined the concentration of the S_S , X_R (defined the authors as representing the concentration of rapidly hydrolyzable particulate matter in wastewater), S_{NH} , S_{ND} and X_{ND} fractions of wastewater, in combination with sludge bio-kinetic coefficients. These methods rely on batch respirometers operating at S_0/X_0 ratios of either $1/200$ or $1/20$. These low ratios allow for negligible biomass growth and decay, which reduces the influence of those processes on the interpretation of the respirograms.

Finally, the respirometric methods which are of greatest interest in the context of this M. Sc. study are those which specifically rely on the use of a RODTOX respirometer (Vanrolleghem, 1994). The most obvious use of this respirometer is the determination of the *stBOD* of influent wastewater. This measurement may be done in many different ways. As can be seen in Figure 1.9, a respirogram produced by the addition of any substrate can be characterized by several characteristics. The first of those features is the maximum peak slope PS , which typically occurs immediately after the substrate addition. The peak height PH , which is the maximum span between the DO concentration and the baseline DO concentration, can also be used. Finally, the respirogram can be described by its peak area PA , which is the

sum of the difference between the baseline and recorded DO throughout the peak. The latter respirogram characteristic can be formally defined using Equation 1.8, where C_e is the baseline DO concentration.

$$PA = \int_{t_0}^{t_f} (C_e - C(t)) dt \quad (1.8)$$

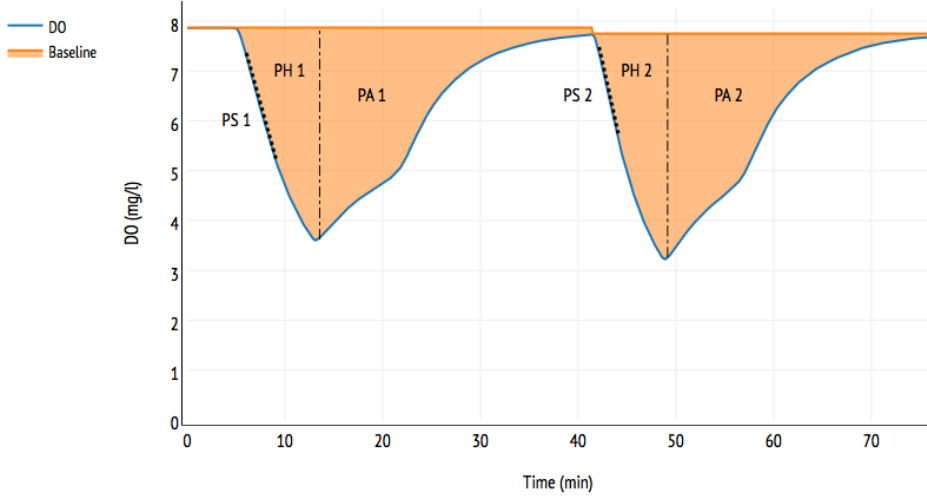


FIGURE 1.9 – Characterization of RODTOX respirogram curves

To obtain a *stBOD* measurement, one can either directly measure the physical characteristics of a sample's respirogram, or compare its characteristic features to those of a 'calibration' respirogram, which is produced by the degradation of a solution for which *stBOD* is known *a priori*. The latter case can be described by Equation 1.9.

$$stBOD_{sample} = \frac{P_{sample}}{P_{calibration}} \frac{stBOD_{calibration} \cdot V_{calibration}}{V_{sample}} \quad (1.9)$$

Here, P is either PA or PH . This method is very convenient, as the only thing needed to calculate a sample's *stBOD* is the calibration substrate's theoretical *stBOD*. Thus, no stoichiometric coefficient pertaining to the biomass is needed, nor is any physical characteristic of the system (e.g., the mass transfer coefficient K_La).

Tangentially, one may note that the very same respirogram used for *stBOD* measurement can yield information on the toxicity of wastewater. Indeed, one may determine the toxicity of a water sample by comparing the slope PS of calibration peaks before and after the addition of a potentially toxic sample to the measurement tank. A decrease in PS indicates that the maximum growth rate of the organisms in the RODTOX sludge has decreased, which means

they are inhibited by a toxicant. Another effect of toxicant doses is an increase in equilibrium DO concentration as the endogenous respiration rate of the sludge decreases.

Digressions aside, one does not have to rely on comparisons with the calibration peaks to estimate *stBOD*. The first step to achieve this is to realize that the oxygen mass balance inside the RODTOX can be calculated using Equation 1.10.

$$\frac{dC}{dt} = K_L a (C_s - C(t)) - (OUR_{end} + OUR_{ex}) \quad (1.10)$$

The endogenous respiration OUR_{end} is assumed constant over the span of the test. This assumption is justified as the test is relatively short (typically 20–40 minutes) and little biomass decay or growth is expected to occur during it. This constant OUR_{end} means that the equilibrium DO concentration C_e can be calculated using Equation 1.11.

$$OUR_{end} = K_L a (C_s - C_e) \quad (1.11)$$

Substituting the OUR_{end} expression from Equation 1.11 into Equation 1.10, Equation 1.12 is obtained.

$$\frac{dC}{dt} = K_L a (C_e - C(t)) - OUR_{ex} \quad (1.12)$$

By integrating Equation 1.12 over the span of the respirogram, one obtains Equation 1.13.

$$C(t_f) - C(t_0) = \int_{t_0}^{t_f} K_L a (C_e - C(t)) dt - \int_{t_0}^{t_f} OUR_{ex} dt \quad (1.13)$$

And since the rightmost term of Equation 1.13 corresponds exactly to the formal definition of *stBOD* (Vanrolleghem et al., 1994), one finds that:

$$stBOD = K_L a \int_{t_0}^{t_f} (C_e - C(t)) dt - (C(t_f) - C(t_0)) \quad (1.14)$$

Assuming that the respirogram begins and ends when $C(t) = C_e$ and substituting Equation 1.8 into Equation 1.14, one finally finds that:

$$stBOD = K_L a PA \quad (1.15)$$

Thus, one is able to calculate *stBOD* by measuring the area of a respirogram, and multiplying the result by the mass transfer coefficient of the RODTOX reactor (Vanrolleghem et al., 1994). This method of *stBOD* measurement is interesting, as it doesn't assume that the biomass reacts the same way to wastewater samples as it does to calibration substrates. However, this method does require the estimation of $K_L a$. Usually, $K_L a$ is measured in clean water (American Society of Civil Engineers, 2006), as this results in the respiration rates figuring in 1.10 being null. However, one can still estimate $K_L a$ in waters containing active biomass using Equation 1.16.

$$C(t) = C(t_0) - C_e e^{-K_L a (t-t_0)} \quad (1.16)$$

This is done by fitting Equation 1.16 to a portion of the respirogram where:

1. The DO equilibrium is disturbed.
2. Only endogenous respiration is taking place ($OUR_{ex} = 0$).
3. The endogenous oxygen uptake rate is constant throughout the reaeration.

Conveniently, these conditions are met in the tail portion of every respirometric test (with some caveats; see below). This is because there always is a point in a respirogram when all the sample's readily biodegradable substrate has been consumed. Therefore, from that point on and up until the equilibrium is reached again, the reaeration curve is mostly only affected by the mass transfer rate and by the background endogenous respiration rate. The tricky part of this K_La estimation method, however, is the detection of the point in the respirogram where exogenous respiration ends. Since any exogenous respiration modifies the reaeration rate away from the model proposed by Equation 1.16, the search for this endogenous-only region can be included into the curve-fitting methods used to estimate the parameters of Equation 1.16 (Vanrolleghem, 1994).

It is important to note, however, that this technique is only valid if the oxygen uptake rate caused by the biodegradation of slowly biodegradable substrate is negligible (or, at the very least, constant, as it can then be lumped into OUR_{end}). This is assumed to be the case in typical respirogram reaeration curve segments because — as the name says — the *slow* biodegradation kinetics of slowly biodegradable substrates means that very little of it can be hydrolyzed and consumed over the span of a single respirogram, thus making its influence on reaeration small.

Thus, given that the contribution of slowly biodegradable substrates to respirograms is thought to be negligible, one can attribute *stBOD* to only the rapidly biodegradable fractions of wastewater. In terms of ASM1 fractions, this means that one may estimate the rapidly biodegradable carbon S_S and rapidly nitrifiable nitrogen $N_{nitr.}$ (which corresponds to the sum of the S_{NH} and S_{ND} fractions) in the following way (Vanrolleghem, 1994):

$$stBOD = stBOD_C + stBOD_N \quad (1.17)$$

$$S_S = \frac{stBOD_C}{1 - Y_H} \quad (1.18)$$

$$N_{nitr.} = \frac{stBOD_N}{1 - Y_A} \quad (1.19)$$

To extract wastewater fractions from *stBOD* then, one needs to know the yield Y of each biomass population (heterotrophic and autotrophic). One may use other respirometric methods to estimate those coefficients (see Section 1.6.3), though average values for these also are easily found in literature ($Y_A = 0.24$ and $Y_H = 0.67$ in typical wastewater (Henze et al., 1987)). However, it is relevant to note that Y_H is not necessarily the same for every substrate. This is especially important in the case of acetate oxidation, as acetate is often used as calibration

substrate for the RODTOX, and it has a heterotrophic yield coefficient of 0.71 (or even higher if storage occurs within the biomass)(Majone et al., 1999).

Beyond the selection of appropriate yield coefficients, however, the main complication one encounters when trying to estimate S_S and S_{NH} is that the oxidation of both substrates takes place at the same time within the respirometer. This synchronicity makes it difficult to tell apart the components of the oxygen uptake rate attributable to either carbon or nitrogen degradation. To counter this, one could merely assess each fraction separately by performing *stBOD* measurements without, and then with, a nitrification inhibitor to isolate *stBOD_C* and *stBOD_N*. However, proceeding in this manner would require a sludge change after each nitrification inhibitor addition, as there is no way to remove the inhibitor from the sludge once it has been added.

A more interesting method than sequential estimation of S_S and $N_{nitr.}$, however, is the prospect of determining both fractions at the same time through the use of dynamic modelling. Thus, by simulating the oxygen uptake rate created by the biodegradation of different amounts of S_S and $N_{nitr.}$, one can determine which concentrations best recreate the oxygen uptake rate which was measured experimentally within the RODTOX. Though this technique requires few experimental manipulations, it still requires one also to characterize the sludge used by the RODTOX to perform its in-sensor experiments (which typically is sludge coming from the plant’s activated sludge units). In-sensor experiments describe experiments aiming to characterize a treatment process which take place outside the system itself. Instead, these experiments are carried out within a ‘hardware simulation’ of the system, of which the RODTOX’s built-in reactor is an example. By using an external reactor instead of the system itself to perform experiments, one has more freedom to disturb the studied system, which enables more detailed characterizations (Vanrolleghem, 1993).

1.6.3 Respirometric methods for sludge characterization

A large number of techniques have been developed to characterize the composition of activated sludge as well as the value of the modelling coefficients needed to describe its behaviour. The determination of those characteristics is crucial if one is to model the sludge’s reaction to different substrates reliably. As previously established, the current project aims to use ASM1 for its modelling needs. Therefore, this review will focus on the determination of the relevant ASM1 biomass characteristics.

Before using a model, one must select values for each of its parameters. In many cases, the default values are more than serviceable, and spending time in the lab to measure an experimental value for them is not worth the trouble regarding the impact it has on the modelling goal. Indeed, not every coefficient greatly affects the outcome of the processes being investigated by the modeller, and in many cases, the lab experiments required to determine a

coefficient value is so different from typical activated sludge unit conditions, that the obtained lab value does not reflect the actual behaviour of the sludge at all (Petersen, 2000). Therefore, before using the RODTOX to model the degradation of rapidly biodegradable substrates, one must answer the following questions:

1. Which ASM1 coefficients actually can be determined via respirometry?
2. Which, among those coefficients, can be determined using the RODTOX respirometer specifically?
3. Which of those have a significant impact on the modelling of the degradation of S_S and S_{NH} ?
4. What is the composition of the sludge inside the RODTOX?

A review of the available literature allows one to answer the first two questions. Table 1.5, which lists all ASM1 components, shows whether any component can be evaluated through respirometry and whether that evaluation can be performed using a respirometer similar to a RODTOX (that is, a batch respirometer with a low S_0/X_0 ratio).

As can be seen in Table 1.5, most ASM1 parameters can be estimated using a RODTOX respirometer in some way or another.

However, during its normal duty as an on-line sensor for a wastewater plant influent, the RODTOX cannot accommodate every experiment needed to evaluate those parameters. In that particular setting, the RODTOX exhibits a S_0/X_0 ratio of 1/20, and is exposed to calibration substrates with known S_S and S_{NH} concentrations, as well as to wastewater doses of unknown composition. By selecting an optimal calibration substrate (Vanrolleghem and Van Daele, 1994), one may use the respirograms generated by the calibration substrates to characterize some sludge parameters. Then, the obtained coefficients may be used in combination with wastewater samples to estimate the fraction of S_S and S_{NH} in the plant influent. A sensitivity analysis may be helpful in determining which parameters should be estimated using the calibration peaks, and which are best left at their default value (De Pauw et al., 2004). Alternatively, a careful analysis of the model structure could be used for the same purpose.

Some parameters which have a direct impact on S_S and S_{NH} degradation are clear targets for parameter estimation. Those include $\hat{\mu}_H$, $\hat{\mu}_A$, K_S and K_{NH} . Because of the short timespan of respirograms, the decay coefficients b_H and b_A become less important, as very little time is available for the biomass to decay during the span of a respirogram. Y_H and Y_A on the other hand, are left at their default value because according to Table 1.5, the value of the yield of each biomass is linked to both the value of the half-saturation concentration and maximum growth rate. Thus, trying to estimate the yield of each biomass would not be productive, as

TABLE 1.5 – ASM1 coefficients and the availability of RODTOX-compatible methods for their determination (adapted from Petersen (2000) and Henze et al. (1987))

| Coefficient | Unit | Typical value (20°C) (Henze et al., 1987) | Estimated through respirometry? | Estimated with RODTOX? | Identifiable combination | Reference |
|--------------------------------------|---|---|------------------------------------|---------------------------|---|----------------------------------|
| <i>Kinetic parameters</i> | | | | | | |
| $\hat{\mu}_H$ | d^{-1} | 6.0 | ✓ | ✓ | $\hat{\mu}_H X_{BH}^{(1-Y_H)}/Y_H$ | Spanjers and Vanrolleghem (1995) |
| K_S | $g\ COD\ m^{-3}$ | 20.0 | ✓ | ✓ | $(1-Y_H)K_S$ | Spanjers and Vanrolleghem (1995) |
| K_{OH} | $g\ O_2\ m^{-3}$ | 0.20 | ✓ | ✓ | K_{OH} | Kappeler and Gujer (1992) |
| K_{NO} | $g\ N-NO_3\ m^{-3}$ | 0.50 | ✗ | ✗ | — | Petersen (2000) |
| b_H | d^{-1} | 0.62 | ✓ | ✓ | b_H | Vanrolleghem et al. (2003) |
| k_h | $g\ COD_{slow\ biod.}$ $(g\ COD_{cell}\ d)^{-1}$ | 3.0 | ✓ | ✓ | k_h | Spanjers and Vanrolleghem (1995) |
| K_X | $g\ COD_{slow\ biod.}$ $(g\ COD_{cell})^{-1}$ | 0.03 | ✓ | ✗ | $(1-Y_H)K_X$ | Ekama et al. (1986) |
| η_g | — | 0.8 | ✓ | ✗ | $\frac{Y_{HD}^C(1-Y_H^C)(CER-CER_{endo})_{anox}}{Y_H^C(1-Y_{HD}^C)(CER-CER_{endo})_{aero}}$ | Sperandio et al. (1999) |
| η_h | — | 0.4 | ✗ | ✗ | — | Petersen (2000) |
| $\hat{\mu}_A$ | d^{-1} | 0.80 | ✓ | ✓ | $\hat{\mu}_A X_{BA}^{(4.57-Y_A)}/Y_A$ | Spanjers and Vanrolleghem (1995) |
| b_A | d^{-1} | 0.132 | ✓ | ✓ | b_A | Vanrolleghem et al. (2003) |
| K_{NH} | $g\ N-NH_3\ m^{-3}$ | 1.0 | ✓ | ✓ | $(1-Y_A)K_{NH}$ | Spanjers and Vanrolleghem (1995) |
| K_{OA} | $g\ O_2\ m^{-3}$ | 0.4 | ✓ | ✓ | K_{OA} | Kappeler and Gujer (1992) |
| k_a | $m^3\ (g\ COD\ d)^{-1}$ | 0.08 | ✓ | ✓ | — | Petersen (2000) |
| <i>Stoichiometric parameters</i> | | | | | | |
| Y_H | $g\ COD_{cellformed}$ $(g\ COD_{oxidized})^{-1}$ | 0.67 | ✓ | ✓ | Y_H | Brands et al. (1994) |
| Y_A | $g\ COD_{cellformed}$ $(g\ N_{oxidized})^{-1}$ | 0.24 | ✓ | ✓ | Y_A | Spanjers and Vanrolleghem (1995) |
| f_P | — | 0.08 | ✓ | ✓ | $f_P b_H$ | Keesman et al. (1998) |
| i_{XB} | $g\ N$ $(g\ COD_{biomass})^{-1}$ | 0.086 | ✓ | ✓ | $\frac{Y_H}{4.57-Y_A} \frac{\Delta COD^{Degraded}}{\Delta \int r_{O_2,exo}^N(t) dt}$ | Petersen (2000) |
| i_{XP} | $g\ N$ $(g\ COD_{cellproducts})^{-1}$ | 0.06 | ✗ | ✗ | — | Petersen (2000) |

any change in its value would directly be compensated by changes in the other two parameters (Vanrolleghem et al., 1995).

1.6.4 Biomass fractionation

As can be seen in Table 1.5, $\hat{\mu}_H$ and $\hat{\mu}_A$ are only identifiable through respirometry in combination with the concentration of X_{BH} and X_{BA} within the sludge. However, the biomass concentrations cannot be estimated using only the data coming from a RODTOX respirogram (Vanrolleghem et al., 1999). Instead, that information may be extracted from the full-scale plant according Equation 1.20 (Vanrolleghem et al., 1999) and Equation 1.21 (Dupont and Sinkjaer, 1994):

$$X_{BH} = Y_H \frac{\theta_X}{\theta_H} \frac{COD^{Degraded}}{1 + b_H \theta_H} \quad (1.20)$$

$$X_{BA} = Y_A \frac{\theta_X}{\theta_H} \frac{f^{Aerobic} N^{Nitrified}}{1 + b_A \theta_H} \quad (1.21)$$

Here, θ_X is the sludge age, θ_H is the hydraulic retention time, $COD^{Degraded}$ is the total amount of COD consumed over the plant, $f^{Aerobic}$ is the proportion of the reactor volume which is aerated and $N^{Nitrified}$ is the total amount of nitrogen nitrified over the plant.

It is good to note that in Equation 1.20, the heterotrophic decay rate b_H does not correspond to the identically-named rate figuring in ASM1. This is because the rate in Equation 1.20 does not take into account the biomass growth which takes place thanks to the biomass's decay products which effectively decreases the ‘apparent’ decay rate of the biomass — this phenomenon is referred to as the ‘death-regeneration concept’ in ASM1 (Henze et al., 1987). Instead, Equation 1.20 only considers the decay process taking place within the sludge. This b_H value can thus be recovered from ASM1's apparent heterotrophic decay coefficient using Equation 1.22.

$$b_H = \frac{b'_H}{1 - Y_H(1 - f_P)} \quad (1.22)$$

Here, b'_H is the ASM1 heterotrophic decay coefficient, and f_P is the mass fraction of the biomass which is turned into particulate products (X_P) after it has decayed. A typical value for f_P is 0.08 (Keesman et al., 1998).

The sludge age within an activated sludge unit is calculated according to Equation 1.23:

$$\theta_X = \frac{XV}{Q_w X_w} \quad (1.23)$$

Here, X is the mixed liquor solids concentration, V is the reactor volume, Q_w is the waste flow rate and X_w is the biomass concentration of the waste flow. This equation is only valid if the mixed liquor concentration is constant in time, however, which is the case when the reactor has reached a steady state (Metcalf & Eddy, 2013).

The hydraulic retention time, on the other hand, is calculated according to Equation 1.24:

$$\theta_H = \frac{V}{Q_{in}} \quad (1.24)$$

Here, Q_{in} is the influent flow rate (Metcalf & Eddy, 2013).

The X_P fraction can also be estimated from a mass balance over the plant in accordance with Equation 1.25 (Vanrolleghem et al., 1999).

$$X_P = \frac{\theta_X}{\theta_H} X_{I, \text{influent}} + f_P b_H X_{BH} \theta_X \quad (1.25)$$

Here, $X_{I, \text{influent}}$ is the amount of inert particulate *COD* in the influent (see Table 1.3). It must be noted that to calculate X_P , one must know the value of b_H (again, not to be confused with b'_H) and f_P , which can either come from respirometric experiments or from default ASM1 values (see Table 1.5).

To obtain the data needed to calculate mass balances over the activated sludge units, one must assay the wastewater plant. This can be done through a measurement campaign where one collects composite wastewater samples over a relevant period of time. Those samples can then be characterized analytically in an *off-line* fashion. For example, Hulsbeek et al. (2002) recommends measurement campaigns lasting 3–7 days for modelling exercises conducted for optimization purposes and lasting at least seven days if the model is to be used for control purposes.

As an alternative to off-line measurements, one may use data which comes from the plant's monitoring equipment. This equipment may include both *in situ* and *on-line* sensors (Plana, 2015). The choice of tools to conduct this measurement campaign will depend on the availability of the sensors and the quality of their data, as well as the overall time and resources available to conduct the assay.

Thus, with methods available for influent fractionation, biomass characterization and the careful selection of model parameters, one possesses all the tools required to use a respirometer efficiently in the setting of a wastewater treatment plant. All that is left, therefore, is to plan carefully how these tools will be put to use.

Chapter 2

Objectives

In light of the above literature review, it becomes clear that respirometry is an invaluable tool for the determination of wastewater fractions and activated sludge characteristics. Further, it is also evident that the information obtained through respirometry can be applied for ASM1-based modelling of activated sludge systems. Up until recently, the pilEAUte wastewater treatment plant, which serves as a pilot-scale, research-oriented treatment plant at Université Laval, was not equipped to perform on-line respirometry experiments, as currently are most treatment plants around the world. In the summer of 2016 however, the pilEAUte plant acquired an on-line respirometer, the Kelma RODTOX. This M. Sc. study, therefore, aims to take advantage of the pilEAUte’s newfound on-line respirometry capabilities in the following ways:

1. Successfully implementing the RODTOX respirometer inside the pilEAUte treatment plant to serve as an on-line respirometer. In the context of this study, a successful implementation is defined as:
 - Installing the RODTOX in an on-line configuration so it can analyze the pilEAUte plant’s influent wastewater.
 - Being able to automatically produce satisfyingly accurate *stBOD* measurements throughout the time periods between regular maintenance operations (e.g., sludge replacements).
 - Automatically forwarding the collected RODTOX raw data to the plant’s data acquisition infrastructure, i.e., the monEAU monitoring station and its associated database — the datEAUbase (Plana, 2015).
2. Developing computing tools to characterize the oxygen mass transfer (K_La) inside the RODTOX’s bioreactor using its on-line data.
3. Developing software and modelling tools that use the RODTOX’s on-line data to determine the rapidly biodegradable substrate (S_S) and nitrifiable nitrogen $N_{Nitr.}$ fractions in the influent of the pilEAUte’s activated sludge units.

4. Fetching the RODTOX's raw data from the plant's data acquisition infrastructure (where it had previously been stored), forwarding it to the above tools, and implementing procedures so that the information produced by those tools is also available to the *pilEAUte*'s data acquisition infrastructure.

To attain those goals, the project must unfold in three distinct phases:

1. The implementation of the RODTOX within the *pilEAUte* plant, the collection of its raw signal and its storage into the *datEAUbase*.
2. The mobilisation or development of the tools required to manipulate and interpret the RODTOX's raw signal.
3. The use of the RODTOX's data for the modelling of its activated sludge and the determination of the composition of its wastewater samples.

Chapter 3

Methodology

3.1 Study site

The study takes place at the *pilEAUte* wastewater treatment plant, so-named because it is a pilot-scale plant which treats wastewater — *eau usée* in French. The facility is installed in a laboratory located inside of the Adrien Pouliot building on the Université Laval campus. Its catchment consists of one student residence building located a few hundred meters South of the pilot hall, two daycare centres, and neighbouring surface drains, as can be seen in Figure 3.1. Water exiting the catchment is directed to a sewer in which a pumping well was set up in 2014. Shredder pumps inside the well collect a portion of the wastewater and redirect it to the *pilEAUte* plant. The shredding action of the pumps breaks large particles in the influent. Moreover, a “Y” strainer sporting a 13 mm mesh is installed upstream of the plant’s influent pump, only allowing into the plant particles small enough not to damage the plant’s corkscrew pumps .

The *pilEAUte* treatment plant was commissioned in early 2015. It consists of primary treatment followed by a biological carbon and nitrogen removal system, which includes both nitrification and denitrification. A schematic drawing of the *pilEAUte* plant’s layout is included in Figure 3.2.

As can be seen in Figure 3.2, the plant stores incoming water inside a 5 m³ tank, which is fitted with a stirrer to avoid particle decantation. From there, the water is pumped to a 2.1 m³ primary clarifier, which removes large particles from the water stream. The water then flows over a weir, directly into a measurement tank which contains an array of sensors. The sensors installed in this location are all listed in Table 3.1. The RODTOX collects water from this location for its influent characterization duties. This is because this sampling point is right upstream of the biological reactors, meaning that any biodegradable carbon or nitrogen detected in the measurement tank will be available to the biomass for treatment. The two biological units — respectively named *pilot* and *co-pilot* — are located directly downstream of

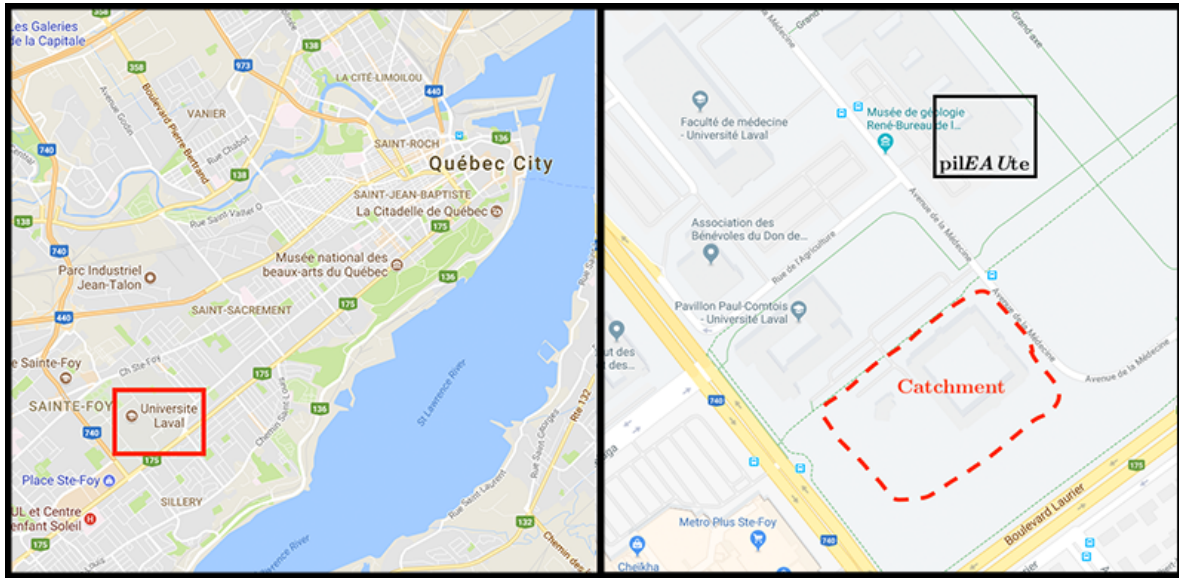


FIGURE 3.1 – Localization of the pilEAUte treatment plant and its catchment (Google Inc., 2017)

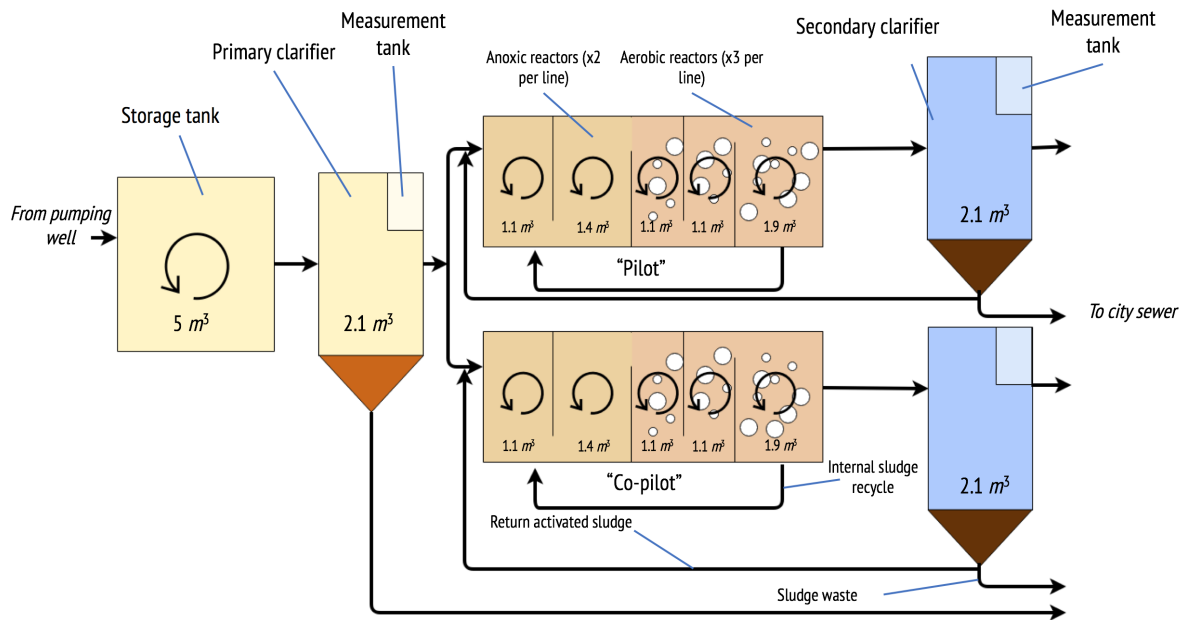


FIGURE 3.2 – Overview of pilEAUte treatment plant

the primary measurement tank, and each is equipped with its own secondary clarifier. Each biological unit functions independently from the other, which enables the study of different operational scenarios in parallel inside the plant.

TABLE 3.1 – Sensor array of the pilEAUte primary effluent

| Maker, brand | Parameter | Unit | Sensing organ |
|------------------------|--------------|------------------------|-------------------------|
| s::can, ammo::lyser | NH_4^+ | $mg\ NH_4^+-N\ l^{-1}$ | Ion-sensitive electrode |
| | K | $mg\ K\ l^{-1}$ | Ion-sensitive electrode |
| | pH | N/A | Electrode |
| | Temperature | $^{\circ}C$ | Thermocouple |
| s::can, spectro::lyser | TSS | $mg\ TSS\ l^{-1}$ | Photo-sensitive array |
| | NO_3^- | $mg\ NO_3^--N\ l^{-1}$ | |
| | COD | $mg\ COD\ l^{-1}$ | |
| | $sCOD$ | $mg\ COD\ l^{-1}$ | |
| WTW, Varion | NH_4^+ | $mg\ NH_4^+-N\ l^{-1}$ | Ion-sensitive electrode |
| | Temperature | $^{\circ}C$ | Thermocouple |
| Kelma, RODTOX | $stBOD$ | $mg\ BOD\ l^{-1}$ | Galvanic-cell DO probe |
| Hach, 3700 | Conductivity | $\mu S\ cm^{-1}$ | Inductive sensor |

As can be seen in Table 3.1, the plant’s primary effluent is fitted with a wide array of sensors to help monitor its composition. This array of sensors is integrated into the plant’s data acquisition infrastructure which consists of the plant’s SCADA (short for Supervisory Control And Data Acquisition) system, a monEAU base station, and the datEAUbase. The monEAU station is a portable monitoring station which includes several connection interfaces to accommodate a large array of water quality sensors (Rieger and Vanrolleghem, 2008). The station collects data from the sensors connected to it and then forwards this data to the datEAUbase, a relational database specifically designed to enable the storage of water quality data and all of its relevant associated metadata (Plana, 2015).

In parallel to the monEAU station, the plant is also equipped with a state-of-the-art SCADA system which enables monitoring of some basic water quality parameters (such as conductivity and temperature) as well as the control of the plant’s actuators, motors and pumps. The SCADA system also forwards its data to the datEAUbase. pilEAUte sensors may either be directly connected to the plant’s SCADA interface or to its monEAU station depending on their associated networking interface. Though the plant’s SCADA system and monEAU station are not completely integrated, plant operators may consult either data source via the SCADA workstation by using commercially-available remote access software to connect to the monEAU station’s built-in PC.

The RODTOX is the latest addition to the plant’s arsenal of sensors. It is the only on-line respirometer found at the plant, although the adjacent laboratory is also equipped with an (off-line) respirometer of the “LSS” type (see Section 1.3).

3.2 Anatomy of the RODTOX

The RODTOX is a complex machine containing several parts which are shown in Figure 3.3. The device is set up inside a 2.5m-tall cabinet which is divided into three main compartments. The “brain” of the machine is held in the top compartment. Inside this compartment, a controller (Siemens, Simatic HMI) is found, which activates and deactivates the electromechanical components of the machine, namely the actuators, the valves, the stirrer, the air supply and the dosing pumps. The controller can operate autonomously. However, the RODTOX also allows the user to interface with the controller using an embedded PC and touchscreen. The different functions found in the interface are discussed in Section 3.4. The top compartment also sports an air rotameter, which lets users set the air flow rate reaching the reactor. The controller toggles the aeration automatically. However, the air flow rate must be adjusted manually using the rotameter.

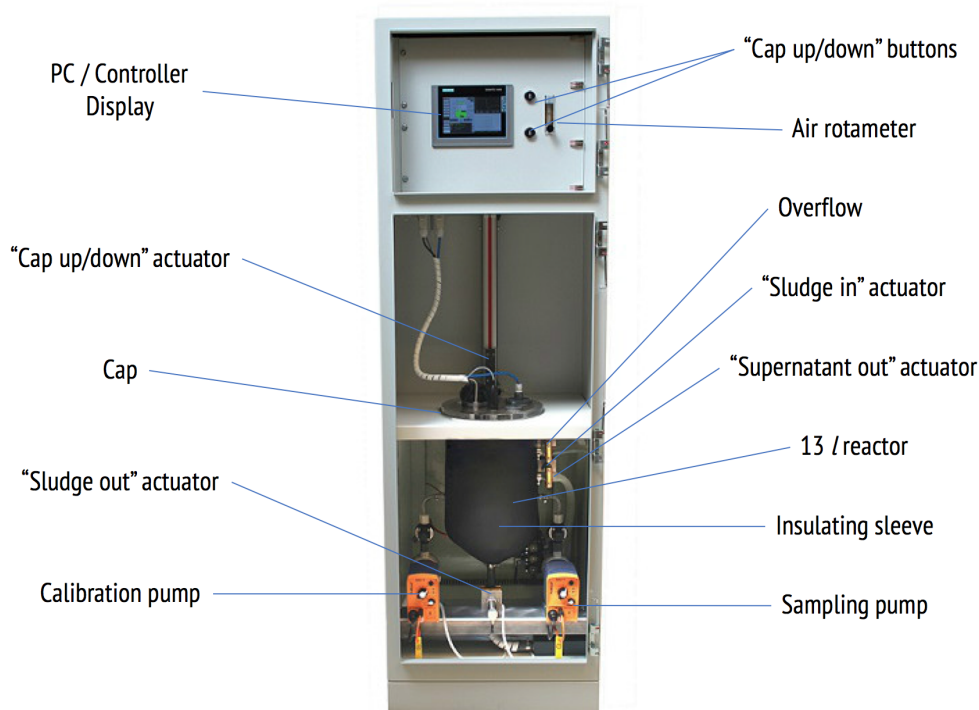


FIGURE 3.3 – Overview of the RODTOX

The lower compartment of the RODTOX cabinet contains the 13 l reactor in which wastewater oxidation is carried out. Its insides are shown in Figure 3.4. The reactor is wrapped in an insulating foam sleeve, under which heating elements are installed. By default, the RODTOX keeps the reactor’s internal temperature near 25°C, though this value is adjustable. It must be noted, however, that the RODTOX contains no cooling device. Therefore the temperature

set point must be above ambient temperature. Inside the reactor are found an electric stirrer, an aeration tube and a dissolved oxygen probe (Oxyguard, model 420).

The RODTOX's DO probe holds an anode and a cathode bathing in an electrolyte solution. A membrane allows oxygen to enter the electrolyte solution from the reactor. When oxygen reaches the anode, the latter is oxidized. For this reaction to occur, however, electrons from the cathode must flow to the anode, thus creating an electric current. By measuring the intensity of this current (which can span 4–20 *mA*), the probe can determine the amount of dissolved oxygen present in the electrolyte, and consequently, inside the reactor. On the outside of the casing, a screw allows for the tuning of the probe's resistance. Doing so modifies the slope of the probe's response curve. One may, therefore, calibrate the probe by adjusting the screw while the membrane is plunged in an oxygen-saturated volume of deionized water at a stable temperature and pressure until the theoretical DO concentration for those conditions is outputted by the probe (Oxyguard International A/S, 2017).

The middle compartment of the cabinet is left empty to allow access to the reactor when the cap is lifted for sludge addition or maintenance.

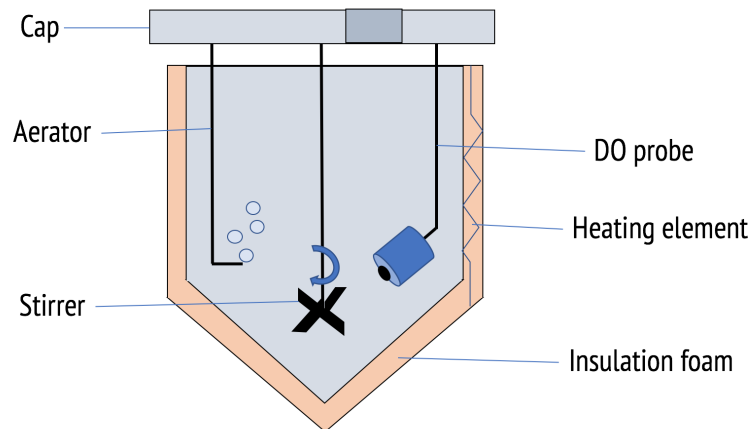


FIGURE 3.4 – Devices inside the RODTOX reactor

Dosing pumps are installed on either side of the reactor (Prominent, model beta/4a 1000). The one on the left is called the “calibration” pump, and it is responsible for feeding the reactor doses of the calibration solution, which is stored in a nearby bottle. The pump on the right, called the “measurement pump”, is the one responsible for feeding fresh wastewater samples to the reactor. The complete reactor setup is shown in Figure 3.5.

The bottom compartment of the RODTOX also contains three actuator valves. The first from the top is called the “sludge in” valve, and it may be used to feed fresh activated sludge to the RODTOX whenever it needs to be replaced. This valve, however, is not used in the context of this study, as all sludge additions are made manually. The next valve, called the

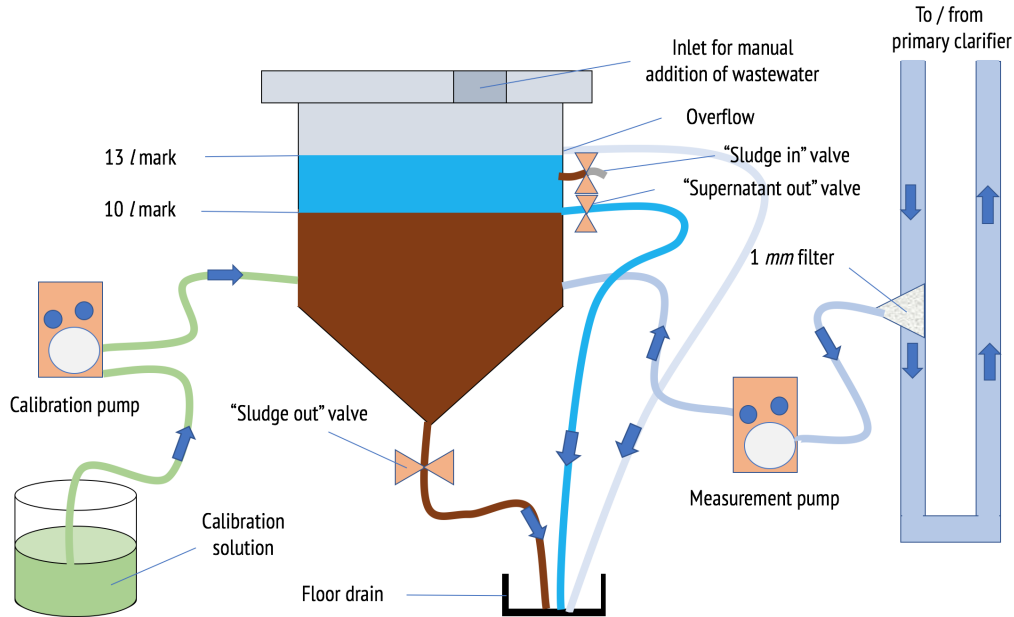


FIGURE 3.5 – Flow diagram of the RODTOX reactor

“supernatant out” valve, is responsible for draining the excess volume of water left inside the RODTOX after several sample additions. To avoid sludge loss, the RODTOX performs a 30 *min* decantation cycle before opening this valve to allow its sludge to settle below the 10 *l* mark (see Section 3.4). The final valve is the “sludge out” valve, and it releases sludge from the reactor at the end of a batch of tests, or whenever the sludge has become intoxicated or inhibited.

3.3 Fast loop

Wastewater samples have to be brought to the RODTOX for it to be able to take measurements. However, since the RODTOX’s cabinet and its sampling point in the measurement tank are a few meters apart, some additional piece of equipment must be installed to enable the easy transport of water samples to the RODTOX’s reactor. Using the built-in measurement pump would not work because its flow rate is very low, so its use would result in a long retention time within the tubing. In turn, this long retention time would allow the composition of the wastewater to change before entering the bioreactor, which must be avoided. Instead, the RODTOX is connected to the primary measurement tank with a *fast loop* system. A design sketch for this system is shown in Figure 3.6.

The *fast loop* pumps water from the measurement tank into a closed loop, the far end of which passes near the inlet of the measurement pump. A 0.5 *mm* particle filter is also fitted onto the loop’s tubing near the measurement pump to prevent damage to the latter and

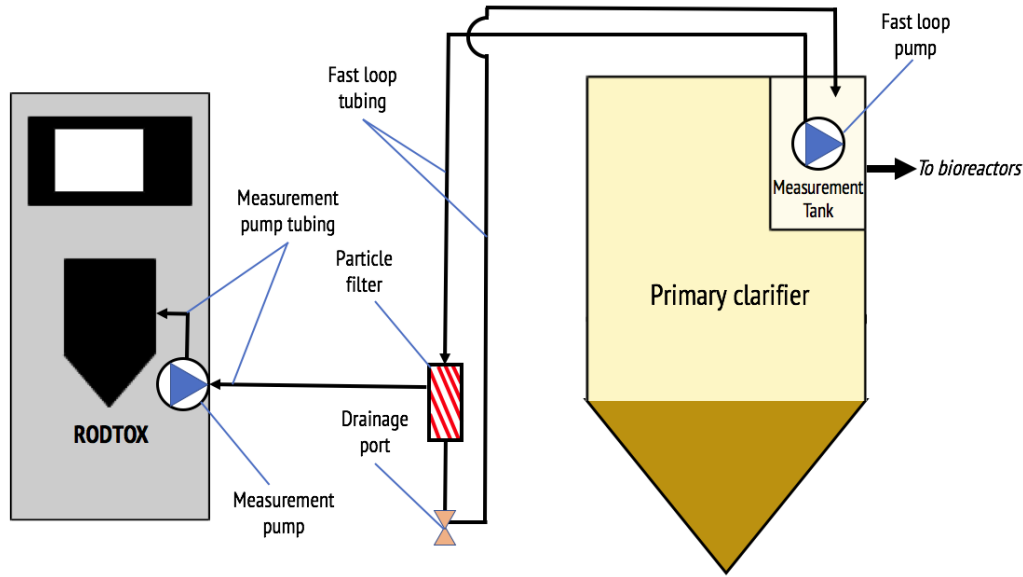


FIGURE 3.6 – Sketch of the essential components of a fast loop system

decrease chances of its narrow tubing clogging. A port located on the outside of this filter serves as the measurement pump’s sampling point. This filter should not greatly affect the RODTOX’s *stBOD* readings, as the compounds contributing to this measurement — *rbCOD* and $N - NH_4^+$ — are mostly soluble (Henze, 1992). Moreover, since the *fast loop* is located downstream of the primary clarifier, most particles large enough to be caught in the filter should already have been removed from the wastewater before entering the loop. To design an adequate *fast loop*, the following criteria were considered:

- The particle filter must be as close to the measurement pump as possible to reduce the amount of wastewater sitting in the pump’s tubing. This reduces the chances of biofilm formation in the tubing and ensures that the measured sample is representative of the measurement tank’s contents.
- The flow velocity within the loop should be sufficient to avoid decantation within the tubing. This ensures that the samples picked up from the loop have a particle size distribution similar to the water found at the sampling point.
- The flow rate within the *fast loop* should not exceed the filter’s manufacturer-recommended maximum flow rate (50 l/min)

3.4 Operation of the RODTOX

With a *fast loop* connecting it to the wastewater stream, the RODTOX can automatically take on-line measurements of the pilEAUte’s primary effluent, though its manual inlet remains available for off-line measurements as well. Before taking measurements, the RODTOX reactor

must be supplied with 10 l of activated sludge. At this point, the sludge is analyzed according to standard methods (Rice et al., 2012) to obtain its total suspended solids (*TSS*) and total volatile solids (*VSS*) concentrations. This information can later be used to calculate the RODTOX's S_0/X_0 ratio. The *TSS* concentration of the sludge may also be measured using the plant's on-line probes.

The RODTOX relies on its built-in programs to perform the correct actions throughout its operation cycle. These programs can be described as follows:

1. **Acclimatization:**

This program turns on the aeration and the stirrer and runs for 1 hour. No addition of wastewater or calibration substrate is made throughout this period. This serves to bring the sludge inside the reactor to a state of constant endogenous respiration, during which the only oxygen being consumed is used to oxidize the cells' energy reserves and decay products.

2. **Calibration 1:**

This program starts with the RODTOX calculating the slope of rising DO concentration inside its reactor. When the slope falls below a threshold value (approximately 10^{-4} mg/l/s or 0.36 mg/l/h), the current DO concentration (which is named the "baseline concentration") is recorded. Then, the RODTOX activates the calibration pump to inject a pre-set amount of calibration solution into the reactor. This produces a reference respirogram as the biomass oxidizes the substrate. The calibration solution contains the following:

- a) A known amount of a simple organic compound. In this study, acetate is used, as it is an easily biodegradable compound often used in respirometry studies (Vanrolleghem (1994), Spanjers and Vanrolleghem (1995)).
- b) A known amount of ammonium nitrogen. Ammonium chloride is used in this study because chloride does not interfere the biomass or with the DO probe inside the RODTOX, which makes ammonium chloride a rather unobtrusive method of supplying nitrogen to the reactor.

The amount of each compound inside the calibration solution is selected to create an optimal experiment (see Vanrolleghem and Van Daele (1994)).

Before adding the calibration solution, the RODTOX calculates the baseline DO concentration. Then, the solution is injected, and once the DO has reached the baseline concentration once more, the area between the respirogram DO and this baseline is calculated. The RODTOX then sets this area as being equal to the solution's theoretical *stBOD* value, which is defined by the user at any point before the test through the controller interface. The theoretical *stBOD* value of the calibration solution is calculated

in accordance with Equations 3.1 – 3.3.

$$stBOD_{th}[mg] = V_{Cal}[l](stBOD_{C,th}[mg/l] + stBOD_{N,th}[mg/l]) \quad (3.1)$$

$$stBOD_{C,th}[mg/l] = COD_{Cal}(1 - Y_H) \quad (3.2)$$

$$stBOD_{N,th}[mg/l] = NH_4^+ - N(4.57 - Y_A) \quad (3.3)$$

Here, V_{Cal} is the injected volume of calibration solution, COD_{Cal} is the chemical oxygen demand of the calibration solution, Y_H is the heterotrophic yield (typical value for acetate is 0.71 (Majone et al., 1999)), $NH_4^+ - N$ is the ammonium nitrogen concentration of the calibration solution (expressed in $g\ N/l$), and Y_A is the autotrophic yield (typical value of 0.24 $g\ COD/g\ N$ (Henze et al., 1987)).

3. Calibration 2:

This program repeats the operations occurring during the Calibration 1 and sets the area of the generated respirogram as corresponding to the theoretical $stBOD$ once more, discarding the value from the first calibration. This is done because the first respirogram may be affected by the sludge organisms being in an endogenous state before it, which may have affected its biodegradation kinetics (Vanrolleghem et al., 1998).

4. Measurement:

Measurements also begin with a baseline DO value being set. This baseline value is recorded when the RODTOX software detects that the slope of the rising DO concentration is below the threshold value mentioned above. Then, the RODTOX activates the measurement pump, which pulls wastewater out of the fast loop, through the radial filter, and feeds it to the reactor. The volume of the sample depends on the strength of the wastewater and may be set by the user through the interface at any time. The RODTOX software also automatically reduces or increases the sample volume by a factor of 2 whenever a peak's area exceeds a maximum or minimum value, respectively. Typically, the sample volume varies between 250–500 ml . The wastewater volume is set to approach the desired S_0/X_0 ratio. Using the measurement tank's on-line soluble COD and ammonium nitrogen measurements, one may estimate the amount of biodegradable substrate S_0 present in the wastewater and adjust the dosage volume accordingly. A respirogram is produced following the wastewater addition, and the $stBOD$ of the sample is calculated based on the sample's respirogram area (PA_{sample}) in relation to the calibration respirogram's area ($PA_{calibration}$) in the way described by Equation 3.4.

$$stBOD_{sample}[mg/l] = stBOD_{calibration}[mg] \frac{PA_{sample}}{PA_{calibration} V_{sample}[l]} \quad (3.4)$$

5. Toxicity check:

Toxicity is monitored continuously by the RODTOX by following the evolution of the “baseline” DO concentration over time. However, the RODTOX also regularly performs more thorough toxicity checks. The number of wastewater samples analyzed by the RODTOX before a toxicity check (let’s call that number of samples “ x ”), may be set by the user through the interface at any time. When a toxicity cycle is started, a pulse of calibration solution is added to the RODTOX, and the software compares the slope of the produced respirogram to that of the latest calibration. Toxicity is then calculated according to Equation 3.5 (Vanrolleghem et al., 1994).

$$Toxicity[\%] = \frac{PS_{latest\ calibration} - PS_{toxicity\ respirogram}}{PS_{latest\ calibration}} 100\% \quad (3.5)$$

6. Endogenous respiration:

This function is built into the RODTOX’s memory. However, it is not part of its automatic workflow (see Figure 3.7). This program turns off aeration and maintains stirring for 30 minutes. During that time, it is hypothesized that the reactor’s oxygen mass transfer coefficient is null ($K_La = 0$) and no substrate is added to the reactor, which means that no oxygen is entering the bioreactor, and the only respiration occurring within it should be endogenous. Since the biomass concentration within the reactor is constant during that time, the DO concentration is expected to decrease linearly (zeroth-order rate). The slope of the generated descending DO profile is calculated to directly obtain the OUR_{end} value using Equation 3.6.

$$OUR_{end} = -\frac{dDO(t)}{dt} \approx \frac{\Delta DO(t)}{\Delta t} \quad (3.6)$$

As was said above, the RODTOX can cycle through its functions automatically. The sequence of actions taken by the RODTOX can be seen in Figure 3.7. In it, one can see that the RODTOX starts by executing the acclimatization, calibrations 1 and 2, and a wastewater addition, in sequence. Then, after each *stBOD* measurement, the controller goes through a series of checks to determine whether it should go ahead with the next *stBOD* measurement, perform a toxicity check, or initiate a decantation cycle. Whenever the controller detects that the sludge is too intoxicated to function effectively, it stops cycling through its programs and prompts the user to tell it what to do next. The user may then either decide to replace the sludge or carry on with the one it already has.

Unless a toxic event is detected, the RODTOX can cycle through its functions indefinitely. However, the operator may also force the RODTOX to stop its current program and start another at any time through the controller interface.

Before draining the sludge, a sample of it can be taken by the user to perform *TSS* and *VSS* measurements. Coupled with the measurements of the same parameters taken when the sludge

was added to the reactor, one can follow the change in the sludge composition throughout its use in the RODTOX.

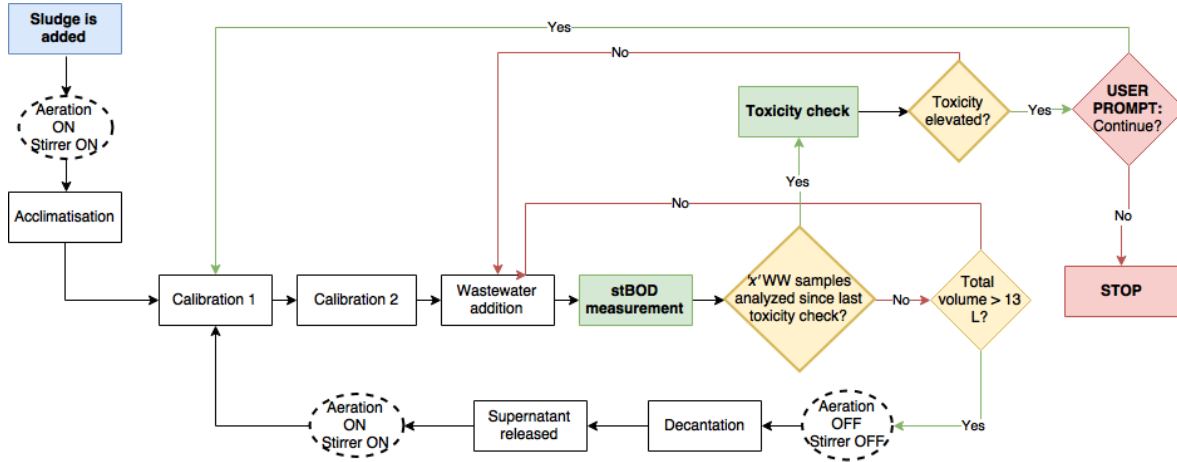


FIGURE 3.7 – RODTOX operation flowchart

3.5 Data processing

After each measurement cycle, the RODTOX generates a *stBOD* and toxicity report in a .pdf file. These reports indicate the value of each measurement along with its associated timestamp. This report, though useful for plant operators, is limited in scope and has a low temporal resolution compared to the other on-line probes sharing the same sampling point (one data point per 30 minutes on average versus one point every 5 seconds). Therefore, to truly take advantage of the RODTOX’s measurements, one must extract the “raw” DO data coming from the RODTOX’s sensors.

The DO signal recorded by the RODTOX is stored along with its timestamp in its internal memory. The temperature measured by the thermocouple, on the other hand, is not stored in memory at all, though the PLC’s programming may be modified to do so. The RODTOX’s default programming does not allow access to any raw probe signal. However, it is possible to modify it to record and store those data in a comma-separated value (.csv) text file. Whenever one wishes to access this data, one may use the RODTOX’s user interface to:

1. Save the data to a flash drive through the use of the controller’s integrated USB port.
2. Send the data to a remote computer using the controller’s built-in Ethernet port. This option does require the modification of the PLC to allow it to interface with the plant’s local area network, however.

Following this data extraction, one may choose to handle the RODTOX’s data in any number of ways. However, in the context of this M. Sc. study, the tools chosen to manipulate and interpret the RODTOX data were based on Python (Python Core Team, 2016). Python is a

programming and scripting language developed in the early 1990's to serve as a cross-platform tool for scientific computing. It has been widely adopted by the scientific community thanks to its modular nature, its legibility and its extensive collection of scientific function libraries.

More specifically, the Python-based tools used in this study are the following:

1. **Pandas:**

Pandas is a Python module used to store data into a structure called “data frame”, which is a type of indexed table (McKinney, 2011). This allows for easy and fast manipulation of data series. This is especially convenient for this project as the data output by the RODTOX is in the form of time series.

2. **SciPy and Numpy:**

SciPy and NumPy are Python function libraries specifically designed for scientific computing (van der Walt et al., 2011). They allow one to conveniently and quickly perform data-driven tasks like linear and non-linear parameter estimation, integration, differentiation, and many more mathematical operations.

3. **Plotly:**

Plotly is a multi-platform framework used to create and publish interactive figures easily. It can easily interact with data stored in structures like Python dictionaries, Numpy arrays and Pandas data frames (Plotly Technologies Inc., 2015), which makes it convenient to represent data produced throughout this study visually.

4. **Jupyter Notebooks and IPython:**

IPython is an interactive command line interface for Python (Pérez and Granger, 2007). It is built into Jupyter Notebooks, a web-based interface which allows one to write and execute Python code inside a type of web document called “notebook”. This means that code can be laid out in a clean, convenient way, and be executed in discreet “chunks” according to its location within the notebook, which is very convenient for debugging. Additionally, code laid out in this way can cohabit with headings, text, equations and figures within the same notebook. Jupyter Notebooks can, therefore, be used for the execution of code as well as for its documentation and publication (Kluyver et al., 2016).

5. **Anaconda:**

Anaconda is a cross-platform environment manager for Python. Once installed, Anaconda manages all the Python environments and libraries installed on a workstation, which allows for simple installation, update and deletion of Python libraries (Continuum Analytics, 2016). Anaconda also takes care of the installation and management of development tools like Jupyter Notebooks.

3.6 Modelling tools

The modelling platform used in this project is the WEST software suite, developed by DHI. WEST contains a user-friendly interface for the modelling of wastewater systems at any scale from reactor to watershed. This software may thus be used to model treatment plants, in which case visual icons and arrows are used to represent the different modelled parts of the wastewater treatment chain as well as the interactions between those parts. The models built into WEST's extensive library allow for the modelling of the hydraulic behaviour of a wastewater treatment plant, of the sedimentation processes occurring within it, and of the biological reactions taking place throughout.

Additionally, WEST contains a model builder which allows one to add, remove or modify model components, which enables one to adapt models easily to one's modelling goals. However, the most crucial element of the suite surely is the simulation module, which allows the simulation of the modelled systems at steady-state as well as in dynamically-changing conditions. The suite also contains tools which allow users to conveniently and easily perform common modelling-related experiments, such as parameter estimations, model calibrations and validations, and sensitivity analyzes.

WEST is used in this study to model the behaviour of the RODTOX's in-sensor experiments. After having described the sensor's activated sludge, WEST is used to estimate the concentration of S_S and S_{NH} contained in the wastewater samples that are fed to the RODTOX's bioreactor. WEST, therefore, has an important role to play in the interpretation of the respirometry data generated by the RODTOX and the attainment of this study's goals.

Chapter 4

Implementation of the RODTOX sensor

4.1 Installation of the *fast loop* system

4.1.1 Design

The *fast loop* consists of a pump and a tube which bring wastewater from the measurement tank and back. It also contains a particle filter which keeps large particles from reaching the RODTOX's measurement pump. For it to work adequately with the RODTOX, the *fast loop* has to conform to the following requirements:

- The particle filter must be positioned as close to the measurement pump as possible to reduce the amount of wastewater standing in the tubing between the filter and the measurement pump. Placing it so reduces the chances of biofilm formation in the measurement pump's tubing by reducing its length (and thus its overall surface) and ensures that the measured sample remains representative of the measurement tank's contents, as little transformations can occur during transit.
- The flow velocity within the loop should be high enough for the particles in the water not to settle inside the loop as they travel through it. This also ensures that the water sampled by the RODTOX is similar in composition to the measurement tank's water.
- The flow rate within the *fast loop* should not exceed the filter's manufacturer-recommended maximum flow rate (50 l/min).

The first physical component to be selected for the loop construction was the pump. Very little room was available to install a pump around the primary measurement tank. Therefore, it was decided to install a submersible pump directly within the tank instead. This positioning meant that the maximum width of the pump had to be under 12 inches. The Optima pump (Ebara, USA) fit into the measurement tank quite nicely, and its rated water flow rate was

more than sufficient to meet the specifications (up to 150 l/min). However, since the maximum acceptable flow rate for the particle filter is of only 50 l/min , a bypass branch was added to the loop design to handle the excess flow. This branch comes out of the main branch a few centimetres above the measurement tank and redirects the water right back into it, as can be seen in Figure 4.1.

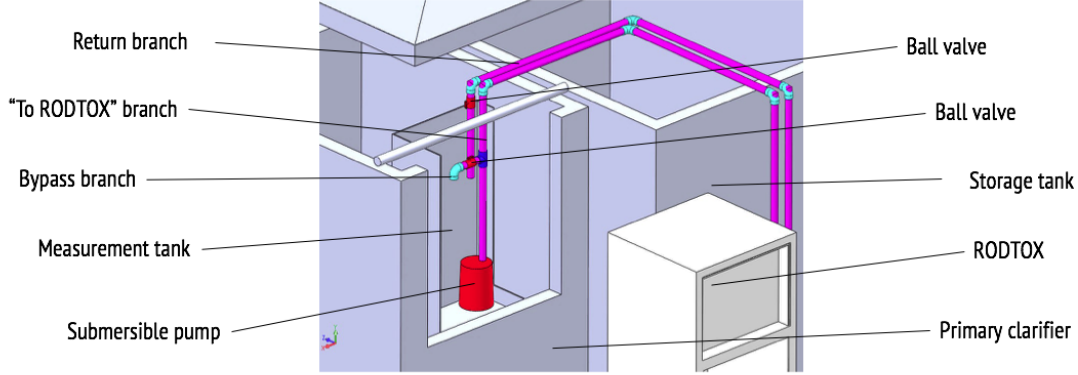


FIGURE 4.1 – Design of the *fast loop* near the measurement tank

Valves on both branches, i.e., the main RODTOX branch and the bypass branch, allow one to adjust the flow going into each and thus let one adjust the flow to the required 50 l/min in the RODTOX branch. Though no flow meter is fitted to the RODTOX branch, the flow through it may be measured by measuring the volume (V) of water exiting its outlet while measuring the time required to collect this volume with a stopwatch. The flow rate is then recovered using Equation 4.1.

$$Q = \frac{V}{\Delta t} \quad (4.1)$$

To check whether the flow within the fast loop is sufficient to avoid decantation, one may use Stoke's law to roughly estimate the “worst-case” terminal settling velocity — that is, the fastest — settling velocity one may expect inside the loop for the particles matching the particle filter's pore size (0.5 mm). This was done by assuming a particle density equal to that of sand (2.65) (Metcalf & Eddy, 2013), a water viscosity of $7 \cdot 10^{-4} \text{ kg/ms}$ — typical for water at room temperature (25°C) — and spherical particles (Vesilind et al., 2013).

$$V_t = \frac{gd_p^2(\rho_p - \rho_w)}{18\mu} \quad (4.2)$$

Using those assumptions, one finds that the maximum settling velocity one should expect within the loop is 0.32 m/s . Thus, to avoid decantation within the loop, the flow velocity in the ascending portion must be superior to this maximum settling velocity estimate. Using Equation 4.3, it was thus possible to check whether a convenient, commercially available tubing

type (schedule-40, 1-inch PVC pipes) would deliver an acceptable flow velocity when coupled with the selected pump.

$$Q = v \cdot A \quad (4.3)$$

The resulting predicted flow velocity (1.5 m/s) thus proved to be adequate, as it exceeded the estimated maximum decantation velocity fivefold. This large safety factor helps to compensate for uncertainty linked to the use of simplifying assumptions made in the assessment of this maximum settling velocity (i.e., particle density and shape may vary, and the water temperature may change within the loop).

A drainage valve was also installed at the loop's lowest point to facilitate cleaning and maintenance of the *fast loop* system. This valve is located directly under the particle filter, as shown in Figure 4.2. Thanks to the filter housing's screw-on connections, cleaning operations are easy to perform on the filter.

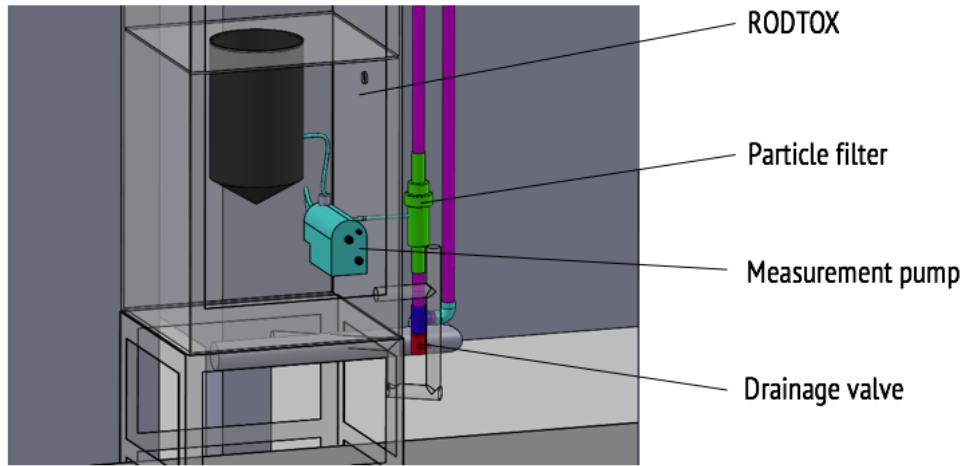


FIGURE 4.2 – Position of the *fast loop* drainage valve.

A sketch of the installed *fast loop* is shown in Figure 4.3. Before the *fast loop* was installed in the pilEAUte plant, a mixer was fitted to the measurement tank to ensure adequate mixing around the sensors. However, the addition of the submersible pump did not leave enough room for the mixer to remain in place. The addition of the bypass branch to the *fast loop* made the issue moot, however, since the turbulent flow created by the water flowing through the loop back into the tank creates plenty of mixing. The fact that the pump serves both as the *fast loop*'s water feeder and as a mixer means that the loop pump has to run at all times, instead of only when the RODTOX is operating. To reduce wear and tear on the pump, this author recommends removing it from the measurement tank and re-installing the original mixer whenever the RODTOX is out of use for an extended period.

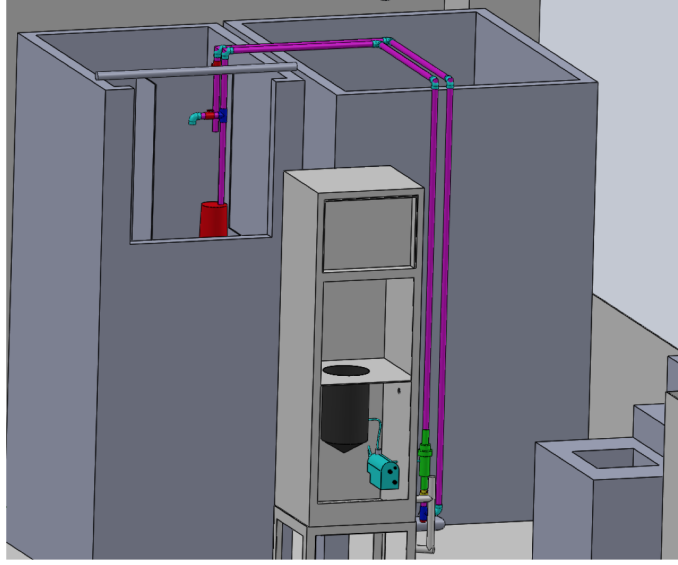


FIGURE 4.3 – Layout of pilEAUte primary treatment installations.

There also are moments during typical RODTOX operation when the pump must be shut down. These cases include maintenance operations within the measurement tank, and whenever the water level in the tank is too low. A floating switch was thus added to the pump's circuitry to automate pump shutdowns to accommodate the latter case. Another case to consider is one where the supply of fresh water to the measurement tank is stopped, which may happen whenever upstream pumps are shut down. In this particular case, the *fast loop* pump, which is usually cooled down by the water flowing into the measurement tank (typically at a flow rate of around $1.1 \text{ m}^3/\text{h}$), starts to heat up the surrounding water. This increase in temperature should be avoided. Therefore, in such cases, one may prefer to stop RODTOX measurements altogether until the usual flow is restored to the measurement tank.

Whenever the RODTOX is shut down for a long period (e.g., longer than a week), the *fast loop's* RODTOX branch may be emptied to avoid the presence of standing water inside the loop, which might cause undesirable biofilm or biogas formation. To drain the RODTOX branch, one may shut its ball valve, and open the drainage valve at the bottom of the loop. This only works, however, if the outlet of the RODTOX branch within the measurement tank is above the water surface. Otherwise, the measurement tank water will be syphoned off into the RODTOX branch.

4.2 Start-up tests of the RODTOX sensor

4.2.1 K_La estimation in clean water

Before using the RODTOX with sludge, aeration tests may be performed in clean water. This helps to gauge later aeration performance characterization, i.e., the mass transfer coefficient K_La obtained in wastewater at a given air flow rate is expected to be similar to the one obtained in clean water. Figure 4.4 shows the mass transfer coefficient of the RODTOX reactor in clean 25°C water.

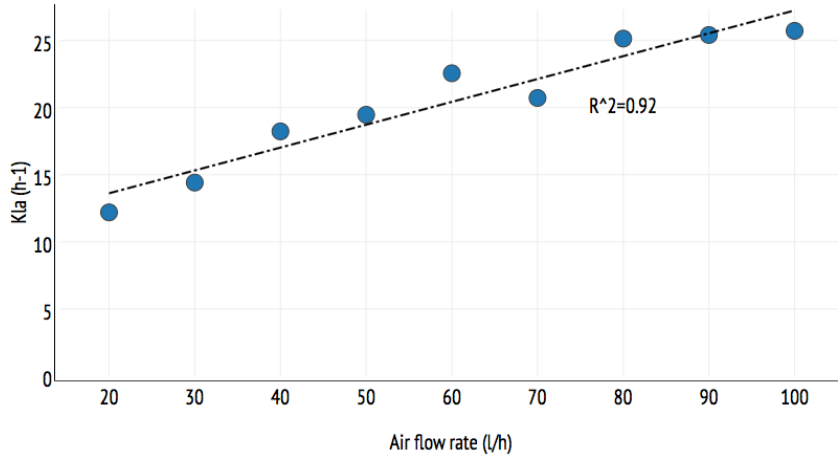


FIGURE 4.4 – Mass transfer coefficient of the RODTOX reactor in clean water at different air flow rates.

In Figure 4.4, it can be seen that an increase in air flow rate increases the mass transfer coefficient, as one expects (Metcalf & Eddy, 2013). Moreover, the increase in mass transfer as air flow rate increases can be modelled linearly with a reasonable level of confidence ($R^2 = 0.92$) in the range of air flow rates typically used during normal RODTOX operation.

4.2.2 Estimation of the oxygen probe's time constant

The dissolved oxygen (DO) probe does not respond instantaneously to changing conditions in the RODTOX's reactor, as it takes some time for the oxygen concentration on both sides of the probe's membrane to reach an equilibrium (Spanjers and Olsson, 1992). The probe signal may thus be modelled using a first-order equation such as Equation 4.4:

$$\tau \frac{dy}{dt} = -y + C \quad (4.4)$$

Here, y represents the probe signal, C is the actual DO concentration, and τ is the time constant of the probe. Thus, to determine the actual DO concentration in the reactor using the probe's signal, the value of the constant τ is needed. This constant may be estimated

rather simply. By immersing the probe inside an oxygen-free solution, waiting for the signal to reach an equilibrium, and then plunging the probe in an oxygen-saturated concentration (or vice-versa), a “step function” is created in the actual dissolved oxygen surrounding the probe. A positive step is described by Equation 4.5, while a negative step is described by Equation 4.6.

$$C(t) = \begin{cases} 0; & t < 0 \\ A; & t \geq 0 \end{cases} \quad (4.5)$$

$$C(t) = \begin{cases} A; & t < 0 \\ 0; & t \geq 0 \end{cases} \quad (4.6)$$

Therefore, the signal of the probe between the two equilibrium points is solely influenced by the probe’s dynamics, which is modelled by Equation 4.4. By solving the latter for a positive step, one obtains Equation 4.7, while one obtains Equation 4.8 for a decreasing step. In each equation, A is a constant.

$$y(t) = A (1 - e^{-t/\tau}) \quad (4.7)$$

$$y(t) = A e^{-t/\tau} \quad (4.8)$$

One may then fit Equations 4.7 and 4.8 to the probe signal generated when alternately immersing the DO probe in oxygen-free and oxygen-saturated solutions using non-linear regression algorithms to obtain values for τ . It is good to note that the value of τ may be different for ascending and descending steps (Spanjers and Olsson, 1992). Figure 4.5 shows the result of such a curve-fitting experiment performed in clean water with the RODTOX’s OxyGuard 420 probe on August 28, 2017. The probe signal between $[0 : 215 \text{ s}]$, $[440 : 630 \text{ s}]$ and $[990 : 1250 \text{ s}]$ is not used, as the solution in which the probe was immersed was not fully anoxic, or the probe was prematurely removed from the water by accident.

It can be seen in Figure 4.5 that the non-linear regression yields a reasonable, albeit not perfect, fit for each curve segment. Analyzing the results of those tests for ascending and descending steps separately, one obtains the following ($\alpha = 5\%$):

- $\tau_{ascending} = 28s \pm 13$
- $\tau_{descending} = 37s \pm 3$

Thus, seeing that $\tau_{descending}$ is not significantly different from $\tau_{ascending}$, they may be analyzed together. One then obtains that $\tau = 32 \pm 13$ ($\alpha = 5\%$). This value is therefore used to model the actual DO concentration within the RODTOX’s reactor moving forward.

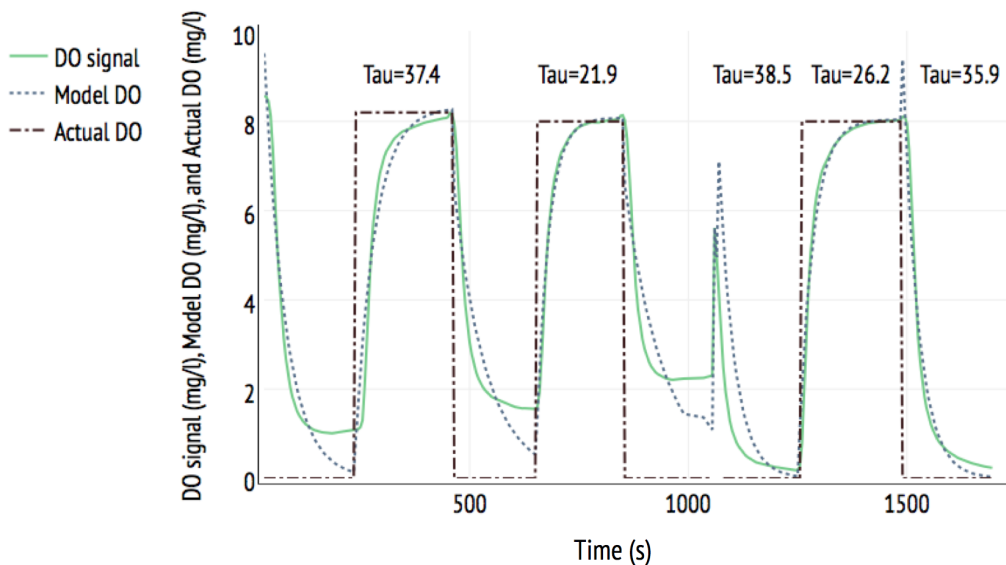


FIGURE 4.5 – Probe dynamics of the OxyGuard 420 when plunged alternately in oxygen-free and oxygen-saturated solutions (τ in seconds).

4.2.3 Calibration solution

As discussed in Section 3.4, the calibration solution must be made in such a way as to create an optimal experiment within the respirometer with each substrate addition (Vanrolleghem and Coen, 1995). That is, the calibration peak should clearly show the separate regimes of biological degradation caused by the consumption of the carbonaceous and nitrogenous substrates. Examples of such respirogram curves are shown in Figure 4.6.

The concentrations of acetate and ammonium used to obtain an optimal calibration respirogram with the pilEAUte’s RODTOX were based on those found in Vanrolleghem and Verstraete (1993), and then tweaked using trial and error. The calibration solution composition used in this study is the following:

- 20 g/l sodium acetate (15.6 g COD/l)
- 7.68 g/l ammonium chloride (2 g N/l)

However, one must keep in mind that the concentrations needed for an optimal experiment vary with the relative activity of the heterotrophic and autotrophic organisms of the sludge, and therefore that calibration solution should be modified when major changes in the sludge activity are noticed. Figure 4.7 shows what can happen when the calibration solution is maladapted to the sludge.

It can be seen in Figure 4.7 that the expected “nitrogen shoulder” does not appear in the respirogram, which means that in this case, carbon and nitrogen degradation are indistin-

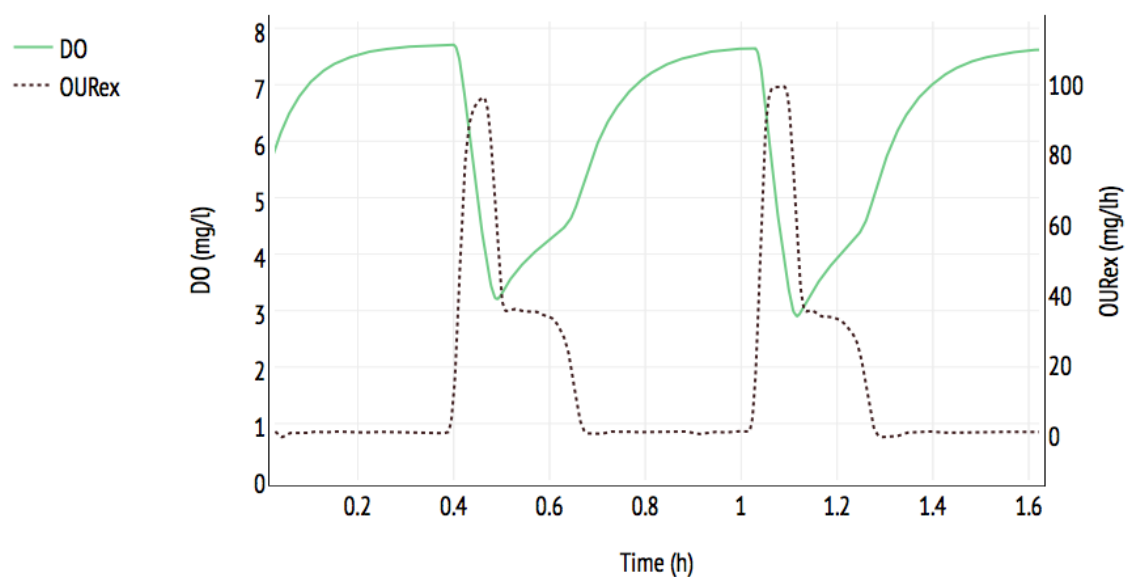


FIGURE 4.6 – Respirograms of two “optimal experiments” created by the calibration solution.

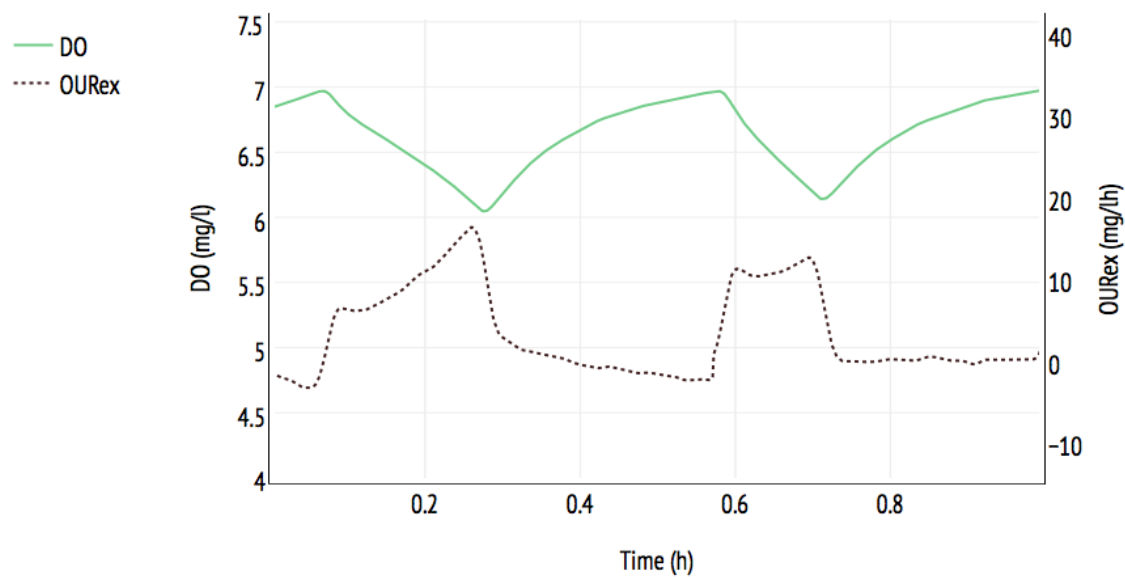


FIGURE 4.7 – Respirogram of a “non-optimal” calibration solution (left)

guishable from one another. This may cause problems at the model parameter estimation stage.

4.2.4 Sludge stability

During normal operation, the sludge inside the RODTOX receives sufficient substrate to stay alive. This means that sludge still produces *stBOD* measurements weeks after its introduction to the RODTOX reactor, as the sludge organisms continue to grow, reproduce and decay. However, this apparent longevity must be taken with a grain of salt. Indeed, because *stBOD* measurements are biological in nature, their accuracy in relation to the treatment plant depends on the similarity of both the sludge within the RODTOX's reactor and the sludge actually carrying out the biodegradation within the plant's bioreactors.

A bioreactor's *VSS* can be split between non-biodegradable solids (*nbVSS*) and active biomass (*X*) using Equation 4.9, where *f* is a *COD*-to-*VSS* conversion factor. The value used for *f* (0.86) comes from Spanjers et al. (1998) (see. p.177, where a typical denitrifying activated sludge system similar to the pilEAUte plant's is characterized).

$$VSS = Xf + nbVSS \quad (4.9)$$

The net growth of sludge organisms ($\frac{dX}{dt}$) can thus be modelled using Equation 4.10 (Metcalf & Eddy, 2013).

$$\frac{dX}{dt} = Y_H \Delta_{COD} - k_d X \quad (4.10)$$

Here, Y_H is the heterotrophic organism synthesis yield coefficient (assumed to correspond to $0.67g^{VSS}/g^{COD}$), k_d is the endogenous decay coefficient (assumed to correspond to $0.1d^{-1}$ (Metcalf & Eddy, 2013), and Δ_{COD} is the average *COD* fed to the RODTOX on a daily basis. Solving Equation 4.10 for $X(t)$ yields Equation 4.11.

$$X(t) = \frac{Y_H \Delta_{COD}}{k_d} - \left(\frac{Y_H \Delta_{COD}}{k_d} - X(t = t_0) \right) e^{-k_d t} \quad (4.11)$$

Thus, the theoretical change in *VSS* concentration within the reactor can be estimated using Equation 4.12.

$$\%VSS_{change} = \frac{X(t = t_f) - X(t = t_0)}{X(t = t_0)} \cdot 100\% \quad (4.12)$$

Table 4.1 shows the *VSS* concentration of RODTOX sludge at the time of its addition to the reactor, and at the time of its removal throughout the summer of 2017.

Table 4.1 shows that the decay of the RODTOX biomass is highly variable across samples. As an example, it can be seen that the sludge sample which was added on July 31 decayed by

TABLE 4.1 – Evolution VSS concentration within the RODTOX.

| Date sludge addition | Sludge use period (d) | VSS initial (mg/l) | VSS final (mg/l) | Actual VSS change (%) | Theoretical VSS change (%) |
|----------------------|-----------------------|--------------------|------------------|-----------------------|----------------------------|
| 2017/07/03 | 14 | 2560 | 1922 | -24.9 | -60 |
| 2017/07/17 | 14 | 3420 | 2959 | -13.5 | -60 |
| 2017/07/31 | 8 | 3110 | 2565 | -17.5 | -40 |
| 2017/08/07 | 8 | 1169 | 1240 | 6.1 | -39 |

17.5% in 8 days, while the one that was added on July 17 and used for 14 days only decayed by 13.5%. It can also be seen that two consecutive sludge additions with identical use periods end up with rather different changes in biomass concentration. Moreover, one of the samples actually bulked up during its use period within the RODTOX, unlike all the others, going against the theory, which predicted a 39% decrease in *VSS*. In fact, it can be seen that the observed changes in the *VSS* concentrations are different from the theoretical predictions for all samples. This could be due to the selected decay coefficient k_d not describing the *pilEAUte* sludge's actual decay rate.

More samples would undoubtedly have helped to extract trends in the behaviour of the RODTOX sludge during its stay in the reactor. However, as it stands, it can clearly be seen that a sizeable amount of biomass is lost when the RODTOX uses the same sludge for more than eight days. In light of this, a weekly sludge replacement would seem to ensure that the RODTOX sludge sufficiently resembles the *pilEAUte* sludge and that the sludge replacement procedure does not become excessively time-consuming.

4.2.5 pH

Several biological processes occur within the RODTOX. Among those processes, nitrification consumes not only oxygen but also alkalinity. If the alkalinity demand from nitrification is superior to what the wastewater samples add to the RODTOX, the contents of the reactor will acidify over time, which can be seen by a decrease of pH within the reactor. Left unchecked, this decrease in pH may impede biological degradation and biomass growth. Indeed, carbon degradation is best performed at pH 6.0 – 9.0, while nitrification rates are optimal at pH 7.5 – 8.0 (Metcalf & Eddy, 2013).

Though the RODTOX had been in operation since February 2017, the pH inside its reactor has only been monitored starting in July of that year. Figure 4.8 shows the evolution of pH inside the RODTOX for different sludge samples. It is plainly seen that sludge samples #1 and #2 are acidified within their first day inside the RODTOX. For the rest of their use period, the sludge samples remain acidic, but their pH never falls below 4. Since nitrification

is known to be inhibited at pH values under 6.8 (Metcalf & Eddy, 2013), the observed drop in pH likely affected the biomass's ability to degrade the samples fed to the RODTOX.

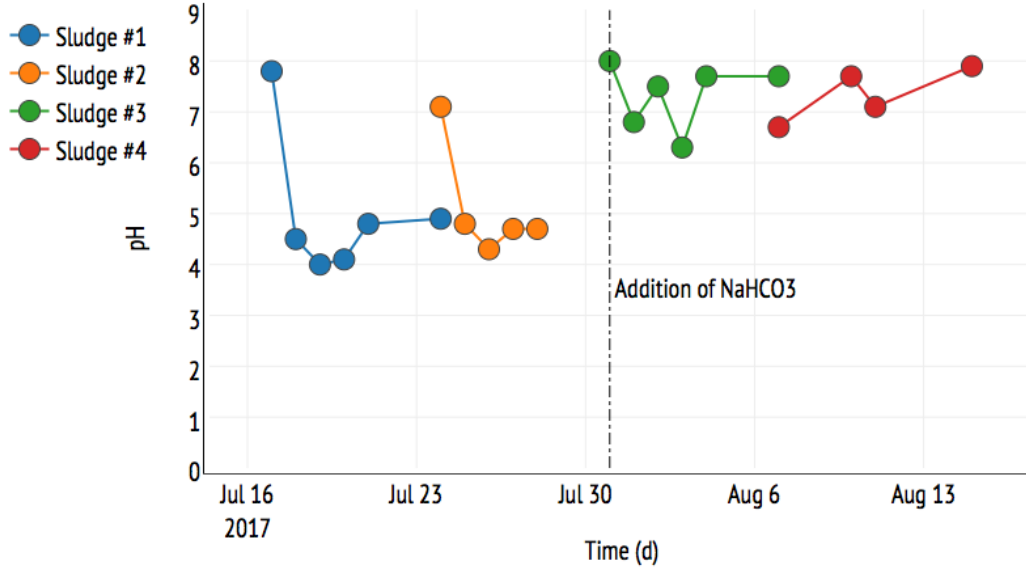
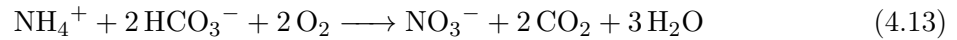


FIGURE 4.8 – Evolution of pH within the RODTOX reactor with and without sodium bicarbonate addition.

Steps were taken to re-establish acceptable pH values inside the RODTOX during standard operation as soon as the problem was noticed in late July 2017. To this end, sodium bicarbonate was added to the calibration solution. The amount of bicarbonate added was calculated according to the typical load of nitrifiable nitrogen to the RODTOX between two decantation cycles. Nitrification was used as a proxy for alkalinity consumption, as it is the biological process responsible for most alkalinity consumption in the activated sludge process (see Equation 4.13).



Thus, for each mole of ammonia nitrogen consumed, the nitrification process consumes two moles of bicarbonate ions. The required sodium bicarbonate concentration in the calibration solution can, therefore, be calculated using:

1. The ammonia nitrogen concentration in the calibration solution itself.
2. The typical ammonia concentration in the measurement tank, where the RODTOX sources its wastewater.

The first item is known from Section 4.2.3, and corresponds to 2 g N/l . Since 24 ml of calibration solution is added during each cycle (2 calibration peaks), this amounts to 45 mg N . The second item, on the other hand, could be estimated using the on-line data available for the measurement tank. Unfortunately, neither the scan analyser $\text{N} - \text{NH}_4^+$ nor the WTW

Varion installed in the primary effluent were properly calibrated during the RODTOX's operation period. Therefore, the nitrogen load was estimated using data from a measurement campaign which took place in July of 2016. During this campaign, the $N-NH_4^+$ concentration in the primary effluent was measured analytically at 1-h intervals during a 24-hour period. It was found that the ammonia nitrogen load to the plant during this period was $38 \pm 6 \text{ mg N/l}$. Adding a safety factor to this value, a 50 mg N/l was used in the alkalinity calculations.

Given that the RODTOX is set to accept 12 samples of 250 ml each during each cycle, the RODTOX can, at most, receive 12 high-ammonia samples (plus the two calibration shots) in a single cycle, which amounts to 195 mg of nitrifiable nitrogen, or 200 mg N , to round things out. Therefore, to compensate the nitrification's alkalinity consumption, the two shots of calibration solution must deliver an amount of alkalinity which can be calculated using Equation 4.14.

$$[NaHCO_3] = \frac{NH_4^{+}{}_{ww+cal} - N(g)}{14g \text{ } NH_4^{+} - N/mol} \cdot \frac{2 \text{ mol } NaHCO_3}{\text{mol } NH_4^{+} - N} \cdot \frac{84.007 \text{ g } NaHCO_3}{\text{mol } NaHCO_3} \cdot \frac{1}{2 \cdot 0.012 \text{ l}} \quad (4.14)$$

In this case, the sodium bicarbonate concentration needed in the calibration solution was calculated to be 25 g/l . The efficacy of this pH control strategy can be judged using the pH measurements performed on sludge samples #3 and #4, which can be seen in Figure 4.8. Those show that sludge samples which are regularly fed sodium bicarbonate *via* the calibration solution maintain a pH between 6 and 8. While this is better than what is observed without sodium bicarbonate, such a pH oscillation is still large enough to negatively affect the degradation rates within the reactor, especially for nitrification. However, despite the fact that the organisms' biodegradation rate is affected by this pH variation, the surface of the respirograms (which corresponds to the mass of O_2 consumed) should remain the same (Kong et al., 1996). The accuracy of the RODTOX *stBOD* measurements, therefore, should not be affected by the remaining changes in pH.

The RODTOX unit installed in the pileEAUte plant cannot monitor pH inside its reactor automatically. Instead, manual measurements have to be taken whenever possible, which lowers the monitoring resolution considerably. The RODTOX does have the necessary communication port to accommodate a pH probe, however, so the implementation of automatic pH monitoring would be trivial if ever desired. One of the advantages of continually monitoring pH is found in the possibility of pairing those measurements to a pH control apparatus, which would allow the RODTOX to collect titrimetric information on the biodegradation process. By replacing the ASM1 model with a related model which accounts for pH more thoroughly, such as the one described in Sin and Vanrolleghem (2007), the stream of data coming from the production and consumption of acids could help better identify the processes occurring within the reactor (Gernaey et al., 2001). Such a titration apparatus would include stock acid and base solutions, each connected to a separate dosing pump. The pumps would supply enough acid or base to compensate any change of pH away from a specific set point, which would have

the effect of both finely regulating pH as well as creating the titrimetric information needed to improve the respirometric data.

4.3 Default output of the RODTOX sensor

With the *fast loop* and the calibration solution in place and the pH well-controlled, the RODTOX is ready to perform measurements. With its internal controller, the RODTOX is capable of calculating a sample's *stBOD* concentration by comparing a sample's respirogram's area (PA) or peak height (PH) to those of the calibration solution (see Section 1.6.2). The RODTOX generates reports after analysing each respirogram on that basis. An example of a typical report is shown in Figure 4.9. For visualization purposes, the report's corresponding respirogram and calibration peak are shown in Figure 4.10. Keep in mind, however, that though the respirograms themselves are shown on the RODTOX's display, the data needed to plot them are not included in the RODTOX's reports when in its factory configuration.


| | | | |
|-----------------------------|---------------------------------|---|----------|
| <u>Text</u> | | 3/30/2017 21:41:18 | |
| <u>Calibration</u> | |  | |
| Dosing volume: | 12 ml | | |
| Minimal slope: | -0.701 mg O ₂ /l.min | Time: | 736 sec |
| Peakheight: | 2.942 mg O ₂ /l | Time: | 750 sec |
| Surface: | 40.740 mg O ₂ .min/l | Time: | 2878 sec |
| Baseline: | 7.282 mg O ₂ /l | | |
| BOD: | 14373 mg BOD/l | | |
| <u>Measurement :</u> | | 0 / 0 | |
| | | Location: 0 | |
| Dosing volume: | 500 ml | | |
| Minimal slope: | -0.233 mg O ₂ /l.min | Time: | 486 sec |
| Peakheight: | 0.928 mg O ₂ /l | Time: | 629 sec |
| BOD: | 109 mg BOD/l | | |
| Surface: | 11.075 mg O ₂ .min/l | Time: | 1708 sec |
| BOD: | 94 mg BOD/l | | |
| Baseline: | 7.282 mg O ₂ /l | | |

FIGURE 4.9 – Typical RODTOX automatically-generated respirometry report.

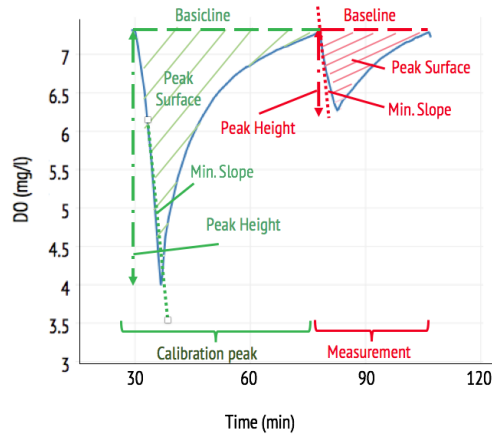


FIGURE 4.10 – DO profile associated with Figure 4.9’s measurement report.

In Figure 4.9, it can be seen that for each addition of wastewater, the following information is outputted by the RODTOX:

1. The dosing volume;
2. The minimal slope of the respirogram (that is, the maximum rate of the decrease of the DO concentration within the reactor);
3. The peak height and the associated calculated *stBOD* concentration;
4. The peak surface and the associated calculated *stBOD* concentration;
5. The “baseline” DO concentration (named “baseline” for the calibration peak);
6. The time elapsed between the beginning of the test and the moment at which the minimal slope, peak height, and peak surface were measured;
7. The value of all those parameters for the associated calibration peak.

Conversely, a similar report is generated for toxic events, as can be seen in Figure 4.11. The respirograms associated with this particular report are also shown in Figure 4.12. In this particular report, the parameters included in the measurement report are coupled with inhibition measurements, which are based on:

1. The change in value of the minimum slope of a calibration solution addition after the addition of a toxic wastewater sample;
2. The change in the value of the peak height;
3. The change in the value of the peak surface.

4.3.1 Extracting the respirogram data for further analysis

While the information described above is useful for operational purposes, it represents only a fraction of the information contained in the RODTOX’s respirograms. Indeed, since a single data point is collected per parameter per respirogram, plotting the respirograms themselves

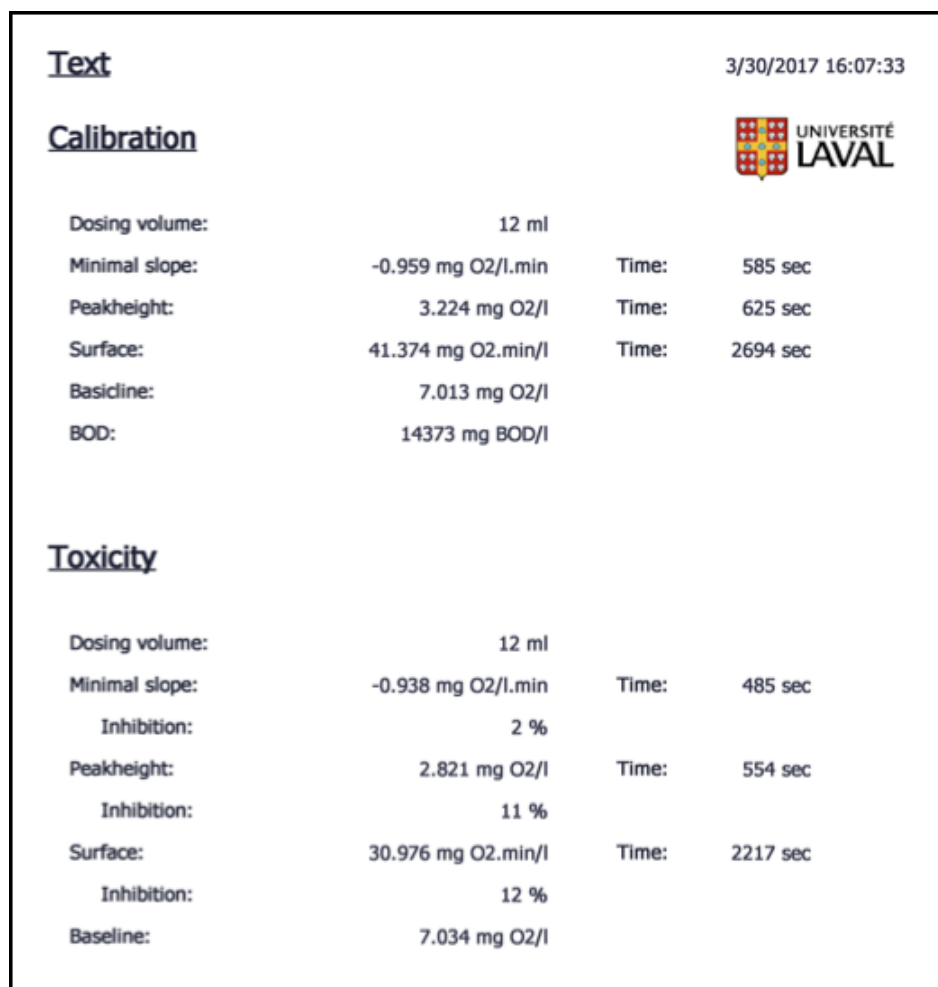


FIGURE 4.11 – Typical automatically-generated RODTOX toxicity report.

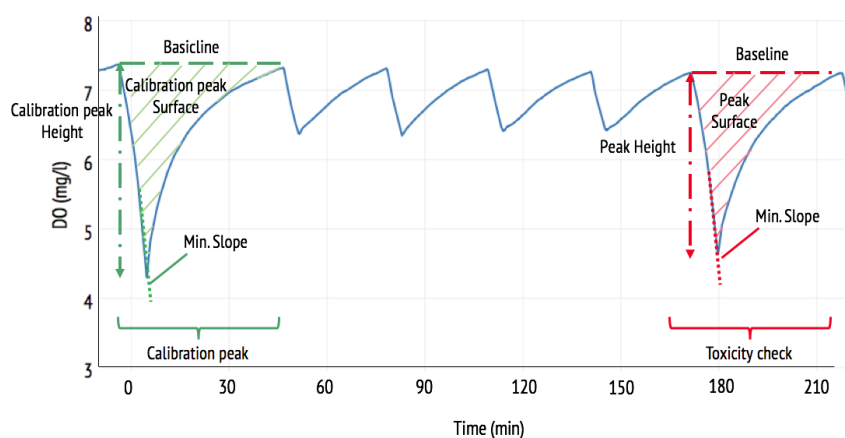


FIGURE 4.12 – DO profile associated with Figure 4.11's toxicity report.

is not possible using those reports. The lack of temperature data also makes it impossible to verify that the temperature inside the reactor stays constant as it should. Additionally, the format in which this data is presented (a .pdf file) is less than ideal, as it is not readily machine-readable. Therefore, gathering the information from several reports has to be done manually. This is a slow and cumbersome process which significantly diminishes the value of the data itself for long-term monitoring.

To facilitate access to the RODTOX's data, changes to the RODTOX's internal program had to be made. The goals of these changes were twofold:

1. To transfer the RODTOX's data to the *pileAUte*'s already existing data acquisition infrastructure — the *monEAU* monitoring station and its associated *datEAUbase*;
2. To export the RODTOX's data in a more convenient format for data storage, manipulation, and analysis.

The connectivity of both the RODTOX and the *monEAU* station has to be taken into consideration to connect them to one another. The *monEAU* station's built-in PC is equipped with two networking cards. One of them is used to connect the station to the wider internet and the *datEAUbase*, while the second is used to communicate with a local area network (LAN). This LAN is composed of a router equipped with Ethernet ports, where other probes may be connected. The RODTOX being equipped with an Ethernet port, it is trivial to connect it to one of the router's free ports. The *monEAU* station can also accommodate probes and probe base stations without Ethernet connectivity through its other serial ports. The network of devices installed in the *pileAUte* plant (including the RODTOX) can be seen in Figure 4.13.

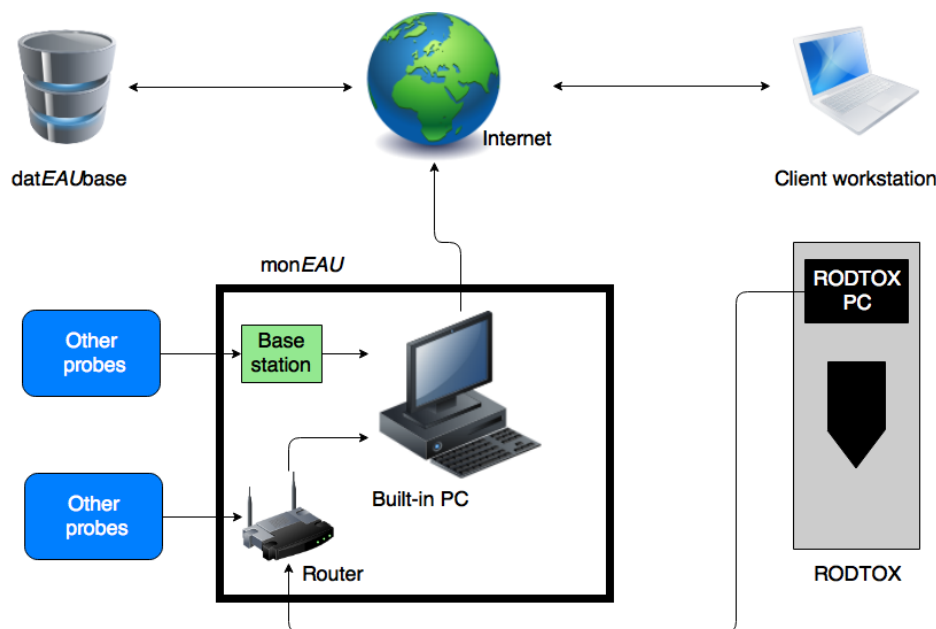


FIGURE 4.13 – *pileAUte* plant data acquisition network

Unfortunately, making the RODTOX communicate with the LAN is somewhat complicated. Indeed, the only way for the RODTOX to communicate with the network is by modifying its controller's TCP/IP settings. Unfortunately, modifying this controller software requires administrator privileges, which only the RODTOX's manufacturer, Kelma NV, has. Incidentally, those same admin privileges are required to modify the RODTOX data's outputs. Therefore, to fix both issues in one fell swoop, Kelma NV's technical support was contacted. To grant them remote access to the RODTOX, the mon EAU station's embedded PC and the LAN router at the same time, TeamViewer was installed both on a laptop directly connected to the RODTOX, and on the mon EAU station's PC (which itself could already communicate with the router).

The following demands were made to Kelma's technical support:

1. Manually assign the following settings to the RODTOX's controller:
 - IP address: 192.168.1.140
 - Subnet: 255.255.255.0
 - Gateway 192.168.1.1
 - DNS: 192.168.1.1
2. Create a button in the RODTOX's interface which would launch the following tasks:
 - Collect all the readings of the RODTOX's DO probe into a .csv file;
 - Collect all the readings of the RODTOX's temperature probe into a separate .csv file;
 - Send both files over the LAN to the mon EAU station's built-in PC and store them inside a specific folder.

Once these actions were performed, the RODTOX could save its raw data to the mon EAU station at the push of a button. An example of the generated files' data structure is shown in Table 4.2. One could conceive of the RODTOX automatically sending its data to the mon EAU station without user intervention. However, this feature has not yet been implemented.

In Table 4.2, it can be seen that most of the rows of the RODTOX's log contain data regarding DO measurements — those are the rows where $VarName = HMI_DO$. Besides the name of the variable used by the RODTOX to store DO data, those rows contain the following:

- The value of the measurement;
- Two separate timestamps (one is a string containing the complete data and time in human-readable format, while the other is an incremental milli-seconds counter)
- A *Validity* value which is either “1” or “2” for valid or invalid data, respectively.

Other rows pertain to other variables relevant to the RODTOX's internal program. A very similar file is generated for the temperature probe.

TABLE 4.2 – Examples of raw probe data coming from the RODTOX.

| VarName | TimeString | VarValue | Validity | Time_ms |
|-------------------|---------------------|----------|----------|-------------|
| HMI_DO | 24.03.2017 14:15:34 | 8.229311 | 1 | 42818594149 |
| HMI_DO | 24.03.2017 14:15:40 | 8.229456 | 1 | 42818594207 |
| ... | ... | ... | ... | ... |
| HMI_DO | 24.03.2017 14:16:00 | 8.219039 | 1 | 42818594443 |
| Reactor_BaseValue | 27.03.2017 13:07:27 | 6.403646 | 1 | 42821546835 |
| ... | ... | ... | ... | ... |
| \$RT_OFF\$ | 30.03.2017 16:06:56 | 0 | 2 | 42824671484 |

| VarName | TimeString | VarValue | Validity | Time_ms |
|------------|---------------------|----------|----------|-------------|
| HMI_Temp | 30.03.2017 16:08:46 | 25,4 | 1 | 4.28247E+14 |
| HMI_Temp | 30.03.2017 16:08:51 | 25,4 | 1 | 4.28247E+14 |
| HMI_Temp | 30.03.2017 16:08:56 | 25,4 | 1 | 4.28247E+14 |
| HMI_Temp | 30.03.2017 16:09:01 | 25,4 | 1 | 4.28247E+14 |
| ... | ... | ... | ... | ... |
| \$RT_OFF\$ | 31.03.2017 16:02:32 | 0 | 2 | 4.28257E+14 |
| \$RT_OFF\$ | 02.04.2017 10:50:40 | 0 | 2 | 4.28275E+14 |

This data is the raw material from which key information concerning the pil EAU te plant's influent, its sludge, and the conditions inside the RODTOX itself may be extracted. The tools which were developed to extract this information are described below.

4.4 Python-based data analysis

4.4.1 Data importation

To manipulate the RODTOX data, they must be imported into a computing environment. The data structure chosen to manipulate the time series is the pandas Data Frame, as it allows for rapid sorting, filtering and selection of data (McKinney, 2011). Ideally, the data coming from the different RODTOX-generated files would be stored into a single structure, and the data should be cleaned up to remove the rows and columns which do not contain relevant information. To do this, the Python function *ImportRdtCSV* was developed. This function can be found in full in Annex A.1. The function takes the following inputs:

- *csvDO*: The name of the RODTOX-generated .csv file containing the desired raw DO data;
- *csvTemp*: The name of the RODTOX-generated .csv file containing the desired raw Temperature data;
- *Existing*: A keyword indicating whether the user wishes to merge the data from those two files with data imported in earlier calls of the function;
- *sourcepath*: The path to get to the source files in the computer's directory structure;

- *destinationpath*: The path to get to the folder where the cleaned-up data should be stored in a .csv format.

To remove useless rows and columns from the raw files, filters are applied so that only rows pertaining to the variables *HMI_DO* and *HMI_Temp*, and only the columns containing the calendar timestamp and the actual measurement values are kept in the final Data Frame. Relevant metadata (i.e., the name of each recorded variable) are also maintained in the Data Frame’s header row. Then, the temperature data are concatenated with the DO data so that every row of the final Data Frame uniquely describes one instant inside the RODTOX. Then, after the data has been cleaned up and concatenated, they can be appended to other Data Frames containing data from other time periods if needed, or it may simply be saved in a .csv file for later use. The function then outputs to the Python environment the following:

- *df*: A Data Frame containing the cleaned-up data;
- *Start*: A string containing the data’s earliest timestamp, and
- *End*: A string containing the data’s latest timestamp;
- A plot of the newly imported DO and temperature data.

The Data Frame can then be used directly by other functions, while the two timestamps, *Start* and *End*, can be to apply the other functions to specific ranges of data.

4.4.2 Evaluation of the actual oxygen concentration and its gradient

After having imported the DO signal, one may begin to prepare that signal for later analysis. The first thing to do is to “recover” the actual DO concentration within the reactor by removing the effect of the probe’s dynamics. As stated in Section 4.2.2, the actual DO concentration DO_{Actual} can be obtained from the raw DO signal y using Equation 4.15.

$$DO_{actual} = \tau \frac{dy}{dt} + y \quad (4.15)$$

Here, τ is the time-delay constant of the probe. As can be seen in Equation 4.15, the derivative of the raw signal is needed to calculate the actual DO concentration. To obtain this derivative, the *Deriv* function calculates an estimate $\hat{\beta}$ of the slope of the DO signal at every point by performing a linear regression on a n -width window of points using Equation 4.16.

$$\hat{\beta}_i = \frac{\sum_{j=i-\frac{n-1}{2}}^{i+\frac{n-1}{2}} (t_j - \bar{t}_i)(y_j - \bar{y}_i)}{\sum_{j=i-\frac{n-1}{2}}^{i+\frac{n-1}{2}} (t_j - \bar{t}_i)^2} \quad (4.16)$$

Since every point’s estimated slope is affected by $n - 1$ of its neighbours and itself, this method of estimation greatly dampens the noise in the signal and helps to obtain a smooth derivative curve. Thus, in order for the time series used later on to be well-behaved and noise-free, the

estimated intercept $\hat{\alpha}$ (calculated using Equation 4.17) and the estimated slope $\hat{\beta}$ are used to calculate an estimation of the “noise-free” signal using Equation 4.18 (Gujer, 2008) (Rowntree, 2004).

$$\hat{\alpha}_i = \overline{y_i} - \hat{\beta}_i \overline{t_i} \quad (4.17)$$

$$\hat{y}_i = \hat{\beta}_i t_i + \hat{\alpha}_i \quad (4.18)$$

Using the smoothened signal \hat{y} and the estimation of the derivative $\hat{\beta}$, the actual DO concentration within the reactor DO_{actual} is calculated by the *Deriv* function using Equation 4.19.

$$DO_{Actual} = \tau \hat{\beta} + \hat{y} \quad (4.19)$$

After DO_{Actual} has been computed, the derivative of the actual DO $\left(\frac{dDO_{Actual}}{dt}\right)$ is calculated using Equation 4.16 once more, replacing \hat{y} by DO_{Actual} .

To perform the tasks described above, the *Deriv* function requires the following inputs:

- *df*: The Data Frame containing the probe signal;
- τ : The probe’s time constant;
- n : The desired window size for the moving-window average.

Applying the function to a Data Frame containing a raw DO probe signal adds to the Data Frame columns containing the following:

- The actual DO concentration within the reactor;
- The estimation of the derivative of the actual DO signal.

Figure 4.14 shows an example of time series generated by the *Deriv* function.

It can be seen in Figure 4.14 that the actual DO is very similar to the probe’s signal, which is to be expected since the probe’s time constant τ is relatively small (32 seconds). Additionally, the derivative of the actual DO appears smooth, which is also expected because the width of the rolling window used is large ($n = 21$), so most of the noise is removed from the time series. Choosing a smaller window ($n = 7$) yields Figure 4.15. As one can see, the smaller window generates a slightly noisier derivative estimation. However, the trend of the time series is still clearly visible.

With the derivative time series generated by the *Deriv* function, it becomes possible to go beyond just processing the probe’s signal, and instead to begin an analysis of what the signal contains — respirograms.

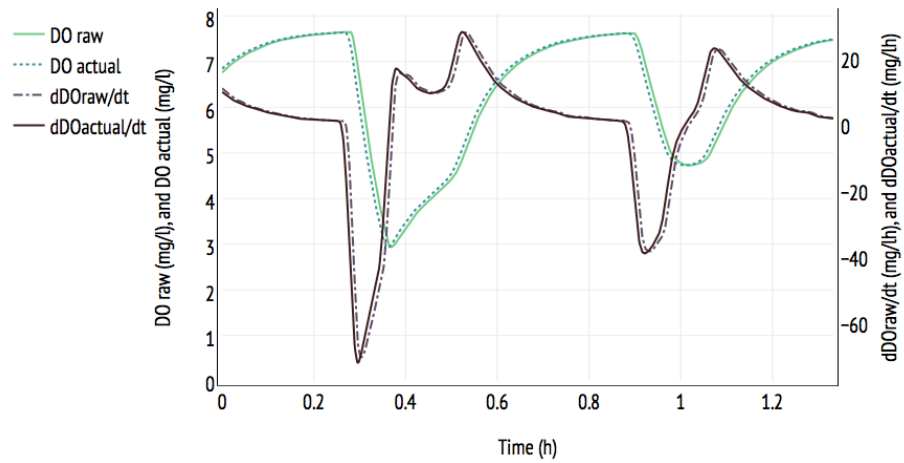


FIGURE 4.14 – Example of the output of the *Deriv* function ($n = 21$).

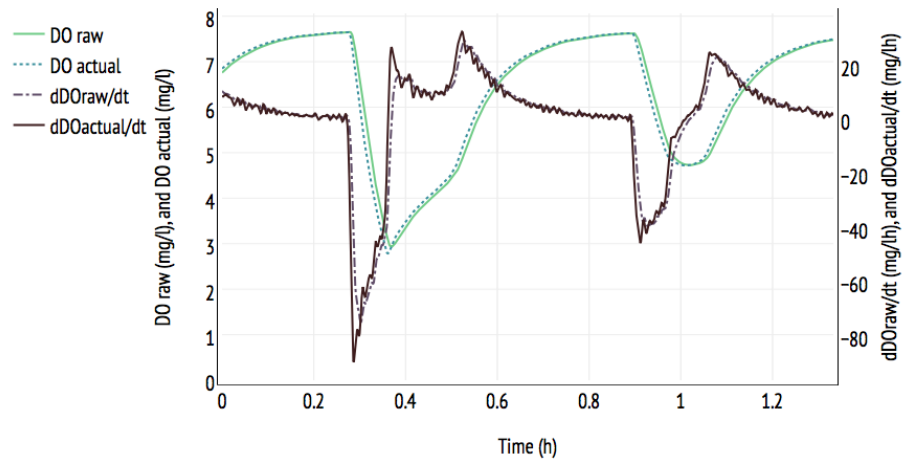


FIGURE 4.15 – Example of the output of the *Deriv* function ($n = 7$).

4.4.3 Respirogram analysis

The *Analyze_Respirograms* Python function is meant to locate respirograms in a time series and extract useful information from them. This is done in several steps:

1. Filters are applied to the time series to find each peak in the signal, and to locate their important features;
2. Each peak's type is identified (decantation, calibration, or wastewater sample), and each is associated with the consequent reactor and sample volumes;
3. The oxygen mass transfer coefficient $K_L a$ is evaluated for each peak, and the associated equilibrium oxygen concentration C_e is estimated;

4. The exogenous oxygen utilization rate OUR_{ex} is evaluated;
5. The surface caught between each respirogram's signal, and its associated equilibrium C_e is evaluated and a $stBOD$ measurement is produced.

The following parameters must be fed to the *Analyze_Respirograms* function to perform these steps listed above:

1. *df*: The variable name of the Data Frame one wishes to analyze;
2. *Start*: A string indicating the beginning of the period one wishes to analyze;
3. *End*: A string indicating the end of the period one wishes to analyze;
4. *Sample_Volume*: The volume of the wastewater samples being added to the RODTOX during the analyzed period (expressed in litres);
5. *Cal_Volume*: The volume of calibration solution added to the RODTOX during the analyzed period (expressed in litres);

In return, the function produces the following outputs:

1. *DF*: The original analyzed Data Frame with additional columns containing the following information:
 - C_e : The estimated equilibrium DO concentration;
 - K_{La} : The estimated oxygen mass transfer coefficient;
 - $stBOD$: The estimated short-term biological oxygen demand;
 - OUR_{ex} : The exogenous respiration rate;
 - *PeakNo*: An incrementally-increasing number identifying each DO peak occurring after a each decantation cycle;
 - *ReacVol*: The reactor volume.
2. *Return_stBOD*: A Data Frame containing the following information for each wastewater respirogram in the analyzed Data Frame:
 - *Start*: The timestamp indicating the beginning of the respirogram;
 - *End*: The timestamp indicating the end of the respirogram;
 - C_e : The estimated equilibrium DO concentration;
 - K_{La} : The estimated oxygen mass transfer coefficient;
 - $stBOD$: The estimated short-term biological oxygen demand;
 - *PeakNo*: The identification number of each respirogram.

Each step of the respirogram analysis process is detailed further below.

Identification of important features

A strength of the Data Frame data structure is the speed at which rows and columns can be parsed, selected and grouped according to logical tests. The *Analyze_Respirograms* function leverages this strength to efficiently group each row of the DO time series which resembles the beginning of a respirogram. To find those rows, both the DO time series itself and its

derivative are analyzed. Looking at both time series (see Figure 4.14), it may be seen that whenever a respirogram starts, the following things happen:

- The DO concentration starts to decrease;
- The derivative goes from positive to negative;
- The DO concentration approaches an equilibrium value C_e .

Thus, the function begins its respirogram localization by finding points where all of those conditions are met. The last condition cannot be filtered for directly, however, as the equilibrium concentration of every respirogram is still unknown at this stage. Instead, the function filters the DO data according to whether the DO concentration is above the 70th percentile of the analyzed range, as the beginning of respirograms tend to occur when DO is relatively high.

The rows of the original Data Frame which meet those conditions are then stored in a new Data Frame — the “filtered” Data Frame — the rows of which each represents a different “peak” in the DO time series, i.e., either a respirogram or a decantation peak. Sometimes, because of the noise in the derivative signal, the start of the same respirogram is captured twice by the above logical tests. To avoid this, the filtered Data Frame is taken through another filter which removes rows whose *Start* timestamp is located less than 5 minutes before the next row’s.

Then, the *End* of each respirogram merely is defined as corresponding to the *Start* of the peak directly after it. This strategy is practical because it removes the need to apply logical tests to locate the end of respirograms. This method, however, has the downside of leaving the last identified respirogram *Start* without an *End* value, so the last row of the filtered respirogram is removed from the Data Frame and thus the last respirogram of the analyzed range is not analyzed by the function.

The lowest DO concentration of the respirogram — its *Bottom* — is found by simply searching for the minimum value of the DO concentration between the *Start* and *End* points found earlier. A test is also applied to verify the length of time between the *Start* and the *Bottom* of a respirogram. This allows the function to identify each respirogram’s type (i.e., to determine whether it is caused by a decantation cycle, a calibration, or a wastewater addition) according to the following rules:

- The peaks where the length of time between the *Start* and the *Bottom* is longer than 28 *min* and shorter than 33 *min* are deemed to be decantation peaks; therefore its associated reactor volume is 10 *l*. This criterion was chosen because the decantation program allows precisely 30 minutes for decantation, during which time aeration is turned off, before turning it back on. Thus, the decantation cycles must always correspond to a period of approximately 30 *min* where DO is decreasing. On the other hand, the *Bottom* value of a respirogram is typically reached during carbon biodegradation,

and this degradation typically takes less than 20 *min*. Therefore the risk of confusing a respirogram for a decantation peak is low.

- The two respirograms directly following the decantation peaks are calibration respirograms. Therefore, their associated sample volume corresponds to the calibration volume parameter provided by the user (typically 12 *ml*), and their associated reactor volumes are respectively $10 + Cal_Volume\ l$ and $10 + 2 \cdot Cal_Volume\ l$.
- Every subsequent respirogram is deemed to be a wastewater sample. Thus, their associated volume corresponds to the following respirogram's reactor volume plus each sample's volume (*Sample_Volume*). Therein lies a limitation of the *Analyze_Respirograms* function: for it to produce accurate *stBOD* measurements, the sample volume must be constant over the investigated range of data. This is less than ideal since the sample volume can either be changed manually by RODTOX users or automatically by the RODTOX itself if the produced respirograms do not have an adequate size. Automatic volume changes modify the sample volume by a factor of 2 to ensure that the respirogram area is adequate to yield reliable measurements. If this happens in the middle of an analyzed range of data, the function currently has no way to detect — and account for — that change. To ensure accurate results, users must, therefore, feed the *Analyze_Respirograms* function data series in which wastewater sample volumes are known to be constant.

One way to handle automatic changes in sample volume would be to modify the *Analyze_Respirograms* function to compare the area of subsequent respirograms. Then, the sample volume could be adjusted accordingly if the function sees that the area of a respirogram is approximately half (or twice) as large as the preceding respirogram's. However, the identification of changes in sample volume was not implemented in the context of this M. Sc. study due to time constraints.

The last important point to be identified (called the *Bend* within the function's namespace) is the point at which the biodegradation of the added pollutants stops. At this point in the respirogram, only re-aeration and endogenous respiration are occurring. This part of the curve can thus be used to estimate K_La . Figure 4.16 shows three respirograms along with their derivative, the calculated exogenous utilization rate OUR_{ex} (see Section 4.4.3), and the position of their respective *Bend* point.

It can be seen in Figure 4.16 that the point where OUR_{ex} falls to zero (and thus the point where only OUR_{end} is non-zero) corresponds with the end of the nitrogen shoulder and with the last peak in the derivative before the end of the respirogram. Therefore, to locate the *Bend*, the function tries to find the last peak of the derivative between the *Bottom* and the *End* of the respirogram.

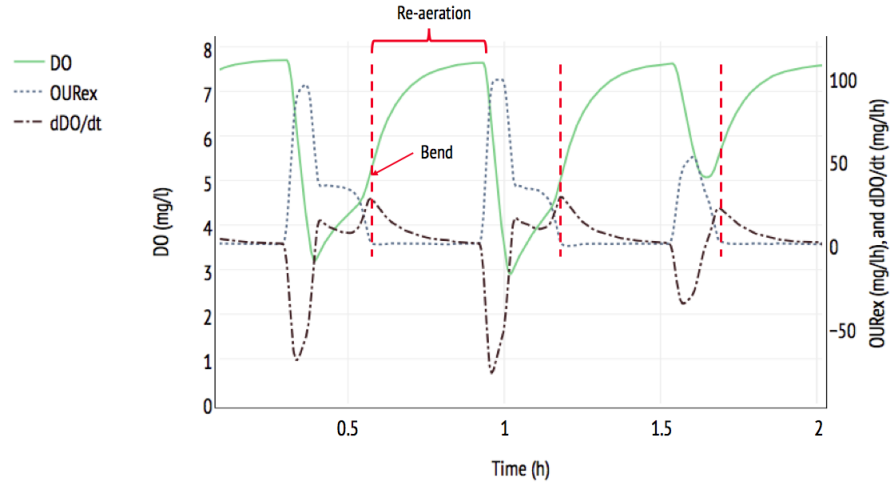


FIGURE 4.16 – Localization of the beginning of the re-aeration phase of respirograms.

After locating those four important points for all DO peaks in the analyzed range of data, the function outputs a plot showing those points' positions. An example of such a plot is found in Figure 4.17.

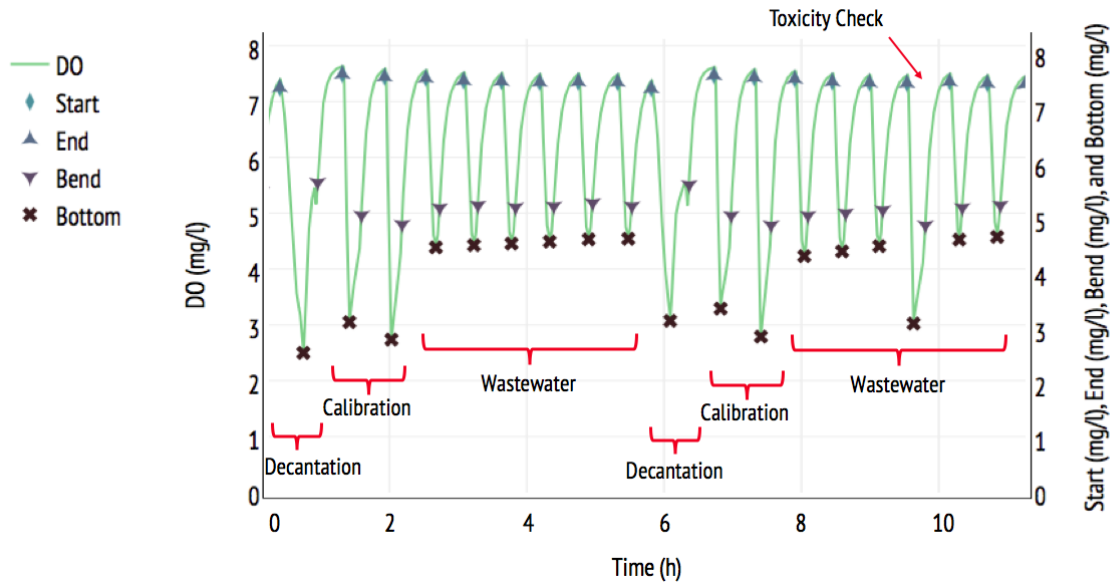


FIGURE 4.17 – Identification of the important points of a series of respirograms.

It can be seen in Figure 4.17 that the algorithm deployed by the *Analyze_Respirograms* can locate the *Start*, *End*, *Bottom* and *Bend* points of several respirograms (18 in this case) in one fell swoop. This points to a major gain in efficiency created by the *Analyze_Respirograms* versus other methods of locating those points, the most obvious being visual identification. By

being able to decode respirograms in large batches such as these, the algorithm thus simplifies and accelerates the analysis of the RODTOX's data in comparison to the manual processing of each respirogram in a spreadsheet program, for example.

Additionally, the brackets in the lower half of Figure 4.17 show the labels assigned by the function to each respirogram. It can thus be seen that one of the respirograms assigned the label “wastewater” is different to the ones surrounding it. This is because this respirogram is actually produced by the RODTOX's “Toxicity check” function, which feeds to the RODTOX the same amount of calibration solution as during a calibration respirogram and no wastewater.

To identify toxicity checks in the future, the function could be modified to accept as a parameter the number of respirograms the RODTOX is set up to measure between toxicity checks. Whenever the peak count reached the same value as this parameter, the current respirogram would be discarded. However, this implies that the analyzed range of data to analyze would need to correspond with the moment the RODTOX begins a new set of experiments, or right after another toxicity check. Though this change in the *Analyze_Respirograms*'s functioning is necessary to avoid toxicity check-related erroneous *stBOD* evaluations, it could not be implemented within the timeframe of this M. Sc. study.

K_La evaluation

The evaluation of K_La is typically done in clean water, where no chemical or biological processes can interfere with the oxygen transfer process. Such conditions, however, are far removed from what is found within the RODTOX, where the changes in the oxygen concentration can be modelled with Equation 4.20 (Vanrolleghem et al., 1994).

$$\frac{dC}{dt} = K_La(C_s - C(t)) - OUR_{end} - OUR_{ex} \quad (4.20)$$

Here, C corresponds to DO_{Actual} . In a reactor containing sludge which is in an endogenous state, the oxygen concentration approaches an equilibrium concentration C_e . This can be modelled with Equation 4.21.

$$OUR_{end} = K_La(C_s - C_e) \quad (4.21)$$

Thus, by substituting Equation 4.21 in Equation 4.20, the latter can be simplified and turned into Equation 4.22:

$$\frac{dC}{dt} = K_La(C_e - C(t)) - OUR_{ex} \quad (4.22)$$

Since K_La estimation is performed in the section of the respirogram where $OUR_{ex} = 0$, the last term of Equation 4.22 can be removed. By solving Equation 4.22 for $C(t)$, one obtains Equation 4.23 which models the DO concentration during the re-aeration phase of a respirogram.

$$C(t) = C_e - (C_e - C_0)e^{-K_La \cdot t} \quad (4.23)$$

Equation 4.23 being a simple exponential equation, one can use a non-linear regression to estimate the parameters K_{La} , C_e and C_0 which best fit the re-aeration phase of the respirograms. However, by naively applying a non-linear regression directly to the entire interval between the *Bend* and the *End* of the respirogram, one may not find a very good fit between the model and the actual DO concentration. This may be caused by residual exogenous respiration by the sludge organisms, or by a change in the organisms' endogenous respiration.

Vanrolleghem(1994, Chapter 3) recommends the use of an algorithm that performs a non-linear regression over a multitude of sub-intervals of the reaeration phase by starting the interval of the non-linear regression one data point later in the time series at every iteration. Then, the standard error of the model produced at each iteration is stored. Finally, the iteration which produced the smallest standard error is kept as the model of the reaeration dynamics for that respirogram. The parameters K_{La} and C_e of this model are then associated with the current respirogram. The *Analyze_Respirograms* function contains an implementation of this algorithm. Figure 4.18 shows several respirograms with the DO concentrations calculated using the models produced by the algorithm described above, and their associated equilibrium DO concentration C_e .

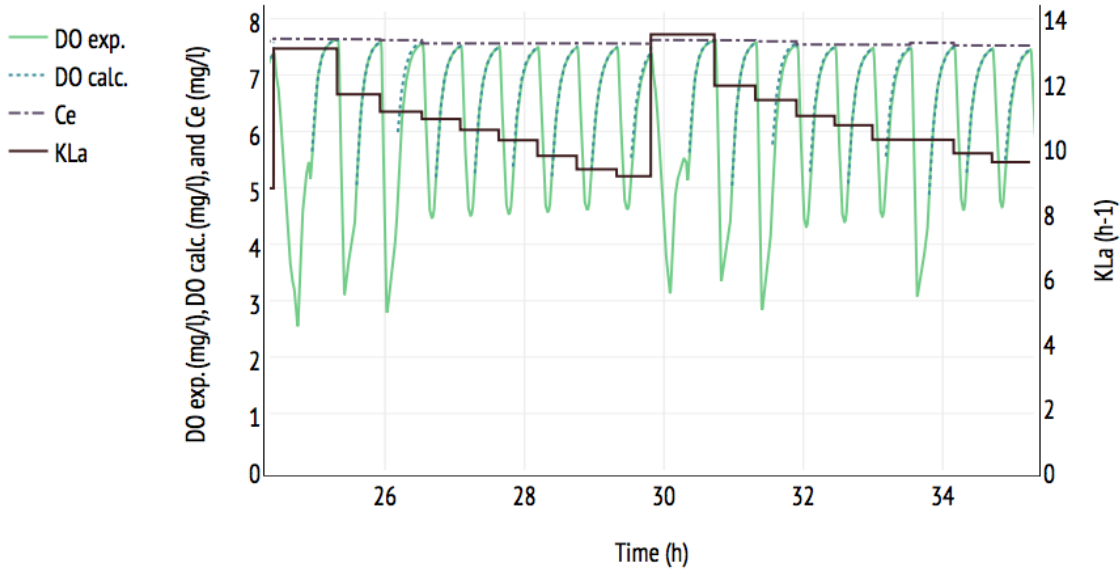


FIGURE 4.18 – Respirograms from Figure 4.17 and their modelled DO concentration in the re-aeration phase and C_e

This algorithm does not always succeed in evaluating K_{La} and C_e , however. This happens when the *Bend* is too close in time to the *End* of the respirogram, and there is not enough data for the algorithm to converge to an adequate model. A logical test was thus implemented to keep the algorithm from attempting a non-linear regression on data spanning less than

8 min. In those cases, the function *Analyze_Respirograms* simply applies the K_La and C_e values obtained in the preceding respirogram to the current one. This is done to ensure that every respirogram is assigned a K_La and C_e value so that *stBOD* may be calculated in the function's next block.

It is good to note here that it might be possible to estimate the saturation DO concentration within the reactor C_s using the information obtained during the K_La evaluation in conjunction with Equation 4.21. In order to do this, however, one more parameter has to be evaluated first — the endogenous respiration rate OUR_{end} . Unfortunately, no moment in the RODTOX's standard automatic operation cycle allows for OUR_{end} evaluation, as this estimation would require the RODTOX not to accept any new samples while turning off the aeration and keeping the stirrer turned on. That way, the only terms on the right side of Equation 4.20 to be non-zero would be OUR_{end} . This would mean that the DO concentration within the reactor follows the linear model of Equation 4.24, where C_0 is the DO concentration at the beginning of the endogenous respiration experiment.

$$C(t) = C_0 - OUR_{end} \cdot t \quad (4.24)$$

While the RODTOX's standard operation cycle does not currently include OUR_{end} estimation experiments, it would be simple to modify the sensor's PLC to perform such an experiment in each operation cycle (that is, in-between two decantations). Figure 4.19 shows a flowchart of what the RODTOX's PLC's logic should look like to generate OUR_{end} measurements along with the rest of the currently-implemented functions.

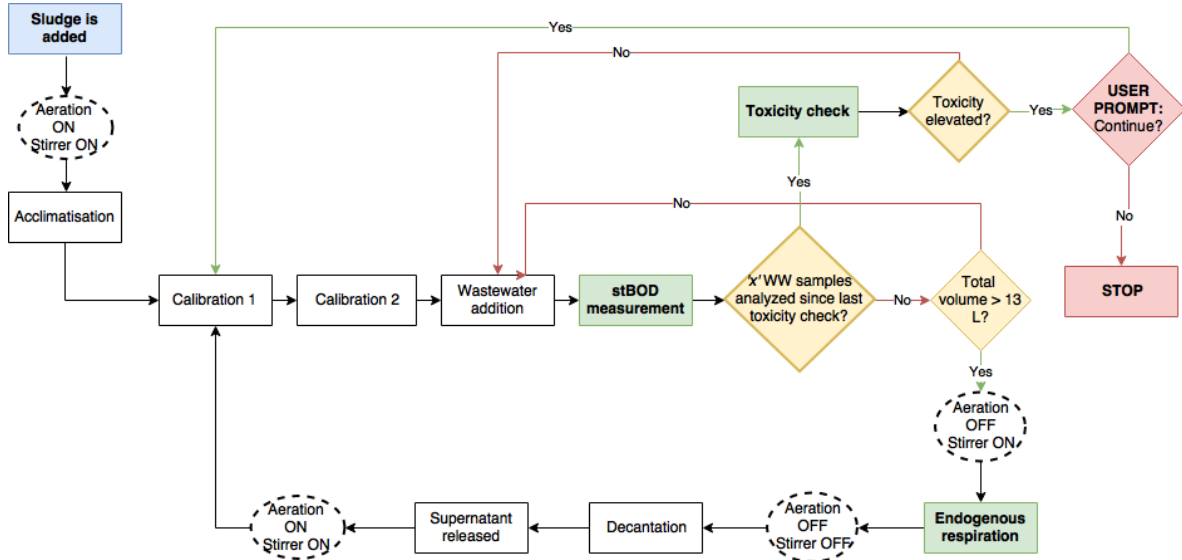


FIGURE 4.19 – Proposed RODTOX PLC flowchart for the inclusion of OUR_{end} estimation.

OUR_{ex} evaluation

Once K_La and C_e have been estimated, all of the terms in Equation 4.22 become known for each respirogram, except for OUR_{ex} . Therefore, by rearranging the terms of Equation 4.22 into Equation 4.25, it is possible to evaluate the exogenous respiration rate over the entire respirogram. Figure 4.16's blue curve is an example of the resulting OUR_{ex} time series.

$$OUR_{ex} = K_La(C_e - C(t)) - \frac{dC}{dt} \quad (4.25)$$

stBOD evaluation

The last section of the *Analyze_Respirograms* function evaluates the *stBOD* of every respirogram. As a reminder, *stBOD* is defined as the respirogram surface multiplied by the reactor's oxygen mass transfer coefficient K_La (see Equation 1.14).

$$stBOD = K_La \int_{t_0}^{t_f} (C_e - C(t)) dt \quad (1.14)$$

The resolution of the DO probe's signal doesn't allow for an exact integration. However, it is possible to perform a finite integration using the trapezoid method, which is described by Equation 4.26 (Gujer, 2008). The last section of the *Analyze_Respirograms* function evaluates the *stBOD* of every respirogram.

$$\int_{t_0}^{t_f} (C_e - C(t)) dt \approx C_e \Delta t \sum_{n=0}^{N+1} \left(\frac{C(t_{i+1}) - C(t_i)}{2} \right) \quad (4.26)$$

Thus, the function calculates an approximation of the area between C_e and the DO signal between the *Start* and *End* points of every respirogram and multiplies it by its associated K_La value. The obtained value corresponds to the *stBOD* concentration inside the reactor at the beginning of the respirogram. However, considering that this *stBOD* comes from the samples being fed to the RODTOX, this measurement becomes more interesting if it is transformed to represent the *stBOD* concentration of the actual samples being fed to the reactor. Therefore, the *Analyze_Respirograms* function transforms the obtained *stBOD* measurements using Equation 4.27.

$$stBOD_{Sample} = stBOD_{RODTOX} \cdot \frac{V_{Reactor}}{V_{Sample}} \quad (4.27)$$

Given the fact that the function tracks the reactor volume for each peak in the DO signal, and adds a different volume to it according to its type (decantation, calibration or wastewater sample), the resulting $stBOD_{Sample}$ values are not skewed by the changing reactor volume.

A plot showing examples of the *Analyze_Respirograms* function's $stBOD_{Sample}$ estimation is seen in Figure 4.20.

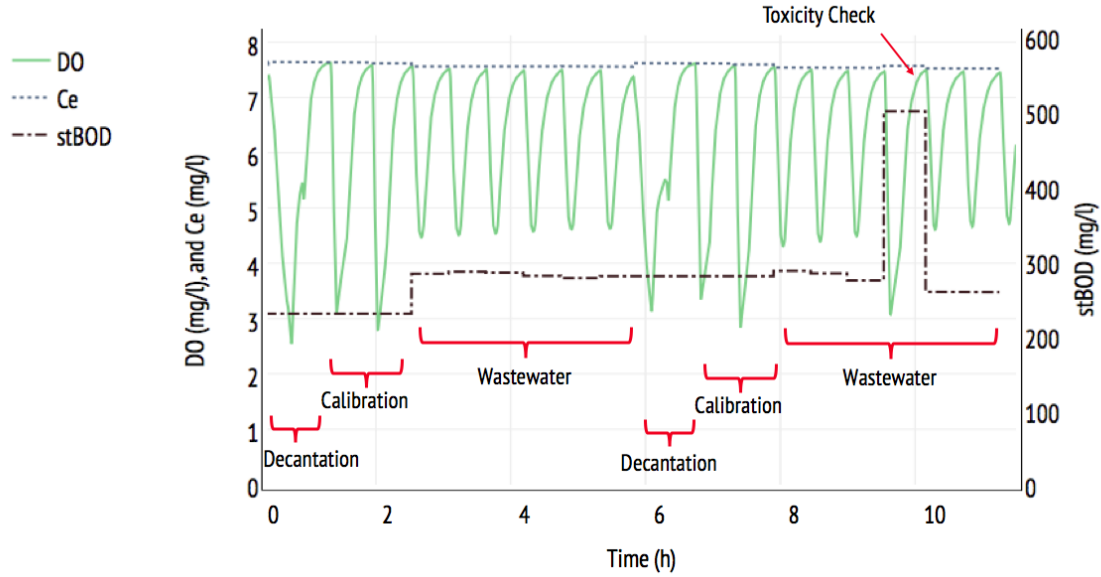


FIGURE 4.20 – Respirograms from Figure 4.17 and their associated $stBOD$ measurements.

As can be seen in Figure 4.20, the produced $stBOD$ measurements are relatively stable across wastewater respirograms. There is one value, however, which is much higher than the others. This outlier is not caused by a wastewater sample but, rather, by a toxicity check. Indeed, respirograms caused by toxicity checks cannot currently be identified by the *Analyze_Respirograms* function, as was explained in Section 4.4.3. Besides this toxicity-check-related outlier, however, the function does successfully detect other possible outliers culprits such as the decantation and calibration peaks. This can be seen in Figure 4.20 since the $stBOD$ does not change during those DO peaks. Instead, the most recent wastewater respirogram's $stBOD$ value is assigned to those peaks.

Interpretation of sample RODTOX data

The average $stBOD$ value associated with the wastewater respirograms shown in Figure 4.20 is 260 mg/l . By looking at a selection of those respirogram's corresponding OUR_{ex} profiles in Figure 4.21, it can be seen that the wastewater sample's OUR_{ex} profiles have a two-lobed shape similar to that of the calibration respirograms on the left of the figure. It is thus probable that those wastewater samples contain two different substrates — just like the calibration solution.

Looking more carefully at the shape of the OUR_{ex} curves, one can see that the decay of the rightmost lobe seems to occur at roughly the same rate for both the calibration solution and the wastewater respirograms. This suggests that the substrate responsible for those lobes in both types of sample is the same — ammonia. The leftmost lobe, on the other hand, is likely caused by rapidly biodegradable organic substrate, on account of it being consumed relatively

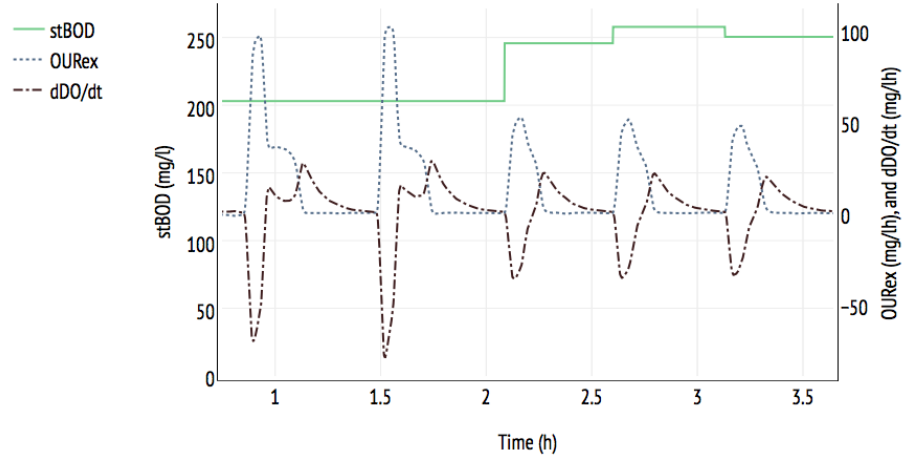


FIGURE 4.21 – $stBOD$ measurements and their associated OUR_{ex} profiles.

quickly (Spanjers and Vanrolleghem, 1995). This rapidly-biodegradable substrate is likely mostly soluble (Metcalf & Eddy, 2013). Therefore, if one wishes to compare the RODTOX’s $stBOD$ measurements to other wastewater quality parameters, the most interesting ones to look at among the pilEAUte’s array of monitored parameters (see Section 3.1) are soluble COD (COD_S) and ammonia nitrogen (NH_4^+-N).

Assuming that most COD_S is rapidly biodegradable and that most nitrifiable nitrogen is in the form of ammonia nitrogen, the expected $stBOD$ value at any given moment can be roughly estimated using Equation 4.28. Here, Y_H and Y_A are left to their ASM1 default values for the sake of the estimation (Henze et al., 1987). This estimation is made with the caveats that some soluble COD is not actually biodegradable (Metcalf & Eddy, 2013) and that some nitrifiable nitrogen in domestic wastewater is, in fact, present in the form of organic nitrogen, not ammonia (Spanjers and Vanrolleghem, 1995).

$$stBOD \approx (1 - Y_H)[COD_S] + 4.57(1 - Y_A)[NH_4^+ - N] \quad (4.28)$$

Figure 4.22 shows the on-line sensor data coming from the RODTOX ($stBOD$), the scanammonolyser (NH_4^+-N), and the scan spectrolyser (COD_S) over the period shown in Figure 4.21. The $stBOD$ estimation based on the latter two measurements is also shown in the figure ($stBOD_{est.}$).

As one can see in Figure 4.22, the estimated $stBOD$ is much lower than the actual recorded $stBOD$. To isolate the source of this discrepancy, one may look at the laboratory measurements taken throughout the summer of 2017 to validate the accuracy of the primary effluent’s

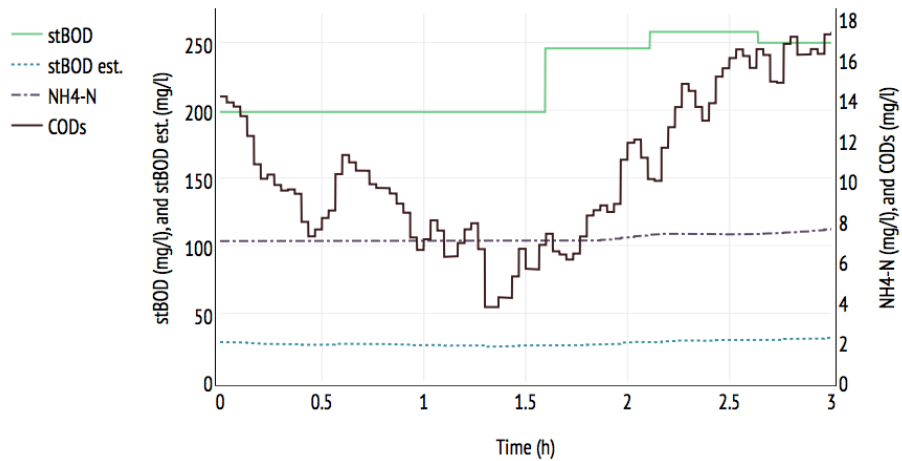


FIGURE 4.22 – On-line sensor output in the primary effluent measurement basin on August 10, 2017.

on-line sensors. The validation measurements performed while the RODTOX was in operation are gathered in Table 4.3.

TABLE 4.3 – Comparison of on-line sensor outputs and their associated laboratory measurement

| Date | COD_S | | | $NH_4^+ - N$ | | |
|-----------------|-----------------|--------------------|-------------|-----------------|--------------------|-------------|
| | On-line mg/l | Laboratory mg/l | % diff % | On-line mg/l | Laboratory mg/l | % diff — |
| July 20, 2017 | 107 | 123 | -13 | 17.2 | 44.6 | -61 |
| August 4, 2017 | 50 | 108 | -54 | 10.8 | 38.1 | -72 |
| August 10, 2017 | 63 | 103 | -39 | 8 | 35.8 | -78 |

As can be seen in Table 4.3, the on-line NH_4^+ measurements are consistently much lower than the laboratory measurements, indicating that the calibration of the scan analyser was not adequate during this period. Therefore, the on-line NH_4^+ measurements of this probe are of little use to evaluate the validity of the RODTOX measurements. Instead, the laboratory values generated during sensor validation itself are used. Table 4.4 shows the results of those measurements and compares them with the corresponding RODTOX-generated $stBOD$ values.

Table 4.4 shows a very good agreement between the lab measurements and the RODTOX-generated value for August 4. This is also true for the August 10 measurement, though to a lesser extent: the RODTOX value is 36% higher than the estimate based on the lab measurements. However, given the fact that the actual yield of the sludge biomass is unknown, and given that the total nitrifiable nitrogen load and the biodegradable fraction of COD_S are unknown, the RODTOX $stBOD$ value is still plausible. This is not necessarily the case for

TABLE 4.4 – Analytical measurements and RODTOX on-line output

| Date | COD_S <i>mg COD/l</i> | NH_4^+-N <i>mg N/l</i> | $stBOD_{est.}$ <i>mg BOD/l</i> | $stBOD$ <i>mg BOD/l</i> | % <i>diff.</i> — |
|-----------------|----------------------------|-----------------------------|-----------------------------------|----------------------------|---------------------|
| 20 July, 2017 | 123 | 44.6 | 235 | 77 | -67 |
| 4 August, 2017 | 108 | 38.1 | 200 | 209 | 4 |
| 10 August, 2017 | 104 | 35.8 | 190 | 258 | 36 |

the July 20 value, however, as the RODTOX value obtained during the validation sampling was three times lower than the estimate suggested by the lab measurements. To understand this result, one may look at that day’s OUR_{ex} profiles, which are shown in Figure 4.23.

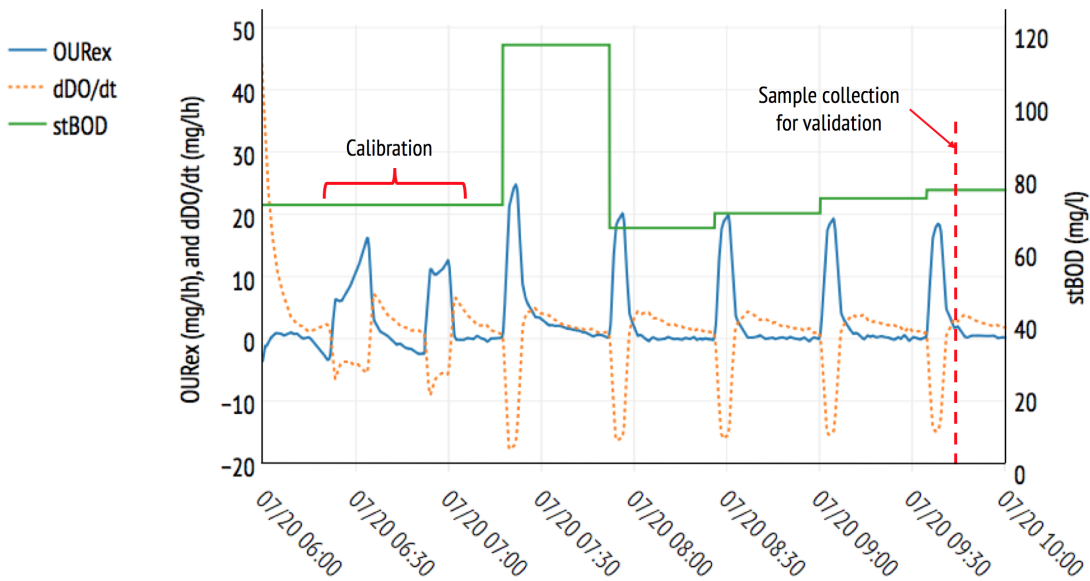
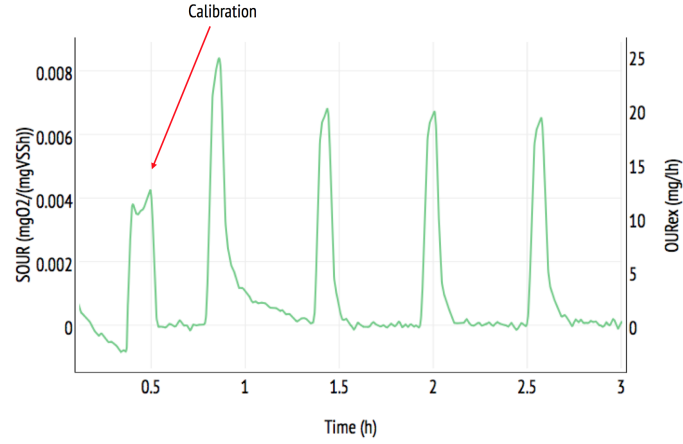


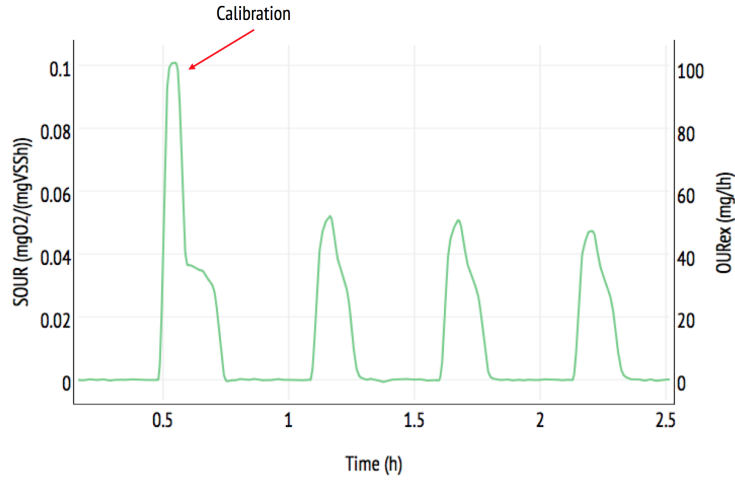
FIGURE 4.23 – Comparison of different water quality measurements relevant to oxygen demand.

Comparing Figure 4.23’s OUR_{ex} profiles to those of other time periods (see Figure 4.21), it can be seen that those of Figure 4.23 do not have the two-lobed shape expected from the degradation of two different samples. Indeed, the OUR_{ex} profiles of each of Figure 4.23’s sample — including the calibration samples — have a single, sharp peak consistent with the biodegradation of rapidly biodegradable organic matter, but no trace of the wider, lower profile created by nitrogen oxidation. Cross-referencing this information with Figure 4.8’s pH measurements, it can be seen that the pH of the reactor at that time was around 5, a value at which nitrification is strongly inhibited (Metcalf & Eddy, 2013).

One may further investigate the possibility of inhibition within the July 20 sludge by comparing the observed specific oxygen uptake rate ($SOUR$) to that of other sludge samples. $SOUR$ indicates the average mass of oxygen consumed by the biomass over a given period, and it is



(a) July 20, 2017



(b) August 10, 2017

FIGURE 4.24 – Comparison of different sludges' exogenous *SOUR*

calculated using Equation 4.29, where *OUR* can either be the total, endogenous or exogenous oxygen uptake rate and *X* is the biomass concentration within the sludge (Spanjers et al., 1998).

$$SOUR = \frac{OUR}{X} \quad (4.29)$$

The exogenous *SOUR* of the RODTOX sludge on July 20, when the pH was around 5, is compared in Figure 4.24 to that of August 10's sludge when the RODTOX's pH was around 7.8. It can thus be seen that the July 20 sludge has a *SOUR* roughly ten times lower than that of the August 10 sludge, which indeed points to inhibition occurring on July 20.

Thus, to evaluate the plausibility of the RODTOX’s *stBOD* measurement on July 20, the oxygen demand coming from nitrification should be disregarded. Doing so yields an estimated *stBOD* value of approximately 40 *mg/l* instead of 235 *mg/l*. Taking this into account, the RODTOX measurement ends up being 47% higher than the *stBOD* value estimated from the *COD_S* measurement. This gap between the two values, though still relatively large, is more reasonable than the earlier estimate, as they both at least lie in the same order of magnitude, and can be explained in part by the fact that *COD_S* is merely an estimate of the available readily biodegradable matter. In fact, some other readily biodegradable substrate is present in wastewater in the form of rapidly-hydrolyzed particulate matter, for example, which is not detectable by a *COD_S* probe (Spanjers and Vanrolleghem, 1995).

4.4.4 Graphical output

To visualize the output of the several Python functions described above, the *Plotit* function was built around the *Plotly* framework (Plotly Technologies Inc., 2015). *Plotly* is a free tool for dynamic plot creation, meaning that the axes and time series can be edited at will after the plot has been generated. To use *Plotly*, a free account must be created, and a validation token must be generated from their website to enable off-line plot generation in the user’s Python environment. Given the fact that all the functions described above include *Plotly*-based data visualization, creating a *Plotly* account is necessary to use this toolset for RODTOX data analysis. The *Plotit* function accepts as its inputs the following parameters:

1. *df*: The Data Frame which contains the data one wishes to plot;
2. *Starttime*: A string containing the timestamp at which one wishes to start plotting time series;
3. *Endtime*: A string containing the timestamp at which one wishes to stop plotting;
4. *List_y1*: A list containing the names of the variables to plot on the first y-axis;
5. *List_y2*: A list containing the names of the variables to plot on the second y-axis;
6. *Label1*: A list containing the names one wishes to show for the data series on the first y-axis;
7. *Label2*: A list containing the names one wishes to show for the data series on the second y-axis;
8. *Units1*: A list containing the units of the quantities plotted on the first y-axis;
9. *Units2*: A list containing the units of the quantities plotted on the second y-axis;
10. *marks*: A list containing the type of marker one wishes to use for the data of either y-axis (either “*line*”, “*markers*” or “*line + markers*”).

The use of the *Plotit* function (reproduced in full in Annex A.4) ensures that every function in the toolset produces visually consistent graphical outputs since the layout parameters of each plot are defined within the *Plotit* function, thus removing the need to re-define them for each data processing function, which is a lengthy and error-prone process.

4.4.5 Retrospective on RODTOX signal analysis

It was shown above that the analysis tools developed in the context of this study can yield $stBOD$, C_e , K_La and OUR_{ex} estimations from the RODTOX's raw DO signal. Moreover, the toolset is able to filter out respirograms which do not contain $stBOD$ measurements for wastewater (e.g., decantations and calibrations). The toolset does not at time of writing include a way to filter out toxicity checks, though it was shown that this feature could be implemented in the future.

It was also shown that the RODTOX's sludge is sensitive to changing conditions within its reactor, such as possible pH drops, and that this leads to changes in its $stBOD$ measurements. Though this problem is largely solved by the addition of alkalinity to the calibration solution (see Section 4.2.5), it was shown that the generated OUR_{ex} profiles are serviceable diagnostic tools which can help interpret the RODTOX's $stBOD$ measurements and identify potential problems with biodegradation within the sensor.

4.5 Building a mathematical model of the RODTOX sensor

The array of analysis tools described in the preceding section have allowed the extraction of $stBOD$ measurements from RODTOX raw data. Conveniently, this $stBOD$ evaluation has had the side effect of unearthing a large amount of information which can be used to describe the features of the respirograms being studied, and also the physical process of oxygen transfer to the RODTOX reactor. One might put this data to use to create a mathematical model of the processes occurring within the RODTOX reactor. Such a modelling exercise would, moreover, extract even more information from the respirograms regarding the analyzed wastewater samples (i.e., determining which portion of each respirogram is caused by either carbonaceous or nitrogenous oxygen demand). This modelling endeavour would also have the side-effect of producing additional information about the RODTOX sludge's biokinetic characteristics (Spanjers and Vanrolleghem, 1995). Thus, this section proposes a procedure for creating a mathematical model of the RODTOX sensor, and then to extract estimations of $stBOD_C$ and $stBOD_N$ using both the WEST modelling suite and the RODTOX data which was processed using the tools described above, and some additional tools described below. The respirograms of the sample injections used in this section are shown in Figure 4.25. In this figure, one may see two calibration experiments, followed by the addition of 3 wastewater samples, one toxicity check (which corresponds to a calibration solution feed), and 3 more wastewater samples.

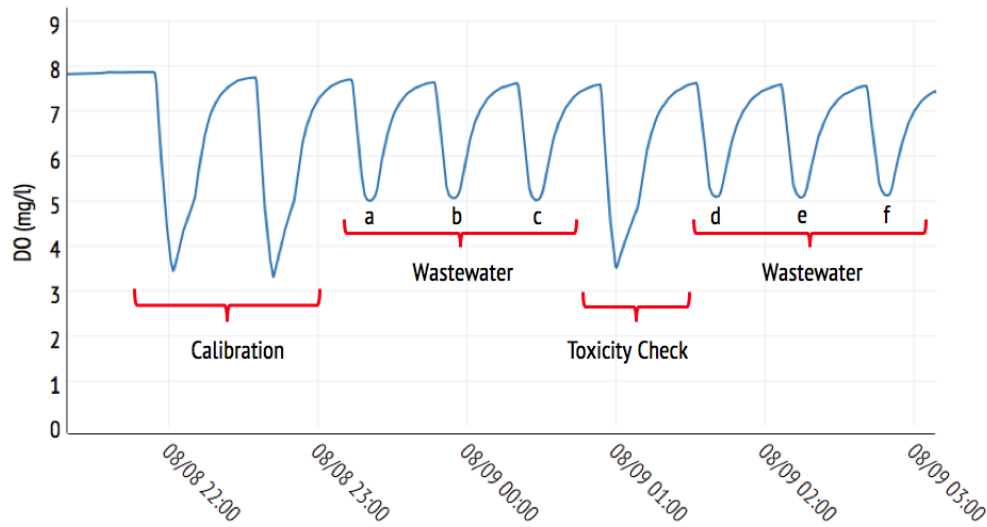


FIGURE 4.25 – Respirograms used in the following modelling exercise

4.5.1 WEST RODTOX model

The first step of this modelling exercise is to represent the RODTOX’s physical setup inside the WEST modelling suite. To do this, one may choose from the many model “blocks” already offered in WEST’s standard library.

For this project’s purposes, the RODTOX is represented by a standard completely mixed activated sludge unit (ASU) block. Of course, the RODTOX’s functioning is closer to a sequencing batch reactor (SBR), as its standard operation cycle includes a decantation phase and an emptying phase, and the RODTOX’s volume varies over time. However, for any single calibration peak, or for a sufficiently small sequence of samples, the change in reactor volume may be considered negligible, and the decantation and emptying phases can be ignored outright since neither occurs during respirograms. Before modelling a time series in which decantation and emptying does occur, however, one should check whether those operations might affect the makeup of the sludge inside the RODTOX by allowing sludge solids to exit the reactor with the supernatant. This may be checked experimentally by measuring the *VSS* and *TSS* contents of the supernatant. However, this verification was not made in the context of this study due to time constraints.

Similarly to the actual RODTOX device, its model may be fed samples in two different ways: either through an injection of calibration solution or by an injection of a dose of wastewater of unknown composition. The former is represented in the model by a “municipality” block, while the latter is represented by a modified “acetate dosage pump” block controlled with a timer. Both of those blocks are described in detail below. Finally, an “effluent” block is used

to allow the “ASU” to release excess water from the reactor. The resulting model layout is presented in Figure 4.26.

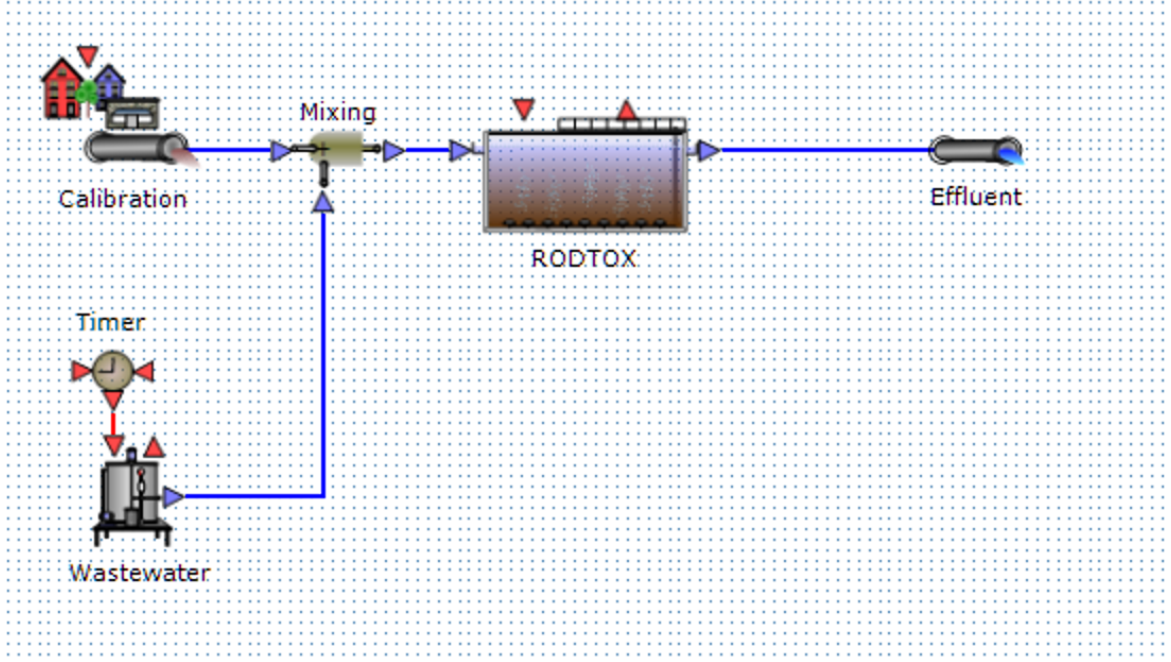


FIGURE 4.26 – Layout of the WEST RODTOX model.

4.5.2 Python-to-WEST bridge

Two types of files can be used with the RODTOX WEST model:

1. **Input files:** These text files describe the influent entering the system. They contain a table which indicates the flow rate of water entering the reactor in conjunction with the concentration of all ASM1 fractions within that influent water at any given time. In this model, input files are used to replicate the injection of a pulse of calibration solution into the reactor. An example of the contents of such a file is found in Table 4.5. When interpolation and extrapolation are turned off in WEST’s “municipality” block, the time series provided are interpreted using a “zero-order hold”, meaning that the value of every input variable (e.g., concentrations or flow rate) is considered constant in the time between each row of the input file. The simulator’s maximum integration time step must thus be set to a value lower than the time difference between two subsequent rows for the integrator to detect every row and thus model the experiment adequately. Null rows are added to the text file to indicate to WEST that the calibration solution injections are restricted to a precise window of time. The null rows following the calibration solution injections must be positioned at time $t_i + \Delta t$ such that Equation 4.30 is respected.

$$(t_{null} - t_i)Q = V_{cal. peak} \quad (4.30)$$

Here, $V_{cal. peak}$ is the total calibration solution volume being fed to the reactor, and Q corresponds to $^{12} ml/5 s$ which, in turn, corresponds to the RODTOX calibration pump's actual flow rate.

TABLE 4.5 – Example of WEST input file representing the injection of two doses of calibration solution.

| .t d | .H2O m3/d | .S_I g/m3 | .S_S g/m3 | .S_O g/m3 | .S_NH g/m3 | ... | .S_NO g/m3 |
|-----------|--------------|--------------|--------------|--------------|---------------|-----|---------------|
| 0.0000000 | 0 | 0 | 0 | 0 | 0 | ... | 0 |
| 0.0173611 | 0.20736 | 0 | 15600 | 0 | 2000 | ... | 0 |
| 0.0174197 | 0 | 0 | 0 | 0 | 0 | ... | 0 |
| 0.0423148 | 0.20736 | 0 | 15600 | 0 | 2000 | ... | 0 |
| 0.0423734 | 0 | 0 | 0 | 0 | 0 | ... | 0 |

A Python function was written (see Annex A.5) to automatically constructs input files for calibration peaks containing either acetate, ammonia, or a combination of the two, with the appropriate timestamps, to create pulses in WEST conforming to the actual calibration substrate injections.

2. **Objective files:** During a parameter estimation experiment, the values of the variables in the objective file are compared to those generated by each subsequent simulation to determine whether the parameters being used are adequate or if they need to be changed. An example of such a file is shown in Table 4.6. The *.RODTOX.C(S_O)* and *.RODTOX.OUR_ASU* are the names WEST assigns to the RODTOX model's DO concentration and *OUR* respectively.

TABLE 4.6 – Example of WEST objective file based on processed RODTOX data.

| .t d | .RODTOX.C(S_O) g/m3 | .RODTOX.OUR_ASU g/(m3.d) |
|-------------|------------------------|-----------------------------|
| 0 | 7.83 | 451 |
| 5.79E-05 | 7.83 | 455 |
| 0.000115741 | 7.84 | 455 |
| 0.000173611 | 7.84 | 450 |
| ... | ... | ... |
| 0.068668981 | 7.73 | 425 |
| 0.068726852 | 7.73 | 432 |

A small Python function was developed to transform the actual DO time series generated while decoding the RODTOX signal into a WEST-ready objective file. In addition to DO, the function also produces a column representing the oxygen utilization rate (*OUR*). However, none of the Python functions created in Section 4.4 to analyze respirograms actually calculate *OUR*. Instead, only OUR_{ex} is calculated since calculating *OUR* requires either knowledge of the oxygen saturation concentration (see Equation 1.11) or of the endogenous respiration rate OUR_{end} itself (see Equation 1.2), both of which are

unknown. Thus, to compare WEST's OUR_ASU variable (which includes both OUR_{ex} and OUR_{end}) and the previously calculated OUR_{ex} occurring inside the RODTOX, the Python objective files function has to add to the measured OUR_{ex} a constant OUR_{end} value. This OUR_{end} value may be drawn from a WEST simulation in which no exogenous substrate is degraded (and thus, where OUR_ASU amounts to pure endogenous respiration). Therefore, the calculated OUR_{ex} time series is transformed according to Equation 4.31. The developed Python function which takes care of the complete objective file creation process is shown in full in Annex A.6.

$$OUR_{ex, calc.} + OUR_ASU_{end, sim.} = OUR_ASU_{calc.} \quad (4.31)$$

These objective files can be used for the following parameter estimation experiments:

- Characterization of the biomass's biodegradation kinetics of a known substrate (i.e., a calibration substrate);
- Characterization of wastewater samples' composition.

Using both of these file types, one would thus be able to bring the data generated by the RODTOX and subsequent Python-based analysis to the WEST environment to perform simulations and parameter estimation experiments. Section 4.5.3 is a description of the steps involved in WEST to ultimately estimate $stBOD_C$ and $stBOD_N$, as well as estimating some of the heterotrophic and autotrophic biomass biokinetic coefficients. Not all biokinetic parameters can be estimated using respirogram data, however, as there are limits to parameter identifiability during such experiments (Vanrolleghem et al., 1998). In some cases, only combinations of parameters may be estimated instead (see Table 1.5, while some parameters may not be estimated at all.

4.5.3 Calibration of the ASM1 model

System characterization

The first step of the calibration of the RODTOX model is to characterize the physical components of the system being studied. In the case of the RODTOX Activated Sludge Unit (ASU), this physical characterization consists of the volume, temperature, and oxygen mass transfer coefficient K_La . Thankfully, this information is already gathered by the Python functions presented in Section 4.4, and it can therefore simply be re-copied into the ASU model at this stage. One should note, however, that though an estimation of K_La was indeed performed in Section 4.4, this estimation may not necessarily correspond to the optimal K_La value for a given respirogram, depending on the accuracy of the Python function's non-linear regression algorithm.

The model blocks used to insert substrates into the system are the “municipality” block and the “acetate dosing pump” block. For this modelling exercise, the “municipality” block is used

to simulate the addition of a calibration substrate into the bioreactor. To do so, an input text file is created where a specific amount of S_S (which is the ASM1 fraction that corresponds to acetate and other easily biodegradable compounds) and S_{NH} (which corresponds to ammonia and other nitrifiable nitrogen compounds). The input file is created such that the injection of those calibration sample pulses occur exactly at the same moment (and for the same duration) as the actual calibration solution addition one wishes to simulate.

Conversely, the wastewater sample additions are made through the use of a modified “acetate dosing pump” block. This block is used for this modelling exercise because it allows one to simulate the addition of a water sample with constant concentrations of easily biodegradable substrates S_S and S_{NH} , which are the ASM1 wastewater fractions which are consumed during a RODTOX respirogram. This S_S and S_{NH} addition is made over a specified period which can be set with a timer. This procedure lets one replicate the RODTOX measurement pump’s sample injection process (which takes approximately 2.5 minutes for a 500 ml sample).

Estimation of biological parameters

The next step of the model calibration is the biological characterization. The laboratory values available for sludge characterization are TSS and VSS measurements of the sludge performed directly before it was added to the RODTOX reactor (see Table 4.1). These tests provide information on the overall sludge solids concentration. However, it does not let one know the composition of the sludge in terms of ASM1 fractions — X_{BH} for heterotrophs and X_{BA} for autotrophs. Therefore, the fractions used in the RODTOX model were based on the ratio of particulates found in the bioreactor of the example problem provided by the authors of the ASM1 model in Henze et al. (2000). This example presents a conventional activated sludge process with nitrification and denitrification, which is analogous to the treatment offered by the *pileAUte* plant. The ratios of the COD fractions X_I , X_{BH} , X_{BA} and X_P used in this example problem and this model calibration are shown in Table 4.7. The other particulate fractions of the ASM1 model, X_S and X_{ND} , are substrate sources for the sludge’s biomasses which the RODTOX is not designed to measure. Therefore, their concentrations were set to values low enough for their effect on respirograms to be negligible, but still high enough not to cause issues with WEST’s numerical engine (which slows down dramatically when computing small floating-point values).

TABLE 4.7 – Typical fractions of particulate components in activated sludge (adapted from Henze et al. (2000))

| Component | Composition (%) |
|-----------|-----------------|
| X_I | 23.1 |
| X_{BH} | 47.8 |
| X_{BA} | 1.8 |
| X_P | 27.2 |

Following this, one may import into the WEST parameter estimation experiment an objective file extracted from a RODTOX calibration experiment and an input file describing the addition of the calibration substrate at the starting point of each respirogram. By running a dynamic simulation over the period covered by the calibration respirograms, one can compare the behaviour of the model to the recorded DO data. The outcome of this preliminary simulation is shown in Figure 4.27.

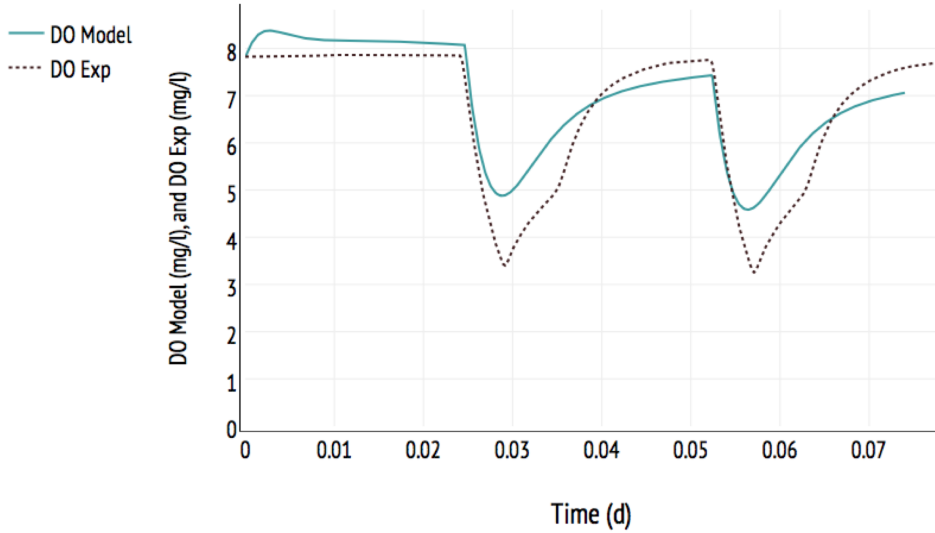


FIGURE 4.27 – RODTOX simulation before any biokinetic calibration

In Figure 4.27, the widening gap between the modelled DO and the RODTOX signal show that the ASM1 standard biokinetic coefficients show the expected trend in DO concentration (i.e., DO drops when substrate is added and comes back up once consumption is over). However, it may be possible to improve upon this result through a calibration procedure.

To improve model performance, the parameters affecting the DO concentration in the reactor must be calibrated. To perform this calibration, one must first focus on the part of the experimental curve where no exogenous substrate is consumed, as fewer processes are at play during those periods. The primary biological process influencing the DO in this period is the endogenous respiration of the sludge. Since the heterotrophic and autotrophic organism concentrations of the sludge have been fixed *a priori* above, however, one may influence the endogenous respiration rate by modifying the two biomass populations' decay coefficients — b_H and b_A . However, given that decay is a long-term process, the short period over which a single respirogram is recorded is not adequate to estimate the value of these coefficients (Vanrolleghem et al., 2003). Other methods, such as the one outlined in Spanjers and Vanrolleghem (1995) could be used to estimate b_H . However, this technique requires prior knowledge of the active biomass fraction X_{BH}/X , which is unknown in the current case.

In addition to the biological process of decay, the physical process of oxygen transfer also affects the DO concentration during endogenous respiration. The saturation oxygen concentration may also be modified to obtain a good fit between the model and the RODTOX measurements since changing this parameter alters the distance between the equilibrium DO concentration and the recorded DO concentration. Modifying C_S thus directly translates to a modification of the driving force which governs the dissolution of oxygen into the mixed liquor (see Equation 4.32, which describes the oxygen mass balance within the RODTOX reactor during steady-state endogenous respiration). Therefore, for a given endogenous respiration rate, a larger C_S concentration directly translates to a higher equilibrium oxygen concentration C_e .

$$OUR_{end} = K_L a (C_s - C_e) \quad (4.32)$$

It is well-known that the oxygen saturation concentration in water changes following the surrounding atmospheric pressure, as well as the temperature, salinity and overall chemical makeup of the water. In the RODTOX, the presence of activated sludge, as well as the repeated addition of a calibration solution with high salinity, may affect the actual DO saturation concentration. The eventual presence of surfactants in the wastewater could also eventually affect oxygen solubility (Painmanakul et al., 2005). However, the default implementation of the ASM1 model used in this modelling exercise only considers water temperature when calculating the oxygen saturation concentration, as it uses Equation 4.33. To compensate for this, a parameter (C_S) was added to the RODTOX model so that the user may set the saturation concentration or include it into a parameter estimation experiment.

$$C_S = 14.65 - 0.41T + 0.00799T^2 - 7.78 \cdot 10^{-5}T^3 \quad (4.33)$$

After the implementation of a new parameter for the oxygen saturation concentration, a parameter estimation experiment was carried out on the parts of the calibration peak DO data which shows endogenous respiration to evaluate the C_S value which provided the best agreement between the WEST simulation of the RODTOX reactor and the experimental data. On the other hand, b_H and b_A were kept to their default values as they could not be estimated anyway. The result was the set of parameters found in Table 4.8. The DO values generated by this model may be compared with the experimental data in Figure 4.28.

TABLE 4.8 – Values of RODTOX model relevant to endogenous respiration.

| ASM 1 parameter | Default value | Calibrated value |
|------------------------------|---------------|------------------|
| $b_H(d^{-1})$ | 0.62 | 0.62 |
| $b_A(d^{-1})$ | 0.15 | 0.15 |
| $C_{s, 28.3^\circ C}(g/m^3)$ | 7.68 | 9.74 |

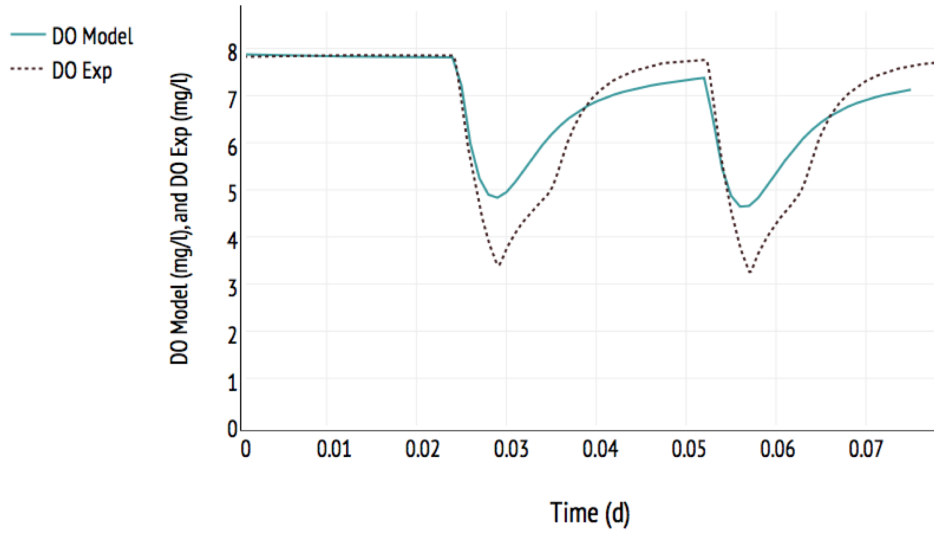


FIGURE 4.28 – RODTOX simulation with calibrated saturation DO concentration (C_S).

It can be seen in Figure 4.28 that the modelled curve shows a good fit during endogenous respiration, as well as the reaeration phases following each respirogram, compared to the modelled curves of Figure 4.27.

At this point, it becomes possible to perform another parameter estimation experiment within WEST to obtain values for the dominant biokinetic parameters during biodegradation of the sodium acetate and ammonium chloride (represented by S_S and S_{NH} respectively). Looking at Equation 4.34 which describes the ASM1's model impact on DO, the following parameters are found to be relevant:

- The heterotrophic yield Y_H ;
- The autotrophic yield Y_A ;
- The heterotrophic maximum specific growth rate $\hat{\mu}_H$;
- The autotrophic maximum specific growth rate $\hat{\mu}_A$;
- The heterotrophic substrate half-saturation coefficient K_S ;
- The autotrophic substrate half-saturation coefficient K_{NH} ;
- The heterotrophic oxygen half-saturation coefficient $K_{O,H}$;
- The autotrophic oxygen half-saturation coefficient $K_{O,A}$.

$$\begin{aligned} \frac{dS_O}{dt} = & -\hat{\mu}_H \left(\frac{1 - Y_H}{Y_H} \right) \left(\frac{S_S}{K_S + S_S} \right) \left(\frac{S_O}{K_{O,H} + S_O} \right) X_{BH} \\ & -\hat{\mu}_A \left(\frac{4.57 - Y_A}{Y_A} \right) \left(\frac{S_{NH}}{K_{NH} + S_{NH}} \right) \left(\frac{S_O}{K_{O,A} + S_O} \right) X_{BA} \end{aligned} \quad (4.34)$$

Not all of these parameters have to be included in the parameter estimation experiment. For instance, since the DO concentration inside the RODTOX is meant not to fall very low during the investigated time period (minimum is around $3.5 \text{ mg O}_2/\text{l}$), the growth of organisms is unlikely to be restricted by lack of oxygen (this typically only occurs at DO concentrations below 2 mg/l (Metcalf & Eddy, 2013)). Thus, $K_{O,H}$ and $K_{O,A}$ may simply be left at their default values.

Furthermore, some of the remaining parameters in the above list are not separately identifiable within the framework of a respirometric experiment (Spanjers and Vanrolleghem, 1995). Indeed, referring to Table 1.5, it can be seen that the autotrophic and heterotrophic maximum growth rates are inexorably linked to the values of their respective biomass concentration and yield coefficient, while the half-saturation concentrations of both biomasses are also linked to their respective yield. Thus, estimating the yield of each biomass at the same time as trying to estimate the half-saturation coefficient and maximum growth rate is counter-productive, as a change in one parameter directly affects the estimation of the other. Thus, given the *a priori* assumptions made above, and the correlation of both yields with each of the other parameters, it was decided that only $\hat{\mu}_H$, K_S , $\hat{\mu}_A$ and K_{NH} would be included in this parameter estimation experiment. Conversely, default values for Y_H and Y_A and the independent estimates of X_{BH} and X_{BA} described above were used.

The heterotrophic yield Y_H was set to 0.71 instead of the default 0.67, as the heterotrophic substrate used is acetate, and the yield of this simple volatile fatty acid is typically higher than that of typical wastewater (Majone et al., 1999).

The choice of the data series over which to optimize the objective function affects the optimization algorithm’s ability to converge to a good solution. Therefore, this decision must be considered carefully. For this parameter estimation, it was decided to calibrate the model over only the second available calibration respirogram instead of using both. This decision was made because, in typical respirographic experiments, the conditions in which the first substrate addition is degraded are less well-defined than for the subsequent ones. Indeed, the period of endogenous respiration preceding the first substrate addition can affect the biomasses’ readiness to consume substrate optimally, as storage products within cells may be depleted, or gene expression may have changed (Vanrolleghem et al., 1998). Thus, choosing to calibrate over the first calibration peak would mean calibrating over a potentially suboptimal biodegradation event, which may lessen the accuracy of the model.

The choice of a good objective function is important as well. Here, the “root mean square error” (*RMSE*) is used (Dochain and Vanrolleghem, 2001). This objective function is calculated using Equation 4.35. In principle, minimizing the value of this function brings the model as close to possible to the values recorded experimentally. *RMSE* is widely used to evaluate the accuracy of models, though it is good to note that it is sensitive to outliers, as each data

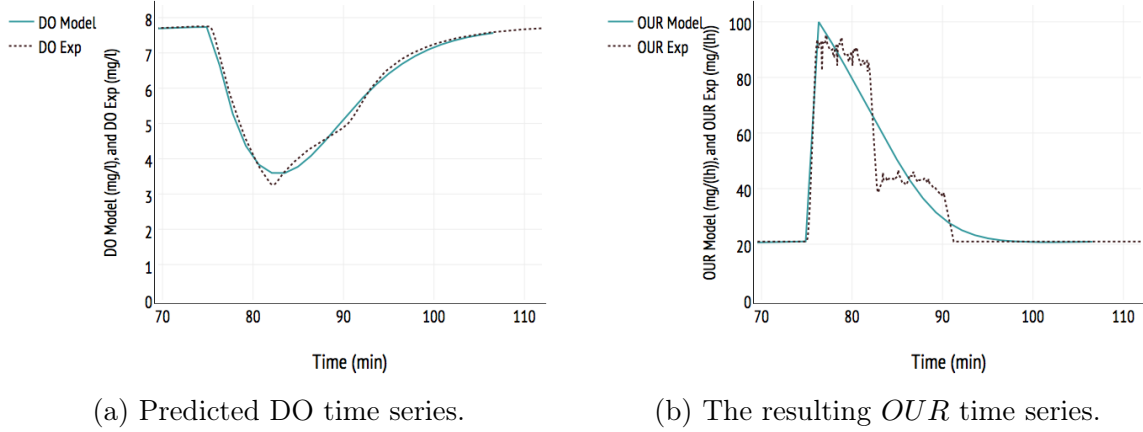


FIGURE 4.29 – Time series from a typical parameter estimation experiment in which $\hat{\mu}_H$, K_S , $\hat{\mu}_A$ and K_{NH} were estimated with DO as the calibration variable. (a) shows the predicted *OUR*, while (b) shows the poor quality of the parameter estimation.

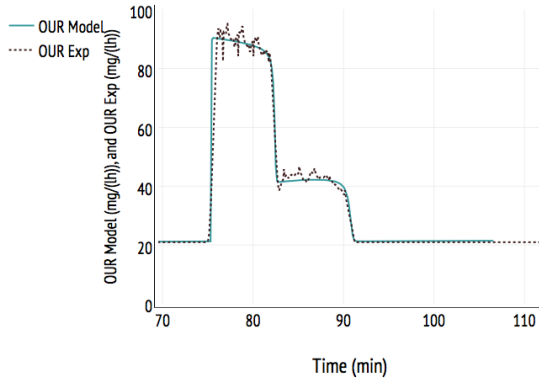
point contributes to the *RMSE* proportionally to the square of its error, so one large error typically has a much larger effect on *RMSE* than several small errors. This proved not to be an issue for this particular modelling exercise, however, as the several time series used have been smoothened and processed beforehand (see Section 4.4.2) and contain no outliers.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_{exp,i} - y_{model,i})^2} \quad (4.35)$$

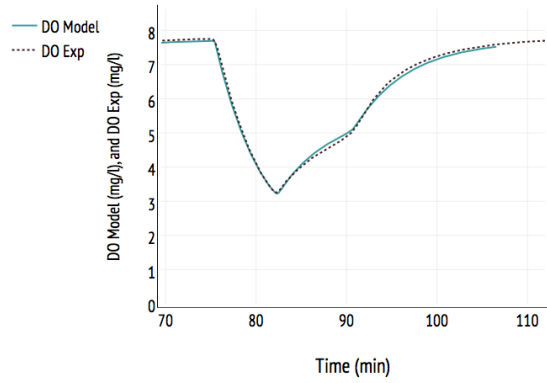
It is also good to note that for any parameter estimation experiment, the choice of an appropriate model variable to calibrate on is crucial. Even though the RODTOX directly measures DO, this variable proves to be numerically more challenging to optimize over for the calibration of biokinetic coefficients of both autotrophic and heterotrophic organisms at the same time. This is because the effect of both organism populations is not easily differentiated using only the DO signal. Indeed, in Figure 4.29 (a), the only distinctive effect of the autotrophic organisms in opposition with heterotrophic organisms is a slight slope change on the right side of the peak which occurs when all readily biodegradable *COD* has been consumed.

As it happens, this change in direction is difficult to detect for the optimization algorithms. This creates a situation in which the optimization algorithm tries to fit the overall shape and size of the DO curve without necessarily describing the separate processes actually occurring during it. This is easily seen in Figure 4.29 (b), which shows the *OUR* curves generated by the calibration performed on the DO concentration from Figure 4.29 (a).

As can be seen in Figure 4.29 (b), the experimentally-obtained *OUR* has a very distinctive, two-lobed shape. The sharp, leftmost lobe is created by carbon oxidation, while the shorter,



(a) Predicted *OUR* time series.



(b) The resulting DO time series.

FIGURE 4.30 – Output of parameter estimation experiment for $\hat{\mu}_H$, K_S , $\hat{\mu}_A$ and K_{NH} using *OUR* as the calibration variable. (a) shows the predicted *OUR*, while (b) shows the good quality of the parameter estimation.

rightmost lobe is formed by nitrification. However, the modelled *OUR* does not exhibit those features, as it contains only one, wide, peak, which clearly shows that the parameter estimation experiment from which this curve comes has failed to accurately model the two separate biomasses acting upon the injected sample.

If, instead of using the DO signal, one elects to use the *OUR* curve to perform the parameter estimation experiment, results like those shown in Figure 4.30 may be obtained.

As can be seen in Figure 4.30, the calibration performed on *OUR* is much more successful at recreating the two separate biomasses' respiration rates. The adequacy of this model calibration may be assessed by the general shape of the respirogram, as well as by comparing the height and timespan associated with both the carbon and nitrogen biodegradation. It can be seen here that both the acetate and the nitrogen lobes of the experimental respirogram is reproduced almost exactly by the model, with both giving the correct height, and lasting for the same amount of time as what was recorded in the experimental data. The overall shape also is well reproduced, with the first transition from endogenous respiration to the acetate peak happening at the correct speed, and the shifts from acetate to nitrogen consumption and from nitrogen consumption to endogenous respiration also occurring at the expected time and with the proper slope.

It can also be seen that the DO curve corresponding to both the experimental data and the model are almost identical to each other, which means that the calibration is successful at reproducing the DO signal even though the latter was not used as the objective function of the optimization. Finally, the low value of the *RMSE* (shown in Table 4.9) indicates that the calibrated model reproduces the experimental data well. Indeed, the obtained *RMSE* value

represents an average error of 3.8% over the calibration respirogram, which is reasonably small. This *RMSE* value and the parameters obtained through this parameter estimation experiment are gathered in Table 4.9.

TABLE 4.9 – Result of the parameter estimation experiment for the calibration substrate biodegradation

| ASM1 parameter | Default value | Calibrated value |
|------------------------------|---------------|------------------|
| $K_S(g\ COD/m^3)$ | 20 | 0.107 |
| $\mu_H(d^{-1})$ | 6.0 | 2.43 |
| $K_{NH}(g\ N/m^3)$ | 1.0 | 0.00774 |
| $\mu_A(d^{-1})$ | 0.8 | 0.941 |
| RMSE (<i>mg/lh</i>) | | 1.45 |

It can be seen in Table 4.9 that the calibrated values for K_S and K_{NH} are three orders of magnitude small than their respective ASM1 default values. In the case of K_S , part of this discrepancy may be due to the default value being meant for use in municipal reactors which are much larger than the RODTOX. Mixing is less uniform in these large reactors, which increases the occurrence of dead zones where organism growth is sub-optimal due to lack of fresh substrate or oxygen, which translates to a larger K_S (Henze et al., 1987). Likewise, ready access by the biomass to the ammonia of the calibration solution might explain part of the gap between the calibrated K_{NH} and its default value. Comparable results were obtained by Spanjers and Vanrolleghem (1995) under similar conditions.

To obtain the results shown in Table 4.9 from the parameter estimation experiment, however, a few additional steps had to be taken. First, the implementation of the ASM1 model used by WEST had to be analyzed to determine the expected effect of each substrate on the heterotrophic and autotrophic biomasses. It was found while doing this that the model implementation of ASM1 used by WEST to calculate the mass balance of heterotrophs followed Equation 4.36, while the published ASM1 model calculates the same quantity using Equation 4.37.

$$\frac{dX_{BH}}{dt} = \mu_H \left(\frac{S_S}{K_S + S_S} \right) \left(\frac{S_O}{K_{O,H} + S_O} \right) \left(\frac{S_{NH}}{K_{NH,H} + S_{NH}} \right) X_{BH} - b_H X_{BH} \quad (4.36)$$

$$\frac{dX_{BH}}{dt} = \mu_H \left(\frac{S_S}{K_S + S_S} \right) \left(\frac{S_O}{K_{O,H} + S_O} \right) X_{BH} - b_H X_{BH} \quad (4.37)$$

Comparing Equation 4.36 to Equation 4.37, it can be seen that WEST adds a Monod saturation kinetic pertaining to the S_{NH} concentration to the published ASM1 model of heterotrophic growth. This fixes a known weakness of the original ASM1 model which allowed numerical solutions to the model which featured negative S_{NH} concentrations in activated sludge (Hauduc et al., 2013). However, this addition to the model intertwines the behaviour of the

heterotrophic and autotrophic biomasses by linking the heterotrophs to both substrates, which impeded the model’s ability to reproduce the experimental *OUR* curve. The half-saturation coefficient $K_{NH,H}$ was thus set to zero to cancel the effect of the S_{NH} concentration on the growth of the heterotrophs.

It is also good to note here that a parameter estimation experiment for this type of problem is far from guaranteed to converge if one does not carefully select the experiment’s initial parameter values. Indeed, the algorithms used to minimize the objective function tend to converge around local minima of the objective function while failing to notice other neighbouring minima, even if those carry a much lower value of the objective function. Thus, it is recommended to begin a parameter estimation by manually changing the value of the parameters to gain an understanding of their effect on the model, and use the manually tuned parameter values to start the automatic parameter estimation experiment with parameters representing as closely as possible the desired model behaviour (Dochain and Vanrolleghem, 2001).

Thus ends the calibration of the RODTOX WEST model. With its physical characterization and the estimation of the biokinetic coefficients relevant to both endogenous respiration and the exogenous respiration related to the biodegradation of S_S and S_{NH} fraction compounds, the model is ready to be applied to wastewater samples using the method described below.

4.5.4 Procedure for the estimation of $stBOD_C$ and $stBOD_N$

During the calibration phase, WEST was used to simulate an experiment for which the substrate concentrations were known, but with biokinetic coefficient values which were unknown. In this new phase of the modelling experiment, the biokinetic coefficient values are now available — they were just estimated —, however, the values of the concentrations of the carbonaceous and nitrogenous substrates present in the wastewater samples being characterized are the new unknowns. Thus, the WEST parameter estimation experiment tool can be re-used here to estimate those concentrations.

As was previously mentioned, the model block used to simulate the addition of wastewater samples into the reactor is an “acetate dosing pump” block modified to accommodate the addition of S_S and S_{NH} . This block releases the sample at a constant flow rate matching that of the RODTOX’s measurement pump, and the flow may be stopped using a timer, which allows for the release of precisely the same volume of wastewater as the RODTOX actually receives for every sample (500 *ml* in this case).

Thus, a parameter estimation experiment may be performed to determine the concentration of each substrate in the wastewater sample. This experiment is done by comparing the experimental *OUR* of each wastewater sample to the *OUR* calculated by the model for the injection of water containing constant concentrations of S_S and S_{NH} by the “acetate dosing pump”. Then, the S_S and S_{NH} that create the best fit with the experimental *OUR* can

be transformed using Equations 4.38, 4.39 and 4.40. to obtain the corresponding $stBOD$ fractions. In Equation 4.39, N_{BH} is the amount of nitrogen assimilated by heterotrophs for cell growth. A method for determining the rate of nitrogen assimilation into heterotrophic biomass r_{NBH} can be determined by evaluating the S_{NH} concentration in the bioreactor during a respirogram, as is shown in Figure 4.31.

$$stBOD_C = S_S(1 - Y_H) \quad (4.38)$$

$$stBOD_N = \frac{S_{NH}V_{Sample} - N_{BH}}{V_{Sample}}(4.57 - Y_A) \quad (4.39)$$

$$stBOD_{tot} = stBOD_C + stBOD_N \quad (4.40)$$

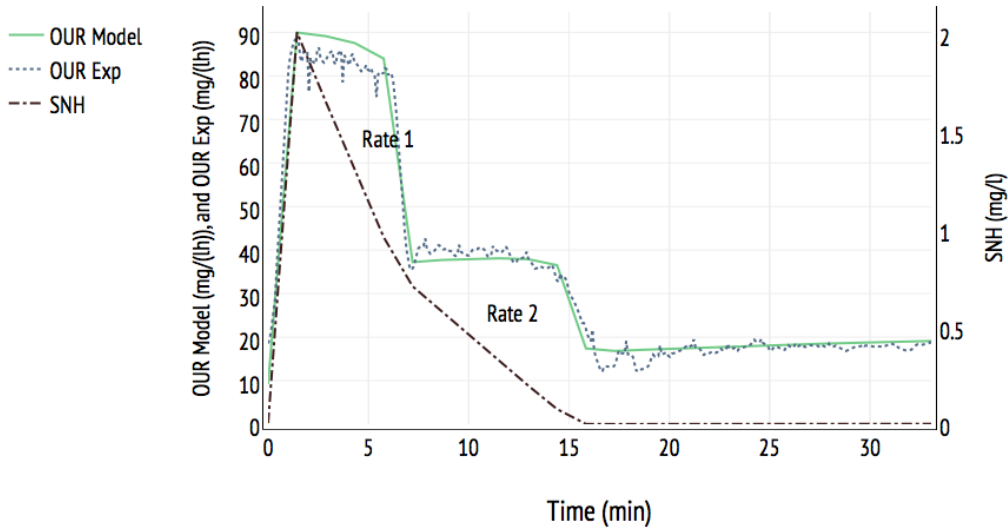


FIGURE 4.31 – S_{NH} concentration in the bioreactor during a calibration respirogram.

It can be seen in Figure 4.31 that the rate of S_{NH} depletion changes from *Rate 1* to *Rate 2* after the rapidly biodegradable substrate has been consumed, and thus when heterotrophic growth stops. Therefore, r_{NBH} can be calculated according to Equation 4.41.

$$r_{NBH} = Rate\ 1 - Rate\ 2 \quad (4.41)$$

If r_{NBH} is assumed constant throughout the section of the respirogram of duration Δt during which S_S is consumed and X_{BH} is produced, the mass of nitrogen included in heterotrophic biomass N_{BH} can be estimated using Equation 4.42.

$$N_{BH} = r_{NBH} \Delta t \quad (4.42)$$

Before simulating the addition of wastewater sample into the RODTOX reactor, one must modify the reactor volume V and mass transfer coefficient $K_L a$ of the RODTOX model to reflect the changing conditions inside the reactor throughout the wastewater analysis. The Python-based data analysis tools described in Section 4.4 already provide a $K_L a$ estimation and a sample volume for each peak, so these values may again merely be recopied into the model. The initial mass of dissolved oxygen $M(S_O)$ may also be modified to reflect the DO concentration inside the RODTOX at the beginning of the respirogram using equation 4.43.

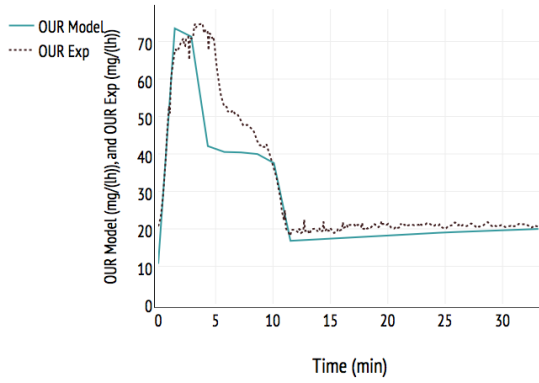
$$M(S_O)[g] = \frac{V[l] \cdot DO(t = t_{Start\ of\ respirogram})[g/m^3]}{1000} \quad (4.43)$$

It is good to note here that the structure of ASM1 is such that whenever a substrate is added to the system, it is immediately available to the organisms that biodegrade it. In reality, however, a delay may be observed between the moment when a substrate is added and the moment when the biomass can degrade it at full speed (Vanrolleghem et al., 1998). It can be seen in Figure 4.32 that the modelled maximum OUR is reached quicker in (a) than in (c), leading to a modelled acetate biodegradation in (a) that finishes earlier than what was experimentally recorded, which isn't the case in (c). Additionally, it can be seen in (b) that the experimental DO concentration decreases over a more extended period than the modelled DO, which is partially corrected in (d) thanks to the time-delay constant.

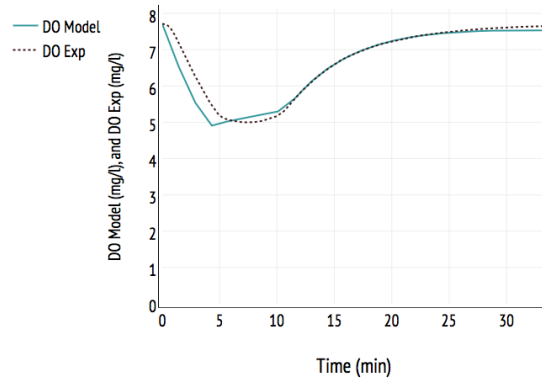
After investigating several potential explanations for the observed delay in reaching the maximum OUR , Vanrolleghem et al. (2004) have concluded that it is likely caused by the chain of metabolic reactions taking place within the organisms which needs to be completed before the actual oxygen consumption occurs. They also have found that the oxygen probe's time constant partially explained this observed delay. However, this is not applicable to the current problem, as the influence of the DO probe has been removed from the signal in Section 4.4.2.

To better reflect this transition period in this modelling exercise, the ASM1 model implementation in WEST was modified. Every instance of the maximum heterotrophic growth rate in the model's system of equations was replaced by a so-called time-delayed maximum growth rate which was defined using the first-order model shown in Equation 4.44 (Vanrolleghem et al., 2004).

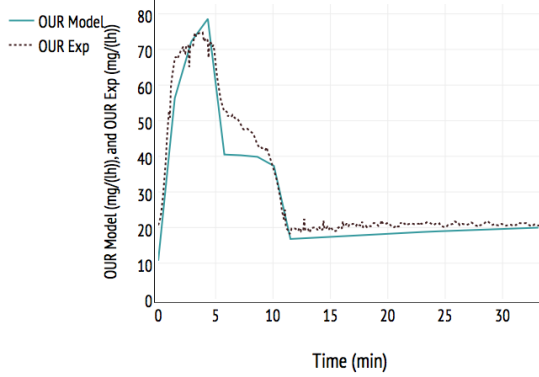
$$\mu_{H, \widehat{delayed}} = \widehat{\mu}_H (1 - e^{-t/\tau}) \quad (4.44)$$



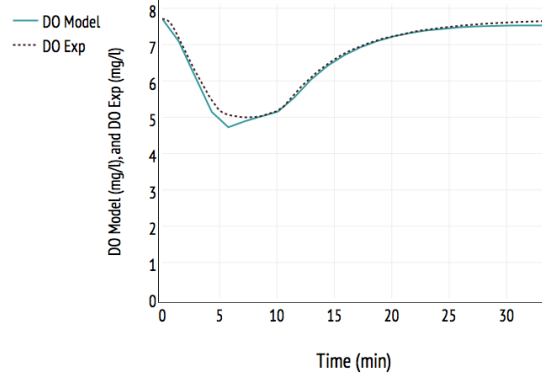
(a) Predicted *OUR* without τ



(b) Resulting DO time series without τ .



(c) Predicted *OUR* with $\tau = 86.4$ s.



(d) Resulting DO time series with $\tau = 86.4$ s.

FIGURE 4.32 – Effect of the time-delay constant for the modelling of the beginning of respirograms

Here, τ is a new parameter introduced into the model to act as a time-delay constant, and t is the time elapsed since the addition of the sample. Its value may be estimated by including it into the parameter estimation experiment being conducted.

As can be seen in Figure 4.32 (c) and (d), the behaviour of the model much more closely resembles reality when a time delay is added to the maximum heterotrophic growth rate. While the time-delay constant does not bring the modelled and experimental *OUR* curves to a perfect agreement, it does help them to espouse much more similar shapes. The same can be said of the DO curves: though the very beginning of the respirogram DO curve still is not perfectly represented, the introduced delay is sufficient for the rest of the decreasing phase of the experimental and modelled curves to proceed in unison, which is the desired effect.

With these modifications to the model in place, the simulation of the wastewater peaks can proceed. A parameter estimation is thus finally performed on each peak to estimate the

concentration of S_S and S_{NH} of the wastewater injected into the RODTOX. The time-delay constant was also added to the parameter experiment, as its value may change depending on the sludge organisms' history (Vanrolleghem et al., 2004).

4.5.5 Evaluation of the modelling results

Figures 4.33 and 4.34 respectively show the *OUR* and DO curves of the six wastewater samples in the investigated series of RODTOX data. To evaluate the results of the modelling exercise, one may analyze the *OUR* curves. Looking at Figure 4.33, it can be seen that the modelled and experimental *OUR* trace similar curves for all six analyzed samples. The maximum recorded *OUR* for each wastewater sample is very close to the modelled *OUR*, which means that the rate at which rapidly biodegradable *COD* is consumed in the wastewater is similar to what was determined using the calibration respirograms. It can also be seen that the end of the exogenous respiration peaks are well described by the model, as the decrease of *OUR* from the peak towards endogenous respiration occurs at the same moment and the same rate for both the model and the experimental data for all analyzed peaks.

For their part, the beginning of the respirograms of Figure 4.33, with their less-than-instantaneous increase towards the maximum respiration rate, though they aren't reproduced perfectly by the model, are reasonably well approximated by it. It may be that the actual dynamics of that increase are not perfectly described by the implemented first-order kinetic function of Equation 4.44 with the time-delay constant. However, the model's behaviour is nonetheless more consistent with the experimental data with this implementation than the default WEST ASM1 implementation, which would predict the biomass reaching its maximum growth rate much earlier.

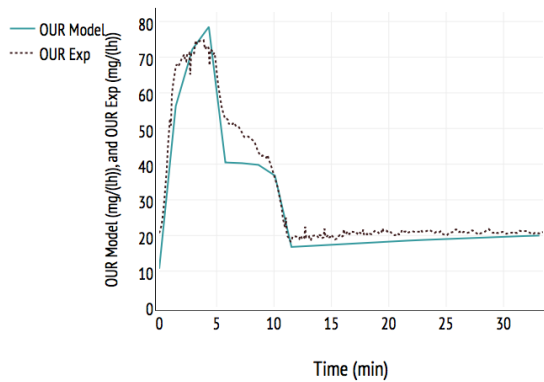
However, it can be seen that for all analyzed samples, a small section of the *OUR* located directly above the nitrogen shoulder is not reproduced by the model. This part of the respiration likely is not due to the biodegradation of nitrogenous substrates, as this would mean that the recorded respiration was superior to the maximum respiration rate of the autotrophic biomass recorded during the calibration respirograms (see Figure 4.30). It is more likely that this respiration is instead due to the biodegradation of a substrate which is not as easily biodegradable as the S_S (which is consumed in the first part of the respirograms) but is still easy enough to biodegrade to be oxidized during the short span of the respirogram. Substrates which may fall into this category include rapid hydrolysis products and organic compounds with a high molecular weight (Spanjers and Vanrolleghem, 1995). In the ASM1 model, hydrolyzable compounds fall into the X_S fraction of wastewater, which is much slower to biodegrade than what is observed here. Therefore, the RODTOX model used in this exercise is incapable of replicating this portion of the experimental *OUR* since the only two substrate fractions accounted for in the model are S_S and S_{NH} .

This “extra” fraction could, in principle, be added to the RODTOX model. To do so, one could adapt ASM1 by defining a new wastewater fraction whose behaviour would resemble that of S_S , but whose kinetics would be slightly slower than the latter’s, as was done in Spanjers and Vanrolleghem (1995). Accounting for this portion of the experimental OUR , however, falls outside the scope of this M. Sc. study, and is thus left for further research.

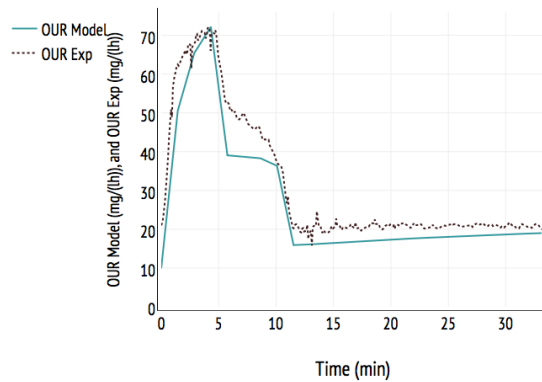
Similarly, it is possible that this portion of the curve represents heterotrophic growth based on the consumption of what is defined as storage products (X_{STO}) in the ASM3 model. In the latter, rapidly biodegradable organic matter is integrated into the biomass’ cellular structure before being available for organism growth. This is in opposition to the ASM1 paradigm in which S_S is directly used for organism growth. This storage function might have helped to fit the unaccounted-for part of the respiration curves described above, as the use of ASM3 for the interpretation of respirographic curves generated by rapidly biodegradable COD has been shown to increase model fit compared to ASM1 (Guisasola et al., 2005). However, the calibration and use of ASM3 to interpret the RODTOX’s respirometric curves also falls outside the scope of this M. Sc. study.

One may also look at the DO curves generated from each of Figure 4.33 to evaluate the performance of the model. It can be seen in Figure 4.34 that all 6 modelled DO curves closely resemble their respective experimental curve in that they start decreasing at the same time, the respirogram’s maximum height is respected, and their re-aeration phase seems to start at the same moment.

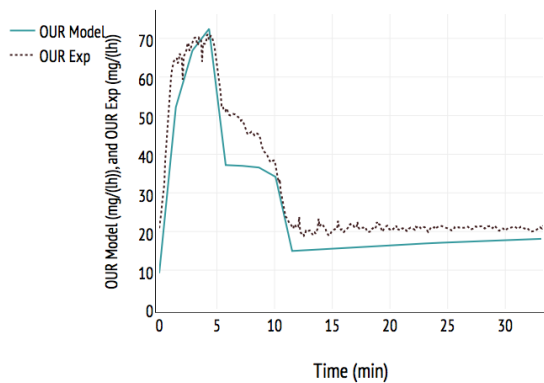
However, it seems that for respirograms (a), (d) and (e), the modelled DO concentration does not reach the same equilibrium as the experimental DO. This may be due to an inadequate evaluation by the Python function described in Section 4.4.3 of the system’s oxygen transfer process. Indeed, according to Equation 1.1, oxygen transfer is influenced by the mass transfer coefficient K_La , the oxygen saturation concentration, and the biomass respiration rate. In the parts of the DO curves where this mismatch is observed, Figure 4.33 shows that the respiration rate has reached its endogenous phase and the modelled OUR closely resembles the experimental OUR , so the respiration rate does not seem to be the culprit. As for the K_La , it has been evaluated using a non-linear regression using the tools described in Section 4.4, and though the resulting fit was quite reasonable, the obtained value may still not be optimal. Conversely, the oxygen saturation concentration has not been evaluated experimentally, but merely estimated during the estimation of the model’s decay coefficients. It is, therefore, possible that this evaluation was not wholly adequate, which might explain why the DO equilibria of the model and of the experiments do not correspond with one another for these three respirograms. This discrepancy could be improved upon in the future by performing a second parameter estimation experiment on the wastewater respirograms to find better estimates of C_S and K_La . Of course, since this would change the DO data, the modelled OUR would be affected, as the latter is calculated using the former (see Equation 4.25).



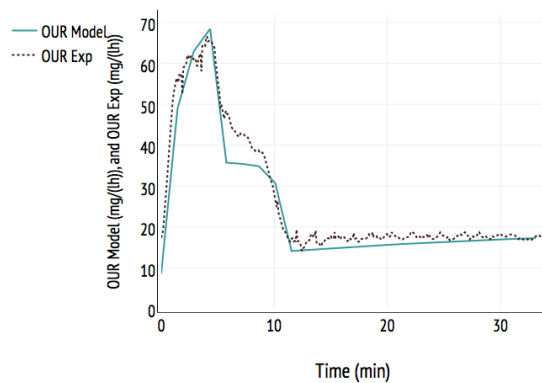
(a) Sample #1 (2017-08-08 23:13:00)



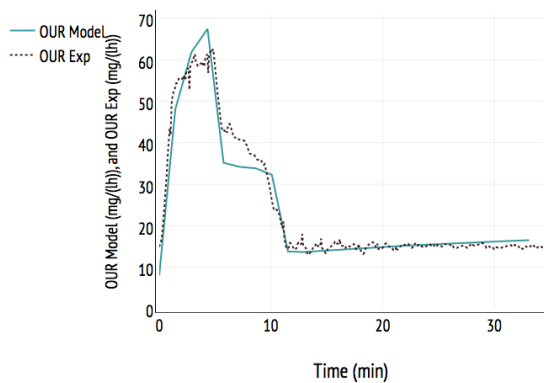
(b) Sample #2 (2017-08-08 23:46:00)



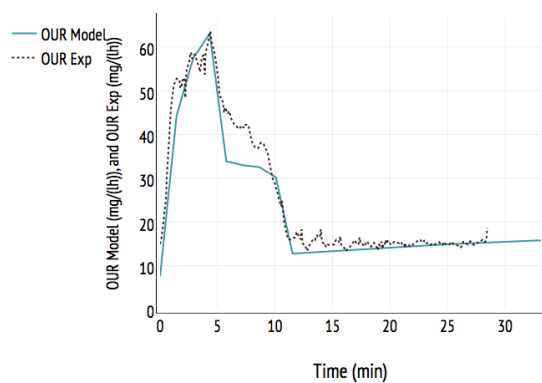
(c) Sample #3 (2017-08-09 00:20:00)



(d) Sample #4 (2017-08-09 01:32:00)

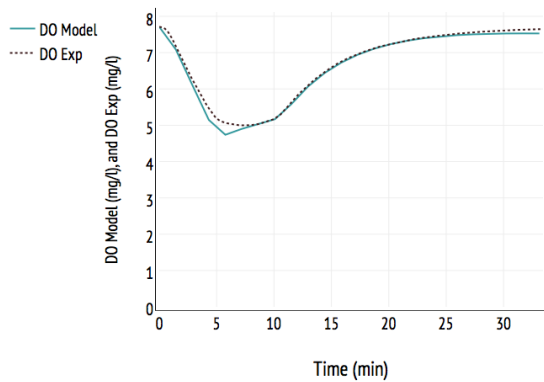


(e) Sample #5 (2017-08-09 02:06:00)

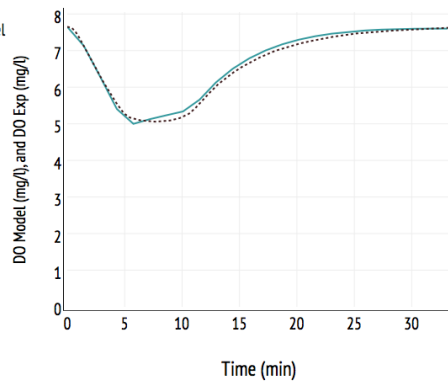


(f) Sample #6 (2017-08-09 02:40:00)

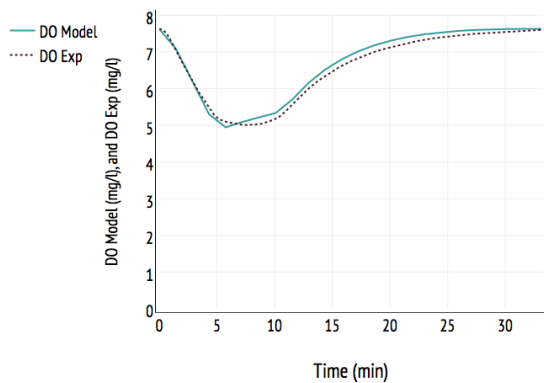
FIGURE 4.33 – Modelled and measured *OUR* of sequential respirograms.



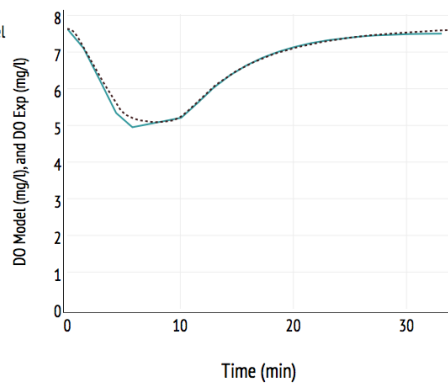
(a) Sample #1 (2017-08-08 23:13:00)



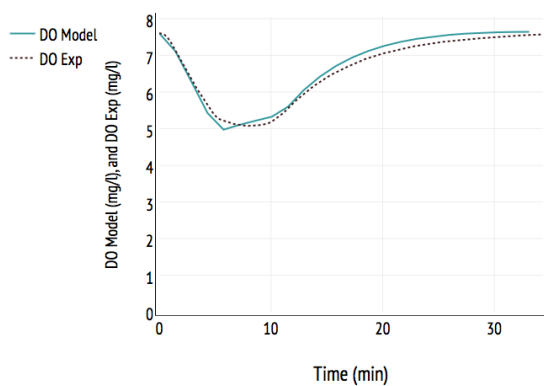
(b) Sample #2 (2017-08-08 23:46:00)



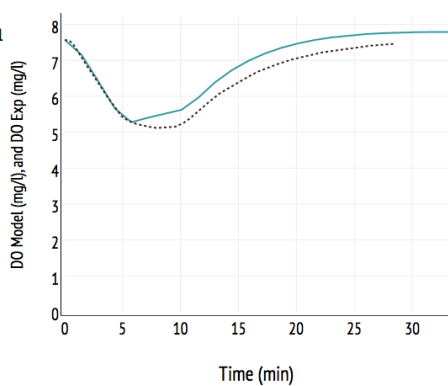
(c) Sample #3 (2017-08-09 00:20:00)



(d) Sample #4 (2017-08-09 01:32:00)



(e) Sample #5 (2017-08-09 02:06:00)



(f) Sample #6 (2017-08-09 02:40:00)

FIGURE 4.34 – Modelled and measured DO of sequential respirograms.

Thus, a final parameter estimation experiment could then be done to re-calculate the S_S and S_{NH} concentrations which provide the best fit given the new K_La and C_S . However, this iterative procedure lies outside the scope of this M. Sc. study.

Based on the simulations described above, concentrations of S_S and S_{NH} and the associated $stBOD_C$ and $stBOD_N$ values were calculated for each wastewater respirogram. The results are shown in Table 4.10. These simulation-based $stBOD$ estimations may be compared with those obtained through the direct analysis of the respirograms with the help of the software tools described in Section 4.4.

TABLE 4.10 – Comparison of $stBOD$ evaluations of Python-based data analysis and WEST simulations

| Sample ID | Direct analysis | Model-based estimation | | | |
|-----------|-------------------------|------------------------|---------------------|-------------------------|--------------------------|
| | $stBOD_{tot}$ (mg/l) | $stBOD_C$ (mg/l) | $stBOD_N$ (mg/l) | $stBOD_{tot}$ (mg/l) | $stBOD_{tot}$ (%diff) |
| a | 122 | 43 | 88 | 131 | 7.0 |
| b | 126 | 38 | 86 | 124 | -1.8 |
| c | 131 | 43 | 91 | 134 | 2.7 |
| d | 126 | 41 | 87 | 129 | 2.0 |
| e | 136 | 47 | 98 | 145 | 6.3 |
| f | 141 | 42 | 93 | 135 | -4.3 |

As can be seen in Table 4.10, the model-based estimations of $stBOD$ are similar to those derived by direct analysis of the respirograms, as the maximum recorded error between the two methods is of only 7% and the average a mere 2%. This good result should be interpreted carefully, however, since, as was discussed above, the two modelled substrates S_S and S_{NH} do not allow the model to account for the rapidly hydrolyzable substrate present in the wastewater samples. Since the latter wastewater fraction has its own oxygen demand, the close agreement between the modelled and directly measured $stBOD$ suggests that the model over-estimates the concentration of at least one substrate. Considering the relatively small size of the non-modelled portions of the respirograms of Figure 4.33, however, this overestimation is bound to be somewhat small. Thus, even if the produced $stBOD_C$ and $stBOD_N$ values do not perfectly describe the experimental respirograms, they could nonetheless be used as robust estimates of those quantities in the analyzed wastewater samples and, by extension, in the influent wastewater.

To further evaluate the performance of the model-based $stBOD$ estimations, one may look at the toxicity check which was performed between wastewater samples (c) and (d). A parameter estimation experiment was performed on this toxicity check to obtain values for the calibration solution's S_S and S_{NH} concentrations while keeping the biokinetic parameters identical to those used for the wastewater samples, and the result is shown in Figure 4.35.

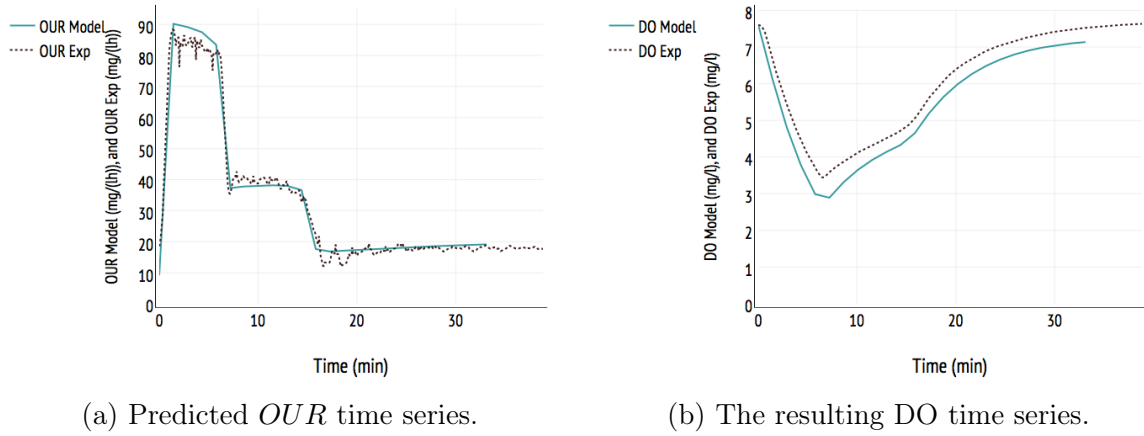


FIGURE 4.35 – Modelled and measured *OUR* and DO during toxicity check (2017-08-09 00:53:00)

As can be seen in Figure 4.35, the model reproduces the respiration rate and DO related to the toxicity check relatively well though it seems to slightly over-estimate the contribution of acetate to the respirogram when one examines it visually. Looking at the numbers tells a different story however, as the obtained values for the S_S and S_{NH} concentrations respectively were $15131 \text{ mg } S_S/l$ and $2173 \text{ mg } N/l$. The actual concentration of acetate in the calibration sample is known to be $15600 \text{ mg } S_S/l$, which amounts to a 3% difference between reality and the model, while the estimated S_{NH} concentration is 9% higher than the actual concentration ($2000 \text{ mg } S_{NH}/l$). Looking at the DO curves, however, it can be seen that the model tends to a different equilibrium concentration than the experimental data, and the modelled DO peak is deeper than the experimental one. This may be caused by an inadequate estimation of the DO saturation concentration C_S or K_{La} . The associated estimated *stBOD* values can be seen in Table 4.11 and compared with the *stBOD* values obtained from other evaluation methods.

TABLE 4.11 – Evaluation of the toxicity check *stBOD* using different techniques.

| Evaluation method | <i>stBOD</i> (mg/l) | <i>stBOD_C</i> (mg/l) | <i>stBOD_N</i> (mg/l) | % _{diff} |
|---------------------|---------------------|---------------------------------|---------------------------------|-------------------|
| Theoretical | 11470 | 4524 | 6946 | — |
| $Area \cdot K_{La}$ | 9244 | — | — | -19 |
| WEST model | 12360 | 4993 | 7367 | 8 |

It can be seen in Table 4.11 that the *stBOD* obtained for the toxicity check using the direct analysis method does not capture the entirety of the theoretical *stBOD*, underestimating it by 19%. This may come from the evaluation of the equilibrium DO concentration C_e , as its value affects the area of the peak being used to evaluate *stBOD* by defining the position of the base of the respirogram surface. An inaccurate evaluation of K_{La} could also contribute to the discrepancy.

Another factor which might contribute to the model's overestimation of $stBOD$ is that the model overestimates the amount of S_{NH} in the wastewater sample. Given the large amount of oxygen needed to oxidize nitrogen compared to carbon ($4.57g\ O_2/g\ N$ vs $1g\ O_2/g\ COD$), any extra S_{NH} has a considerable impact on the global $stBOD$ estimation. In turn, the S_{NH} overestimation may in part be due to the nitrogen assimilation not being accounted for thoroughly enough when using the method outlined in this section. This could have led to an underestimation of the assimilated S_{NH} , and therefore to an over-estimation of the initial S_{NH} concentration in the calibration substrate.

Despite all this, the results of the toxicity check estimation are encouraging, since the estimated substrate concentrations, though not exact, are quite close to their actual measured value. This means that the model can distinguish between the two biomasses' behaviour well and that their metabolism is relatively well characterized.

Retrospective on WEST modelling

It was shown above that it has been possible to extract information on both the $stBOD_C$ and $stBOD_N$ concentrations of the analyzed wastewater samples using the WEST modelling suite in conjunction with respirogram data processed with the computing tools shown in Section 4.4. It was also shown that these tools allow one to estimate biokinetic coefficients $\hat{\mu}_H$, K_S , $\hat{\mu}_A$ and K_{NH} relevant to readily biodegradable substrate consumption in activated sludge. The steps taken to get to these results have also allowed the determination of the time-delay constant τ , which allows for a better modelling of the start-up phase of respirograms. It can, therefore, be said that mathematical modelling is a powerful tool with which one may extract a trove of information from simple respirograms, which further increases the value of the RODTOX's online in-sensor-experiments. Indeed, the obtained biokinetic coefficients may be used for further modelling of the plant's biological treatment. Furthermore, the obtained $stBOD_N$ proves to be an interesting water quality parameter, as it provides a global measurement for all types of nitrifiable nitrogen, which no analytical method can measure. Finally, the obtained $stBOD_C$ measurements provide an estimation of the organic matter which is immediately available to the plant's denitrifiers and (potentially) to phosphorus accumulating organisms, which may help to operate these processes better.

Chapter 5

Conclusions and perspectives

In this M. Sc. thesis, a multitude of tools and methods were developed to make sure that the pilEAUte plant's RODTOX sensor can be fully taken advantage of. With this in mind, it may serve to look back at what was achieved and draw conclusions from the findings. A discussion may then occur in which further steps can be proposed to make sure that the sensor's capabilities are exploited further.

5.1 Conclusions

The implementation of the RODTOX sensor had three main components: the physical installation of the machine into the pilEAUte plant, its setup to allow it to perform reliable measurements over long periods of time, and the development of analysis tools to decode the information contained in its signal. The following conclusions can be drawn from this process:

1. The RODTOX is now online (and on-line):

With the construction and installation of a *fast loop*, the collection of fresh wastewater samples from the primary effluent is now possible for the RODTOX. It which enables this on-line respirometer to monitor the quality of the influent wastewater over time.

2. pH control is important:

It was shown that alkalinity depletion might occur within the RODTOX bioreactor during normal operation. This causes a drop in reactor pH and biomass inhibition. A practical solution to this problem was found by including an appropriate amount of alkalinity in the calibration solution to avoid drops in reactor pH and, consequently, inhibition of the biomass involved in sample biodegradation.

3. The RODTOX's outputs can be modified — but not by its users:

With help from Kelma NV's technical support team, it was possible to extract the DO and temperature sensors' raw data from the RODTOX. However, it was determined that it is not possible for RODTOX users to modify its onboard software, and thus its

output files. The desired outputs must, therefore, be discussed with the respirometer’s manufacturer for customers to get the most use out of their machine.

4. The proof is in the pudding signal:

Only having access to the RODTOX’s raw probe signal, it was shown to be possible within this study to automatically identify most of the different experiments carried out by the RODTOX at any given moment (i.e., either decantation, calibration solution injections, wastewater injections). Though toxicity checks are not yet identified by the analysis tools, a method to do so was proposed in Section 4.4.3.

5. K_{La} can be estimated from a respirogram’s re-aeration phase:

The procedure described in Chapter 3 of Vanrolleghem (1994) can be (and was) applied successfully for the estimation of the mass transfer coefficient K_{La} describing oxygen transfer during each respirogram. This has enabled the estimation of each sample’s *stBOD* without relying on the calibration peaks at all, contrary to the RODTOX’s default method of *stBOD* estimation. K_{La} estimations are also useful by themselves for monitoring, as K_{La} can be affected by certain wastewater components present in the influent (e.g., detergents) (Painmanakul et al., 2005). Thus, estimating K_{La} and monitoring changes in it using the RODTOX upstream of the plant’s bioreactor can be useful for the best operation of the aeration system of the full-scale plant.

6. *stBOD* estimations made through direct analysis of respirograms are plausible:

By comparing the results of *stBOD* estimations made through direct analysis of respirograms with laboratory measurements of ammonia nitrogen and soluble *COD*, it was found that the obtained *stBOD* measurements were in the right order of magnitude to likely be correctly estimated, which confirms that the RODTOX implementation was successful.

7. OUR_{ex} profiles are excellent diagnostic tools for activated sludge:

As can be seen in Section 4.4.3, looking at the OUR_{ex} profiles produced during sample biodegradation can help determine whether one or both organism populations in a sludge sample are inhibited. Depending on the source of this inhibition, this information could have consequences for the plant’s biological treatment. At the very least, though, this information lets RODTOX users know that they should look into what is going wrong with the sludge and take corrective action.

8. Calibration solution injections can successfully be used for sludge characterization:

Using data from calibration peak injections, laboratory measurements and the data processing functions described in Section 4.4, the value of the ASM1 biokinetic parameters which define the RODTOX sludge’s ability to biodegrade readily biodegradable substrate and nitrifiable nitrogen can be estimated. It was also found that expert knowledge was

vital in initializing the parameter estimation process. Moreover, it was noted that the oxygen uptake rate *OUR* is a much more powerful variable to estimate parameters from than the dissolved oxygen concentration. This is consistent with similar parameter estimation studies found in literature (Spanjers and Vanrolleghem, 1995).

9. ***stBOD_C* and *stBOD_N* can be successfully distinguished:**

It is possible to estimate the origin of the oxygen demand of wastewater samples — readily biodegradable substrate and nitrifiable nitrogen— using the biokinetic parameters found during the sludge characterization process. This finding greatly increases the usefulness of the RODTOX’s *stBOD* measurements, as they allow the estimation of substrate concentrations used by two distinct biomasses within the bioreactors. Eventually, knowing the amount of substrate available to each organism population could allow the deployment of independent control actions to optimize both of these populations’ treatment performance, i.e., nitrification, denitrification and enhanced biological phosphorus removal.

These conclusions, when taken together, tell the story of a successful RODTOX implementation within the *pileAUte* plant. Through this implementation, a lot was learned about the best way to operate a RODTOX and how to handle the data it produces, which is what this study was ultimately aiming for.

5.2 Paths for improvement

Though the tools developed in this study increase the usability and usefulness of the RODTOX and its data, there are some further modifications which could be made to increase the quality of the processed data and the reliability of the computing tools. Some of those proposed changes are outlined here:

1. **Uniformize the RODTOX output:**

The RODTOX’s built-in PLC could be modified to send its raw data automatically to the *monEAU* station before each decantation cycle instead of only sending it when the user pushes a button in the interface (see Section 4.3.1). This would not only automate the data transfer process, but it would also uniformize the data series of each file created by the RODTOX, which would help to increase the reliability of the Python scripts. By breaking down the data in “chunks” bookended by decantation cycles, Section 4.3.1’s algorithms would need minimal changes to create vastly more reliable outputs. This “chunking” would also help keep track of the number of wastewater samples analyzed between toxicity checks, which would allow the *Analyze_Respirograms* function to filter them out easily (see Section 4.4.3).

2. **Measure the endogenous oxygen uptake rate:**

The RODTOX’ PLC could also be modified so that its “Endogenous respiration” function is activated once every measurement cycle, as is suggested in Section 4.4.3. Doing so would not only provide more information on the sludge behaviour, but it would also increase the accuracy of the RODTOX’s mathematical model since the endogenous respiration rate used in the objective files could then be based on real measurements instead of being generated by the model itself (see Section 4.5.2).

3. Track the changes in sample volume:

As was discussed in Section 4.4.3, the *Analyze_Respirograms* function assumes that the volume of the wastewater samples is constant over the data series being analyzed. In fact, however, the sample volume can be changed at any time by the RODTOX’s built-in software — it does this to ensure that the respirograms are of adequate size to warrant measurement quality. This occurs when the preceding respirogram’s area is either too large or too small. When this happens, the sample volume is either halved or doubled. This feature of the RODTOX sensor could be accounted for in future versions of the *Analyze_Respirograms* function to produce more reliable *stBOD* measurements.

4. Integrate the analysis tools to the monEAU station:

Now that the Python-based computing tools from Section 4.4 have been developed and tested, they could be integrated to the monEAU monitoring station, where they could be run automatically as soon as a new series of raw data is received from the RODTOX. This would remove the need for manually fetching the data out of the monEAU station before they can be processed. This would only work, of course, if the Python functions were modified such that they could be run one after the other without user intervention. This process would definitely be helped by item (1), as more standardized files would reduce the risk of errors during the functions’ execution.

5. Export data to the datEAUbase:

Given the monEAU station’s already-existing ability to send the data coming from several probes to the datEAUbase, it would be very beneficial to extend the station’s data export process to include the RODTOX’s data. Depending on the realization of item (4), the exported data could include either just the raw data, or both the raw and processed data. This would not only increase the data’s accessibility (which plays a large part in its effective usability), but also its security — redundancy is key to prevent data loss, and the datEAUbase, with its regular backups, provides the redundancy which is lacking in the monEAU station itself.

6. Automate the modelling process:

Finally, another way in which the results of this study could be enhanced concerns the RODTOX’s mathematical model. As it stands, users wishing to estimate the *stBOD_C* and *stBOD_N* of wastewater samples must manually go through each step of the modelling process. This includes inputting the correct physical characterization into the

model, selecting the appropriate data series to use as objective files and carrying out the parameter estimation process for both the sludge being investigated and the wastewater samples that are fed to it. Given the amount of time and effort needed to carry out these steps, the usability of those estimates is significantly reduced at the moment.

To make the process easier, one could work on the automation of most, if not all, of the steps of this modelling process. By developing Python scripts which would interface with the WEST modelling suite, one could automatically feed the RODTOX model the right information for every simulation, as this information comes for the most part from the data processing which was done upstream with the tools presented in Section 4.4. The results of these automated parameter estimation experiments could then be sent to the *datEAU*base along with the rest of the RODTOX's data. Not only would the automation of the modelling process make available on-line estimates of the $stBOD_C$ and the $stBOD_N$, but it would also make available the obtained biokinetic coefficients describing the sludge, which could be put to use for other modelling and control studies on the *pilEAU*te plant.

Together, these changes would make the RODTOX much more user-friendly and more reliable, in addition to creating richer outputs and saving the generated data to a safer location. Regardless of whether or not these features are implemented promptly, however, it can be said that the work performed in this M. Sc. study has laid the groundwork for using the RODTOX in the *pilEAU*te plant for on-line estimations of $stBOD$ and the collection and analysis of respirometric data.

Bibliographie

- American Society of Civil Engineers (2006). *Measurement of Oxygen Transfer in Clean Water : ASCE Standard, ASCE/EWRI 2-06*. Reston, VA, United States : ASCE.
- Brands, E., M. Liebeskind, and M. Dohmann (1994). Parameters for dynamic simulation of wastewater treatment plants with high-rate and low-rate activated-sludge tanks. *Water Science & Technology* 30(4), 211–214.
- Continuum Analytics (2016). Anaconda. <https://www.anaconda.com>.
- Copp, J. B. and P. L. Dold (1998). Confirming the nitrate-to-oxygen conversion factor for denitrification. *Water Research* 32(4), 1296–1304.
- Copp, J. B., H. Spanjers, and P. A. Vanrolleghem (2002). *Respirometry in Control of the Activated Sludge Process : Benchmarking Control Strategies*. London, UK : IWA Publishing.
- De Pauw, D. J. W., G. Sin, G. Insel, S. W. H. Van Hulle, V. Vandenberghe, and P. A. Vanrolleghem (2004). Discussion of "Assessing parameter identifiability of activated sludge model number 1" by Pedro Afonso Cunha and Maria da Conceicao Cunha. *Journal of Environmental Engineering* 130(1), 110–112.
- Dochain, D. and P. A. Vanrolleghem (2001). *Dynamical Modelling & Estimation in Wastewater Treatment Processes*. London, UK : IWA Publishing.
- Dupont, R. and O. Sinkjaer (1994). Optimization of waste-water treatment plants by means of computer-models. *Water Science & Technology* 30(4), 181–190.
- Ekama, G. A., P. L. Dold, and G. v. R. Marais (1986). Procedures for determining influent COD fractions and the maximum specific growth rate of heterotrophs in activated sludge systems. *Water Science & Technology* 18(6), 91 — 114.
- Gapes, D., S. Pratt, Z. Yuan, and J. Keller (2003). Online titrimetric and off-gas analysis for examining nitrification processes in wastewater treatment. *Water Research* 37(11), 2678–2690.

- Gernaey, K. A., B. Petersen, J. P. Ottøy, and P. A. Vanrolleghem (2001). Activated sludge monitoring with combined respirometric - titrimetric measurements. *Water Research* 35(5), 1280–1294.
- Google Inc. (2017). Map data - Université Laval Localization.
- Gouvernement du Québec (2003). *Loi sur la qualité de l'environnement : L.R.Q., c. Q-2, à jour au 9 septembre 2003*. Lois et règlements du Québec. Québec : Éditeur officiel du Québec.
- Guisasola, A., G. Sin, J. A. Baeza, J. Carrera, and P. A. Vanrolleghem (2005). Limitations of ASM1 and ASM3 : a comparison based on batch oxygen uptake rate profiles from different CD full-scale wastewater treatment plants. *Water Science & Technology* 52(10-11), 69–77.
- Gujer, W. (2008). *Systems Analysis for Water Technology* (Illustrate ed.). Springer Berlin Heidelberg.
- Gujer, W., M. Henze, T. Mino, T. Matsuo, M. C. Wentzel, and G. Vonmarais (1995). The Activated Sludge Model No.2 - Biological phosphorous removal. *Water Science & Technology* 31(2), 1–11.
- Gujer, W., M. Henze, T. Mino, and M. C. M. Van Loosdrecht (1999). Activated Sludge Model No. 3. *Water Science & Technology* 39(12), 183–193.
- Guo, L. and P. A. Vanrolleghem (2014). Calibration and validation of an activated sludge model for greenhouse gases no. 1 (ASMG1) : prediction of temperature-dependent N₂O emission dynamics. *Bioprocess and Biosystems Engineering* 37(2), 151–163.
- Hauduc, H., L. Rieger, A. Oehmen, M. C. M. van Loosdrecht, Y. Comeau, A. Heduit, P. A. Vanrolleghem, and S. Gillot (2013, jan). Critical review of activated sludge modeling : State of process knowledge, modeling concepts, and limitations. *Biotechnology and Bioengineering* 110(1), 24–46.
- Henze, M. (1992). Characterization of wastewater for modelling of activated sludge processes. *Water Science & Technology* 25(6), 1 — 15.
- Henze, M., C. P. L. Grady Jr, W. Gujer, G. v. R. Marais, and T. Matsuo (1987). Activated Sludge Model No. 1. *IWAPRC Scientific and Technical Report No.1*, 38. London, UK : IWA Publishing.
- Henze, M., W. Gujer, T. Mino, T. Matsuo, M. C. Wentzel, G. V. R. Marais, and M. C. M. Van Loosdrecht (1999). Activated Sludge Model No.2d, ASM2d. *Water Science & Technology* 39(1), 165–182.

- Henze, M., W. Gujer, T. Mino, and M. Van Loosdrecht (2000). *Activated Sludge Models* (Reprint ed.). Scientific and Technical Report Series. London, UK : IWA Publishing.
- House of Commons of the Parliament of Great Britain (1912). *Royal Commission on Sewage Disposal. Eighth Report of the Commissioners Appointed to Inquire and Report what Methods of Treating and Disposing of Sewage (including Any Liquid from Any Factory Or Manufacturing Process) May Properly be Adopted. Standards an.* Reports of Commissioners. Great Britain. Parliament. Session 1912-13. House of Commons.
- Hulsbeek, J. J. W., J. Kruit, P. J. Roeleveld, and M. C. M. van Loosdrecht (2002). A practical protocol for dynamic modelling of activated sludge systems. *Water Science & Technology* 45(6), 127 — 136.
- Kappeler, J. and W. Gujer (1992). Estimation of kinetic parameters of heterotrophic biomass under aerobic conditions and characterization of wastewater for activated sludge modelling. *Water Science & Technology* 25(6), 125 – 139.
- Keesman, K. J., H. Spanjers, and G. Van Straten (1998). Analysis of endogenous process behavior in activated sludge. *Biotechnology and Bioengineering* 57(2), 155–163.
- Kluyver, T., B. Ragan-Kelley, F. Pérez, B. Granger, M. Bussonnier, J. Frederic, K. Kelley, J. Hamrick, J. Grout, S. Corlay, P. Ivanov, D. Avila, S. Abdalla, C. Willing, and J. development team [Unknown] (2016). Jupyter Notebooks : a publishing format for reproducible computational workflows. IOS Press : Amsterdam, The Netherlands.
- Kong, Z., P. A. Vanrolleghem, P. Willems, and W. Verstraete (1996). Simultaneous determination of inhibition kinetics of carbon oxidation and nitrification with a respirometer. *Water Research* 30(4), 825–836.
- Majone, M., K. Dircks, and J. J. Beun (1999). Aerobic storage under dynamic conditions in activated sludge processes : The state of the art. *Water Science & Technology* 39(1), 61–73.
- McKinney, W. (2011). Pandas : a foundational Python library for data analysis and statistics. <http://pandas.pydata.org>.
- Mesquita, D. P., C. Quintelas, A. Luis Amaral, and E. C. Ferreira (2017). Monitoring biological wastewater treatment processes : recent advances in spectroscopy applications. *Reviews in Environmental Science and Bio-Technology* 16(3), 395–424.
- Metcalf & Eddy (2013). *Wastewater Engineering : Treatment and Resource Recovery* (5, Illustr ed.). McGraw-Hill series in civil and environmental engineering. Columbus, OH, United States : McGraw-Hill Education.
- Oxyguard International A/S (2017). Brochure of OxyGuard model 420. <http://www.oxyguard.dk/products/probes/model-420/>.

- Painmanakul, P., K. Loubiere, G. Hebrard, M. Mietton-Peuchot, and M. Roustan (2005). Effect of Surfactants on Liquid-Side Mass Transfer Coefficients. *Chemical Engineering Science* 60(22), 6480–6491.
- Pérez, F. and B. E. Granger (2007). IPython : A system for interactive scientific computing. *Computing in Science & Engineering* 9(3), 21–29.
- Petersen, B. (2000). *Calibration, Identifiability and Optimal Experimental Design of Activated Sludge Models*. Ph. D. thesis, Ghent University, Belgium.
- Plana, Q. (2015). *Automated Data Collection and Management at Enhanced Lagoons for Wastewater Treatment*. Master’s thesis, Université Laval, Québec, QC, Canada.
- Plosz, B. G., K. H. Langford, and K. V. Thomas (2012). An activated sludge modeling framework for xenobiotic trace chemicals (ASM-X) : Assessment of diclofenac and carbamazepine. *Biotechnology and Bioengineering* 109(11), 2757–2769.
- Plotly Technologies Inc. (2015). Collaborative data science. <https://plot.ly>.
- Pratt, S., Z. Yuan, D. Gapes, M. Dorigo, R. J. Zeng, and J. Keller (2003). Development of a novel titration and off-gas analysis (TOGA) sensor for study of biological processes in wastewater treatment systems. *Biotechnology and Bioengineering* 81(4), 482–495.
- Python Core Team (2016). Python : A dynamic, open-source programming language. <https://www.python.org>.
- Rice, E. W., R. B. Baird, A. D. Eaton, and L. S. Clesceri (2012). *Standard Methods for the Examination of Water and Wastewater* (22 ed.). Denver, CO, United States : American Public Health Association, American Water Works Association, Environment Federation.
- Rieger, L. and P. A. Vanrolleghem (2008). monEAU : a platform for water quality monitoring networks. *WATER SCIENCE AND TECHNOLOGY* 57(7), 1079–1086.
- Roeleveld, P. J. and M. C. M. van Loosdrecht (2002). Experience with guidelines for wastewater characterisation in The Netherlands. *Water Science & Technology* 45(6), 77–87.
- Rowntree, D. (2004). *Statistics Without Tears : A Primer for Non-mathematicians*. Allyn and Bacon classics. London, UK : Pearson/A & B.
- Sin, G. and P. A. Vanrolleghem (2007). Extensions to modeling aerobic carbon degradation using combined respirometric-titrimetric measurements in view of activated sludge model calibration. *Water Research* 41(15), 3345–3358.
- Spanjers, H. and G. Olsson (1992). Modelling of the dissolved oxygen probe response in the improvement of the performance of a continuous respiration meter. *Water Research* 26(7), 945–954.

- Spanjers, H. and P. A. Vanrolleghem (1995). Respirometry as a tool for rapid characterization of wastewater and activated sludge. *Water Science & Technology* 31(2), 105–114.
- Spanjers, H. and P. A. Vanrolleghem (2016). *Respirometry. Experimental Methods in Wastewater Treatment*. Eds. van Loosdrecht, M. C. M., Nielsen, P. H., Lopez-Vazquez, C. M and D. Brdjanovic. London, UK : IWA Publishing.
- Spanjers, H., P. A. Vanrolleghem, G. Olsson, and P. L. Dold (1998). *Respirometry in Control of the Activated Sludge Process*. Scientific and Technical Report No. 7. London, UK : IWA Publishing.
- Sperandio, M., V. Urbain, J. M. Audic, and E. Paul (1999). Use of carbon dioxide evolution rate for determining heterotrophic yield and characterising denitrifying biomass. *Water Science & Technology* 39(1), 139–146.
- STOWA (1996). Methods for wastewater characterization, inventory and guidelines. In *STOWA report 97-23*. Zwijndrecht, The Netherlands : Hageman Fulfilment.
- Surmacz-Gorska, J., K. A. Gernaey, C. Demuynck, P. A. Vanrolleghem, and W. Verstraete (1996). Nitrification monitoring in activated sludge by oxygen uptake rate (OUR) measurements. *Water Research* 30(5), 1228–1236.
- US-EPA (2010). *National Pollutant Discharge Elimination System (NPDES) - Permit Writers' Manual*. Washington, DC, United States : DIANE Publishing.
- van der Walt, S., S. C. Colbert, and G. Varoquaux (2011). The NumPy Array : A structure for efficient numerical computation. *Computing in Science & Engineering* 13(2), 22–30.
- van Rijswick, H. F. M. W. and H. J. M. Havekes (2012). *European and Dutch Water Law*. Deventer, The Netherlands : Europa Law Publishing.
- Vanrolleghem, P. A. (1993). Parameter estimation of wastewater treatment models : An academic perspective. In *Workshop on Research Projects in Modelling and Control of Wastewater Treatment Plants.*, Copenhagen, Denmark. June 10-23 1993.
- Vanrolleghem, P. A. (1994). *On-line Modelling of Activated Sludge Processes : Development of an Adaptive Sensor*. Ph. D. thesis, Gent University, Belgium.
- Vanrolleghem, P. A. and F. Coen (1995). Optimal-design of In-Sensor-Experiments for online modeling of nitrogen removal processes. *Water Science and Technology* 31(2), 149–160.
- Vanrolleghem, P. A., K. A. Gernaey, B. Petersen, B. De Clercq, F. Coen, and J. P. Ottoy (1998). Limitations of short-term experiments designed for identification of activated sludge biodegradation models by fast-dynamic phenomena. In *Proceedings 7th IFAC Conference on Computer Applications in Biotechnology CAB7*, Osaka, Japan, pp. 567–572.

- Vanrolleghem, P. A., Z. Kong, G. Rombouts, and W. Verstraete (1994). An on-line respirographic biosensor for the characterization of load and toxicity of wastewaters. *Journal of Chemical Technology & Biotechnology* 59(4), 321–333.
- Vanrolleghem, P. A., B. Petersen, D. J. W. De Pauw, H. Dovermann, S. Weijers, and K. A. Gernaey (2003). A comprehensive model calibration procedure for activated sludge models. In *Proceedings of the Water Environment Federation, WEFTEC 2003*, Los Angeles, CA, United States, October 11-15, 2003. (on CD-ROM).
- Vanrolleghem, P. A., G. Sin, and K. V. Gernaey (2004). Transient response of aerobic and anoxic activated sludge activities to sudden substrate concentration changes. *Biotechnology and Bioengineering* 86(3), 277–290.
- Vanrolleghem, P. A., H. Spanjers, B. Petersen, P. Ginestet, and I. Takacs (1999). Estimating (combinations of) Activated Sludge Model No. 1 parameters and components by respirometry. *Water Science & Technology* 39(1), 195–214.
- Vanrolleghem, P. A. and M. Van Daele (1994). Optimal experimental design for structure characterization of biodegradation models - online implementation in a respirographic biosensor. *Water Science and Technology* 30(4), 243–253.
- Vanrolleghem, P. A., M. Van Daele, and D. Dochain (1995, nov). Practical identifiability of a biokinetic model of activated-sludge respiration. *Water Research* 29(11), 2561–2570.
- Vanrolleghem, P. A. and W. Verstraete (1993). Simultaneous biokinetic characterization of heterotrophic and nitrifying populations of activated sludge with an online respirographic biosensor. *Water Science & Technology* 28(11-12), 377–387.
- Vesilind, P. A., J. J. Peirce, and R. F. Weiner (2013). *Environmental Engineering*. Elsevier Science.
- Volskay, V. T. and C. P. L. Grady (1990). Respiration inhibition kinetic analysis. *Water Research* 24(7), 863–874.

Annexe A

RODToX python functions

A.1 Function to import RODToX data into a python environment

```
1 def ImportRdtxCsv(csvDO, csvTemp, Existing, sourcepath, destinationpath):
2     import pandas as pd
3
4     #Create the complete filepath to get to the data
5     filepath = sourcepath + '/' + csvDO
6     #Read the.csv file and turn it into a DataFrame
7     RawDO = pd.read_csv(filepath, sep=';')
8     #Remove invalid entries from the data
9     RawDO = RawDO[RawDO.Validity == 1]
10    #Only keep rows where the probe sends a DO measurement
11    RawDO = RawDO[RawDO.VarName == 'HMI_DO']
12    #Replace commas by dots so that DO values are treated as floats instead of
    strings
13    RawDO.replace({' ','': '.'}, regex=True, inplace = True)
14    #Transform the timestamps into machine-readable DateTime objects
15    RawDO['Time'] = pd.to_datetime(RawDO.TimeString, format='%d.%m.%Y %H:%M:%S')
16    #Assign new name to the columns
17    RawDO.columns = ['VarName', 'TimeString', 'DO', 'Validity', 'Time_ms', 'Time']
18    #Set the 'Time' column as the DataFrame's index
19    RawDO.set_index(RawDO.Time, inplace = True) # set the Time column as the
    DataFrame's index
20    #Transform the DO values from strings into numbers
21    RawDO.DO = pd.to_numeric(RawDO.DO)
22
23    #Repeat the same process for Temperature data.
24    filepath = sourcepath + '/' + csvTemp
25    RawTemp = pd.read_csv(filepath, sep=';')
26    RawTemp = RawTemp[RawTemp.Validity == 1]
27    RawTemp = RawTemp[RawTemp.VarName == 'HMI_Temp']
28    RawTemp.replace({' ','': '.'}, regex=True, inplace = True)
```

```

29 RawTemp['Time']=pd.to_datetime(RawTemp.TimeString,format='%d.%m.%Y %H:%M:%S')
30 RawTemp.columns = ['VarName', 'TimeString', 'Temp', 'Validity', 'Time_ms', '
Time']
31 RawTemp.set_index(RawTemp.Time, inplace = True)
32 RawTemp.Temp = pd.to_numeric(RawTemp.Temp)
33
34 #Concatenate the temperature and dissolved oxygen data, and drop useless
columns and empty rows
35 df = pd.concat([RawDO.DO, RawTemp.Temp], axis = 1, keys = ['DO', 'Temp'])
36 df.dropna(inplace = True)
37 df.sort_index(inplace=True)
38
39 #If we do not want to merge the new data with existing files, save the new
data to a new file
40 if Existing == 'None':
41     #Find the earliest timestamp
42     Start = str(df.first_valid_index())
43     #Find the latest timestamp
44     End = str(df.last_valid_index())
45     #Save the resulting file to the correct directory
46     dest_filepath=destinationpath+'/' +str(Start)[:10]+'_'+str(End)[:10]+'.csv'
47     df.to_csv(dest_filepath, sep = ';')
48
49 #If we do want to merge the new data with an existing file:
50 else:
51     #read the existing file
52     filepath=destinationpath+'/' +Existing
53     df_ex = pd.read_csv(filepath, sep=';')
54     df_ex.Time = pd.to_datetime(df_ex.Time)
55     df_ex.DO = pd.to_numeric(df_ex.DO)
56     df_ex.Temp = pd.to_numeric(df_ex.Temp)
57     df_ex.set_index(df_ex.Time, drop=True, inplace=True)
58
59     #Append one data frame to another
60     df_ex = df_ex.append(df)
61     df_ex.drop_duplicates(inplace = True)
62     df_ex.sort_index(inplace=True)
63
64     #Find the earliest timestamp
65     Start = str(df_ex.first_valid_index())
66     #Find the latest timestamp
67     End = str(df_ex.last_valid_index())
68     #Save the resulting file to the correct directory
69     dest_filepath=destinationpath+'/' +str(Start)[:10]+'_'+str(End)[:10]+'.csv'
70     df_ex.to_csv(dest_filepath, sep = ';')
71     df = df_ex
72
73 #Plot the imported DO and Temperature data to make sure it looks OK

```

```
74 Plotit(df, Start, End, ['DO'], ['Temp'], ['DO'], ['Temperature'], ['mg/l'], ['  
    °C'], ['line', 'line'])  
75  
76 #Return the resulting DataFrame containg the imported DO and temperature data,  
77 #as well as the earliest and lates timestamps  
78 return df, Start, End
```


A.2 Function to smooth the RODTOX signal and to obtain the first and second derivative

```
1 def Deriv(df, tau, n):
2
3     #The rolling_window function helps to calculate moving-window statistics. ©
4     #2016 Erik Rigtorp
5     #http://www.rigtorp.se/2011/01/01/rolling-statistics-numpy.html
6     def rolling_window(a, window):
7         shape = a.shape[:-1] + (a.shape[-1] - window + 1, window)
8         strides = a.strides + (a.strides[-1],)
9         return np.lib.stride_tricks.as_strided(a, shape=shape, strides=strides)
10
11     DF = df.copy(deep=True)
12     if n%2==0:
13         n+=1
14     NewColumns = ['DOdt_A', 'DOdt_smooth', 'DO_A', 'DO_smooth', 'seconds']
15     for name in NewColumns:
16         if name not in DF.columns:
17             DF.insert(len(DF.columns), name, np.nan)
18     DF.reset_index(inplace=True)
19     DF['seconds'] = (DF['Time']-DF['Time'].iloc[0]).dt.total_seconds()
20     dist=int((n-1)/2)
21     ##### Calculate the Derivative of the raw signal
22     x = DF['seconds']
23     y = DF['DO']
24
25     RolledX=rolling_window(x,n)
26     RolledY=rolling_window(y,n)
27     x_mean = np.mean(RolledX,axis=1)
28     y_mean = np.mean(RolledY,axis=1)
29     x_xmean = np.transpose(np.transpose(RolledX)-x_mean)
30     y_ymean = np.transpose(np.transpose(RolledY)-y_mean)
31
32     prodxy=np.multiply(x_xmean,y_ymean)
33     prodxx = np.multiply(x_xmean,x_xmean)
34
35     Sumxy = np.sum(prodxy, axis=1)
36     Sumxx = np.sum(prodxx, axis=1)
37     Slope= np.divide(Sumxy,Sumxx)
38     Intercept = y_mean - Slope * x_mean
39     y_hat = Intercept+Slope*x[dist:-(dist)]
40
41     ##### Assign the results to the right columns
42     DF = DF.iloc[dist:-(dist)]
43     DF['DOdt_smooth'] = Slope*3600
44     DF['DO_smooth'] = y_hat
45     DF['DO_A'] = DF['DO_smooth'] + tau*DF['DOdt_smooth']/3600
```

```

45
46     ### Calculate the derivative of the DO_Actual time series
47     x = DF['seconds']
48     RolledX=rolling_window(x,n)
49     x_mean = np.mean(RolledX,axis=1)
50     x_xmean = np.transpose(np.transpose(RolledX)-x_mean)
51     prodx = np.multiply(x_xmean,x_xmean)
52     Sumxx = np.sum(prodx, axis=1)
53
54     DOA = DF['DO_A']
55     RolledDOA=rolling_window(DOA,n)
56     DOA_mean = np.mean(RolledDOA,axis=1)
57     DOA_DOAmean = np.transpose(np.transpose(RolledDOA)-DOA_mean)
58     proxdOA=np.multiply(x_xmean,DOA_DOAmean)
59     SumxDOA = np.sum(proxdOA, axis=1)
60     SlopeDOA= np.divide(SumxDOA,Sumxx)
61     DF = DF.iloc[dist:-(dist)]
62     DF['DOdt_A'] = SlopeDOA * 3600
63
64     DF.set_index(DF['Time'], inplace=True)
65     DF = DF.drop(['Time'], 1)
66     return DF

```

A.3 Function to analyze every respirogram to determine K_{La} and $stBOD$

```

1 def Analyze_Respirograms(df, Start, End, Sample_Volume, Cal_Volume): #Volumes in
    liters
2     #Import libraries used within the function
3     import scipy.stats
4     import scipy.optimize
5     import numpy as np
6     import scipy.integrate
7
8     #The rolling_window function helps to calculate rolling statistics. © 2016
    Erik Rigtorp
9     #http://www.rigtorp.se/2011/01/01/rolling-statistics-numpy.html
10    def rolling_window(a, window):
11        shape = a.shape[:-1] + (a.shape[-1] - window + 1, window)
12        strides = a.strides + (a.strides[-1],)
13        return np.lib.stride_tricks.as_strided(a, shape=shape, strides=strides)
14
15    ##### Make a copy of the DataFrame to work on #####
16    DF = df[Start:End].copy(deep=True)
17
18    ##### Add the new columns we need to the DataFrame #####
19    NewColumns = ['Co', 'Ce', 'Kla', 'Area', 'stBOD', 'OURex', 'PeakNo', 'ReacVol',
20                  'DO_Start', 'DO_End', 'DO_Bend', 'DO_Bottom', 'x_plot', 'y_plot']
21    for name in NewColumns:
22        if name not in DF.columns:
23            DF.insert(len(DF.columns), name, np.nan)
24
25    ##### Find important points in peaks (Start, End, Bend, Bottom) #####
26    #Filter the DataFrame to find the starting point of a respirogram
27    Percentile = np.percentile(DF['DO_A'], 70)
28    #Filtered= DF.loc[(DF.index==DF.first_valid_index()) | ((DF['DO_A'] >
    Percentile) & (DF['DOdt_A'].diff(-1) > 0) & (DF['DOdt_A'] > 0) & (DF['DOdt_A
    '].shift(-1) < 0) & (DF['DO_A'] > DF['DO_A'].shift(-1)))]
29    Filtered= DF.loc[(DF.index==DF.first_valid_index()) | ((DF['DO_A'] >
    Percentile) & (DF['DOdt_A'] > 0) & (DF['DOdt_A'].shift(-1) < 0))]
30
31    Filtered['Start']=pd.to_datetime(Filtered.index)
32    Filtered['DO_Start'] = Filtered['DO_A']
33    #remove false positives by removing the Start points that are too close
    together in time
34    i=0
35    while i < len(Filtered)-1:
36        if pd.to_timedelta(Filtered['Start'].iloc[i+1]-Filtered['Start'].iloc[i]).
    total_seconds() <= pd.to_timedelta('5 minutes').total_seconds():
37            Filtered=Filtered.drop(Filtered.index[i])
38            i=0

```

```

39     else:
40         i+=1
41     #Add columns to the filtered DataFrame for the other key points in the
42     respirogram
43     Filtered = Filtered.reset_index(drop = True)
44     Columns = ['End', 'Bend', 'Bottom']
45     for name in Columns:
46         Filtered.insert(len(Filtered.columns), name, np.nan)
47
48     for i in range(len(Filtered)-1):
49         # Defining the end of each respirogram
50         Filtered['End'].iloc[i] = Filtered['Start'].iloc[i+1]
51         Filtered['DO_End'].iloc[i] = DF['DO_A'][Filtered['End'].iloc[i]]
52
53         S=Filtered['Start'].iloc[i]
54         E=Filtered['End'].iloc[i]
55         # Find the lowest point of each respirogram
56         Bottom = pd.to_datetime(DF[S:E].loc[DF[S:E]['DO_A']== DF[S:E]['DO_A'].min
57         ()].index)
58         Filtered['Bottom'].iloc[i] = pd.to_datetime(DF.iloc[[DF.index.get_loc(
59         Bottom, method='nearest')]].index)
60         Filtered['DO_Bottom'].iloc[i] = DF['DO_A'][Filtered['Bottom'].iloc[i]]
61
62         ##### Finding the "Bend" in each respirogram - the point at which
63         biodegradation is complete
64         # The Bend corresponds to the last peak of the derivative over the span of
65         the respirogram
66         bendS = Filtered['Bottom'].iloc[i]
67         bendE = Filtered['End'].iloc[i]
68         benddf = DF[bendS:bendE].copy(deep=True)
69         benddf.reset_index(inplace=True)
70         n=51
71         dist=int((n-1)/2)
72         x = benddf['Time'].iloc[dist:-dist]
73
74         if len(benddf) <= 1.5*n:
75             Filtered['Bend'].iloc[i] = bendS
76             Filtered['DO_Bend'].iloc[i] = np.nan
77         else:
78             RolledDO = rolling_window(benddf['DOdt_A'], n)
79             RolledMax = np.max(RolledDO, axis=1)
80
81             Compa = pd.DataFrame({'DOdt':benddf['DOdt_A'].iloc[dist:-dist],
82                                 'Max':RolledMax,
83                                 'Time':x})
84
85             Compa=Compa.loc[Compa['DOdt']==Compa['Max']]

```

```

82
83     Compa.reset_index(inplace=True,drop=True)
84     if len(Compa)==0:
85         Filtered['Bend'].iloc[i] = bendS
86         Filtered['DO_Bend'].iloc[i] = np.nan
87     else:
88         LastMax = Compa.iloc[Compa.last_valid_index()]
89         Filtered['Bend'].iloc[i] = LastMax['Time']
90         Filtered['DO_Bend'].iloc[i] = DF['DO_A'].loc[Filtered['Bend'].iloc
91 [i]]
92
93     # If the interval between two respirograms is too large, the End of the
94     first respirogram is set (arbitrarily) to 15 minutes after the end of
95     biodegradation.
96     if pd.to_timedelta(Filtered['End'].iloc[i]-Filtered['Start'].iloc[i])>pd.
97     to_timedelta('2 hours'):
98         Set= pd.to_datetime(Filtered['Bend'].iloc[i])+pd.to_timedelta('15
99 minutes')
100         Filtered['End'].iloc[i] = pd.to_datetime(DF.iloc[[DF.index.get_loc(Set
101 , method='nearest')]].index)[0]
102         Filtered['DO_End'].iloc[i] = DF['DO_A'][Filtered['End'].iloc[i]]
103
104
105     S=DF.first_valid_index()
106     E=Filtered['Bottom'].iloc[0]
107     Top = pd.to_datetime(DF[S:E].loc[DF[S:E]['DO_A']== DF[S:E]['DO_A'].max()]).
108     index)
109     Filtered['Start'].iloc[0] = pd.to_datetime(DF.iloc[[DF.index.get_loc(Top,
110 method='nearest')]].index)
111     Filtered['DO_Start'].iloc[0] = DF['DO_A'][Filtered['Start'].iloc[0]]
112
113     #Drops the last row of the filtered DataFrame, which only contains a Start
114     point with no corresponding End, Bottom or Bend
115     Filtered = Filtered[:-1]
116
117     ##### Identifying Decantation peaks and assigning Peak ID's
118     #####
119
120
121     for i in range(len(Filtered)):
122         #Check if a respirogram's DO falls below the threshold under which only
123         decantation cycles fall
124         #if Filtered['DO_Bottom'].iloc[i] <= float(Decthresh):
125         if ((Filtered['Bend'].iloc[i]-Filtered['Start'].iloc[i]).total_seconds()
126 >= 1700) & ((Filtered['Bend'].iloc[i]-Filtered['Start'].iloc[i]).total_seconds
127 () <= 1920) :
128
129
130         #Reset the Peak identification number and the reactor volume

```

```

117         Filtered['PeakNo'].iloc[i] = 0
118         Filtered['ReacVol'].iloc[i] = 10.0
119
120     #Assign the Important points to the appropriate rows in the original
    DataFrame
121
122     DF['DO_Start'][Filtered['Start'][i]]= Filtered['DO_Start'][i]
123     DF['DO_End'][Filtered['End'][i]]= Filtered['DO_End'][i]
124     DF['DO_Bend'][Filtered['Bend'][i]]= Filtered['DO_Bend'][i]
125     DF['DO_Bottom'][Filtered['Bottom'][i]]= Filtered['DO_Bottom'][i]
126
127     ##### Plotting the Important points #####
128     Label1 = ['DO']
129     Label2 = ['Start','End','Bend','Bottom']
130     Units1 = ['mg/l']
131     Units2 = ['mg/l','mg/l','mg/l','mg/l']
132     marks = ['line','markers']
133     Plotit(DF, Starttime, Endtime, ['DO_A'], ['DO_Start','DO_End','DO_Bend','
    DO_Bottom'], Label1, Label2, Units1, Units2, marks)
134
135     ##### Assinging peak identification numbers and volumes for each respirogram
    #####
136     for i in range(1, len(Filtered)):
137         if Filtered['PeakNo'].iloc[i] == 0:
138             continue
139         elif Filtered['PeakNo'].iloc[i-1]==0:
140             Filtered['PeakNo'].iloc[i] = 1
141             Filtered['ReacVol'].iloc[i] = 10+Cal_Volume
142         elif Filtered['PeakNo'].iloc[i-1]==1:
143             Filtered['PeakNo'].iloc[i] = 2
144             Filtered['ReacVol'].iloc[i] = 10+2*Cal_Volume
145         else:
146             Filtered['PeakNo'].iloc[i] = Filtered['PeakNo'].iloc[i-1]+1
147             Filtered['ReacVol'].iloc[i] = Filtered['ReacVol'].iloc[i-1]+
    Sample_Volume
148
149     ##### Kla estimation #####
150
151     #Define the re-aeration function under conditions without biodegradation
152     def f(t, Kla, Ce, Co):
153         return Ce-(Ce-Co)*np.exp(-Kla*(t))
154
155     #Add relevant columns to the Filtered DataFrame
156     Kla_columns = ['Co','Ce','Kla','std_err']
157     for name in Kla_columns:
158         if name not in Filtered.columns:
159             Filtered.insert(len(Filtered.columns),name, np.nan)
160

```

```

161     for i in range(len(Filtered)):
162         if ((Filtered['Bend'].iloc[i] == Filtered['End'].iloc[i])|(pd.to_timedelta
(Filtered['End'].iloc[i]-Filtered['Bend'].iloc[i]).total_seconds() <= pd.
to_timedelta('8 minutes').total_seconds())):
163             print ('Peak %s of %s is skipped: Interval too short.'
164                     %(i+1,len(Filtered)))
165             continue
166         else:
167             #Define a DataFrame containing only the re-aeration phase of
respirogram i
168             T00 = Filtered['Bend'][i]
169             Tf = Filtered['End'][i]
170             print ('Peak %s of %s. Start is at %s'
171                     %(i+1,len(Filtered),T00))
172
173             Timedel =pd.to_timedelta(Tf-T00)
174             rollingdf = DF[T00 : T00+0.9*Timedel]
175             rollingdf['tvalue'] = rollingdf.index
176
177             #Create a DataFrame to contain the results of non-linear regressions
operated on sub-sections of the re-aeration phase
178             SubResult = pd.DataFrame(index=range(0,int(0.5*len(rollingdf))),
columns=['Start','End','Ce','Co','Kla','std_err'])
179
180             #Try to perform a non-linear regression on several sub-section of the
re-aeration phases
181             for j in range(len(rollingdf)):
182                 #The beginning of the range of the non-linear regression shifts
one data point forward at each iteration
183                 #The final value in the range stays constant
184                 T0 = rollingdf['tvalue'].iloc[[j]][0]
185                 y_given = DF['DO_A'][T0:Tf].copy()
186                 y_frame = y_given.to_frame()
187                 y_frame.reset_index(drop = True, inplace = True)
188                 x_given = (DF[T0:Tf].index-T0).total_seconds()
189
190                 try:
191                     #Define the bounds of the non-linear regression for parameters
Kla, Ce, Co
192                     param_bounds=([0,0,-20],[100/3600,20,20])
193                     #Try to perform the non-linear regression
194                     params, cov = scipy.optimize.curve_fit(f, x_given, y_frame['
DO_A'], bounds = param_bounds)
195                     #Move on to the next sub-section if the non-linear regression
fails
196                     except RuntimeError:
197                         continue
198

```

```

199         #If it doesn't fail, assign the results of the non-linear
regression to the SubResult DataFrame
200     else:
201         SubResult['Start'][j] = pd.to_datetime(T0)
202         SubResult['End'][j] = pd.to_datetime(Tf)
203         SubResult['Ce'][j] = params[1]
204         SubResult['Co'][j] = params[2]
205         SubResult['Kla'][j] = params[0]
206         perr = np.sqrt(np.diag(cov))
207         sterr = perr[0]
208         SubResult['std_err'][j]= sterr
209     #And delete stored variables which are no longer needed
210     finally:
211         del y_given
212         del y_frame
213         del x_given
214     #If the non-linear regressions have succeeded at least once over all
the sub-sections of the re-aeration phase:
215     if len(SubResult) != 0:
216         #Select the values of Ce, Co and Kla for the iteration with the
lowest standard error
217         index_position = ((SubResult['std_err']**2).idxmin())
218         Filtered['Ce'][i] = SubResult['Ce'][index_position]
219         Filtered['Co'][i] = SubResult['Co'][index_position]
220         Filtered['Kla'][i] = SubResult['Kla'][index_position]*3600
221         Filtered['std_err'][i] = SubResult['std_err'][index_position]
222         #Calculate the corresponding DO saturation concentration
223
224     #If all non-linear regressions failed, move on to the next respirogram
without assigning values
225     else:
226         continue
227     #Fill the Filtered DataFrame with values from the nearest successful non-
linear regression
228     Columns = ['Ce','Co','Kla','std_err']
229     for name in Columns:
230         Filtered[name] = Filtered[name].fillna(method='ffill')
231     # Assign the obtained Kla, Ce, Co and Cs values to the original DataFrame
232     Columns = ['Kla','Ce','Co']
233     for name in Columns:
234         for i in range(len(Filtered)):
235             DF[name][Filtered['Start'][i]:Filtered['End'][i]] = Filtered[name][i]
236
237     #Re-calculate the expected DO concentrations during re-aeration using the
found Kla, Ce and Co values and plot them
238     def f2(t, Ce, Co, Kla, to):
239         return Ce-(Ce-Co)*np.exp(-Kla*pd.to_timedelta(t-to).astype('timedelta64[s]
'))

```



```

240
241 DF['x_plot']=pd.to_datetime(DF.index)
242 for i in range(len(Filtered)):
243     DF['y_plot'].loc[Filtered['Bend'][i]:Filtered['End'][i]]= f2(DF['x_plot'],
244         DF['Ce'], DF['Co'], DF['Kla']/3600, Filtered['Bend'][i] )
245
246 Label1 = ['DO exp.','DO calc.','Ce']
247 Label2 = ['KLa']
248 Units1 = ['mg/l','mg/l','mg/l']
249 Units2 = ['h-1']
250 marks = ['line','line']
251 Plotit(DF, Starttime, Endtime, ['DO_A','y_plot','Ce'], ['Kla'], Label1, Label2
252 , Units1, Units2, marks)
253
254 ##### Calculate OURex #####
255
256 DF['OURex'] = DF['Kla']*(DF['Ce']-DF['DO_A'])-DF['DOdt_A']
257
258 ##### Calculate the surface of the respirograms #####
259 #Create the needed columns in the Filtered DataFrame
260 Columns = ['Area','stBOD','Mass','intdt']
261 for name in Columns:
262     if name not in Filtered.columns:
263         Filtered.insert(len(Filtered.columns),name, np.nan)
264
265 # Calculate the difference between the equilibrium DO concentration and the
266 actual Do concentration throughout the Data
267 DF['Deficit'] = DF['Ce']-DF['DO_A']
268
269 for i in range(len(Filtered)):
270     Si = Filtered ['Start'][i]
271     Ei = Filtered ['End'][i]
272
273     #Calculate the Area of each respirogram, and integrate the DO
274     concentration gradient throughout the respirogram (should be near zero)
275     Filtered['Area'].iloc[i] = scipy.integrate.trapz(DF['Deficit'][Si:Ei].
276         values, DF['Deficit'][Si:Ei].index.astype(np.int64) / 10**9)
277     Filtered['intdt'].iloc[i] = scipy.integrate.trapz(DF['DOdt_A'][Si:Ei].
278         values/3600, DF['DOdt_A'][Si:Ei].index.astype(np.int64) / 10**9)
279
280     #Calculate the stBOD mass indicated by each respirogram, and convert it to
281     a sample concentration according to the volume of the sample
282     Filtered['Mass'].iloc[i] = (-Filtered['intdt'].iloc[i]+Filtered['Area'].
283         iloc[i]* Filtered['Kla'].iloc[i]/3600)*Filtered['ReacVol'].iloc[i]
284
285     #Return different stBOD concentrations for the respirogram's associated
286     sample according their volume (which changes with the peak identification
287     number)

```

```

278         if Filtered['PeakNo'].iloc[i] == 0:
279             Filtered['stBOD'].iloc[i] = np.nan
280         elif (Filtered['PeakNo'].iloc[i] == 1) | (Filtered['PeakNo'].iloc[i] == 2):
281             Filtered['stBOD'].iloc[i] = np.nan#Filtered['Mass'].iloc[i]/
Cal_Volume
282         else:
283             Filtered['stBOD'].iloc[i] = Filtered['Mass'].iloc[i]/Sample_Volume
284
285     #Assign stBOD measurements to the main DataFrame
286     stBODResults = Filtered.loc[Filtered['PeakNo'] >= 3.0]
287     stBODResults = stBODResults.reset_index(drop=True)
288
289     length = len(stBODResults)-1
290     for i in range(length):
291         DF['stBOD'][stBODResults['Start'][i]:stBODResults['Start'][i+1]] =
stBODResults['stBOD'][i]
292     DF['stBOD'][stBODResults['Start'][length]:stBODResults['End'][length]] =
stBODResults['stBOD'][length]
293
294     ##### Plot the resulting stBOD measurements #####
295
296     Label1 = ['DO', 'Ce']
297     Label2 = ['stBOD']
298     Units1 = ['mg/l', 'mg/l']
299     Units2 = ['mg/l']
300     marks = ['line', 'line']
301     Plotit(DF, Start, End, ['DO_A', 'Ce'], ['stBOD'], Label1, Label2, Units1,
Units2, marks)
302
303     #Select the dataframe rows to return to the user in the Results DataFrame (
with the relevant data for each respirogram)
304     Return_stBOD = ['Start', 'End', 'Ce', 'Kla', 'stBOD', 'PeakNo']
305     Return_DF = ['DO', 'Temp', 'DOdt_A', 'DOdt_smooth', 'DO_A', 'DO_smooth', 'seconds', '
Ce', 'Kla', 'stBOD', 'OURex', 'PeakNo', 'ReacVol']
306     return DF.loc[:, Return_DF], stBODResults.loc[:, Return_stBOD]

```

A.4 Function to plot respirograms and their associated parameters

```
1 def Plotit(df, Starttime, Endtime, List_y1, List_y2, Label1, Label2, Units1,
2           Units2, marks):
3     import plotly.offline as offline
4     dashes = ['', 'dot', 'dashdot', '', 'dot', 'dashdot']
5
6     symbollist=[22,23,19,20,4,17,1]
7     colorlist_1 = ['#5385A1']
8     colorlist_2 = ['#70CC87', '#3E1D1E']
9     colorlist_3 = ['#70CC87', '#546D8A', '#3E1D1E']
10    colorlist_4 = ['#70CC87', '#3A909D', '#5F4C68', '#3E1D1E']
11    colorlist_5 = ['#70CC87', '#3A909D', '#546D8A', '#5F4C68', '#3E1D1E']
12    colorlist_6 = ['#70CC87', '#34AA9D', '#458298', '#5D5977', '#5A3648', '#3E1D1E']
13    colorlist_7 = ['#70CC87', '#39B09A', '#3A909D', '#546D8A', '#5F4C68', '#563140', '#3E1D1E']
14    colorlist_8 = ['#70CC87', '#3FB499', '#34999E', '#4A7C95', '#5B5F7D', '#5F445C', '#542D3B', '#3E1D1E']
15
16    markerlist=[]
17    Number_of_lines = len(List_y1)+len(List_y2)
18    for i in range(Number_of_lines):
19        if Number_of_lines ==1:
20            dictA={'color':colorlist_1[i], 'size': '10', 'symbol':symbollist[i]}
21            markerlist.append(dictA)
22        elif Number_of_lines ==2:
23            dictA={'color':colorlist_2[i], 'size': '10', 'symbol':symbollist[i]}
24            markerlist.append(dictA)
25        elif Number_of_lines ==3:
26            dictA={'color':colorlist_3[i], 'size': '10', 'symbol':symbollist[i]}
27            markerlist.append(dictA)
28        elif Number_of_lines ==4:
29            dictA={'color':colorlist_4[i], 'size': '10', 'symbol':symbollist[i]}
30            markerlist.append(dictA)
31        elif Number_of_lines ==5:
32            dictA={'color':colorlist_5[i], 'size': '10', 'symbol':symbollist[i]}
33            markerlist.append(dictA)
34        elif Number_of_lines ==6:
35            dictA={'color':colorlist_6[i], 'size': '10', 'symbol':symbollist[i]}
36            markerlist.append(dictA)
37        elif Number_of_lines ==7:
38            dictA={'color':colorlist_7[i], 'size': '10', 'symbol':symbollist[i]}
39            markerlist.append(dictA)
40        elif Number_of_lines ==8:
41            dictA={'color':colorlist_8[i], 'size': '10', 'symbol':symbollist[i]}
42            markerlist.append(dictA)
```

```

43 x = df[Starttime:Endtime].index
44 #x = pd.to_timedelta(x-x[0]).total_seconds()
45 #x=x/3600
46
47 Lab1 = ''
48 for i in range(len(Label1)):
49     if len(Label1)==1:
50         Lab1 += Label1[i] + ' (%s) '%Units1[i]
51     elif i == len(Label1)-1:
52         Lab1 += 'and '+Label1[i] + ' (%s) '%Units1[i]
53     else:
54         Lab1 += Label1[i] + ' (%s), '%Units1[i]
55
56 Lab2=''
57 for i in range(len(Label2)):
58     if len(Label2)==1:
59         Lab2 += Label2[i] + ' (%s) '%Units2[i]
60     elif i == len(Label2)-1:
61         Lab2 += 'and '+Label2[i] + ' (%s) '%Units2[i]
62     else:
63         Lab2 += Label2[i] + ' (%s), '%Units2[i]
64 layout = go.Layout(
65     font=dict(
66         family='PT Sans Narrow',
67         size=20,
68         color='#000000'
69     ),
70     yaxis=dict(
71         title=Lab1,
72         ticks='-',
73         titlefont=dict(
74             color='#000000',
75             size=20
76         ),
77         autorange=True,
78         rangemode='tozero', #normal or nonnegative
79         linecolor='#000000',
80         zeroline=False,
81         showline=True,
82     ),
83     xaxis=dict(
84         showticklabels=True,
85         showline=True,
86         ticks='-',
87         title='Time (h)',
88         titlefont=dict(
89             color='#000000',
90             size=20

```

```

91         ),
92         #tickangle=45,
93         #tickformat='%m/%d %H:%M'
94     ),
95
96     legend=dict(
97         x=-0.35,
98         y=1,
99         traceorder='normal',
100         font=dict(
101             family='PT Sans Narrow',
102             size=20,
103             color='#000000'
104         ),
105     ),
106     autosize=False,
107     width=900,
108     height=500,
109     margin=go.Margin(
110         l=50,
111         r=100,
112         b=100,
113         t=50,
114         pad=4
115     ),
116     paper_bgcolor='#FFFFFF',
117     plot_bgcolor='#FFFFFF'
118
119 )
120 yaxis2 = {'yaxis2': {'title': Lab2,
121                     'ticks': '-',
122                     'titlefont': {'color': '#000000', 'size': 20},
123                     'rangemode': 'nonnegative',
124                     'autorange': True,
125                     'linecolor': '#000000',
126                     'zeroline': False,
127                     'showline': True,
128                     'anchor': 'x',
129                     'overlying': 'y',
130                     'side': 'right',
131                     'rangemode': 'tozero',
132                     'showgrid': False
133                     }
134         }
135 if len(List_y2) != 0:
136     layout.update(yaxis2)
137
138 data = []

```

```

139     for i in range(len(List_y1)):
140         trace1 = go.Scatter(
141             x=x,
142             y=df.loc[Starttime:Endtime, List_y1[i]],
143             mode=marks[0], #markers, line, or lines+markers
144             marker = markerlist[i],
145             name='{}'.format(Label1[i]),
146             yaxis = 'y1',
147             line = dict(
148                 dash = dashes[i])
149         )
150         data.append(trace1)
151
152     series2 = ['stBOD']
153     for i in range(len(List_y2)):
154         trace2 = go.Scatter(
155             x=x,
156             y=df.loc[Starttime:Endtime, List_y2[i]],
157             mode=marks[1], #markers, line, or lines+markers
158             marker = markerlist[i+len(List_y1)],
159             name='{}'.format(Label2[i]),
160             yaxis = 'y2',
161             line = dict(
162                 dash = dashes[i+len(List_y1)])
163         )
164         data.append(trace2)
165
166     fig = go.Figure(data=data,layout=layout)
167
168     py.offline.iplot(fig)

```

A.5 Function to create WEST-compatible input files

```

1 def WestInputFile(df, Start, End, WType, Tstamp, name):
2     params = ['keep', '.t', '.H2O', '.S_I', '.S_S', '.S_O', '.S_NH', '.S_ND', '.S_ALK', '.X_I', '.X_S', '.X_BH', '.X_BA', '.X_P', '.X_ND', '.S_NO']
3     units = ['-','d','m3/d','g/m3','g/m3','g/m3','g/m3','g/m3','g/m3','g/m3','g/m3','g/m3','g/m3','g/m3','g/m3','g/m3']
4     In = pd.DataFrame(index = df[Start:End].index, columns = params)
5     In.columns = pd.MultiIndex.from_arrays([params,units], names = ['var','units'])
6
7     Dex=[]
8     Empty=[]
9     for i in Tstamp:
10         Dex.append(In.index.get_loc(pd.to_datetime(i),method='nearest'))
11         Next = pd.to_datetime(pd.to_datetime(i)+pd.to_timedelta('5 seconds'))
12         Empty.append(In.index.get_loc(Next,method='nearest'))
13
14     In.reset_index(inplace = True)
15     In['.t'] = pd.to_timedelta(In['Time']-In['Time'].iloc[0]).fillna(0)
16     In['.t'] = In['.t'].astype('timedelta64[s]')/(3600*24)
17     del In['Time']
18
19     In=In.fillna(0)
20     In.ix[0,'keep']='yes'
21     for item in Empty:
22         In.set_value(item,('keep','-'),'yes')
23     if WType == 'Ac':
24         for item in Dex:
25             In.set_value(item,('keep','-'),'yes')
26             In.set_value(item,('.S_S','g/m3'),15600)
27
28     if WType == 'AcN':
29         for item in Dex:
30             In.set_value(item,('keep','-'),'yes')
31             In.set_value(item,('.S_S','g/m3'),15600)
32             In.set_value(item,('.S_NH','g/m3'),2000)
33
34     if WType == 'N':
35         for item in Dex:
36             In.set_value(item,('keep','-'),'yes')
37             In.set_value(item,('.S_NH','g/m3'),2000)
38
39     In=In.loc[In[('keep','-')]=='yes']
40     In[(('.H2O','m3/d')).loc[(In[(('.S_S','g/m3'))]!=0) | (In[(('.S_NH','g/m3'))]!=0)]=0.20736
41
42     del In['keep']
43     datapath = "WESTIn/%s.txt" %name

```

```
44     In.to_csv(datapath, sep='\t', index=False)
45     return In
```


A.6 Function to create WEST-compatible objective DO and *OUR* time series for calibration

```
1 def ExportWestObj(df, Start, End, name, OURend):
2     #OURend is in mg/lh
3     import datetime
4     import pandas as pd
5     import numpy as np
6
7     Out = pd.DataFrame(index = df[Start:End].index, columns = [".t", ".ASU_1.C(S_O)
8     ", ".ASU_1.OUR_ASU"])
9     Out.columns = pd.MultiIndex.from_arrays([Out.columns, ["d", "g/m3", "g/(m3.d) "
10     ]], names=['var', 'unit'])
11     Out[".ASU_1.C(S_O)"] = df['DO_A']
12     Out[".ASU_1.OUR_ASU"] = df['OURex'] * (3600*24) + OURend
13
14     Out.reset_index(inplace = True)
15     Out['.t'] = pd.to_timedelta(Out['Time'] - Out['Time'].iloc[0]).fillna(0)
16     Out['.t'] = Out['.t'].astype('timedelta64[s]') / (3600*24)
17     del Out['Time']
18
19     Avg_temp = np.average(df[Start:End]['Temp'])
20     print('Average temperature is %s'%Avg_temp)
21
22     datapath = "WESTObj/%s.txt" %name #C:/Users/Jean-David/Dropbox/Rodtox/
23     raw_output/
24     Out.to_csv(datapath, sep='\t', index=False)
25     return Out
```

A.7 Function to automatically create WEST-compatible objective DO and *OUR* time series from wastewater respirograms

```
1 def WestRespirogram(df,resultsdf, newdirname, filenames, PeakVol, OUREnd):
2     import os
3     newpath = '/Users/therrien/Dropbox/Rodtox/raw_output'+ '/' + '%s'%newdirname
4     if not os.path.exists(newpath):
5         os.makedirs(newpath)
6
7     params = ['.t','.ASU_1.C(S_O)','ASU_1.OUR_ASU']
8     units = ['d','g/m3', "g/(m3.d)"]
9     for i in range(len(resultsdf)):
10         Start=resultsdf.loc[i,'Start']
11         End=resultsdf.loc[i,'End']
12         Respirodf = pd.DataFrame(index = df[Start:End].index, columns =params)
13         Respirodf.columns = pd.MultiIndex.from_arrays([params,units], names = ['
var','units'])
14         Respirodf[('.ASU_1.C(S_O)','g/m3')] = df.loc[Start:End,'DO_A']
15         Respirodf[('.ASU_1.OUR_ASU','g/(m3.d)')] = df.loc[Start:End,'OURex'
]*3600*24+OUREnd
16         Respirodf.reset_index(inplace = True)
17         Respirodf['.t'] = pd.to_timedelta(Respirodf['Time']-Respirodf['Time'].iloc
[0]).fillna(0)
18         Respirodf['.t'] = Respirodf['.t'].astype('timedelta64[s]')/(3600*24)
19         del Respirodf['Time']
20         PeakNo = resultsdf.loc[i,'PeakNo']
21         dfpath = newpath+'/'+'%s_PeakNo%s.txt'%(filenames, PeakNo)
22         Respirodf.to_csv(dfpath, sep='\t', index=False)
23     return
```