CLOSED LOOP CONTROL OF THE MEDIUM FEEDING FOR YEAST CULTURES IN A PILOT BIOREACTOR BASED ON EXHAUST GAS ANALYSIS
Acknowledgements

I want to thank everyone that has made this work possible, a work that concludes one phase of my life and that will hopefully lead to an even better one.

Special thanks go to my family and my friends, without their help none of that would have been possible.

I would also like to thank my Thesis’ Director, Philippe Bogaerts, for his continuous support and will to help, always understanding that as a foreigner student there were procedures I was not aware of.

Finally, I want to thank all of the 3-BIO department personal, specially Anne Richelle for investing the time in teaching me the basics of the lab and always being helpful if any questions arised.
Summary

In order to control one bioprocess we first need to know it, we need to be able to describe it, to anticipate what is going to happen, what is going to trigger the desired effect.

But in order to do that, in order to ensure that our results are reproducible and that they are optimal, firstly, we need to determine all the parameters involved.

This work proposes to study the culture of *Saccharomyces cerevisiae* in a pilot bioreactor, obtaining all the required data for the construction of a mathematical model, which in turn will allow us to design a controller that ensures the maximum productivity of the yeast, all that by regulating the ethanol concentration, setting a small fix setpoint.

To that end, we propose that, instead of using an expensive and unreliable ethanol probe, we should employ the analysis of the exhaust gas of the bioreactor, equipment already available most of the time and much more responsive.
Content

1 Introduction ...........................................................................................................................................5
   1.1 Context and motivation ..................................................................................................................5
   1.2 The organism: Saccharomyces cerevisiae ....................................................................................6
   1.3 Kinetic model .....................................................................................................................................7
   1.4 Optimal running point for the culture ............................................................................................8
2 Description of the math modelling of the bioprocess .................................................................9
   2.1 Introduction .......................................................................................................................................9
   2.2 Mass balance for the batch mode operation of the bioreactor ......................................................9
3 Estimation of the ethanol concentration ..................................................................................12
   3.1 Introduction ......................................................................................................................................12
   3.2 Material and methods ...................................................................................................................12
   3.3 Preparation of the batch culture .....................................................................................................13
   3.4 Results ............................................................................................................................................14
      3.4.1 First approach: lineal regression ..............................................................................................15
      3.4.2 Second approach: fminsearch ..................................................................................................16
      3.4.3 Third approach: fminsearch with normalization of the data .................................................16
      3.4.4 Validation of the results ...........................................................................................................17
      3.4.5 Robustness of non-linear approach .........................................................................................18
      3.4.6 Linear no-negative values constrained regression .................................................................20
4 Closed loop control of the ethanol concentration .................................................................22
   4.1 Introduction ......................................................................................................................................22
   4.2 Mass balance equations for a fed-batch system .........................................................................22
   4.3 Boundary case ...................................................................................................................................22
   4.4 Respirofermentative pathway ......................................................................................................23
   4.5 Results ............................................................................................................................................25
      4.5.1 Results of the first fed-batch culture ....................................................................................25
      4.5.2 Results of the second fed-batch culture ..............................................................................28
      4.5.3 Results of the third fed-batch culture ....................................................................................30
      4.5.4 Results of the fourth fed-batch culture ...............................................................................32
5 Conclusions and perspectives ........................................................................................................37
6 Bibliography .........................................................................................................................................38
7 Annexes ................................................................................................................................................39
1 Introduction

1.1 Context and motivation
Since ancient times humanity has tried to understand and master the biological processes around it. For instance: bread, wine, cheese, beer and many others are all common products whose production is based on the use of microorganisms. Nowadays new uses of these microorganisms are possible thanks to genetic engineering, the production of recombinant proteins for various products of biotechnological interest (production of insulin, vaccines or even proteins that could help in the treatment of some cancer) is becoming more and more important, and a good control to optimize the culture conditions is necessary.

Closed-loop regulation has many advantages over open-loop regulation, but there are some inherent difficulties to that kind of controller, as for example the need for precise and quick probes (which can be very expensive or even not available) or the need for a mathematical model representative of the bioprocess that can be translated into control algorithms. Obtaining such a model is not an easy task and the complexity of the system (three metabolic pathways, influence of a high number of variables, non-linearity of the process, etc.) makes some simplifications of the model indispensable, thus requiring the assumption of some hypothesis and the restriction of the domain of validity for the obtained functions.

Another clear disadvantage is the need for precise and often expensive probes

The microorganism studied will be *Saccharomyces Cerevisiae*, which is a yeast that has been studied for a very long time. The aim of that Master’s Thesis is to obtain the data needed for the modelling of the bioprocess in a bioreactor and, with that information, develop a closed-loop controller that, instead of using an expensive and slow ethanol probe, uses the analysis of the exhaust gas from the bioreactor to achieve the regulation of the ethanol concentration in the culture media.
1.2 The organism: Saccharomyces cerevisiae
The organism used for that Master’s Thesis, *Saccharomyces cerevisiae*, is a species of yeast that has been used since ancient times production of bread (that is why it is also known as baker’s yeast) and is also responsible for the production of alcoholic beverages (thus also known as brewer’s yeast).

This species can reproduce by budding or by sexual reproduction (*Lodder et al.*, 1952) and has a series of characteristics that make its study interesting:

- It was the first eukaryote organism to have its genome sequenced.
- It is a unicellular organism with a small size (5-10 μm).
- It has a quick generation time (the time needed to double the amount of cells is approximately 2 hours) and is easy to cultivate (*Leveau, Kreger-Van Rij*, 1993; *Boekhout, Robert*, 2003).
- Its structure is similar to those of the cells in the superior organisms, but it does not have so much of the non-coding DNA that those organisms have.

For all these reasons, it is a model organism and is widely used in both the food industry and the pharmaceutical industry, where it can be used for the production of different recombinant proteins (*Renard, Vande Wouver*, 2007).

The growth of the yeast depends on many parameters such as the pH of the media, the temperature, and the availability of organic carbon or the presence of water. In order to satisfy the energetic needs and provide the materials needed for the cell synthesis, the elements required are basically carbon and nitrogen, but some others such as phosphorus, sulphur, potassium, magnesium, calcium, zinc, manganese, etc. are needed in lesser amounts (*Leveau, Kreger-Van Rij*, 1993).

As all the living systems, the source of energy employed is the ATP, obtained by oxidation of the carbon compounds. The yeast is not thermoresistant so it is destroyed at high temperatures, and its optimal growth temperature is around 30ºC. It is not a strict anaerobic (so it can develop in presence or absence of oxygen), and it has the highest ethanol tolerance of all of the yeasts.

For the yeasts, there are two important effects on the glucose metabolism that are to be taken into account:

- **The Pasteur effect**: the sugars that the yeasts use can be metabolized either in an aerobic way (respirative pathway) or in a anaerobic way (fermentative pathway). It has been demonstrated that respirative pathway inhibits the fermentation (*Pasteur, 1861*). It must be taken into account that the Pasteur effect is absent or very feeble for *S. Cerevisiae*.
- **The Crabtree effect**: the presence of glucose inhibits the synthesis of multiple respirative enzymes. Therefore, even in the presence of oxygen a high glucose concentration inhibits the catabolic respirative pathway and favours the ethanol production (fermentation) (*Leveau, Kreger-Van Rij*, 1993).
In order to achieve our aim of controlling the system, a kinetic model able to reflect the particularities of the metabolism of the yeast is required. The model that we have chosen is the one proposed by Sonnleitner and Käppeli (Sonnleitner, Käppeli, 1986), as its description of the different phenomena is the most accepted in the literature (Renard, 2006).

Sonnleitner and Käppeli base their model on the hypothesis that the respirative catabolic pathway has a limit (bottleneck assumption), and that the use of the yeast’s respiratory capacity tends to be always optimized.

This yeast, when in aerobic conditions, has two metabolic pathways: oxidative (respiration) and oxidoreductive (respiration and fermentation).

The preferred pathway is the oxidative one so, as long as there is enough glucose and the respiratory capacity of the yeast is not surpassed, the yeast will oxidise the glucose and if there is still some respiratory capacity left it will be used to consume the ethanol that may be present in the media.

If all the respiratory capacity is used, then any surplus of glucose will be fermented and ethanol will be obtained.

The reaction scheme used to define the processes involved in the metabolism of the *Saccharomyces cerevisiae*, as defined by Sonnleitner and Käppeli (Sonnleitner, Käppeli, 1986), is the following one:

**Table 1. Notation of the reaction scheme**

<table>
<thead>
<tr>
<th>Sign</th>
<th>G</th>
<th>O</th>
<th>P</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Compound</strong></td>
<td>Glucose</td>
<td>Oxygen</td>
<td>Product: carbon dioxide</td>
<td>Ethanol</td>
</tr>
<tr>
<td><strong>Sign</strong></td>
<td>X</td>
<td>r_i</td>
<td>k_i</td>
<td></td>
</tr>
<tr>
<td><strong>Compound</strong></td>
<td>Biomass</td>
<td>Reaction rate</td>
<td>Pseudo-stoichiometric coefficient</td>
<td></td>
</tr>
</tbody>
</table>

**Glucose oxidation (respiration):** 
\[ G + k_5 O \overset{r_1}{\rightarrow} k_1 X + k_7 P \]

**Glucose fermentation:** 
\[ G \overset{r_2}{\rightarrow} k_2 X + k_8 P + k_4 E \]

**Ethanol oxidation (respiration):** 
\[ E + k_6 O \overset{r_3}{\rightarrow} k_3 X + k_9 P \]

The kinetics of the reactions associated to each of the metabolic pathways can be explained with a Monod-like kinetic (Monod, 1942), where the substrate is the oxygen (O) for the respiration pathway and the glucose (G) for the fermentation pathway:

\[ r_O = \mu_O^{\text{max}} \cdot \frac{O}{O + K_O} \]

\[ r_G = \mu_G^{\text{max}} \cdot \frac{G}{G + K_G} \]
Introduction

\[
r_1 = \min \left( r_G, \frac{r_O}{k_5} \right)
\]

\[
r_2 = \max \left( 0, r_G - \frac{r_O}{k_5} \right)
\]

\[
r_3 = \max \left( 0, \frac{r_O - k_5 \cdot r_G}{k_6}, \frac{E}{E + K_E} \right)
\]

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>(k_1)</td>
<td>0.49</td>
<td>g of X / g of G</td>
</tr>
<tr>
<td>(k_2)</td>
<td>0.05</td>
<td>g of X / g of G</td>
</tr>
<tr>
<td>(k_3)</td>
<td>0.72</td>
<td>g of X / g of E</td>
</tr>
<tr>
<td>(k_4)</td>
<td>0.48</td>
<td>g of E / g of G</td>
</tr>
<tr>
<td>(k_5)</td>
<td>0.3968</td>
<td>g of O(_2) / g of G</td>
</tr>
<tr>
<td>(k_6)</td>
<td>1.104</td>
<td>g of O(_2) / g of E</td>
</tr>
<tr>
<td>(k_7)</td>
<td>0.5897</td>
<td>g of CO(_2) / g of G</td>
</tr>
<tr>
<td>(k_8)</td>
<td>0.4621</td>
<td>g of CO(_2) / g of G</td>
</tr>
<tr>
<td>(k_9)</td>
<td>0.6249</td>
<td>g of CO(_2) / g of E</td>
</tr>
<tr>
<td>(\mu_0)</td>
<td>0.256</td>
<td>g of CO(_2) / g of X / h</td>
</tr>
<tr>
<td>(\mu_G)</td>
<td>3.5</td>
<td>g of G / g of X / h</td>
</tr>
<tr>
<td>(K_O)</td>
<td>0.1</td>
<td>g of O(_2) / l</td>
</tr>
<tr>
<td>(K_G)</td>
<td>0.1</td>
<td>g of G / l</td>
</tr>
<tr>
<td>(K_E)</td>
<td>0.1</td>
<td>g of E / l</td>
</tr>
</tbody>
</table>

1.4 Optimal running point for the culture

If the objective is to maximize the biomass production, the optimal point would be the limit point between the respirative and the respiro-fermentative metabolic pathway, as the highest biomass production for gram of glucose is achieved there.

We will define that point as \(G_{\text{crit}}\), the concentration of glucose where the respiratory capacity is just saturated: \(G_{\text{crit}} = \frac{K_G r_O}{K_S \mu_G - r_O}\).

While that would be the optimal solution, the control is not precise enough to guarantee that we stay on that point all the time, and as changing between different metabolic pathways is harmful for the yeast, we will choose a sub-optimal solution: to stay in the respiro-fermentative pathway but with a low level of ethanol production, in order to maintain a concentration of about 1g/l of ethanol in the media.
2 Description of the math modelling of the bioprocess

2.1 Introduction

One of the most challenging issues that must be addressed is the fact that every strain of yeast is different, so a way to obtain the characteristic parameters needed to model the bioprocess is required.

In order to do that we need to define a mathematical model that will allow us to establish the relationship between the different variables, so with the kinetic model and the mass balances of the system we will be able to achieve that end.

Working in batch mode (no inwards or outwards liquid flow) simplifies matters considerably (no dilution rate has to be taken into account), so we will start by expressing the mass balances for the bioreactor in batch mode.

2.2 Mass balance for the batch mode operation of the bioreactor

There is no liquid flow in the system, but there exists an air flow that passes through the liquid and provides the oxygen needed for the culture and takes out the carbon dioxide produced. Those transfers are accounted for in the mass balances by introducing the variables OTR (Oxygen Transfer Rate) and CTR (Carbon dioxide Transfer Rate), which are defined as follows:

\[
OTR = \frac{Air\ flow \ \left[\frac{nl}{min}\right]}{Volume \ [l]} \cdot \frac{60 \ \left[\frac{min}{h}\right]}{} \cdot \frac{1}{22.4 \ \left[\frac{mol}{nl}\right]} \cdot M_{O_2} \left[\frac{g}{mol}\right] \cdot \left( y_{O, in} - y_{O, out} \cdot \frac{y_{N, in}}{y_{N, out}} \right)
\]

\[
CTR = \frac{Air\ flow \ \left[\frac{nl}{min}\right]}{Volume \ [l]} \cdot \frac{60 \ \left[\frac{min}{h}\right]}{} \cdot \frac{1}{22.4 \ \left[\frac{mol}{nl}\right]} \cdot M_{CO_2} \left[\frac{g}{mol}\right] \cdot \left( y_{CO_2, out} \cdot \frac{y_{N, in}}{y_{N, out}} - y_{CO_2, in} \right)
\]

Once those concepts have been introduced, using the reaction scheme:

\[
G + k_5 O \xrightarrow{r_1} k_4 X + k_7 P
\]

\[
G \xrightarrow{r_2} k_2 X + k_6 P + k_4 E
\]

\[
E + k_6 O \xrightarrow{r_3} k_3 X + k_9 P
\]

The mass balances are the following ones:
Description of the math modelling of the bioprocess

If we only consider the metabolic pathway of our interest (in this case the respiro-fermentative pathway) then we can assume there is no reconsumption of ethanol and thus \( r_3 = 0 \).

That way:

\[
\frac{dX}{dt} = k_1 \cdot r_1 \cdot X + k_2 \cdot r_2 \cdot X + k_3 \cdot r_3 \cdot X
\]

\[
\frac{dG}{dt} = -r_1 \cdot X - r_2 \cdot X
\]

\[
\frac{dE}{dt} = k_4 \cdot r_2 \cdot X - r_3 \cdot X
\]

\[
\frac{dO}{dt} = -k_5 \cdot r_1 \cdot X - k_6 \cdot r_3 \cdot X + OTR
\]

\[
\frac{dP}{dt} = k_7 \cdot r_1 \cdot X + k_8 \cdot r_2 \cdot X + k_9 \cdot r_3 \cdot X - CTR
\]

If we only consider the metabolic pathway of our interest (in this case the respiro-fermentative pathway) then we can assume there is no reconsumption of ethanol and thus \( r_3 = 0 \).

If we develop that, trying to express all of the equations as a function of OTR and CTR:

\[
r_1 \cdot X = \left( OTR - \frac{dO}{dt} \right) \cdot \frac{1}{k_5}
\]

\[
r_2 \cdot X = \left( \frac{dP}{dt} - \frac{k_7}{k_5} \cdot \left( OTR - \frac{dO}{dt} \right) + CTR \right) \cdot \frac{1}{k_8}
\]

\[
= \frac{1}{k_8} \cdot \left( \frac{dP}{dt} + CTR \right) - \frac{k_7}{k_5 \cdot k_8} \cdot \left( OTR - \frac{dO}{dt} \right)
\]

\[
\frac{dE}{dt} = \frac{k_4}{k_8} \cdot \left( \frac{dP}{dt} + CTR \right) - \frac{k_4 \cdot k_7}{k_5 \cdot k_8} \cdot \left( OTR - \frac{dO}{dt} \right)
\]
If we apply the hypothesis of quasi-stationary state for the oxygen and the carbon dioxide (all of the oxygen is consumed instantaneously and all of the CO₂ is evacuated instantaneously, there is no accumulation):

\[
\frac{dO}{dt} \equiv 0, \quad \frac{dP}{dt} \equiv 0
\]

\[
\frac{dE}{dt} = \frac{k_4}{k_8 \cdot CTR} - \frac{k_4 \cdot k_7}{k_5 \cdot k_8 \cdot OTR}
\]

\[
\frac{dX}{dt} = \frac{k_1 \cdot k_8 - k_2 \cdot k_7}{k_5 \cdot k_8} \cdot OTR + \frac{k_2}{k_8} \cdot CTR
\]

\[
\frac{dG}{dt} = \frac{k_7 - k_8}{k_5 \cdot k_8} \cdot OTR - \frac{1}{k_8} \cdot CTR
\]

And if we integrate the resulting equations:

\[
E - E_0 = \frac{k_4}{k_8} \int_0^t CTR(\tau) \cdot d\tau - \frac{k_4 \cdot k_7}{k_5 \cdot k_8} \int_0^t OTR(\tau) \cdot d\tau
\]

\[
X - X_0 = \frac{k_1 \cdot k_8 - k_2 \cdot k_7}{k_5 \cdot k_8} \int_0^t OTR(\tau) \cdot d\tau + \frac{k_2}{k_8} \int_0^t CTR(\tau) \cdot d\tau
\]

\[
G - G_0 = \frac{k_7 - k_8}{k_5 \cdot k_8} \int_0^t OTR(\tau) \cdot d\tau - \frac{1}{k_8} \int_0^t CTR(\tau) \cdot d\tau
\]

We can see that the equations we obtain link all of the variables with the integral of OTR and CTR, so it is possible to perform experiments to determine the parameters that we will need for our controller.
3 Estimation of the ethanol concentration

3.1 Introduction
As seen in the previous chapter, in order to determine the pseudo-stoichiometric coefficients we will need to perform a number of experiments to obtain the required data.

The procedure followed and the results obtained will be presented in this section.

3.2 Material and methods
In order to perform the various experiments required for the development of the studied subject, a 20l bioreactor BIOSTAT C-DCU3 (Sartorius B. Braun Biotech International) was used.

That bioreactor has similar technical specifications to those of the ones used in the industry, with multiple probes to measure and control the different variables that affect the culture. In order to allow the sampling of the media it is also equipped with two valves that can be sterilized after taking each sample.

In addition to the standard equipment there is also a digitally controlled pump that allows the user to define a feeding profile, a non-dispersive infrared gas analyser that determines the composition of the exhaust gas, a pH probe and a built-in pump that allow pH control and, lastly, another built-in pump that can feed anti-foam when needed.

All of the devices connected or built into the bioreactor can be supervised and controlled directly with the integrated control unit (via a touchscreen), or remotely through a computer, thanks to an interface called MFCS/win (Sartorius B. Braun Biotech International). Using that interface we can see the evolution of the different variables, plot them, change the setpoints, etc. Moreover, on the installed computer it is possible to interact with MATLAB, so we can write a program to establish the feeding profile based on the different values of the state variables, changing the setpoint of the pumping speed (Sartorius) as necessary.

So as to assure an optimal growth of the yeast, some variables are controlled for all of the cultures:

- The pressure of the vessel is set to 500 mbar (gauge pressure).
- The temperature of the media is set to 30ºC (as that is the optimal growth temperature for *S. Cerevisiae*) and is regulated by changing the temperature of the water flowing through the heating jacket of the bioreactor.
- The dissolved oxygen (pO$_2$) must be higher than 20% so, in order to maintain that value, the stirring speed increases when needed.
Estimation of the ethanol concentration

Figure 1: Schematic representation of the bioreactor system

The variables that cannot be measured directly by the bioreactor are measured off-line. To determine the OD (optical density) a spectrometer at a wavelength of 650 nm was used.

For the glucose and the ethanol analysis, centrifugation and removal of the biomass in the sample media was done, followed by enzymatic tests designed to obtain the concentration of the two components in the media.

3.3 Preparation of the batch culture

All of the equipment must be properly sterilized in order to avoid a possible contamination inside the bioreactor. To sterilize the media itself, the bioreactor has its own sterilization program that runs at 121ºC for 30 min, and the sampling valves can be sterilized using vapour after each use.

The composition of the 7 l that make the media is: 50 g/l of glucose (to ensure that we get to a respiro-fermentative state), 20 g/l of tryptone and 10 g/l of yeast extract.

In order to prepare the inoculum (the liquid actually containing the organism that we are going to cultivate) for the bioreactor, this procedure was followed (always in duplicate to ensure the reproducibility of the results):

- Preparation of a Petri plate with the studied yeast strain, leaving it in an incubator at 30ºC for 48h.
- Preculture in small shake-flasks with 100 ml of media for 12 hours.
- Second preculture using 1ml of the previous preculture in 400 ml of media for 12 hours.
3.4 Results

In order to obtain the pseudo-stoichiometric coefficients for the studied yeast two batch experiments were done.

As determined earlier, the mass balance equations for the system can be expressed as a function of the integrals of OTR and CTR:

\[
E - E_0 = \frac{k_4}{k_8} \int_0^t CTR(\tau) \cdot d\tau - \frac{k_4 \cdot k_7}{k_5 \cdot k_8} \int_0^t OTR(\tau) \cdot d\tau
\]

\[
X - X_0 = \frac{k_1 \cdot k_8 - k_2 \cdot k_7}{k_5 \cdot k_8} \int_0^t OTR(\tau) \cdot d\tau + \frac{k_2}{k_8} \int_0^t CTR(\tau) \cdot d\tau
\]

\[
G - G_0 = \frac{k_7 - k_8}{k_5 \cdot k_8} \int_0^t OTR(\tau) \cdot d\tau - \frac{1}{k_8} \int_0^t CTR(\tau) \cdot d\tau
\]

However, the equations were obtained for a batch system, and taking the necessary samples in order to obtain the required data modifies the volume (even if slightly). To take into account that fact, the mass balance is reformulated:

\[
-\frac{d(G \cdot V)}{dt} = F_{out} \cdot G + r_1 \cdot X \cdot V + r_2 \cdot X \cdot V
\]

\[
\frac{dV}{dt} = -F_{out}
\]

\[
-V \cdot \frac{dG}{dt} - G \cdot \frac{dV}{dt} = F_{out} \cdot G + r_1 \cdot X \cdot V + r_2 \cdot X \cdot V
\]

\[
-V \cdot \frac{dG}{dt} - G \cdot (-F_{out}) = F_{out} \cdot G + r_1 \cdot X \cdot V + r_2 \cdot X \cdot V
\]

\[
\frac{dG}{dt} = -r_1 \cdot X - r_2 \cdot X
\]

As observed, the differential equation that is obtained is the same as in the batch system, and that is the case for the other variables too. The same expressions are valid for that case.

Another point to take into account is how to choose the range of data that will actually be used for the model: in that case, a good way to ensure that we remain in fermentative mode is to check that ethanol is being produced.

For the detailed results check the Annexes, but as the readings of the second experiment were the best, those were the ones chosen for the following steps.
3.4.1 First approach: lineal regression

Seeing that model equations are expressed as a linear combination of the different variables, the first reasonable step seems to be the multiple linear regression of the data, thus obtaining the factors that are a combination of the different pseudo-stoichiometric coefficients.

In order to do that a simple MATLAB script which uses the least squares method has been made, using the backslash ("\") operator.

That method, when used with the obtained data gives coefficients with a good fit. But the problem is that some of the pseudo-stoichiometric coefficients obtained that way are negative. As that has no physical meaning, some constraints must be applied in order to avoid that situation.

<table>
<thead>
<tr>
<th>k₁ (OD/g G)</th>
<th>k₂ (OD/g G)</th>
<th>k₄ (g E/g G)</th>
<th>k₅ (g O₂/g G)</th>
<th>k₇ (g CO₂/g G)</th>
<th>k₈ (g CO₂/g G)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0,120758</td>
<td>0,154876</td>
<td>0,210106</td>
<td>-0,303266</td>
<td>0,417742</td>
<td>0,412226</td>
</tr>
</tbody>
</table>

**Figure 2.** Results of the linear regression (using second culture as data)

If the results of the first culture are to be used, these problems also appear and, moreover, the fitting is not nearly as good (see Figure 3):

<table>
<thead>
<tr>
<th>k₁ (OD/g G)</th>
<th>k₂ (OD/g G)</th>
<th>k₄ (g E/g G)</th>
<th>k₅ (g O₂/g G)</th>
<th>k₇ (g CO₂/g G)</th>
<th>k₈ (g CO₂/g G)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0,819622</td>
<td>17,620510</td>
<td>9,324747</td>
<td>0,287891</td>
<td>-1,902853</td>
<td>40,769583</td>
</tr>
</tbody>
</table>
3.4.2 Second approach: *fminsearch*

MATLAB has a function called *fminsearch* that implements the Nelder-Mead simplex algorithm (Lagarias et al., 1998).

While it does not explicitly allow for constraints in the variables, if a non-negativity criteria is to be applied one option is to use as input the natural logarithm of the variables and later utilize the exponential function to recover the original variables.

The cost function is the sum of squares of the difference of the results obtained with the mass balance equations and the real data.

While the fit for two of the components (biomass and glucose) is acceptably good ($R^2=0.9875$ and $R^2=0.9932$), the fit for the ethanol is not so good ($R^2=0.6554$).

The possible solution to that is to work with normalized data.

3.4.3 Third approach: *fminsearch* with normalization of the data

To solve the problem we can try normalizing the data so the difference in scale between the different data doesn’t have a big effect on the results of the cost function.

By doing that, we see an improvement in the ethanol fit, a slight decrease on the goodness of the fit of the biomass and a slight improvement in the glucose fit ($R^2=0.8447$; $R^2=0.9486$ and $R^2=0.9939$ respectively).

The possible solution to that is to work with normalized data.

\[ k_1 \text{ (OD/g G)} \quad k_2 \text{ (OD/g G)} \quad k_4 \text{ (g E/g G)} \quad k_5 \text{ (g O}_2\text{/g G)} \quad k_7 \text{ (g CO}_2\text{/g G)} \quad k_8 \text{ (g CO}_2\text{/g G)} \]

<table>
<thead>
<tr>
<th>$k_1$ (OD/g G)</th>
<th>$k_2$ (OD/g G)</th>
<th>$k_4$ (g E/g G)</th>
<th>$k_5$ (g O$_2$/g G)</th>
<th>$k_7$ (g CO$_2$/g G)</th>
<th>$k_8$ (g CO$_2$/g G)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.430635</td>
<td>0.059809</td>
<td>0.513405</td>
<td>0.916803</td>
<td>0.354722</td>
<td>0.433010</td>
</tr>
</tbody>
</table>
3.4.4 Validation of the results

In addition to the simple validation shown in Figure 4, in Figure 5 we can see the results of the reconstructed signal for the different variables of the culture.

It can be seen that the model succeeds reconstructing the signal in a way that resembles the original data quite well, with the best fit for the biomass ($R^2=0.8730$), followed by the glucose ($R^2=0.6290$), and lastly the ethanol ($R^2=0.1195$).

If the high peak for ethanol is removed the fitting increases greatly, leaving coefficients of $R^2_{\text{ethanol}} = 0.7627$; $R^2_{\text{biomass}} = 0.9194$; $R^2_{\text{glucose}} = 0.8792$. That way, if the values for those points are considered not valid because of sampling problems and not because some unexplained stress triggered an unexpected response in the yeast behavior, the fitting can be considered good enough and thus the coefficients obtained valid.
3.4.5 Robustness of non-linear approach

The main difference between the linear and the non-linear approach is that, while the linear approach always reaches a unique best fit, needing no estimated values for the variables, the non-linear approach is dependent on the initial guess for the estimated coefficients.

Even though it is reasonable enough using the bibliographic values provided by Sonnleitner and Käppeli (Sonnleitner, Kappeli, 1986), in order to see if the values of the different pseudo-stoichiometric coefficients could be better by initializing the search at a different point (randomly chosen) instead of initializing with the literature values, one thousand different values randomly generated with MATLAB (from 0 to 1) have been used, and the result for the 20 best fits is shown in the Table 3.

- The first column value corresponds to the index of the essay (between 1 and 1000).
- The columns 2-7 are the values of initialization, randomly generated.
- The columns 8-10 correspond to the goodness of fit for each variable (E, X and G).
- The column 11 is the result of the cost function for this analysis, defined as: 
$$3 - R_E^2 - R_X^2 - R_G^2$$
- The columns 11-16 correspond to the obtained values for the different $k_i$.

From the results obtained we can see that while the values for $k_2$, $k_4$ and $k_8$ are similar for the different values, they play an important value in the fit, but $k_1$, $k_5$ and $k_7$ can take different values without affecting much of the fit.

That presents some questions: $k_2$, $k_4$ and $k_8$ seem robust, but the others not that much, is there a way to verify that? To try to give a better response to that question, we will return to the first linear regression but adding some constraints.
Table 3. Results of the randomized initialization point test

<table>
<thead>
<tr>
<th>Index</th>
<th>( k_1 _r )</th>
<th>( k_2 _r )</th>
<th>( k_4 _r )</th>
<th>( k_5 _r )</th>
<th>( k_7 _r )</th>
<th>( k_8 _r )</th>
<th>( R_2 _E )</th>
<th>( R_2 _X )</th>
<th>( R_2 _G )</th>
<th>Cost</th>
<th>( k_1 )</th>
<th>( k_2 )</th>
<th>( k_4 )</th>
<th>( k_5 )</th>
<th>( k_7 )</th>
<th>( k_8 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>635</td>
<td>0.49</td>
<td>0.05</td>
<td>0.48</td>
<td>0.397</td>
<td>0.59</td>
<td>0.462</td>
<td>0.845</td>
<td>0.949</td>
<td>0.994</td>
<td>0.213</td>
<td>0.431</td>
<td>0.06</td>
<td>0.513</td>
<td>0.9168</td>
<td>0.355</td>
<td>0.433</td>
</tr>
<tr>
<td>231</td>
<td>0.5823</td>
<td>0.4064</td>
<td>0.8321</td>
<td>0.9734</td>
<td>0.4929</td>
<td>0.1422</td>
<td>0.9333</td>
<td>0.9873</td>
<td>0.9939</td>
<td>0.8555</td>
<td>0.0066</td>
<td>0.2028</td>
<td>0.4314</td>
<td>4.7210</td>
<td>0.0686</td>
<td>0.4302</td>
</tr>
<tr>
<td>241</td>
<td>0.8329</td>
<td>0.9294</td>
<td>0.8624</td>
<td>0.7166</td>
<td>0.7037</td>
<td>0.0090</td>
<td>0.9219</td>
<td>0.9879</td>
<td>0.9941</td>
<td>0.0960</td>
<td>0.0248</td>
<td>0.1891</td>
<td>0.3712</td>
<td>492.4720</td>
<td>1.1932</td>
<td>0.4052</td>
</tr>
<tr>
<td>57</td>
<td>0.6457</td>
<td>0.5697</td>
<td>0.3058</td>
<td>0.5546</td>
<td>0.3122</td>
<td>0.1023</td>
<td>0.9158</td>
<td>0.9932</td>
<td>0.9941</td>
<td>0.0969</td>
<td>0.2984</td>
<td>0.1702</td>
<td>0.4563</td>
<td>2.1745</td>
<td>0.3219</td>
<td>0.4202</td>
</tr>
<tr>
<td>174</td>
<td>0.7682</td>
<td>0.2536</td>
<td>0.5519</td>
<td>0.4856</td>
<td>0.1358</td>
<td>0.1078</td>
<td>0.9179</td>
<td>0.9895</td>
<td>0.9927</td>
<td>0.0999</td>
<td>0.4707</td>
<td>0.1408</td>
<td>0.4958</td>
<td>2.1830</td>
<td>0.0951</td>
<td>0.4590</td>
</tr>
<tr>
<td>176</td>
<td>0.9189</td>
<td>0.2940</td>
<td>0.7501</td>
<td>0.3338</td>
<td>0.0186</td>
<td>0.6216</td>
<td>0.9136</td>
<td>0.9910</td>
<td>0.9928</td>
<td>0.1026</td>
<td>0.5545</td>
<td>0.1657</td>
<td>0.3839</td>
<td>10.7974</td>
<td>1.3330</td>
<td>0.3774</td>
</tr>
<tr>
<td>276</td>
<td>0.8448</td>
<td>0.0373</td>
<td>0.7898</td>
<td>0.5645</td>
<td>0.6896</td>
<td>0.1307</td>
<td>0.9104</td>
<td>0.9928</td>
<td>0.9942</td>
<td>0.1027</td>
<td>0.7506</td>
<td>0.1414</td>
<td>0.3959</td>
<td>4.0893</td>
<td>0.3792</td>
<td>0.4080</td>
</tr>
<tr>
<td>469</td>
<td>0.9550</td>
<td>0.7513</td>
<td>0.2333</td>
<td>0.7027</td>
<td>0.3520</td>
<td>0.5591</td>
<td>0.9077</td>
<td>0.9910</td>
<td>0.9938</td>
<td>0.1075</td>
<td>0.8333</td>
<td>0.1404</td>
<td>0.3851</td>
<td>4.5470</td>
<td>0.0009</td>
<td>0.4363</td>
</tr>
<tr>
<td>998</td>
<td>0.8466</td>
<td>0.4651</td>
<td>0.4598</td>
<td>0.0285</td>
<td>0.9065</td>
<td>0.1863</td>
<td>0.9025</td>
<td>0.9918</td>
<td>0.9936</td>
<td>0.1121</td>
<td>0.8155</td>
<td>0.1668</td>
<td>0.3703</td>
<td>9.3095</td>
<td>0.8852</td>
<td>0.3887</td>
</tr>
<tr>
<td>108</td>
<td>0.2152</td>
<td>0.3000</td>
<td>0.7073</td>
<td>0.9748</td>
<td>0.3004</td>
<td>0.3131</td>
<td>0.8985</td>
<td>0.9935</td>
<td>0.9905</td>
<td>0.1174</td>
<td>0.2091</td>
<td>0.1794</td>
<td>0.5817</td>
<td>0.9853</td>
<td>0.2307</td>
<td>0.4897</td>
</tr>
<tr>
<td>769</td>
<td>0.7018</td>
<td>0.5190</td>
<td>0.5655</td>
<td>0.4343</td>
<td>0.3421</td>
<td>0.2517</td>
<td>0.8991</td>
<td>0.9909</td>
<td>0.9923</td>
<td>0.1177</td>
<td>0.3883</td>
<td>0.1634</td>
<td>0.4301</td>
<td>2.7997</td>
<td>0.6733</td>
<td>0.3713</td>
</tr>
<tr>
<td>392</td>
<td>0.5594</td>
<td>0.8793</td>
<td>0.1930</td>
<td>0.9405</td>
<td>0.0788</td>
<td>0.3298</td>
<td>0.9101</td>
<td>0.9830</td>
<td>0.9821</td>
<td>0.1249</td>
<td>0.0340</td>
<td>0.2880</td>
<td>0.6570</td>
<td>0.8201</td>
<td>0.1358</td>
<td>0.5780</td>
</tr>
<tr>
<td>669</td>
<td>0.7062</td>
<td>0.5359</td>
<td>0.4446</td>
<td>0.6576</td>
<td>0.6734</td>
<td>0.1330</td>
<td>0.8894</td>
<td>0.9914</td>
<td>0.9931</td>
<td>0.1260</td>
<td>0.4961</td>
<td>0.1243</td>
<td>0.4573</td>
<td>1.8362</td>
<td>0.5234</td>
<td>0.3809</td>
</tr>
<tr>
<td>486</td>
<td>0.0713</td>
<td>0.3000</td>
<td>0.8672</td>
<td>0.9718</td>
<td>0.3079</td>
<td>0.3147</td>
<td>0.8869</td>
<td>0.9936</td>
<td>0.9927</td>
<td>0.1268</td>
<td>0.2103</td>
<td>0.1790</td>
<td>0.5507</td>
<td>0.9942</td>
<td>0.2936</td>
<td>0.4588</td>
</tr>
<tr>
<td>345</td>
<td>0.2662</td>
<td>0.4505</td>
<td>0.7823</td>
<td>0.9265</td>
<td>0.5880</td>
<td>0.0321</td>
<td>0.8934</td>
<td>0.9852</td>
<td>0.9906</td>
<td>0.1308</td>
<td>0.2939</td>
<td>0.1344</td>
<td>0.5681</td>
<td>0.9327</td>
<td>0.2465</td>
<td>0.4887</td>
</tr>
<tr>
<td>595</td>
<td>0.8055</td>
<td>0.3742</td>
<td>0.0668</td>
<td>0.0739</td>
<td>0.9655</td>
<td>0.3010</td>
<td>0.8844</td>
<td>0.9936</td>
<td>0.9896</td>
<td>0.1323</td>
<td>0.7607</td>
<td>0.1326</td>
<td>0.4415</td>
<td>3.5582</td>
<td>0.9536</td>
<td>0.3517</td>
</tr>
<tr>
<td>4</td>
<td>0.8313</td>
<td>0.4065</td>
<td>0.7493</td>
<td>0.3865</td>
<td>0.1280</td>
<td>0.7118</td>
<td>0.8878</td>
<td>0.9897</td>
<td>0.9839</td>
<td>0.1386</td>
<td>1.0614</td>
<td>0.1479</td>
<td>0.3957</td>
<td>7.0972</td>
<td>2.0719</td>
<td>0.3272</td>
</tr>
<tr>
<td>734</td>
<td>0.4808</td>
<td>0.3486</td>
<td>0.4925</td>
<td>0.0814</td>
<td>0.2716</td>
<td>0.4997</td>
<td>0.8783</td>
<td>0.9886</td>
<td>0.9942</td>
<td>0.1389</td>
<td>0.2724</td>
<td>0.1883</td>
<td>0.3423</td>
<td>33.9870</td>
<td>0.1500</td>
<td>0.4082</td>
</tr>
<tr>
<td>448</td>
<td>0.6270</td>
<td>0.0454</td>
<td>0.7490</td>
<td>0.6915</td>
<td>0.5341</td>
<td>0.2597</td>
<td>0.8780</td>
<td>0.9880</td>
<td>0.9937</td>
<td>0.1403</td>
<td>0.3318</td>
<td>0.1290</td>
<td>0.5621</td>
<td>1.0795</td>
<td>0.3277</td>
<td>0.4385</td>
</tr>
<tr>
<td>203</td>
<td>0.0402</td>
<td>0.3766</td>
<td>0.8998</td>
<td>0.0438</td>
<td>0.0007</td>
<td>0.6074</td>
<td>0.9039</td>
<td>0.9700</td>
<td>0.9765</td>
<td>0.1496</td>
<td>0.0000</td>
<td>0.3503</td>
<td>0.6936</td>
<td>0.6940</td>
<td>0.1426</td>
<td>0.6302</td>
</tr>
</tbody>
</table>

Estimation of the ethanol concentration
3.4.6 Linear no-negative values constrained regression

\[
E - E_0 = \frac{k_4}{k_8} \int_0^t CTR(\tau) \cdot d\tau - \frac{k_4 \cdot k_7}{k_5 \cdot k_8} \int_0^t OTR(\tau) \cdot d\tau
\]

\[
X - X_0 = \frac{k_1 \cdot k_9 - k_2 \cdot k_7}{k_5 \cdot k_8} \cdot \int_0^t OTR(\tau) \cdot d\tau + \frac{k_2}{k_8} \int_0^t CTR(\tau) \cdot d\tau
\]

\[
G - G_0 = \frac{k_7}{k_5 \cdot k_8} \cdot \int_0^t OTR(\tau) \cdot d\tau - \frac{1}{k_8} \int_0^t CTR(\tau) \cdot d\tau
\]

If we treat each equation separately and we apply constraints that prevent negative values of the pseudo-stoichiometric coefficients, the values for the values that were of the wrong sign now become zero.

Analyzing closely the data at hand, we can see that if:

\[
\frac{k_7 - k_8}{k_5 \cdot k_8} \approx 0 \quad \text{and} \quad \frac{k_4 \cdot k_7}{k_5 \cdot k_8} \approx 0 \rightarrow \frac{k_7}{k_5} \cdot \frac{1}{k_8} - \frac{1}{k_5} \approx 0, \quad \text{and as} \quad \frac{k_7}{k_5} \approx 0 \rightarrow \frac{1}{k_5} \approx 0
\]

That value suggests that \( k_5 \) is very big, and that could be related with a problem with the readings of oxygen (\( k_5 \) is the coefficient that links the consumption of oxygen with the production of glucose for the respirative metabolic pathway).

The fit obtained via this method is the best fit because it is determined via linear regression, and that can be further confirmed comparing the values for \( R^2 \).

With the information obtained in that way the only coefficients that can be precisely determined are \( k_1, k_4 \) and \( k_8 \), and so the problem of the robustness of the values isn’t solved, but at least it is known that the best fits will be those in which the value of \( k_7/k_5 \) and \( k_4/k_5 \) are as low as possible.

The values obtained are as follows:

**Table 4.** Results of the constrained linear regression

<table>
<thead>
<tr>
<th>OTR coefficient</th>
<th>CTR coefficient</th>
<th>Ethanol</th>
<th>Biomass</th>
<th>Glucose</th>
<th>R² ethanol</th>
<th>R² biomass</th>
<th>R² glucose</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0,98610684</td>
<td>0,11933665</td>
<td>0</td>
<td>0</td>
<td>0,93458594</td>
<td>0,9937765</td>
<td>0,99415756</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>k₁ (OD/g G)</th>
<th>k₂ (OD/g G)</th>
<th>k₄ (g E/g G)</th>
<th>k₅ (g O₂/g G)</th>
<th>k₇ (g CO₂/g G)</th>
<th>k₈ (g CO₂/g G)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>0,15298793</td>
<td>0,40154245</td>
<td>-</td>
<td>-</td>
<td>0,40719975</td>
</tr>
</tbody>
</table>
Estimation of the ethanol concentration

Figure 6. Obtained data with non-negative linear regression
4 Closed loop control of the ethanol concentration

4.1 Introduction
After obtaining the required parameters of the model, what we have left to do is to obtain a control law that will allow us to implement a closed-loop regulation of the system. In order to do that we will go back to the mass balances for a fed-batch system.

4.2 Mass balance equations for a fed-batch system
For a fed-batch system the dilution rate has to be taken into account:

\[ D = \frac{F^{in}}{V} \]

\[ \frac{dX}{dt} = k_1 \cdot r_1 \cdot X + k_2 \cdot r_2 \cdot X + k_3 \cdot r_3 \cdot X - D \cdot X \]

\[ \frac{dG}{dt} = -r_1 \cdot X - r_2 \cdot X + D \cdot (G^{in} - G) \]

\[ \frac{dE}{dt} = k_4 \cdot r_2 \cdot X - r_3 \cdot X - D \cdot E \]

\[ \frac{dO}{dt} = -k_5 \cdot r_1 \cdot X - k_6 \cdot r_3 \cdot X + OTR - D \cdot O \]

\[ \frac{dP}{dt} = k_7 \cdot r_1 \cdot X + k_8 \cdot r_2 \cdot X + k_9 \cdot r_3 \cdot X - CTR - D \cdot P \]

\[ \frac{dV}{dt} = D \cdot V = F^{in} \]

4.3 Boundary case
If the respiratory capacity is all used, there is no ethanol and there is no fermentation, then:

\[ r_2 = r_3 = 0 \]

\[ r_1 = r_6 = \frac{r_0}{k_5} = \frac{\mu_{\text{max}}}{k_5} \cdot \frac{O}{O + K_0} = r^* \]

\[ \frac{d(X \cdot V)}{dt} = k_1 \cdot r^* \cdot X \cdot V \Rightarrow \frac{d(X \cdot V)}{X \cdot V} = k_1 \cdot r^* \cdot dt \Rightarrow \ln(X \cdot V) = k_1 \cdot r^* \cdot t + C \Rightarrow \]

\[ X \cdot V = e^{k_1 \cdot r^* \cdot t} \cdot e^C \]

\[ \lim_{t \to 0} x_{V=X_0 \cdot V_0} \quad X \cdot V = X_0 \cdot V_0 \cdot e^{k_1 \cdot r^* \cdot t} \]

With the hypothesis of no accumulation of glucose (everything that is added is consumed instantaneously):

\[ \frac{d(G \cdot V)}{dt} \approx 0 \rightarrow F^{in} \cdot G^{in} - r^* \cdot V \cdot X = 0 \rightarrow \]

\[ F^{in} = \frac{r^* \cdot V \cdot X}{G^{in}} = \frac{r^* \cdot X_0 \cdot V_0 \cdot e^{k_1 \cdot r^* \cdot t}}{G^{in}} \]
4.4 Respirofermentative pathway

A control law for the ethanol concentration of the following form is to be obtained:

\[
\frac{d(E^* - E)}{dt} = -\lambda \cdot (E^* - E)
\]

If the ethanol has a constant setpoint (dE*/dt=0):

\[
\frac{dE}{dt} = \lambda \cdot (E^* - E) = k_4 \cdot r_2 \cdot X - D \cdot E \Rightarrow \lambda \cdot V \cdot (E^* - E) = k_4 \cdot r_2 \cdot X \cdot V - F^{in} \cdot E
\]

If \(d(G \cdot V)/dt \approx 0\):

\[
0 = -r_1 \cdot V \cdot X - r_2 \cdot V \cdot X + F^{in} \cdot G^{in} \Rightarrow r_2 \cdot X \cdot V = F^{in} \cdot G^{in} - r_1 \cdot X \cdot V
\]

\[
\lambda \cdot V \cdot (E^* - E) = k_4 \cdot \left[F^{in} \cdot G^{in} - r_1 \cdot X \cdot V\right] - F^{in} \cdot E
\]

\[
\lambda \cdot V \cdot (E^* - E) = k_4 \cdot F^{in} \cdot G^{in} - k_4 \cdot r_1 \cdot X \cdot V - F^{in} \cdot E
\]

\[
\lambda \cdot V \cdot (E^* - E) = F^{in} \cdot \left[k_4 \cdot G^{in} - E\right] - k_4 \cdot r_1 \cdot X \cdot V
\]

\[
F^{in} = \frac{\lambda \cdot V \cdot (E^* - E) + k_4 \cdot r_1 \cdot X \cdot V}{k_4 \cdot G^{in} - E}; \text{ if } \frac{d(O \cdot V)}{dt} \approx 0 = -k_5 \cdot r_1 \cdot X \cdot V + OTR \cdot V \Rightarrow
\]

\[
\Rightarrow F^{in} = \frac{\lambda \cdot V \cdot (E^* - E) + k_4 \cdot OTR \cdot V}{k_4 \cdot G^{in} - E}
\]

If OTR and CTR are to be used instead of E, then:

\[
\frac{d(X \cdot V)}{dt} = k_1 \cdot r_1 \cdot X \cdot V + k_2 \cdot r_2 \cdot X \cdot V
\]

\[
\frac{d(G \cdot V)}{dt} \approx 0 \Rightarrow F^{in} \cdot G^{in} - X \cdot V \cdot (r_1 + r_2) = 0
\]

\[
\frac{dE}{dt} = k_4 \cdot r_2 \cdot X - D \cdot E
\]

If we accept the hypothesis of instantaneous consumption of the O\(_2\) and instantaneous transfer of CO\(_2\):

\[
\frac{d(O \cdot V)}{dt} \approx 0 \Rightarrow -k_5 \cdot r_1 \cdot X \cdot V + OTR \cdot V = 0 \Rightarrow r_1 \cdot X \cdot V = \frac{OTR \cdot V}{k_5} = \alpha \cdot OTR \cdot V
\]

\[
\frac{d(P \cdot V)}{dt} \approx 0 \Rightarrow (k_7 \cdot r_1 + k_8 \cdot r_2) \cdot X \cdot V - CTR \cdot V \Rightarrow
\]

\[
\Rightarrow k_8 \cdot r_2 \cdot X \cdot V = CTR \cdot V - k_7 \cdot r_1 \cdot X \cdot V = CTR \cdot V - \frac{k_7}{k_5} \cdot OTR \cdot V \Rightarrow
\]
\[ r_2 \cdot X \cdot V = \frac{CTR \cdot V - k_2 \cdot OTR \cdot V}{k_8} = -\frac{k_7}{k_5 \cdot k_8} \cdot OTR \cdot V + \frac{1}{k_8} \cdot CTR \cdot V \]
\[ = \beta \cdot OTR \cdot V + \gamma \cdot CTR \cdot V \]

That way:

\[ \frac{d(E \cdot V)}{dt} = k_4 \cdot r_2 \cdot X \cdot V = k_4 \cdot \beta \cdot OTR \cdot V + k_4 \cdot \gamma \cdot CTR \cdot V \]

\[ \int_{E_0 \cdot V_0}^{E \cdot V} d(E \cdot V) = \int_{0}^{t} k_4 \cdot \beta \cdot OTR \cdot V \cdot d\tau + \int_{0}^{t} k_4 \cdot \gamma \cdot CTR \cdot V \cdot d\tau \]

\[ E \cdot V - E_0 \cdot V_0 = k_4 \cdot \beta \int_{0}^{t} OTR \cdot V \cdot d\tau + k_4 \cdot \gamma \cdot \int_{0}^{t} CTR \cdot V \cdot d\tau \]

We see that, in fact, it is possible to reconstruct the ethanol signal:

\[ E = \frac{1}{V} \cdot \left[ k_4 \cdot \beta \cdot \int_{0}^{t} OTR \cdot V \cdot d\tau + k_4 \cdot \gamma \cdot \int_{0}^{t} CTR \cdot V \cdot d\tau + E_0 \cdot V_0 \right] \]

The mass balance for the other components can be expressed as a function of OTR and CTR too:

\[ \frac{d(X \cdot V)}{dt} = k_1 \cdot \alpha \cdot OTR \cdot V + k_2 \cdot \beta \cdot OTR \cdot V + k_2 \cdot \gamma \cdot CTR \cdot V \]

\[ \frac{d(G \cdot V)}{dt} \approx 0 \rightarrow F^{in} \cdot G^{in} - \alpha \cdot OTR \cdot V \cdot \beta \cdot OTR \cdot V \cdot \gamma \cdot CTR \cdot V \]

\[ = F^{in} \cdot G^{in} - (\alpha + \beta) \cdot OTR \cdot V - \gamma \cdot CTR \cdot V = 0 \]

Given that the setpoint for (EV) is equal to the setpoint for E times V:

\[ (E \cdot V)^* = E^* \cdot V \]

\[ \frac{d[(E \cdot V)^* - (E \cdot V)]}{dt} = -\lambda \cdot [(E \cdot V)^* - (E \cdot V)] \]

\[ E^* \cdot F^{in} \cdot \frac{d(E \cdot V)}{dt} = -\lambda \cdot [(E \cdot V)^* - (E \cdot V)] \]

\[ E^* \cdot F^{in} = -\lambda \cdot [(E \cdot V)^* - (E \cdot V)] + k_4 \cdot \beta \cdot OTR \cdot V + k_4 \cdot \gamma \cdot CTR \cdot V \]

\[ F^{in} = \frac{k_4}{E^*} \cdot (\beta \cdot OTR \cdot V + \gamma \cdot CTR \cdot V) - \lambda \cdot \left[ V - \frac{k_4 \cdot \beta \cdot \int_{0}^{t} OTR \cdot V \cdot d\tau + k_4 \cdot \gamma \cdot \int_{0}^{t} CTR \cdot V \cdot d\tau + E_0 \cdot V_0}{E^*} \right] \]
As $\beta \cdot OTR \cdot V + \gamma \cdot CTR \cdot V \approx F^{in} \cdot G^{in} - \alpha \cdot OTR \cdot V$:

$$E^{*} \cdot F^{in} = k_{4} \cdot (F^{in} \cdot G^{in} - \alpha \cdot OTR \cdot V) - \lambda$$

$$\cdot \left[ E^{*} \cdot V - k_{4} \cdot \beta \cdot \int_{0}^{t} OTR \cdot V \cdot d\tau - k_{4} \cdot \gamma \cdot \int_{0}^{t} CTR \cdot V \cdot d\tau - E_{0} \cdot V_{0} \right]$$

$$F^{in} \cdot (E^{*} - k_{4} \cdot G^{in})$$

$$= -k_{4} \cdot \alpha \cdot OTR \cdot V - \lambda$$

$$\cdot \left[ E^{*} \cdot V - k_{4} \cdot \beta \cdot \int_{0}^{t} OTR \cdot V \cdot d\tau - k_{4} \cdot \gamma \cdot \int_{0}^{t} CTR \cdot V \cdot d\tau - E_{0} \cdot V_{0} \right]$$

If the setpoint is low enough, and the inlet concentration of glucose is high:

$$k_{4} \cdot G^{in} - E^{*} \approx k_{4} \cdot G^{in}$$

$$F^{in} \cdot k_{4} \cdot G^{in} = k_{4} \cdot \alpha \cdot OTR \cdot V + \lambda$$

$$\cdot \left[ E^{*} \cdot V - k_{4} \cdot \beta \cdot \int_{0}^{t} OTR \cdot V \cdot d\tau - k_{4} \cdot \gamma \cdot \int_{0}^{t} CTR \cdot V \cdot d\tau - E_{0} \cdot V_{0} \right]$$

$$F^{in} = \frac{\alpha}{G^{in}} \cdot OTR \cdot V + \frac{\lambda}{k_{4} \cdot G^{in}} \cdot \left[ E^{*} \cdot V - k_{4} \cdot \beta \cdot \int_{0}^{t} OTR \cdot V \cdot d\tau - k_{4} \cdot \gamma \cdot \int_{0}^{t} CTR \cdot V \cdot d\tau - E_{0} \cdot V_{0} \right]$$

## 4.5 Results

### 4.5.1 Results of the first fed-batch culture

The first controlled fed-batch culture of *S. Cerevisiae* was done under the following conditions:

- 5 liters of a media with a composition of 5 g/l of glucose, 20 g/l of tryptone and 10 g/l of yeast extract was used.
- The setpoints for the bioreactor were: 300 rpm for the stirring, increased if the PO2 went under 20%; temperature of 30ºC and pressure of 500 mbar.
- One slightly improved version (compared to the first one used to run the preliminary tests) of the MATLAB controller was used for determining the feeding profile, consisting of a 250g/l glucose solution.

That experiment addressed some of the problems noted during the preliminary tests, for example the precision of the calibration for the feeding with a small needle (which was previously not good enough for all pump values), or the problem that represents using a balance with a sensitivity of only 0,5 g/l (that was solved doing a check of what the pump is supposed to have fed, and the accumulated feeding measured via the weight of the bottle).

Some problems that occurred during the experiment were as follows:

- The readings for the O2 and CO2 were very unstable at the beginning, leading to a higher concentration of O2 than the starting one, thus keeping the value of OTR to 0 or very low values for most of the time.
- Even when the setpoint for the feeding was set to 0%, some pumping was still done, with the bioreactor reporting a pumping rate of 0,1%.
The variables relevant to the performance of the controller are represented in the following graphics:

![Figure 7. Evolution of the OTR and CTR according to time](image)

![Figure 8. Detail of the feeding rate calculated via the control law](image)

As we can see, the feeding rate starts fairly high (the first five minutes of running the controller acquires the data needed for the calculation of the feeding rate, so it starts as 0), increasing a bit due to the increment of the volume in the bioreactor caused by the feeding, but otherwise remaining fairly constant until the CTR starts to increase in a noticeable way.

Here we can see what I believe is the main problem: the system doesn’t have a sensitivity high enough to ensure that very small variations on the CO$_2$ concentration are registered and,
moreover, the response of the system is too slow compared to the speed of change the controller tries to impose, so a very high amount of glucose is added the first hours to try to get to the ethanol setpoint, but by the time it is achieved the amount of glucose on the media is too high and the system can’t be controlled anymore, thus reverting to a batch system in a high glucose concentration media.

\[
F^{in} = \frac{\alpha}{G^m} \cdot OTR \cdot V + \frac{\lambda}{k_4 \cdot G^m} \cdot \left[ E^* \cdot V - k_4 \cdot \beta \cdot \frac{\lambda}{k_4 \cdot G^m} \int_0^t OTR \cdot V \cdot d\tau - k_4 \cdot \gamma \cdot \frac{\lambda}{k_4 \cdot G^m} \int_0^t CTR \cdot V \cdot d\tau - E_0 \cdot V_0 \right]
\]

If the two parts of the control law are compared, we can see that the first one, the one associated with the needs of glucose that the yeast needs to support a respirative metabolic pathway, is very small compared to the second one (the one that tries to control the amount of ethanol). Moreover, as in that experiment the readings for OTR are not accurate, its contribution is even less important.

As that control law is conceived for a respiro-fermentative metabolic pathway, if the ethanol concentration increases past the setpoint the system will just stop feeding, but as ethanol reconsumption is not taken into account, when the surplus of glucose is finally consumed the system will switch to a respirative pathway, eventually consuming all of the glucose in the media.

Some further graphics:

![Graph](image)

**Figure 9.** Theoretical feeding (red) and real feeding (green)

The feeding works as commanded but does not stop at zero because of the feeding bottle is higher than the bioreactor vessel.
It can be noted that the feeding is higher than the consumption of glucose (the maximum valued obtained is higher than 42 g/l, but further analysis have not been done), and that there is a latency phase of a considerable length for the culture.

4.5.2 Results of the second fed-batch culture
The 18th July 2011 the second controlled fed-batch culture of S. Cerevisiae was done under the following conditions:

- 5 liters of a media with a composition of 5 g/l of glucose, 20 g/l of tryptone and 10 g/l of yeast extract was used.
- The setpoints for the bioreactor were: 300 rpm for the stirring, increased if the PO2 went under 20%; temperature of 30°C and pressure of 500 mbar.
- In order to better adjust the speed of the controller to that of the system, the lambda parameter has been decreased 4 times, until \( \lambda = 0.025 \).
- To try to achieve more noticeable changes in the parameters as well as speeding up the process (shorten the latency), the initial OD for that culture is of 1.

The aim of that experiment was to determine if decreasing the speed with which the controller acts could ensure a proper running of the system. During the execution of the program some problems were encountered:

- The pump stopped from time to time, needing a manual restart each time, it was fixed for the most part after an hour of running, tightening the bolts of the pump.
- The pump stopped once again after two more hours of running, but the influence is minimal because the feeding was going to stop anyway.
- Some foam was forming, so after taking the third sample 5 ml of antifoam were added.
- A crash occurred when finalizing the execution of the MATLAB program, so the data related to the running of the controller was lost; measures were taken to ensure the storage of the data in fixed time intervals, so as to avoid future problems.

![Graph](image1.png)

**Figure 11.** Feeding profile for the second fed-batch culture

As we can see, the amount of glucose fed is lower than in the first experiment, and there is no residual feeding after the controller stops it, as the set-up of the equipment has been improved.

![Graph](image2.png)

**Figure 12.** Optical density and glucose profiles for the second fed-batch experiment

The culture presents a latency phase of about the same length ($\approx 3$ h) as usual, so increasing the initial OD has not helped in that regard.

The glucose has the same problem seen in the first experiment, its concentration rises at the beginning, in a way that when the controller thinks that the system has surpassed the setpoint, the feeding stops but there is enough glucose to continue the fermentation, so it continues, and later the glucose is consumed until the respiratory capacity is not all used, and the shift to a respirative only metabolic pathway occurs.
4.5.3 Results of the third fed-batch culture

The 20th July 2011 the third controlled fed-batch culture of \textit{S. Cerevisiae} was done under the following conditions:

- 5 liters of a media with a composition of 20 g/l of glucose, 20 g/l of tryptone and 10 g/l of yeast extract was used.
- The setpoints for the bioreactor were: 300 rpm for the stirring, increased if the PO2 went under 20%; temperature of 30ºC and pressure of 500 mbar.
- The controller implemented a function to store all relevant data to an Excel file every 5 minutes.
- There was a wait of approximately 3 hours until the controller was launched, in order to skip the latency phase and ensure we were in a respiro-fermentative mode.

\textbf{Figure 13.} Feeding profile for the third fed-batch culture

Note the difference between the real feeding and the planned feeding, there is some error at the time of stopping the pump, probably because the feeding pipe is emptied by gravity, adding some final glucose.
After waiting for the end of the latency phase, at $t=3.74$ h the feeding starts. Its length is much shorter than in the previous experiments, as the setpoint (see Figure 9) is rapidly approached. The problem is that the glucose concentration in the media for the batch phase is higher than the required to attain a concentration of 1 g/l, so the setpoint is surpassed and the previously discussed problems arise.

We should note that the ethanol signal increases until the glucose concentration is at a level under the one of the respirative bottleneck.
Closed loop control of the ethanol concentration

Figure 16. Evolution of the OTR and CTR for the third fed-batch culture while the controller is running

4.5.4 Results of the fourth fed-batch culture

The 22\textsuperscript{nd} July 2011 the fourth controlled fed-batch culture of \textit{S. Cerevisiae} was done under the following conditions:

- 5 liters of a media with a composition of 5 g/l of glucose, 20 g/l of tryptone and 10 g/l of yeast extract was used.
- The setpoints for the bioreactor were: 300 rpm for the stirring, increased if the PO2 went under 20%; temperature of 30\degree C and pressure of 500 mbar.
- The controller implemented a function to store all relevant data to an Excel file every 5 minutes.
- There was a wait of approximately 3 hours until the controller was launched, in order to skip the latency phase and ensure we were in a respiro-fermentative mode.

The reason for the decrease in the glucose concentration in the media is to try to ensure that the ethanol concentration does not increase past the setpoint because of an excess of glucose.

The procedure followed is the same than in the third culture, but in this case the feeding started at \( t=4h \) until \( t=5,47h \). It was later restarted manually following the drop in CTR that occurred (and that signals a metabolic switch) at \( t=6,92h \) but it finally stopped at \( t=8,73h \).
Closed loop control of the ethanol concentration

**Figure 17.** Feeding profile for the fourth fed-batch culture

**Figure 18.** Feeding rate an reconstructed ethanol signal for the fourth fed-batch culture

It can be seen there that the feeding stops the moment the setpoint is reached. The reconstructed signal suddenly drops to zero as a consequence of manually restarting the control program, which is programmed to set zero as the initial value for the ethanol concentration.
Figure 19. Evolution of the OTR and CTR values for the fourth fed-batch culture

The evolution of the OTR and CTR values is shown in this Figure 19. As we can see, the moment there is a switch in the metabolic pathway of the yeast, the CTR decreases dramatically, so the change can be used to detect when that is happening (but it is not a good idea to establish the control around that point, as the constant change between the two states is harmful for the yeast).

Figure 20. Evolution of the respiratory quotient

It is worth noting that, as we can see in Figure 20, when the ethanol production stops, the respiratory quotient (defined as RQ=CER/OUR, almost CTR/OTR for our system) drops below the unity, and when the fermentation continues it rises again.

We can also see that the maximum ethanol concentration is at the peak of the CTR, and after the first feeding, when the glucose is scarce (see Figure 21), following the change in metabolism most of the ethanol is consumed, and, in fact, when resuming the feeding the ethanol production is also restarted (as can be seen through the increase in CTR), but the
Closed loop control of the ethanol concentration

Feeding is not capable of keeping the concentration of ethanol stable, as when the setpoint is reached the feeding stops in order to avoid more production of ethanol, but if we desire the system to remain in a respirofermentative metabolic pathway, it is necessary to add more glucose, which in turn means more ethanol production.

The only method to regulate the concentration of ethanol in the media while remaining in that state is the increase in the volume of the media but, given the high glucose concentration of the feeding that cannot be changed it is not possible to operate like that, so the system eventually gets to a point of no return where the specifications are not met. To change that it would be necessary to introduce means to change the glucose concentration, trying to remain in a state where the production of ethanol would be the same as the dilution rate.

Figure 21. Evolution of the optical density, glucose and ethanol concentration for the fourth fed-batch culture

One final verification that needs to be done is comparing the ethanol signal that the controller reconstructs with the actual ethanol concentration in the media:

Figure 22. Reconstructed ethanol signal versus ethanol concentration for the fourth fed-batch culture
As we can see, the ethanol signal is actually about 5 times bigger than the real concentration for those conditions, but it follows the same tendency as the real one, and we have to take into account that after a change in the metabolic pathway the yeast probably needs an adaptation time. If that fact is due to an error in the measurement of the data remains to be seen, and it is one of the issues that should be addressed if there is a further study on the matter.
5 Conclusions and perspectives

That Master’s Thesis presents the procedure followed to obtain the needed parameters for the modelling of the bioprocess of the culture for the yeast *Saccharomyces cerevisiae*, explaining the required steps for that: from the mathematical development required, to the experiments needed to obtain the data, and the analysis of these data.

After obtaining those parameters, we have developed a control law and coded it for use in a bioreactor expecting to succeed in the control of the ethanol concentration, but there have been some problems with the acquisition of the ethanol signal. Because the modelling was not meant for the use in a respirative metabolic pathway the reconsumption of ethanol is not taken into account, and thus, even if the closed-loop regulation is robust by design, without a proper signal it cannot work properly.

There is also room for improvement in other areas, as for example:

- The controller response time should be tuned according to the actual response time of the system more accurately if good results are to be expected.
- If we want to work with a very small feeding rate to guarantee that the ethanol production is low, more precise equipment is needed to feed the exact quantities and to measure the small variance in the exit gases.
- In order to maintain the setpoint, controlled reconsumption of ethanol could be programmed, and means to control the concentration of the feeding could be implemented, adjusting the dilution rate as necessary to guarantee an optimal growth.

If all of these points are addressed, the development of a functional controller for that microorganism should be possible, achieving the aim of developing a regulation system that works without the need of an expensive, insensitive and slow ethanol probe.
6 Bibliography


LODDER, JACOMINA. and KREGER-VAN RIJ, N. J. W. The yeasts; a taxonomic study, by J. Lodder and N.J.W. Kreger-van Rij North-Holland Publishing Co., Amsterdam, 1952


RENARD F., Commande robuste de bioprocédés opérés en mode fed-batch : Applications industrielles à des cultures de levures Saccharomyces cerevisiae. Thèse de doctorat, Faculté Polytechnique de Mons, 2006
7 Annexes

Data from the first batch experience

<table>
<thead>
<tr>
<th>n</th>
<th>t (h)</th>
<th>DO</th>
<th>Glucose (g/l)</th>
<th>Ethanol (g/l)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0,41</td>
<td>0,24</td>
<td>44,4</td>
<td>0,1164</td>
</tr>
<tr>
<td>2</td>
<td>1,58</td>
<td>0,255</td>
<td>46,1</td>
<td>0,2396</td>
</tr>
<tr>
<td>3</td>
<td>2,86</td>
<td>0,359</td>
<td>41,6</td>
<td>0,5516</td>
</tr>
<tr>
<td>4</td>
<td>3,87</td>
<td>0,502</td>
<td>31,6</td>
<td>0,9455</td>
</tr>
<tr>
<td>5</td>
<td>4,82</td>
<td>0,697</td>
<td>28,8</td>
<td>2,6258</td>
</tr>
<tr>
<td>6</td>
<td>5,97</td>
<td>1,346</td>
<td>23,4</td>
<td>12,5585</td>
</tr>
<tr>
<td>7</td>
<td>6,88</td>
<td>2,58</td>
<td>23,8</td>
<td>38,1970</td>
</tr>
<tr>
<td>8</td>
<td>7,86</td>
<td>3,75</td>
<td>21,5</td>
<td>10,3028</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>4,7</td>
<td>10,1</td>
<td>13,1243</td>
</tr>
<tr>
<td>10</td>
<td>9,91</td>
<td>5,12</td>
<td>6,1</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>10,83</td>
<td>7,25</td>
<td>5,3</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>11,66</td>
<td>5,57</td>
<td>3,1</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>12,93</td>
<td>6,66</td>
<td>7,9</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>13,86</td>
<td>6,64</td>
<td>6,7</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>14,84</td>
<td>6,48</td>
<td>21,4</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>15,83</td>
<td>5,61</td>
<td>16,6</td>
<td></td>
</tr>
</tbody>
</table>

Data from the second batch experience

<table>
<thead>
<tr>
<th>n</th>
<th>t (h)</th>
<th>DO</th>
<th>Glucose (g/l)</th>
<th>Ethanol (g/l)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0,99</td>
<td>0,229</td>
<td>38,87</td>
<td>0,4047</td>
</tr>
<tr>
<td>2</td>
<td>2,07</td>
<td>0,316</td>
<td>38,89</td>
<td>0,4759</td>
</tr>
<tr>
<td>3</td>
<td>3,18</td>
<td>0,441</td>
<td>37,78</td>
<td>0,7841</td>
</tr>
<tr>
<td>4</td>
<td>4,32</td>
<td>0,707</td>
<td>36,58</td>
<td>2,0982</td>
</tr>
<tr>
<td>5</td>
<td>5,08</td>
<td>1,34</td>
<td>33,94</td>
<td>3,5739</td>
</tr>
<tr>
<td>6</td>
<td>6,12</td>
<td>2,34</td>
<td>30,61</td>
<td>6,0132</td>
</tr>
<tr>
<td>7</td>
<td>6,74</td>
<td>3,08</td>
<td>24,52</td>
<td>5,5314</td>
</tr>
<tr>
<td>8</td>
<td>7,16</td>
<td>3,52</td>
<td>21,77</td>
<td>7,1372</td>
</tr>
<tr>
<td>9</td>
<td>7,64</td>
<td>4,14</td>
<td>17,35</td>
<td>9,1023</td>
</tr>
<tr>
<td>10</td>
<td>8,28</td>
<td>4,76</td>
<td>14,41</td>
<td>9,7599</td>
</tr>
<tr>
<td>11</td>
<td>9,28</td>
<td>5,44</td>
<td>8,65</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>10</td>
<td>5,99</td>
<td>6,36</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>10,88</td>
<td>6,08</td>
<td>3,05</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>11,78</td>
<td>6,2</td>
<td>1,57</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>12,48</td>
<td>6,18</td>
<td>0,56</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>13,35</td>
<td>6,4</td>
<td>0,37</td>
<td></td>
</tr>
</tbody>
</table>
Regression code

First attempt:

```matlab
% Input arguments: X,Y
% Output arguments: coef,R2,yhat
% Uses backslash operator to resolve least square problem
function [coef,R2,yhat]=reglineal(X,Y)
coef=X\Y;
yhat=X*coef;%Regressed data
ymean=mean(Y);
%ymean=sum(Y)/length(Y);
sserr=sum((Y-yhat).*(Y-yhat));
sstot=sum((Y-ymean).*(Y-ymean));
R2=1-sserr/sstot;%Coefficient of determination

% The inputs are:
% bdata, which has the following structure (vertical vectors):
% time,ethanol,
% biomass,glucose,OTR and CTR
% out_sheet, which is the path to the excel file where the output
% results
% are to be stored (optional)
% sheet, which is the name of the sheet of the excel file where the results
% will be written (optional, if not used defaults to the first sheet,
% overwriting the present data!)

function [k,coefreg,R2,yhat]=regbatch(bdata,out_sheet,sheet)
if nargin==2
    sheet=1;
end
% Note that the function cumtrapz exists and is implemented in MATLAB,
% but it is not used, although it could
OTRint=intacum(bdata(:,1),bdata(:,5));
CTRint=intacum(bdata(:,1),bdata(:,6));
coefreg=ones(2,3);
R2=ones(3,1);
yhat=ones(length(bdata(:,1)),3);
for n=1:3
    V=(bdata(:,n+1)-bdata(1,n+1));%Variable regressed, like E-E0,
i.e.
    y(:,n)=V;
    [coefreg(:,n),R2(n,1),yhat(:,n)]=reglineal([OTRint CTRint],V);
end
% Obtention of the different pseudo-stoichiometric coefficients
k8=-1/coefreg(2,3);
k2=coefreg(2,2)*k8;
k4=coefreg(2,1)*k8;
k5=-k4*k8/(k4*k8*coefreg(1,3)+k8*coefreg(1,1));%Cramer
k7=k8*coefreg(1,1)*(-k8)/(k4*k8*coefreg(1,3)+k8*coefreg(1,1));
k1=(coefreg(1,2)*k5*k8+k2*k7)/k8;
k=[k1 k2 k4 k5 k7 k8];
plot(bdata(:,1),y,bdata(:,1),yhat) %Plot of the regressed data and the experimental data
% Various command to write to excel file, can be refined to speed up
% execution
```
if nargin==1
    return
else
    output_data1=k;
    output_data2=coefreg;
    output_data3=R2';
    output_data4=bdata(:,1);
    output_data5=[y yhat bdata(:,5) bdata(:,6) OTRint CTRint];
    xlswrite(out_sheet,{'k1' 'k2' 'k4' 'k5' 'k7' 'k8'},sheet,'A1');
    xlswrite(out_sheet,{'reg eth' 'reg biom' 'reg gluc'},sheet, 'H1');
    xlswrite(out_sheet,output_data2,sheet, 'H2');
    xlswrite(out_sheet,{'R2 eth' 'R2 biom' 'R2 gluc'},sheet, 'L1');
    xlswrite(out_sheet,output_data3,sheet, 'L2');
    xlswrite(out_sheet,'t',sheet, 'A5');
    xlswrite(out_sheet,output_data4,sheet, 'A6');
    xlswrite(out_sheet,{'E-E0' 'X-X0' 'G-G0' 'Ehat-Ehat0' 'Xhat-Xhat0' 'Ghat-Ghat0' 'OTR' 'CTR' 'OTRint' 'CTRint'},sheet, 'B5');
end

Second attempt:

%Theta is a vector [k1 k2 k4 k5 k7 k8]

function [minsq,yhat]=funfmin_nonorm(theta)

global bdata
Y(:,1)=intacum(bdata(:,1),bdata(:,5));%OTRint
Y(:,2)=intacum(bdata(:,1),bdata(:,6));%CTRint
theta=exp(theta);
k1=theta(1);
k2=theta(2);
k4=theta(3);
k5=theta(4);
k7=theta(5);
k8=theta(6);
test=[];
yhat(:,1)=k4/k8*Y(:,2)-k4*k7/k5/k8*Y(:,1);
%yhat(:,1)=(theta(3)/theta(6)*Y(:,2)-theta(3)*theta(5)/theta(4)/theta(6)*Y(:,1));
test(:,1)=(bdata(:,2)-bdata(1,2))-yhat(:,1);
%testmean(1)=mean(test(:,1));%Fixed?-%Normalize, doesn't seem to give
good%results so far
yhat(:,2)=(k1*k8-k2*k7)/(k5*k8)*Y(:,1)+k2/k8*Y(:,2);
%yhat(:,2)=((theta(1)*theta(6)-theta(2)*theta(5))/theta(4)/theta(6)*Y(:,1)+theta(2)/theta(6)*Y(:,2));
test(:,2)=(bdata(:,3)-bdata(1,3))-yhat(:,2);
%testmean(2)=mean(test(:,2));

yhat(:,3)=(k7-k8)/k5/k8*Y(:,1)-1/k8*Y(:,2);
%yhat(:,3)=((theta(5)-theta(6))/theta(4)/theta(6)*Y(:,1)-1/theta(6)*Y(:,2));
test(:,3)=(bdata(:,4)-bdata(1,4))-yhat(:,3);
%Normalization of results
%testmean(3)=mean(test(:,3));
%[testmean ind]=sort(testmean);
%for t=1:3
```matlab
% coef(ind(t))=testmean(t)/testmean(end);
% testnorm(:,t)=test(:,t)*coef(t);
% end
testnorm=test;%Use only if not normalizing

if nargout==1
elseif nargout==2
    yhat=yhat;
end
testnorm=testnorm.*testnorm;
mins==sum(sum(testnorm));

%If a variable named bdata exists, it is rewritten
%function [k,coefreg,R2,yhat]=regbatch2(out_sheet,sheet)
function [theta]=regbatch2(theta0)
global bdata
theta=fminsearch(@funfmin_nonorm,log(theta0),optimset('MaxFunEvals',4000));
theta=exp(theta);

%Calculates the coefficient of determination of a set of batch
bioreactor
%data using one of three methods (the method is specified via the
toggle):
%No toggle or 0 -> The first input must be an initial guess for the value
%of the pseudo-stoichiometric coefficients, the program tries to obtain the
%best values for them
%1 -> The program uses the provided values of the coefficients
%2 -> Instead of using the coefficients, the fitted data is the input (the
%structure must be a (n,3) matrix with the values of E,X and G (one in each
%column)
%In addition to those calculations, the program can write the data to an
%Excel file, using two additional inputs (note that in that case the toggle
%0 is obligatory if that mode of operation is desired):
%out_sheet, which is the path to the excel file where the output results
%are to be stored (optional)
sheet, which is the name of the sheet of the excel file where the results
will be written (optional, if not used defaults to the first sheet,
overwriting the present data!)

function [R2test,theta]=estadreg_nonorm(theta0,toggle,out_sheet,sheet)
global bdata
%if nargout==2
%    if nargin==2
%        disp('Error, no theta')
%        return
```
%end
%end
if nargin==1;
    theta=regbatch2_nonorm(theta0);
else if nargin>=2;
    if toggle=='1'
        theta=theta0;
    elseif toggle=='2'
        yhat=theta0;
    elseif toggle=='0'
        theta=regbatch2_nonorm(theta0);
    else
        disp('Error, unrecognized toggle')
        return
    end
end

Y(:,1)=intacum(bdata(:,1),bdata(:,5)); %OTRint
Y(:,2)=intacum(bdata(:,1),bdata(:,6)); %CTRint

if nargin==1 || (nargin==2 && (toggle=='1') || (toggle=='0'))
    yhat=[];
    yhat(:,1)=(theta(3)/theta(6)*Y(:,2)-theta(3)*theta(5)/theta(4)/theta(6)*Y(:,1));
    yhat(:,2)=((theta(1)*theta(6)-theta(2)*theta(5))/theta(4)/theta(6)*Y(:,1)+theta(2)/theta(6)*Y(:,2));
    yhat(:,3)=((theta(5)-theta(6))/theta(4)/theta(6)*Y(:,1)-1/theta(6)*Y(:,2));
end

test1mean=mean(yhat(:,1));
test2mean=mean(yhat(:,2));
test3mean=mean(yhat(:,3));
E=(bdata(:,2)-bdata(1,2));
X=(bdata(:,3)-bdata(1,3));
G=(bdata(:,4)-bdata(1,4));
sserr1=sum((E-yhat(:,1)).*(E-yhat(:,1)));
sserr2=sum((X-yhat(:,2)).*(X-yhat(:,2)));
sserr3=sum((G-yhat(:,3)).*(G-yhat(:,3)));
sstot1=sum((E-test1mean).* (E-test1mean));
sstot2=sum((X-test2mean).* (X-test2mean));
sstot3=sum((G-test3mean).* (G-test3mean));
R21=1-sserr1/sstot1;
R22=1-sserr2/sstot2;
R23=1-sserr3/sstot3;
R2test=[R21 R22 R23];
plot(bdata(:,1),E,bdata(:,1),yhat(:,1),bdata(:,1),X,bdata(:,1),yhat(:,2),bdata(:,1),G,bdata(:,1),yhat(:,3));

%Various command to write to excel file, can be refined to speed up execution
if nargin<3
    return
else
    if nargin==3
        sheet=1;
    end
    y=[E X G];
    output_data1=theta;
    output_data3=R2test';
    output_data4=bdata(:,1);
output_data5=[y yhat bdata(:,5) bdata(:,6) Y(:,1) Y(:,2)];
xlswrite(out_sheet,['k1' 'k2' 'k4' 'k5' 'k7' 'k8'],sheet,'A1');
xlswrite(out_sheet,output_data1,sheet,'A2');
xlswrite(out_sheet,'R2 eth' 'R2 biom' 'R2 gluc',sheet,'L1');
xlswrite(out_sheet,output_data3,sheet,'L2');
xlswrite(out_sheet,'t',sheet,'A5');
xlswrite(out_sheet,output_data4,sheet,'A6');
xlswrite(out_sheet,['E-E0' 'X-X0' 'G-G0' 'Ehat-Ehat0' 'Xhat-Xhat0'
'Ghat-Ghat0' 'OTR' 'CTR' 'OTRint' 'CTRint'],sheet,'B5');
xlswrite(out_sheet,output_data5,sheet,'B6');
end

Third attempt:

Same as second but normalizing data as commented in the code.

Fourth attempt:

%Input arguments: X,Y
%Output arguments: coef,R2,yhat
%Uses backslash operator to resolve least square problem
function [coef,R2,yhat]=reglineal(X,Y)
coef=lsqnonneg(X,Y);
yhat=X*coef;%Regressed data
ymean=mean(Y);
%ymean=sum(Y)/length(Y);
sserr=sum((Y-yhat).*(Y-yhat));
sstot=sum((Y-ymean).*(Y-ymean));
R2=1-sserr/sstot;%Coefficient of determination

%The inputs are:
%bdata, which has the following structure (vertical vectors):
time,ethanol,
%biomass,glucose,OTR and CTR
%out_sheet, which is the path to the excel file where the output
%are to be stored (optional)
%sheet, which is the name of the sheet of the excel file where the
%will be written (optional, if not used defaults to the first sheet,
%overwriting the present data!)

function [k,coefreg,R2,yhat]=regbatch_nonneg(bdata,out_sheet,sheet)
if nargin==2
    sheet=1;
end
%Note that the function cumtrapz exists and is implemented in MATLAB, but
%it is not used, although it could
OTRint=intacum(bdata(:,1),bdata(:,5));
CTRint=intacum(bdata(:,1),bdata(:,6));
coefreg=ones(2,3);
R2=ones(3,1);
yhat=ones(length(bdata(:,1)),3);
for n=1:3
    E=(bdata(:,2)-bdata(1,2)); %Variable regressed, like E-E0, i.e.
y(:,1)=E;
end
[coefreg(:,1),R2(1,1),yhat(:,1)]=reglineal_nonneg([-OTRint CTRInt],E);
X=(bdata(:,3)-bdata(1,3));
y(:,2)=X;
[coefreg(:,2),R2(2,1),yhat(:,2)]=reglineal_nonneg([OTRint -CTRint],X);
G=(bdata(:,4)-bdata(1,4));
y(:,3)=G;
[coefreg(:,3),R2(3,1),yhat(:,3)]=reglineal_nonneg([OTRint -CTRint],G);
end
%Correcting the signs
coefreg(1,1)=-coefreg(1,1);
coefreg(2,3)=-coefreg(2,3);

%Obtention of the different pseudo-stoichiometric coefficients
k8=-1/coefreg(2,3);
k2=coefreg(2,2)*k8;
k4=coefreg(2,1)*k8;
k5=-k4*k8/(k4*k8*coefreg(1,3)+k8*coefreg(1,1));%Cramer
k7=k8*coefreg(1,1)*(k8)/(k4*k8*coefreg(1,3)+k8*coefreg(1,1));
k1=(coefreg(1,2)*k5*k8+k2*k7)/k8;
k=[k1 k2 k4 k5 k7 k8];
plot(bdata(:,1),y,bdata(:,1),yhat) %Plot of the regressed data and the experimental data

%Various command to write to excel file, can be refined to speed up
%execution
if nargin==1
    return
else
    output_data1=k;
    output_data2=coefreg;
    output_data3=R2';
    output_data4=bdata(:,1);
    output_data5=[y yhat bdata(:,5) bdata(:,6) OTRInt CTRInt];
xlswrite(out_sheet,['k1' 'k2' 'k4' 'k5' 'k7' 'k8'],sheet,'A1');
xlswrite(out_sheet,output_data1,sheet,'A2');
xlswrite(out_sheet,{'reg eth' 'reg biom' 'reg gluc'},sheet,'H1');
xlswrite(out_sheet,output_data2,sheet,'H2');
xlswrite(out_sheet,['R2 eth' 'R2 biom' 'R2 gluc'],sheet,'L1');
xlswrite(out_sheet,output_data3,sheet,'L2');
xlswrite(out_sheet,'t',sheet,'A5');
xlswrite(out_sheet,output_data4,sheet,'A6');
xlswrite(out_sheet,['E-E0' 'X-X0' 'G-G0' 'Ehat-Ehat0' 'Xhat-Xhat0' 'Ghat-Ghat0' 'OTR' 'CTR' 'OTRInt' 'CTRInt'],sheet,'B5');
xlswrite(out_sheet,output_data5,sheet,'B6');
end
Controller code:

Pump calibration:

%------Connection au bioréacteur------
RUN=0;
hr=mxOPC('open','BBI.MFCSSOPCS.1','localhost',1000);
hr=mxOPC('setdoublecache','BIOSTAT C 15l.start.Value',1,0); % for read
hr=mxOPC('Startdoublenotify','BIOSTAT C 15l.start.Value'); % program waits for
change of BIOSTAT C 15l.RUN.Value
hr=mxOPC('setdoublecache','BIOSTAT C 15l.SUBS.Value',1,0); %read
hr=mxOPC('setdoublecache','BIOSTAT C 15l.SUBS.Setpoint',1,1); %readwrite
hr=mxOPC('setdoublecache','BIOSTAT C 15l.WEIGF.Value',1,0); %readwrite
hr = mxOPC('readcache');
[weigfold,hr]=mxOPC('readdouble','BIOSTAT C 15l.WEIGF.Value');
hr = mxOPC('readcache');
[RUN,hr]=mxOPC('readdouble','BIOSTAT C 15l.start.Value')
while (RUN==1)
   hr=mxOPC('wait','BIOSTAT C 15l.start.Value'); % program waits here
   hr = mxOPC('readcache');
   [RUN,hr]=mxOPC('readdouble','BIOSTAT C 15l.start.Value');
end
%------Formation de la droite de calibration------
tcalib=timer('StartDelay',300,'TimerFcn','disp(1)');
h = mxOPC('readcache');
t(1)=300;
h = mxOPC('readcache');
[weigf,hr]=mxOPC('readdouble','BIOSTAT C 15l.WEIGF.Value')
for n=1:11
   hr = mxOPC('readcache');
   Q(n)=weigfold-weigf
   [weigf,hr]=mxOPC('readdouble','BIOSTAT C 15l.WEIGF.Value')
   subs(n)=-10+10*n
tic
   hr=mxOPC('writedouble','BIOSTAT C 15l.SUBS.Setpoint',subs(n));
   hr=mxOPC('writecache');
   start(tcalib)
   wait(tcalib)
   hr = mxOPC('readcache');
   [weigf,hr]=mxOPC('readdouble','BIOSTAT C 15l.WEIGF.Value')
t(n+1)=toc
end
Q(12)=weigfold-weigf;
for n=1:12
   f(n)=Q(n)/t(n);
end
hr=mxOPC('writedouble','BIOSTAT C 15l.SUBS.Setpoint',0);
hr=mxOPC('writecache');
plot(subs,f)
p = polyfit(subs,f,1)
K=p(1)
Control function:

```matlab
function testtimer(obj, event)
    global data
    persistent n weigfold vol c wfeed OTRV CTRV intOTRV intCTRV
    if n>=1
        n=n+1
    else
        n=1
        c=1;
    end
   dens0=1000; %Density of the liquid for initial conditions of the bioreactor
densf=1000; %Density of the feeding liquid
Gin=250; %Concentration of fed glucose in g/l
alpha=1/0.916803236743176;
lambda=0.1;
k4=0.513404732913368;
beta=(0.354721529013343)/(0.916803236743176*0.43300976660018);
gamma=1/0.43300976660018;
Estar=1; %Setpoint for the ethanol signal
K=0.0053617; %Obtained by previous calibration of the pump
o2i=19.9375;
co2i=0.01;
n2i=100-o2i-co2i;
t_ech=60; %Sampling time
hr=mxOPC('readcache');
[weigh,hr]=mxOPC('readdouble','BIOSTAT C 15l.WEIGH.Value')
[weigf,hr]=mxOPC('readdouble','BIOSTAT C 15l.WEIGF.Value')
[o2,hr]=mxOPC('readdouble','BIOSTAT C 15l.O2.Value')
%o2=data(n,1);
%co2=data(n,2);
[airfl,hr]=mxOPC('readdouble','BIOSTAT C 15l.AIRFL.Value');
if n=1
    wsto=weigf
    wfeed=0
    vol=weigh/dens0*1000
    weigfold=weigf
    OTRV=[]
    CTRV=[]
    intOTRV=0
    intCTRV=0
elseif n>1
    wsto=weigfold
    vol=vol+(weigfold-weigf)/densf
    weigfold=weigf
end
OTR=airfl*60/22.4*32/vol*(o2i-o2*n2i/(100-o2-co2))/100%20.95 and 79.011 values of ac, g/l/h
if OTR<0
    OTR=0
end
CTR=airfl*60/22.4*44/vol*(co2*n2i/(100-o2-co2)-co2i)/100%0.039 value of ac, g/l/h
if CTR<0
    CTR=0
end
if mod(n,5)==0
    c=1
end
```
OTRVacum=trapz([1:1:4], OTRV)
CTRVacum=trapz([1:1:4], CTRV)
intOTRV=intOTRV+OTRVacum
intCTRV=intCTRV+CTRVacum
else
  OTRV(c)=OTR/60*vol %To have it in g/l/min
  CTRV(c)=CTR/60*vol
  c=c+1
end
if n<5
  Fin=0
else
  Fin=alpha/Gin*OTR/60*vol+lambda/(k4*Gin)*(Estar*vol-k4*beta*intOTRV-k4*gamma*intCTRV)
end
wfeed=Fin/60*densf*t_ech+(-wsto+weigf+wfeed) %g, revisar, no seguro
if wfeed<0
  wfeed=0
end
subs=1/K*(wfeed/t_ech)
if subs>100
  disp('Pump capacity insufficient, using 100 instead')
  subs=100
elseif subs<0
  disp('Negative pump setpoint, using 0 instead')
  subs=0
end
hr=mxOPC('writedouble','BIOSTAT C 15l.SUBS.Setpoint',subs);
hr=mxOPC('writecache');
data(n,1)=n;
data(n,2)=OTR;
data(n,3)=CTR;
data(n,4)=vol;
data(n,5)=Fin;
data(n,6)=wfeed;
data(n,7)=subs;
data(n,8)=weigf;
data(n,9)=intOTRV;
data(n,10)=intCTRV;
data(n,11)=o2;
data(n,12)=co2;
data(n,13)=airfl;

Initialization script:

RUN=0;
hr=mxOPC('open','BBI.MFCSSOPCS.1','localhost',1000);
pause(5);
if (hr ~=0)
  disp('Error connecting to OPC server');
  break;
end
[batchID,hr]=mxOPC('readstring','BIOSTAT C 15l.Process Unit Information.Batch');
if strcmp('OffVal',batchID)==1
  disp('Batch not defined : start a batch in MFCS');
  break;
elseif
  batchID
hr=mxOPC('setdoublecache','BIOSTAT C 15l.start.Value',1,0); % for read
hr=mxOPC('Startdoublenotify','BIOSTAT C 15l.start.Value');
hr=mxOPC('wait','BIOSTAT C 15l.start.Value'); % program waits for
change of BISTAT C 15l.RUN.Value
hr=mxOPC('setdoublecache','BIOSTAT C 15l.SUBS.Value',1,0); %read
hr=mxOPC('setdoublecache','BIOSTAT C 15l.SUBS.Setpoint',1,1); %readwrite
hr=mxOPC('setdoublecache','BIOSTAT C 15l.WEIGF.Value',1,0);
hr=mxOPC('setdoublecache','BIOSTAT C 15l.CO2.Value',1,0);
hr=mxOPC('setdoublecache','BIOSTAT C 15l.AIRFL.Value',1,0);
hr=mxOPC('setdoublecache','BIOSTAT C 15l.WEIGH.Value',1,0);
hr = mxOPC('readcache');
[wmI,hr]=mxOPC('readdouble','BIOSTAT C 15l.WEIGF.Value')
t=timer('Period',60.0,'TimerFcn',@testtimer,'ExecutionMode','fixedSpacing'); %Timer that calls the function f_echant every minute
hr = mxOPC('readcache');
[RUN,hr]=mxOPC('readdouble','BIOSTAT C 15l.start.Value')
while (RUN~=1)
    hr=mxOPC('wait','BIOSTAT C 15l.start.Value'); % program waits here
    hr = mxOPC('readcache');
    [RUN,hr]=mxOPC('readdouble','BIOSTAT C 15l.start.Value');
end
start(t)