



A comprehensive mechanistic yeast model able to switch metabolism according to growth conditions

Yusmel González-Hernández, Emilie Michiels, Patrick Perré

► To cite this version:

Yusmel González-Hernández, Emilie Michiels, Patrick Perré. A comprehensive mechanistic yeast model able to switch metabolism according to growth conditions. *Fermentation*, 2022, 8 (12), pp.710. <10.3390/fermentation8120710>. <hal-04441254>

HAL Id: hal-04441254

<https://hal.science/hal-04441254v1>

Submitted on 9 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

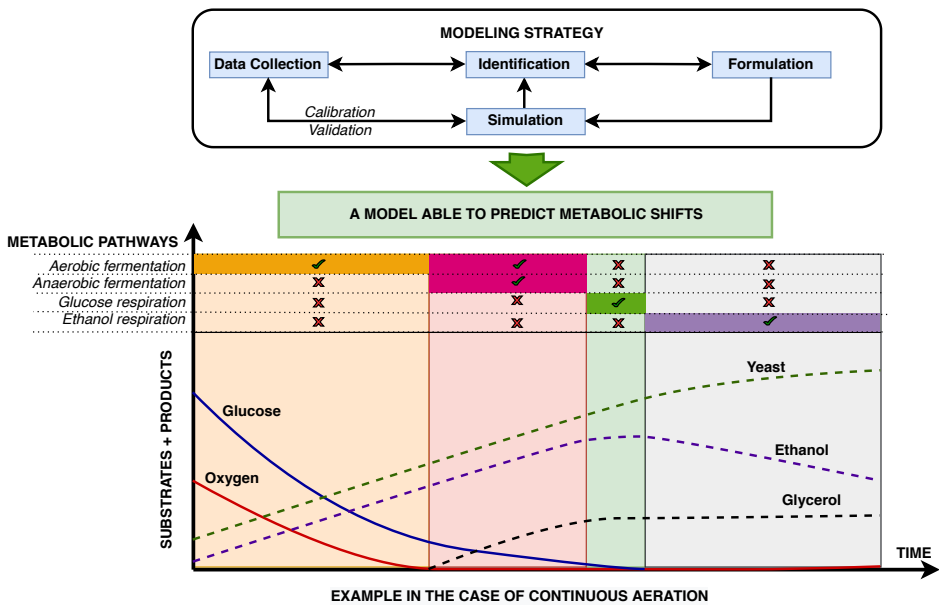


HAL Authorization

Graphical Abstract

A comprehensive mechanistic yeast model able to switch metabolism according to growth conditions

Yusmel González-Hernández, Emilie Michiels, Patrick Perré



A comprehensive mechanistic yeast model able to switch metabolism according to growth conditions

Yusmel González-Hernández^a, Emilie Michiels^a, Patrick Perré^{a,b}

^a*LGPM, CentraleSupélec, Centre Européen de Biotechnologie et de Bioéconomie (CEBB), 3 rue des
Rouges Terres, 51 110 Pomacle, France*

^b*LGPM, CentraleSupélec, Université Paris-Saclay, 3 rue Joliot Curie, F-91 190 Gif-sur-Yvette, France*

Abstract

This paper proposes a general approach for building a mechanistic model that is able to predict the shift of metabolic pathways. This general approach is applied to a commercial strain of *Saccharomyces cerevisiae*. The mechanistic model accounts for the coexistence of several metabolic pathways whose activation depends on growth conditions. Stoichiometry and yeast kinetics were experimentally determined to the extent that was possible under aerobic and completely anaerobic conditions. Known parameters were taken from the literature, and the remaining parameters were estimated by inverse analysis using particle swarm optimization method.

The mechanistic model switches from aerobic glucose fermentation to glucose respiration when the glucose concentration is below 0.10-0.15 g/l and to ethanol respiration after glucose depletion. Different affinities of yeast for glucose and ethanol tolerance under aerobic and anaerobic conditions were obtained. Finally, the model was successfully tested against a validation test, a batch fermentation process without gas injection, with an overall mean relative error of 7 %.

This model represents a useful tool for the control and optimization of yeast fermentation systems. More generally, the modeling framework proposed here is intended to be used as a building block of a digital twin of any bioproduction process.

Highlights

- A set of original yeast cultures with contrasting growth conditions,
- A complete formulation of yeast growth with co-current metabolic pathways and adapted shift functions,
- A particle swarm optimization method was used to determine the most difficult parameters,
- Different glucose affinity and ethanol tolerance under aerobic and anaerobic conditions,
- Model successfully validated using a commercial yeast strain under different operation conditions,

Email address: yusmel.gonzalez-hernandez@centralesupelec.fr (Yusmel González-Hernández)

- The formulation and the whole set of parameters are provided for future use.

Keywords: yeast, fermentation, Crabtree effect, switching metabolism, modeling, calibration

Nomenclature

δ	Safety factor of switching function for the ethanol respiration process activation
$\mu_{max(i)}$	Maximum growth rate for process $-i$ in the matrix model representation (h^{-1})
$\rho(T)$	Water density at temperature T (mol/l)
$v_{(i,j)}$	Stoichiometric coefficient corresponding to the biological process $-i$ and the component $-j$
$b_{(i)}$	Oxygen stoichiometric coefficient for biological process $-i$ in the matrix model representation (mol)
$c_{(i)}$	Ammonia stoichiometric coefficient for biological process $-i$ in the matrix model representation (mol)
$C_{(j)}$	Concentration of the component $-j$ in the matrix model representation (g/l)
$d_{(i)}$	Yeast stoichiometric coefficient for biological process $-i$ in the matrix model representation (mol)
$e_{(i)}$	Ethanol stoichiometric coefficient for biological process $-i$ in the matrix model representation (mol)
$f_{(i)}$	Glycerol stoichiometric coefficient for biological process $-i$ in the matrix model representation (mol)
$g_{(i)}$	Carbon dioxide stoichiometric coefficient for biological process $-i$ in the matrix model representation (mol)
$h_{(i)}$	Water stoichiometric coefficient for biological process $-i$ in the matrix model representation (mol)
$H_{(T)}$	Henry constant at temperature T (mol/l/atm)
$K_{E,Ax}^{inh}$	Half-saturation coefficient for ethanol inhibition under anaerobic conditions (g/l)
K_E	Half-saturation coefficient for ethanol (g/l)
$K_{Glc,Ax}$	Half-saturation coefficient for glucose under anaerobic conditions (g/l)
$K_{Glc,Ox}$	Half-saturation coefficient for glucose under aerobic conditions (g/l)
$K_{iE,Ox}^{inh}$	Half-saturation coefficient for ethanol inhibition under aerobic conditions (g/l)
$k_L a$	Overall oxygen transfer rate (h^{-1})
K_{O_2}	Half-saturation coefficient for oxygen (mg/l)
$K_{S(j)}$	Half-saturation coefficient for the limiting substrate $S_{(j)}$ in the matrix model representation (g/l)
$K_{S(j)}^{inh}$	Half-saturation coefficient for the inhibiting by-product $S_{(j)}$ in the matrix model representation (g/l)

K_S	Half-saturation coefficient for substrate S (g/l)
M_E	Ethanol molar mass (g/mol)
M_{Glc}	Glucose molar mass (g/mol)
M_{Gly}	Glycerol molar mass (g/mol)
M_{O_2}	Oxygen molar mass (g/mol)
M_{X_Y}	Yeast molar mass (g/mol)
P_{O_2}	Oxygen partial pressure in the air (atm)
$R_{(i)}$	Specific growth rate of the biological process $-i$ in the matrix model representation (g/l/h)
S	Substrate concentration (g/l)
$S_{O_2}^*$	Saturation oxygen concentration (g/l)
$S_{(j)}$	Limiting substrate $-j$ in the matrix model representation (g/l)
$S_{(j)}^{inh}$	Inhibiting by-product $-j$ in the matrix model representation (g/l)
S_E	Ethanol concentration (g/l)
$S_{Glc,c}$	Critical glucose concentration for switching from aerobic fermentation to respiration (g/l)
S_{Glc}	Glucose concentration (g/l)
S_{Gly}	Glycerol concentration (g/l)
S_{O_2}	Dissolved oxygen concentration (g/l)
X_Y	Yeast concentration (g/l)
$y_{e(j,k)}^{(i)}$	Experimental data value $-k$ of variable $-j$ in the experiment $-i$
$y_{m(j,k)}^{(i)}$	Model output data value $-k$ of variable $-j$ in the experiment $-i$

1. Introduction

Following the growing popularity of digital twins in bio-production ([Muldbak et al., 2022](#); [Sokolov et al., 2021](#); [Babi et al., 2022](#)), mechanistic modeling has received renewed attention by the scientific community. Whatever the complexity of digital twins at the plant level, mechanistic models remain important, as they provide an excellent summary of available process knowledge. Although scientists and industry experts use these models efficiently, they can be further improved by Machine Learning, either using data taken from online sources or existing databases. In addition, such models are useful for planning experiments and determining which critical process variables need to be monitored and controlled tightly ([Lencastre Fernandes et al., 2012](#)). More precisely, these mechanistic models allow a better understanding, description, and quantification of the phenomena involved in highly important and complex bioprocesses, such as alcoholic fermentation using the yeast *Saccharomyces cerevisiae*.

Saccharomyces cerevisiae is a Crabtree positive yeast of great importance for various biotechnological applications, some of which date back to several thousands of years ([Parapouli et al., 2020](#)). This yeast strain is commonly used for its capacity to rapidly convert

sugars to ethanol and carbon dioxide under both aerobic and anaerobic conditions (Hagman et al., 2014). Although the Crabtree effect has been extensively studied, much remains to be well understood about this phenomenon. Under aerobic conditions, the alcoholic fermentation process occurs when the glucose concentration exceeds 0.10 - 0.15 g/l (Verduyn et al., 1984), switching to the respiration process when glucose concentration is below these values. Once the glucose is depleted, ethanol respiration takes place. Even under aerobic conditions, the anaerobic metabolic pathway can also be activated when the rate of biological oxygen uptake exceeds the rate of oxygen supply, which is identifiable by the production of glycerol. In the absence of molecular oxygen, *S. cerevisiae* carries out the anaerobic fermentation process, producing glycerol for the cytosolic redox balance (Bakker et al., 2001).

The above observations confirm the complexity of yeast metabolism, and its use requires precise control of the process to obtain maximum productivity and quality products. Modeling has proven to be a powerful ally in explaining yeast metabolism and a useful tool for optimizing and controlling fermentation processes under aerobic and anaerobic conditions.

Several mechanistic yeast models have been developed using the typical Monod-type expression ($S/(K + S)$). This mathematical expression considers the limiting substrates as glucose, nitrogen and oxygen. Glucose is particularly important to *Saccharomyces cerevisiae*, as it is by far the yeast’s preferred carbon source. Yeast cells can sense glucose and utilize it efficiently over a broad range of concentrations, from a few micromolar to even a few molar (Johnston and Kim, 2005). Nitrogen is also an essential element in *S. cerevisiae* composition, since it is mandatory for protein synthesis and represents 9 % (w/w) of yeast biomass (Verduyn et al., 1990b). Oxygen is required to regenerate NAD^+ used in the glycolytic pathway of biomass formation, closing the redox balance for the co-enzyme system NAD^+/NADH . The oxidation of cytosolic NADH into NAD^+ can occur through mitochondrial respiration with external NADH dehydrogenase (Bakker et al., 2001; Overkamp et al., 2000). Oxygen is also important for the synthesis of yeast membrane compounds (sterols and unsaturated fatty acids) (Sablayrolles and Barre, 1986), though this process could be neglected since the required amount is very weak at between 0.3 and 1.5 $\text{mgO}_2 \text{gDW}^{-1}$ (Rosenfeld et al., 2003). *Saccharomyces cerevisiae* is a superb ethanol producer yet is also sensitive to higher ethanol concentrations, especially under high gravity or very high gravity fermentation conditions. Ethanol tolerance is associated with the interplay of complex networks at the genome level. Although significant efforts have been made to study ethanol stress response in past decades, mechanisms of ethanol tolerance are not well known (Ma and Liu, 2010; Sostaric et al., 2021). Eukaryotic cells have developed diverse strategies to combat the harmful effects of a variety of stress conditions. In the model yeast *Saccharomyces cerevisiae*, the increased concentration of ethanol as the primary fermentation product will influence the membrane fluidity and be toxic to membrane proteins, leading to cell growth inhibition and ultimately death (Ding et al., 2009).

These limiting substrates and by-products’ inhibition effect on yeast growth rate should be considered in any model describing the metabolism of this complex yeast in order to adequately describe the fermentation process under both aerobic and anaerobic conditions. Each limiting substrate can be easily included in the model with the Monod-like function, and several alternative expressions exist for the inhibiting by-products’ effects (Ghose and Tyagi, 1979; Carteni et al., 2020; Pham et al., 1998; Scheiblaue et al., 2018;

Veloso et al., 2019; Amenaghawon et al., 2012; Luong, 1985). However, the Monod-like function and even the conventional expression used by computer scientists to shift from a limiting substrate to an inhibiting one ($K/(K + S)$) are not suitable to model the shift between aerobic fermentation and respiration in the particular case of the Crabtree effect. This phenomenon requires a more precise switching function that allows aerobic fermentation to be turned off and glucose respiration to be turned on when glucose concentration falls under 0.10-0.15 g/l. Several models have already been developed to simulate the fermentation process with *Saccharomyces cerevisiae* achieving significant advances in its metabolism description (Hanly and Henson, 2013; Sulieman et al., 2018; Thierie, 2004, 2019; La et al., 2020; Pham et al., 1998; Sonnleitner and Käppli, 1986; Scheiblaue et al., 2018). However, much work remains to be done for increasing the accuracy of the models in terms of triggering/inhibiting metabolism pathways when the environmental conditions change.

The main objective of this work is to propose an accurate mechanistic model, capable of predicting the metabolic shift from glucose aerobic fermentation to glucose respiration when glucose concentration is lower than 0.10-0.15 g/l and to ethanol respiration once glucose is depleted. This mechanistic model is developed to be as simple as possible, easy to use, and adaptable to the conditions of each system. The model can be adapted to a significant number of existing mutated yeast strains used currently in the industry. In addition, the model activates/deactivates simultaneous anaerobic fermentation in the absence/presence of dissolved oxygen, respectively. This integral model will be calibrated by combining experimental data generated using a commercial yeast strain *Saccharomyces cerevisiae* used for wine production and modeling in an inverse method where the metabolic pathways' stoichiometry and kinetics are determined independently when possible.

2. Materials and methods

2.1. Experimental setup

A Sartorius Biostat B Plus bioreactor (Fig. 1) of 5 liters in batch mode was used for the experimental tests.

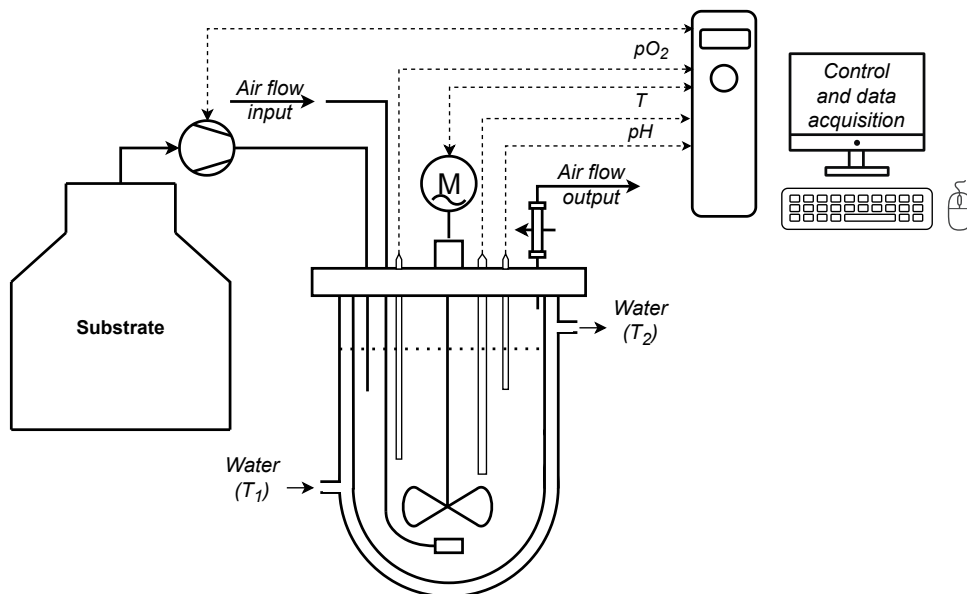


Figure 1: Experimental setup of the bioreactor.

The yeast used in this study is a strain commercialized by the Institut Oenologique de Champagne under the name IOC Fizz+. This yeast strain is presented in the form of Active Dry Yeast (ADY). It is a *Saccharomyces cerevisiae* yeast resistant to alcohol up to 14 % vol. (approximately 110 g/l of ethanol) and possessing the killer factor (K2). These properties allow it to grow while blocking wild yeasts that do not possess this K2 factor. After rehydration, the number of living yeast is more than 10 billion cells/g with a purity of less than 10 wild yeasts per million cells. The culture medium was composed of glucose (10-35 g/l), peptone (20 g/l), yeast extract (20 g/l), and silicone antifoaming. The initial temperature was set at 25 °C. The bioreactor was continuously stirred at 295 rpm. It was operated under aerobic and anaerobic conditions ensured by continuous injection of 500 ccm of sterilized air and nitrogen gas flow respectively. A cooling finger was installed on the gas output to reduce the ethanol and water losses caused by vaporization. The pH and dissolved oxygen values in the liquid phase were continuously measured using internal Hamilton electrodes, while a Pt100 probe was used for temperature measurement.

2.2. Analytical methods

2.2.1. Yeast, glucose, ethanol and glycerol concentration measurements

The yeast concentration was determined using samples of known volumes taken from the bioreactor. These samples were centrifuged in a Centrifuge 5804 R Eppendorf at 2 °C and 5800 g during ten minutes, dried in a Memmert oven at 105 °C during seven days, and weighed. Glucose, ethanol, and glycerol concentrations were measured by high-pressure liquid chromatography with refractive index detection. A Thermo Scientific UltiMate 3000 (HPLC-RI) equipped with a Bio-Rad Aminex HPX-87H column performed the measurements. The mobile phase had a flowrate of 0.5 ml/min with 2 mM H₂SO₄, and the column temperature was controlled at 30 °C. The supernatants were filtered by HPLC-Certified Syringe Filter Whatman Spartan 0.20 µm/30 mm of regenerated cellulose before the HPLC analysis.

2.2.2. Determination of parameter k_La

The procedure proposed by [Garcia-Ochoa and Gomez \(2009\)](#) was used for the k_La determination, which consisted in filling the bioreactor with the culture medium without yeast inoculation. The dynamic technique of absorption consists of producing the elimination of oxygen in the liquid phase by means of bubbling nitrogen until the oxygen concentration is equal to zero. Bubbling was then changed to air, and the time-increase of oxygen concentration was recorded. The k_La value can be calculated as the slope of the resulting straight line representing the $\ln \left(\frac{S_{O_2}^* - S_{O_2}}{S_{O_2}^*} \right)$ versus time:

$$\ln \left(\frac{S_{O_2}^* - S_{O_2}}{S_{O_2}^*} \right) = -k_La \cdot t \quad (1)$$

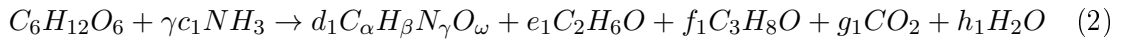
3. Theoretical formulation

We applied the following assumptions in our model:

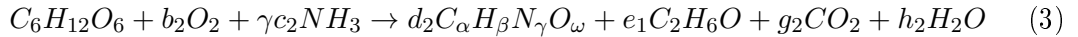
1. A perfectly stirred bioreactor.
2. Constant yeast kinetics parameter values without considering the temperature effect.
3. Constant stoichiometric parameter values.
4. A constant chemical composition of yeast, independent of the metabolic process and substrate nature (glucose or ethanol). The yeast elementary composition obtained by [Rieger et al. \(1983\)](#) expressed per mol of carbon ($C_\alpha H_\beta N_\gamma O_\omega$, in which $\alpha = 1$, $\beta = 1.79$, $\gamma = 0.15$ and $\omega = 0.57$) was used.
5. Glycerol and ethanol formation pathways which are included; however, pathways for the production of other by-products, such as fusel alcohols, were not considered.
6. A constant oxygen transfer rate coefficient k_La without including the influence of medium properties variation throughout the study.

3.1. Stoichiometry

In the absence of molecular oxygen, *S. cerevisiae* carries out anaerobic fermentation (Eq. 2), producing glycerol for the cytosolic redox balance ([Bakker et al., 2001](#)).



It is well known that *S. cerevisiae* is a Crabtree positive yeast that exhibits fermentation in aerobic conditions producing ethanol (Eq. 3) when the glucose concentration exceeds 0.10 - 0.15 g/l ([Verduyn et al., 1984](#)).



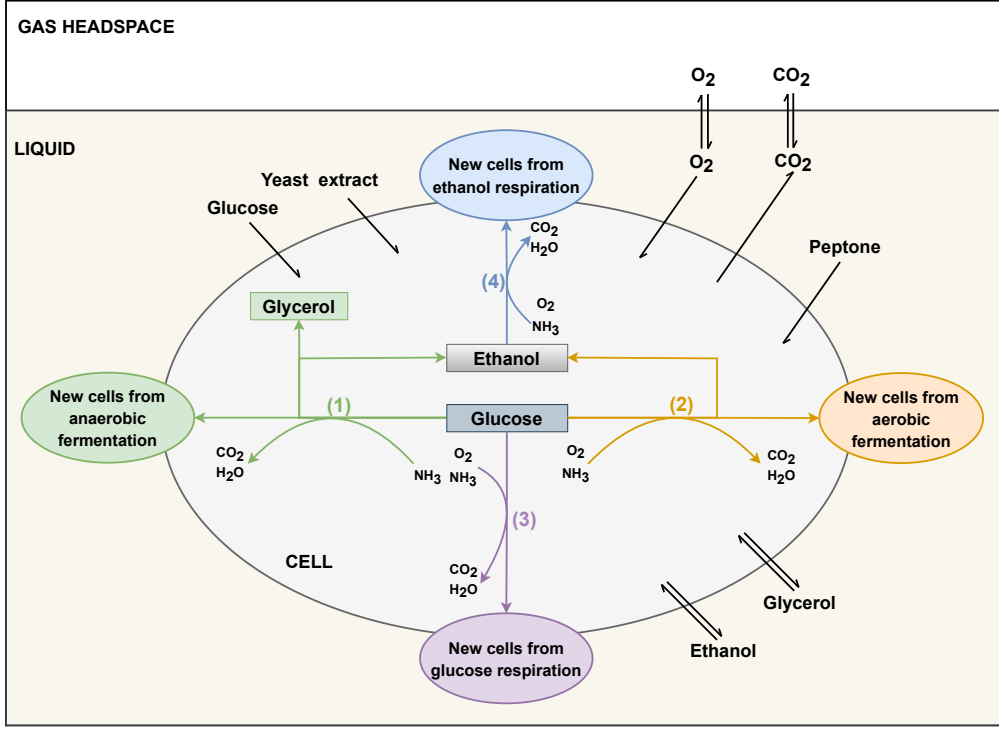
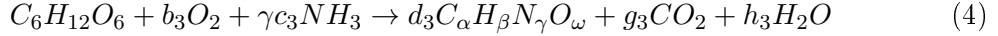


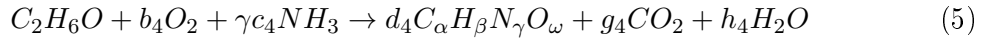
Figure 2: Metabolic pathways of *Saccharomyces cerevisiae* considering the Crabtree effect. (1) Aerobic fermentation (glucose concentration $> 0.10 - 0.15$ g/l), (2) Respiration based on glucose (glucose concentration $< 0.10 - 0.15$ g/l), (3) Anaerobic fermentation and (4) Respiration based on ethanol.

Once the glucose concentration reaches values below this threshold, the metabolism switches from aerobic fermentation to respiration (Eq. 4):



Even when aerobic conditions are established in our model, the anaerobic fermentation pathway is still included. In some cases, the specific respiration rate in a bioreactor is limited by low concentrations of dissolved oxygen. The anaerobic fermentation pathway is then partly activated, in which case both ethanol and glycerol are produced. In anaerobic cultures of wild-type *Saccharomyces cerevisiae*, glycerol production is essential to reoxidize NADH produced in biosynthetic processes. Glycerol consequently becomes a major by-product during anaerobic production of ethanol by *S. cerevisiae*, the largest fermentation process in industrial biotechnology (Nissen et al., 2000; Medina et al., 2010).

Once glucose is completely depleted, ethanol degradation by yeast occurs under aerobic conditions:



These four metabolic pathways are included in the model shown in figure 2. Depending on the growth conditions, each pathway can be inhibited or partially/fully activated. The following section describes the activation/deactivation of these metabolic pathways via kinetics expressions.

3.2. Kinetics

As the YPD medium used here supplies enough nitrogen, no nitrogen limitation was taken into account in the metabolic pathways. For the other substrates, the limiting effect on specific growth rate was described as the product of the switching mathematical expressions proposed by Monod, which is represented in the first product term of Eq. 6. In contrast the inhibiting by-products' effect and the metabolism switching from aerobic to anaerobic conditions and glucose to ethanol respiration were described as a product of a conventional mathematical expression used in most of the biological models, which is represented in the second product term of Eq. 6

$$\frac{dX_Y}{dt} = \mu_{max(i)} \prod_{j=1}^n \frac{S_{(j)}}{K_{S(j)} + S_{(j)}} \prod_{j=n+1}^k \frac{K_{S(j)}^{inh}}{K_{S(j)}^{inh} + S_{(j)}^{inh}} X_Y, \quad (6)$$

where $\mu_{max(i)}$ is the maximum growth rate for process $-i$, $S_{(j)}$ is the limiting substrate $-j$, $K_{S(j)}$ is the half-saturation coefficient for the limiting substrate $S_{(j)}$, $S_{(j)}^{inh}$ is the inhibiting by-product $-j$, and $K_{S(j)}^{inh}$ is the half-saturation coefficient for the inhibiting by-product $S_{(j)}^{inh}$.

In the Monod-like function, S_{Glc} , is a generic substrate concentration, K_{Glc} defines the steepness of the function near zero, and the rate is divided by 2 when $S_{Glc} = K_{Glc}$.

As the Monod-like function is not suitable to model the shift between fermentation and respiration, a smoothed C^∞ stepwise function proposed by La et al. (2020) was used instead:

$$\lambda_r = \frac{1 + \tanh(\alpha(S_{Glc}/S_{Glc,c} - 1))}{2} \quad (7)$$

The stepwise function is defined by two parameters: the shift value $S_{Glc,c}$ defines the concentration value at which transition occurs, and the α parameter defines the sharpness of this transition. In this case, taking $\alpha = 150$ and $S_{Glc,c} = 0.125$ g/l (the mean value of the glucose transition range), the metabolism transition occurs in the interval 0.10-0.15 g/l of glucose (Fig. 3).

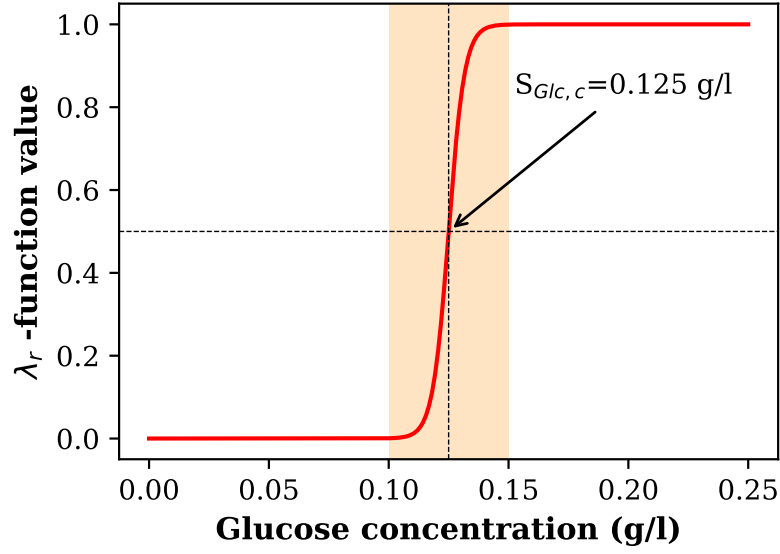


Figure 3: Shift function used to obtain a metabolism transition between aerobic fermentation and respiration processes over the glucose concentration range of 0.10 - 0.15 g/l.

3.3. Matrix model representation

A detailed description of the stoichiometry and kinetics of the model's biological processes is shown in Table 1.

Table 1: Matrix representation of the biological model

i ↓	Component → j Process	1 S_{Glc}	2 S_{O_2}	3 S_E	4 S_{Gly}	5 X_Y	Rate
1	Anaerobic fermentation based on glucose	$-\frac{M_{Glc}}{d_1 M_{X_Y}}$		$\frac{e_1 M_E}{d_1 M_{X_Y}}$	$\frac{f_1 M_{Gly}}{d_1 M_{X_Y}}$	1	R_1
2	Aerobic fermentation based on glucose	$-\frac{M_{Glc}}{d_2 M_{X_Y}}$	$-\frac{b_2 M_{O_2}}{d_2 M_{X_Y}}$	$\frac{e_2 M_E}{d_2 M_{X_Y}}$		1	R_2
3	Respiration based on glucose	$-\frac{M_{Glc}}{d_3 M_{X_Y}}$	$-\frac{b_3 M_{O_2}}{d_3 M_{X_Y}}$			1	R_3
4	Respiration based on ethanol		$-\frac{b_4 M_{O_2}}{d_4 M_{X_Y}}$	$-\frac{M_E}{d_4 M_{X_Y}}$		1	R_4
5	Aeration		1				R_5

The process expression rates are shown in Appendix A

The system of differential equations that describes the biological system is represented in matrix form, establishing the corresponding balance equations

$$\frac{dC_{(j)}}{dt} = \sum_{i=1}^5 \sum_{j=1}^5 v_{(i,j)} R_{(i)}, \quad (8)$$

where $C_{(j)}$ is the concentration of component $-j$, $v_{(i,j)}$ the stoichiometric coefficient corresponding to biological process $-i$ and component $-j$, and $R_{(i)}$ is the rate of the biological process $-i$.

3.4. Model calibration

The yeast stoichiometry and kinetics of each metabolic pathway were studied independently to the extent that was possible. Experiments with the injection of nitrogen or air were performed to achieve complete anaerobic fermentation (anaerobic sub-model) or aerobic conditions (aerobic sub-model) respectively (step 1, Fig. 4). Six batch fermentation experiments (A-F, Table 2) with different initial glucose, ethanol and yeast concentrations were performed to observe the effect of glucose, and ethanol on the yeast kinetics during calibration, while experiment G was used to test the calibrated model. The initial conditions of the batch experiments are shown in Table 2.

Table 2: Initial conditions of batch fermentation experiments performed under anaerobic and aerobic conditions

Exp.	S_{Glc} (g/l)	S_{O_2} (mg/l)	S_E (g/l)	S_{Gly} (g/l)	X_Y (g/l)	T (°C)	Conditions
A	13.75	-	52.13	0.26	2.43	26.10	N ₂ injection
B	15.37	-	10.60	0.12	0.30	25.20	N ₂ injection
C	31.06	-	10.99	0.15	0.81	26.50	N ₂ injection
D	10.24	6.08	49.44	0.16	0.81	24.60	Air injection
E	15.93	5.84	9.11	0.10	0.50	26.10	Air injection
F	30.79	2.86	9.84	0.25	1.31	25.20	Air injection
G	12.41	-	1.27	0.15	0.40	23.80	-

Most of model parameters in this work were experimentally determined to the extent that was possible. The remaining model parameters were taken either from the literature or estimated by inverse analysis using the particle swarm optimization method. The oxygen and ethanol half-saturation coefficients were considered the same for all aerobic metabolic pathways. However, for the first time to our knowledge, different glucose affinity and ethanol tolerance for aerobic and anaerobic conditions were applied in the model, establishing different half-saturation coefficients for glucose and ethanol inhibition. This modeling process was guided by the identification procedure and is notably different from what the literature reports, i.e., that average values of these parameters are used for all metabolic pathways and both for aerobic and anaerobic conditions. The modeling strategy used in the calibration of the developed model is explained in figure 4. This strategy requires advancing step by step towards maximum complexity, starting with the analysis of experiments under nitrogen injection, which ensures that one single metabolic pathway exists (Eq. 2).

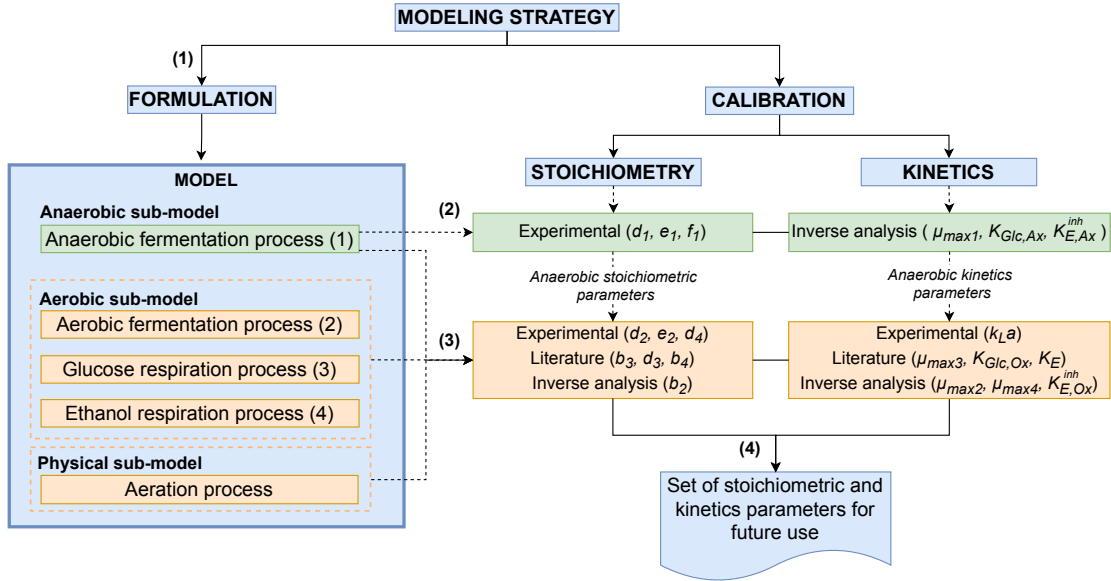


Figure 4: Modeling strategy.

The procedure varies according to the specific metabolic pathway:

Anaerobic fermentation process

The stoichiometry of the anaerobic fermentation process was initially determined experimentally using three experiments (A, B, and C). Kinetics parameters (maximum growth rate, half-saturation coefficients for glucose and ethanol inhibition under anaerobic conditions) were then estimated by inverse analysis (step 2, Fig. 4).

Aerobic fermentation processes

Three experiments (D, E, and F) were used to determine the stoichiometry and kinetics of the different aerobic metabolic pathways. The inverse analysis was performed, applying the previously calibrated anaerobic sub-model. This application was required, since anaerobic conditions were observed during the aerobic fermentation process when the rate of biological oxygen uptake was higher than the rate of oxygen dissolution provided by the aeration process (step 3, Fig. 4).

- For the aerobic fermentation process ($S_{Glc} > 0.10 - 0.15g/l$), all stoichiometric coefficients were determined experimentally, with the exception of the oxygen stoichiometric coefficient determined by inverse analysis. The inverse analysis was required due to the impossibility of performing mass balances caused by the lack of information about the specific form of nitrogen assimilated by the biomass, as supplied by the YPD medium. Some kinetics parameters such as oxygen and glucose half-saturation coefficients under aerobic conditions were taken from the literature (Fig. 4), while the maximum growth rate and half-saturation coefficient for ethanol inhibition under aerobic conditions were determined by inverse analysis.
- As the glucose respiration process ($S_{Glc} < 0.10 - 0.15g/l$) has a short duration involving low values of glucose concentration, it is sometimes difficult to estimate its

stoichiometry and kinetics experimentally. Therefore, these parameters were taken entirely from the literature (step 3, Fig. 4).

- The stoichiometry of the ethanol respiration process was also experimentally determined using the three experiments. The oxygen stoichiometric coefficient is also difficult to determine experimentally; this coefficient was also taken from the literature, as was the half-saturation coefficient for ethanol. The maximum growth rate was obtained by inverse analysis (step 3, Fig. 4).

The inverse analysis was performed by means of an optimization procedure using the mean relative error as the objective function between the experimental data (yeast, ethanol, glycerol, and glucose concentrations) and the model prediction

$$MRE = \frac{1}{lmn} \sum_{i=1}^{\ell} \sum_{j=1}^m \sum_{k=1}^n \left| \frac{y_{e(k,j)}^{(i)} - y_{m(k,j)}^{(i)}}{y_{e(k,j)}^{(i)}} \right|, \quad (9)$$

where $y_{e(j,k)}^{(i)}$ is the experimental data value $-k$ of variable $-j$ in experiment $-i$ and $y_{m(j,k)}^{(i)}$ is the model output data value $-k$ of variable $-j$ in experiment $-i$.

The model was implemented in Python, where the ODE system was numerically solved using the *solve_ivp* method from the *integrate* SciPy package. The *LSODA* integration method was used for solving the ODE system; this method is a wrapper of the Fortran solver from ODEPACK that consists of an Adams/BDF method with automatic stiffness detection and switching (Hindmarsh, 1982; Petzold, 1983).

The calibration process was carried out using the Particle Swarm Optimization (PSO) method from PySwarm package. The PSO algorithm is based on a simplified social model that is closely tied to swarming theory. A physical analogy for the PSO algorithm might be a swarm of birds searching for a food source. In this analogy, each bird (referred to as the particle) makes use of its own memory as well as knowledge gained by the swarm as a whole to find the best available food source (Venter and Sobieszczanski-Sobieski, 2003). This algorithm is mainly described by the relative weights of three different mechanisms (inertia, memory and socialization), which in this case were set to 0.6, 0.5, and 0.5 respectively, leading to fast convergence. The calibration parameter ranges were defined using the values reported in the literature.

The quality of the model in terms of experimental data fitting was also evaluated by estimating Pearson’s correlation. This correlation was computed using the *pearsonr* method from the *stats* SciPy package in Python.

4. Results and discussion

4.1. Modeling yeast activity under completely anaerobic conditions

Table 3 summarizes the stoichiometry and kinetics parameters of the biological model under completely anaerobic conditions (anaerobic sub-model), which were obtained from the calibration procedure. As explained above, stoichiometry was obtained experimentally under completely anaerobic conditions, while kinetics was determined by inverse analysis. The inverse analysis combining experimental data and modeling, using three experimental data sets obtained for different glucose, ethanol and yeast concentrations (experiments A, B, and C).

Table 3: Stoichiometry and kinetics parameters of the biological model under complete anaerobic conditions

Parameters name	Symbol	Value	Source
Yeast stoichiometric coefficient in anaerobic fermentation process (mol)	d_1	1.05 ± 0.04	Experimental
Ethanol stoichiometric coefficient in anaerobic fermentation process (mol)	e_1	1.70 ± 0.01	Experimental
Glycerol stoichiometric coefficient in anaerobic fermentation process (mol)	f_1	0.10 ± 0.00	Experimental
Half-saturation coefficient for glucose under anaerobic conditions (g/l)	$K_{Glc,Ax}$	1.72 ± 0.01	Calibrated
Half-saturation coefficient for ethanol inhibition under anaerobic conditions (g/l)	$K_{E,Ax}^{inh}$	202.83 ± 0.97	Calibrated
Maximum growth rate for anaerobic fermentation process (h^{-1})	μ_{max1}	0.31 ± 0.00	Calibrated

The optimization process using the PSO method converged to a unique minimum, which is demonstrated by the low standard deviation values obtained for each parameter during the optimization process (below 0.5 %, as can be seen in Table 3). Using the optimal parameters, the model was able to reproduce all the experiments with mean relative errors under 7 % and with a unique set of parameters (Table 4).

Table 4: Mean relative errors reported between model output and experimental data under completely anaerobic conditions

No.	Experiment	$S_{Glc}(\%)$	$S_E(\%)$	$S_{Gly}(\%)$	$X_Y(\%)$	Mean value(%)
1	A	21.37	1.14	0.96	2.65	6.53
2	B	9.76	3.55	5.19	7.94	6.61
3	C	17.44	3.74	4.07	4.32	7.39
	Mean value(%)	16.19	2.81	3.41	4.97	6.84

Interestingly, the maximum growth rate obtained during the calibration process is the same reported by Verduyn et al. (1990a) (0.31 h^{-1}), which validates the quality of the calibration process. On the other hand, the value of the half-saturation coefficient obtained for glucose (1.72 g/l) reveals a larger need for this source of carbon in relation to what is reported in the literature (1×10^{-3} to 0.5 g/l) (Hanly and Henson, 2013; Sulieman et al., 2018; Thierie, 2004, 2019; La et al., 2020; Pham et al., 1998; Sonnleitner and Käppeli, 1986; Scheiblaue et al., 2018). Similarly, the high value obtained for the half-saturation coefficient for ethanol inhibition (202.83 g/l) indicates that *S. cerevisiae* has a high tolerance to ethanol under anaerobic conditions, significantly higher than those reported in the literature (10-26.97 g/l) (Scheiblaue et al., 2018; Pham et al., 1998).

The model provides very good results in the prediction of the variables behavior (Table 4, Fig. 5); however the half-saturation coefficient for ethanol inhibition obtained in this model might not be adequate to simulate anaerobic fermentation processes when dealing with higher ethanol concentrations. According to the inhibition function used in the model

and the value of the half-saturation coefficient found for ethanol inhibition, the growth rate would be divided by a factor of two for an ethanol concentration of 202.83 g/l. Yet according to Arroyo-López et al. (2010), no yeast activity should be observed at this concentration.

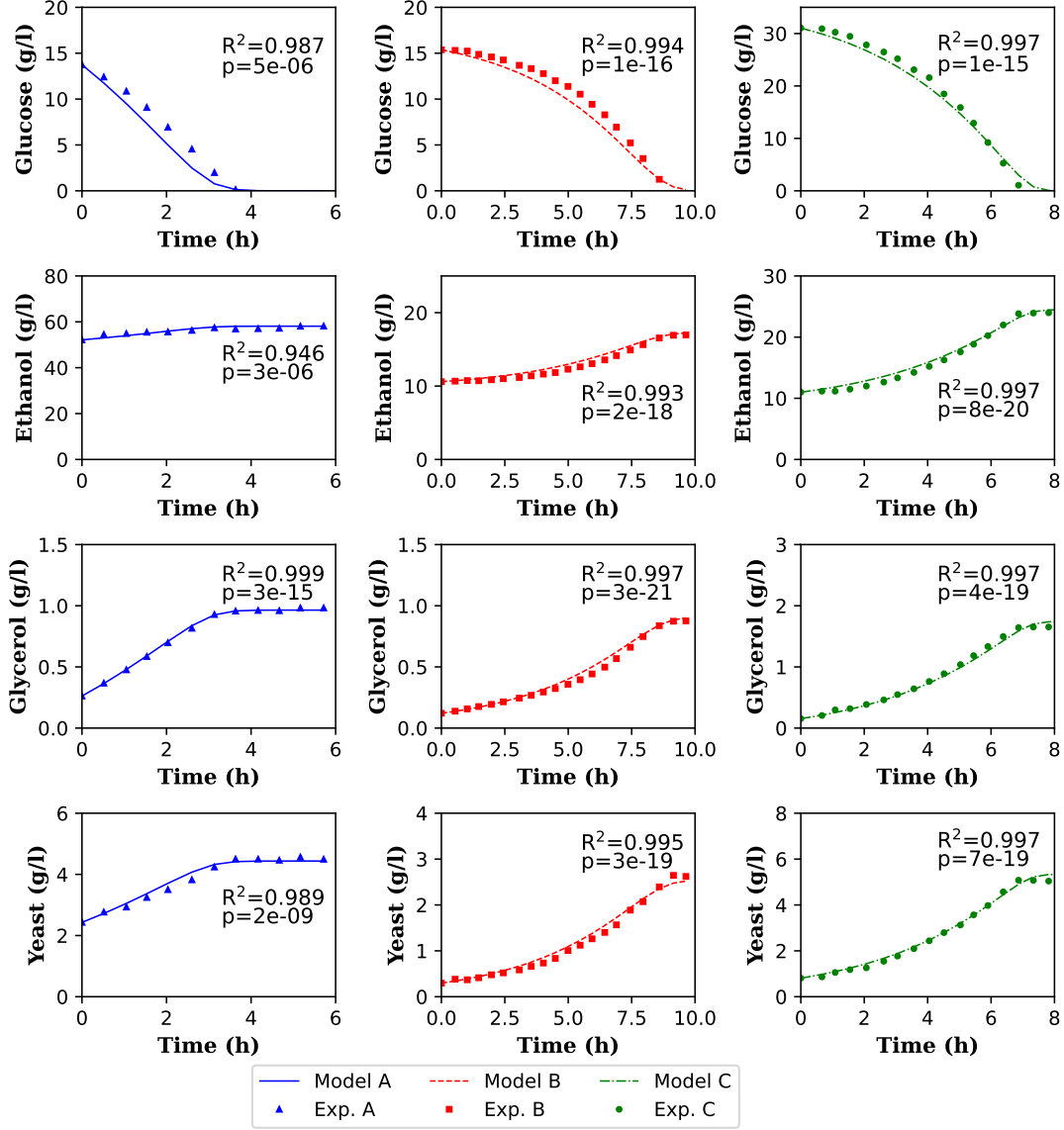


Figure 5: Yeast activity simulation under complete anaerobic conditions. Comparison between model output and experimental data.

The high value of half-saturation coefficient for ethanol inhibition results from the combination of both the experimental range of ethanol tested here and the shape of the inhibition function, which is too loose to account for toxicity. This problem could be solved with a recalibration process using a wider ethanol concentration range and a different mathematical expression for the inhibition function. For example, applying the λ_r function (Eq. 7), would involve a critical value (the ethanol concentration value above which no yeast activity is observed) and the range of inhibitory effect (α). This approach could be useful

for future modeling research.

In general, the values obtained for the half-saturation coefficients for glucose and ethanol inhibition through the inverse analysis are outside the ranges reported in the literature. Yet these values describe very well the behavior of the main variables of the model. This is an interesting result since it is the first time, to our knowledge, that these parameters are differentiated in yeast metabolic models for aerobic and anaerobic conditions. The model is consequently more accurate in describing the phenomena, since it uses different substrate affinities and inhibitions for each operating condition.

The accuracy of the model for predicting the experimental data was quite good, reporting the highest MRE in predicting glucose concentration and satisfying the adequate description of the trends of all variables in the experimental data (Table 4, Fig. 5).

In addition, the correlation analysis between the experimental data and the output variables of the model performed in this study is statistically significant, as its p-values are less than $1 \times 10^{-3}\%$, which is well below what is expected to achieve acceptable quality ($p \leq 5\%$). Fig. 5 confirms that the model is able to reproduce the behavior of the main experimental variables very well, thus validating the description of the phenomena considered in the model. Even the lowest MRE value obtained for ethanol, the worst variable trend of the model reporting the lowest R-squared value, is still good (Fig. 5).

In general, the agreement between the three sets of experimental data and the model outputs is very strong, reporting an overall MRE value in the prediction of all the variables studied of less than 7 %. In addition the model is able to predict yeast metabolism under fully anaerobic conditions at different concentrations of glucose, ethanol, and yeast. These results confirm the robustness of the model and the quality of the calibration method.

4.2. Modeling yeast activity under aerobic conditions

Modeling the fermentation process under aerobic conditions is more complex, since it involves different metabolic pathways such as aerobic fermentation and respiration processes, typical of a Crabtree effect positive yeast. Anaerobic fermentation pathway could be also included: even under aerobic conditions, the oxygen uptake can be larger than the rate of dissolved oxygen, partially triggering the anaerobic pathway. This metabolic pathway is verified by the data concerning glycerol production. Dissolved oxygen deficiency also worsens with time due to decreased oxygen solubility in the medium, which is caused by an increased temperature resulting from metabolic heat release and the increase in yeast population. Fortunately, the model accounts for these two effects of the anaerobic pathway and the influence of temperature on dissolved oxygen. Different glucose affinity and ethanol inhibition coefficients under aerobic and anaerobic conditions were established in the model, assuming that yeast tolerance to ethanol and affinity for glucose is different under aerobic and anaerobic conditions.

Three test combinations with different initial concentrations of glucose, ethanol, and yeast were performed to consider their influence on yeast kinetics (experiments D, E, and F). Most of the model parameters were taken from the literature, so only four parameters had to be calibrated, considerably simplifying the model calibration process. This simplified calibration process demonstrates the phenomenological character of the model, its universality, and its versatility for use in production.

The optimization process using the PSO method converged to a single minimum (Table 5), reporting a standard deviation in parameter optimization less than 0.25 %.

Table 5: Stoichiometry and kinetics parameters of the biological model under aerobic conditions

Parameters name	Symbol	Value	Source
Oxygen stoichiometric coefficient in aerobic fermentation process (mol)	b_2	0.37 ± 0.00	Calibrated
Yeast stoichiometric coefficient in aerobic fermentation process (mol)	d_2	1.40 ± 0.25	Experimental
Ethanol stoichiometric coefficient in aerobic fermentation process (mol)	e_2	1.26 ± 0.20	Experimental
Oxygen stoichiometric coefficient in glucose respiration process (mol)	b_3	2.50 ± 0.06	Thierie (2004)
Yeast stoichiometric coefficient in glucose respiration process	d_3	3.33 ± 0.06	Thierie (2004)
Oxygen stoichiometric coefficient in ethanol respiration process (mol)	b_4	1.61 ± 0.10	Scheiblaue et al. (2018)
Yeast stoichiometric coefficient in ethanol respiration process (mol)	d_4	1.17 ± 0.29	Experimental
Half-saturation coefficient for oxygen (mg/l)	K_{O_2}	0.10	Sonnleitner and Käppli (1986)
Half-saturation coefficient for glucose under aerobic conditions (g/l)	$K_{Glc, Ox}$	1×10^{-3}	Thierie (2004)
Half-saturation coefficient for ethanol (g/l)	K_E	0.10	Sonnleitner and Käppli (1986)
Half-saturation coefficient for ethanol inhibition under aerobic conditions (g/l)	$K_{E, Ox}^{inh}$	19.70 ± 0.01	Calibrated
Maximum growth rate for aerobic fermentation process (h^{-1})	μ_{max2}	0.51 ± 0.00	Calibrated
Maximum growth rate for glucose respiration process (h^{-1})	μ_{max3}	0.20	Thierie (2004)
Maximum growth rate for ethanol respiration process (h^{-1})	μ_{max4}	0.11 ± 0.00	Calibrated
Overall oxygen transfer rate (h^{-1})	k_{La}	27.27 ± 0.42	Experimental

Stoichiometric parameters for the respiration process were taken as the mean values reported by Thierie (2004) using three glucose concentrations values (5, 15 and 30 g/l) for the respiration process at the maximum growth rate of $0.20 h^{-1}$.

The model was able to reproduce all experiments with mean relative errors of less than 11 % (Table 6). In this case, a half-saturation coefficient for ethanol inhibition (19.70 g/l) was obtained under aerobic conditions, which was lower than the same value of half-saturation coefficient for ethanol inhibition under anaerobic conditions. The differences between these values supports this work's assumption, i.e., yeast tolerance to ethanol is different under anaerobic and aerobic conditions (10-fold more tolerant under anaerobic conditions). Arroyo-López et al. (2010) studied the inhibitory effect of ethanol using the

Lambert and Pearson (2000) methodology for the estimation of the Minimum Inhibitory Concentration (MIC) and Non-Inhibitory Concentration (NIC) of a compound using Optical Density (OD) measurements. MIC is related to the resistance or tolerance of the microorganism to the compound and is the lowest concentration which results in maintenance or reduction of an inoculum’s viability (marks the concentration above which no growth is observed). In contrast, NIC is related to the susceptibility of the microorganism to the compound, and it is the concentration above which the inhibitor begins to have a progressive and negative effect on growth (Lambert and Pearson, 2000). The authors studied some yeast strains, and in the particular case of *S. cerevisiae* they obtained values of NIC and MIC in the ranges of (36.7-73.9 g/l) and (95.6 - 141.4 g/l), respectively. In our experiments, the ethanol inhibition effect was only observed in the data set in which an initial ethanol concentration over 50 g/l concentration was used, which lies well within inhibitory concentration ranges reported by Arroyo-López et al. (2010).

According to the stoichiometry and kinetics obtained in the present study, the aerobic fermentation process has a lower ethanol production yield, a higher yeast yield, and a higher growth rate than the anaerobic fermentation process. According to Thierie (2004), stoichiometry varies as a function of the specific growth rate of the yeast. In contrast, our work uses constant stoichiometric parameters corresponding to the mean values of the experiments performed for the individual metabolic pathways under aerobic and anaerobic conditions (Tables 3 and 5). Yet in all batch fermentation processes carried out in this study, the apparent stoichiometry varies, due to the varying partition of the metabolic pathways (Fig. 7). Accounting for the varying metabolic pathways partition confers a much better predictive capability on the model. Our model thus has the potential to increase ethanol yield with reduced consumption of time and resources, which would be useful for all ethanol producers.

Table 6: Mean relative errors reported between model output and experimental data under aerobic conditions

No.	Experiment	$S_{Glc}(\%)$	$S_{O_2}(\%)$	$S_E(\%)$	$S_{Gly}(\%)$	$X_Y(\%)$	Mean value(%)
1	D	7.77	9.95	0.83	24.19	4.56	9.46
2	E	11.13	17.19	3.87	10.96	7.36	10.10
3	F	9.23	14.27	4.70	22.70	7.09	11.60
	Mean value(%)	9.38	13.80	3.13	19.28	6.34	10.39

Figure 6 summarizes the results obtained with the final set of parameters.

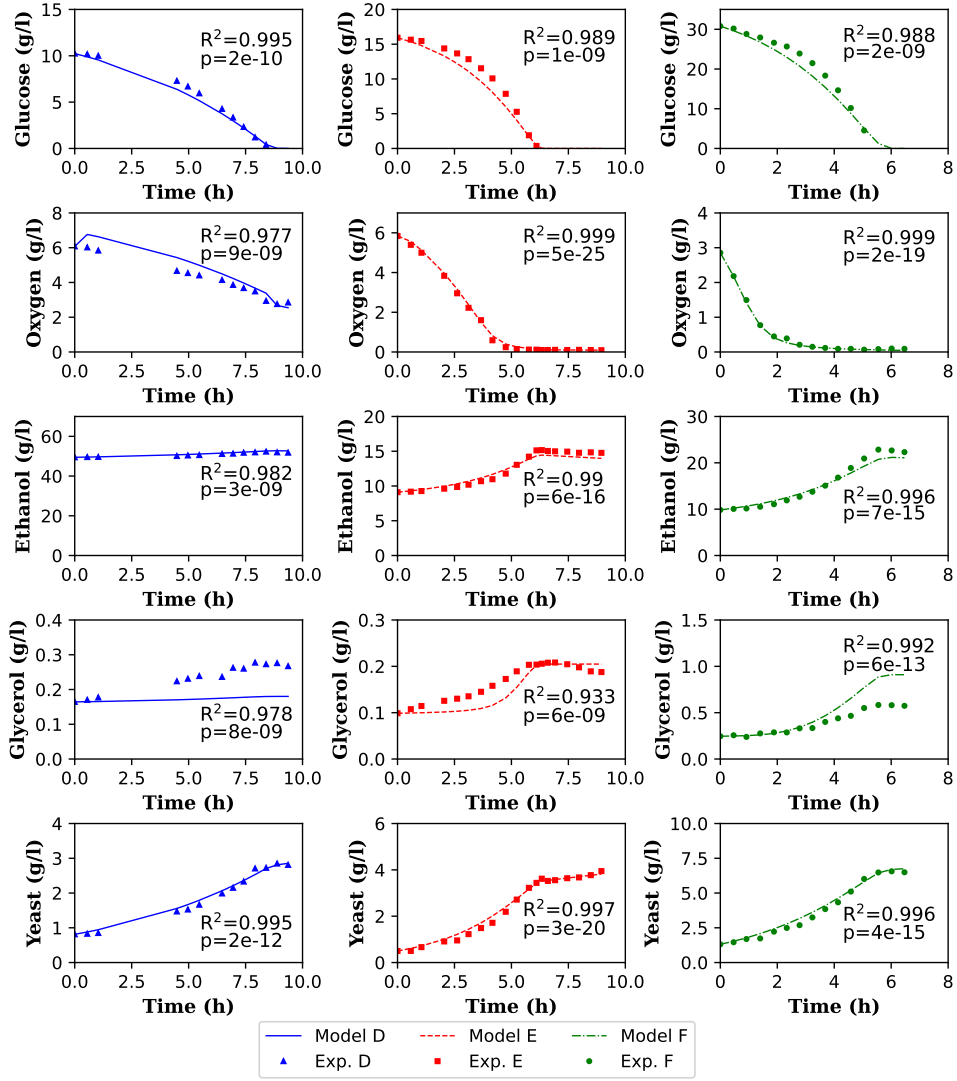


Figure 6: Yeast activity simulation under aerobic conditions. Comparison between model output and experimental data.

In most cases, the model is able to predict the concentrations over time with accuracy. The highest MRE values and the lowest R-squared values were found for glycerol (experiment D, 24.19 % and 93.40 %, respectively)(Table 6, Fig. 6). Once again, the correlation analysis between the experimental data and the output variables of the model performed in this study is statistically significant, as its p-values are less than 1×10^{-6} %. As *S. cerevisiae* has already been shown to grow by using glycerol as a carbon source under aerobic conditions at low specific growth rates (0.01 - 0.20 h^{-1}) (Ochoa-Estopier et al., 2011). Interestingly glycerol consumption was observed in experiments E and F beyond 6 hours. Yeast growth based on glycerol under aerobic conditions was not included in the model, which might significantly affect the model balance. Improving this yeast model by including this metabolic pathway would be a novel approach, and would require the identification of its main metabolites, their stoichiometry, and kinetics.

4.2.1. Individual metabolic pathways contribution under aerobic conditions

As stated in section 4.2, all metabolic pathways are always kept in the model. However, their relative importance is triggered during the culture as a function of growth conditions. The rates of all these pathways can then be obtained from the simulation. The relative rates of each metabolic pathway as integrated over time during the process are depicted in Fig. 7 for glucose and ethanol consumption. In the model, ethanol production and utilization are expressed with positive and negative rate values, respectively, because ethanol is a fermentation product that becomes a carbon source for yeast when glucose is depleted. In all cases, the aerobic alcoholic fermentation process is the predominant metabolic pathway.

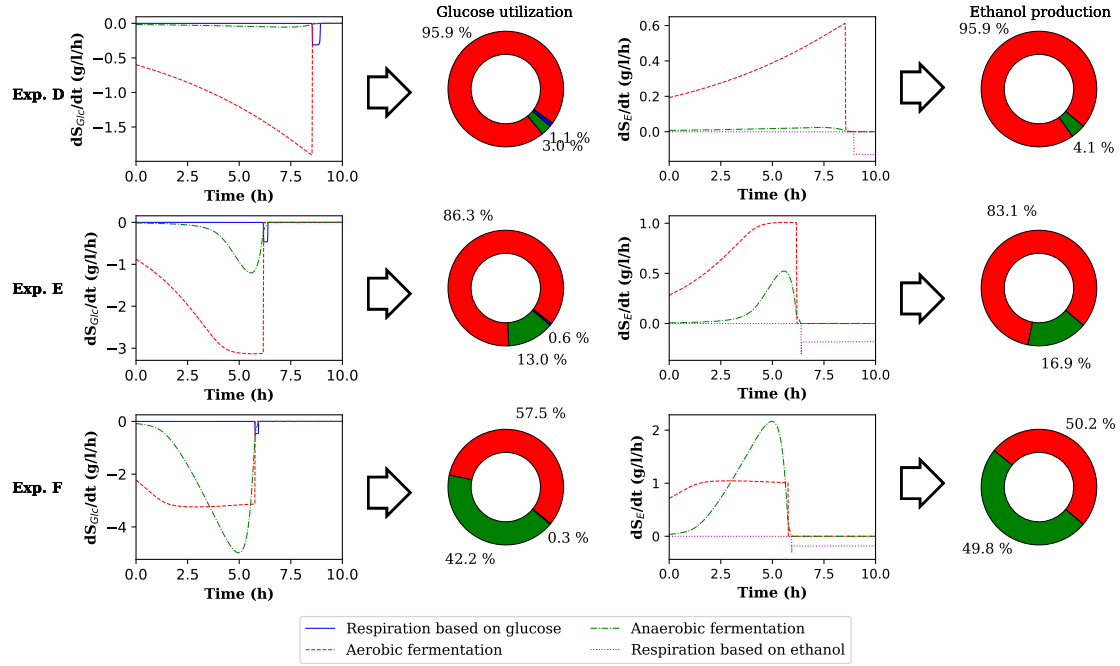


Figure 7: Individual metabolic pathways contribution to glucose and ethanol (Line graphs). Stoichiometry of glucose utilization and ethanol production during glucose degradation process (Pie graphs).

The highest glucose consumption rates were predicted by the model for experiment F, in which the maximum initial concentrations of glucose and yeast were used. Indeed, an increase in the initial yeast concentration with a constant air supply rate favors the anaerobic fermentation pathway because it increases oxygen consumption rates above the oxygen supply rate, generating anaerobic conditions. However, although a higher initial yeast concentration was used in experiment D than in experiment E, the anaerobic fermentation pathway was activated to a lesser extent in experiment D. These differences are explained by the higher initial ethanol concentration, which inhibits the aerobic fermentation pathway and thus reduces the rate of oxygen consumption. The glucose respiration process in our experiments was almost negligible compared to the other metabolic pathways.

4.3. Model validation: yeast culture without gas injection

Our model was validated using a batch fermentation process without gas injection (experiment G) instead of the experiments from the learning database. In this experiment, the gas headspace of the bioreactor was maintained at a constant atmospheric pressure.

Even without air injection, oxygen is still transferred from the bioreactor gas headspace to the liquid volume through the agitation process. This oxygen transfer occurs through the oxygen concentration gradient maintained by the biological oxygen consumption in the liquid phase. Consequently, both anaerobic and aerobic metabolic pathways take place simultaneously. In the simulation, the k_La value was replaced by the oxygen transfer mass from the gas headspace, taking the value already estimated by [La et al. \(2020\)](#) for the same installation and operations conditions ($2 \times 10^{-4} \text{ s}^{-1}$).

Table 7: Model validation: mean relative errors reported between model output and experimental data for fermentation without gas injection

No.	Experiment	$S_{Glc}(\%)$	$S_E(\%)$	$S_{Gly}(\%)$	$X_Y(\%)$	Mean value(%)
1	G	10.69	5.19	8.60	3.65	7.03

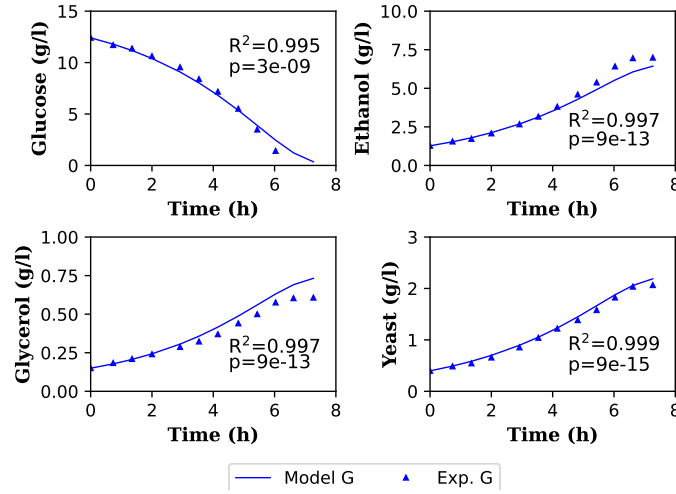


Figure 8: Model validation: yeast activity simulation without gas injection. Comparison between model output and experimental data.

As can be seen in Table 7 and Fig. 8, the model performs very nicely for this validation test. It allows all variables' evolution to be predicted with a MRE less than 11 %, reporting a global MRE for all variables of 7 %. The trends of all variables were well described, showing a very good correlation coefficient value above 99 % (Fig. 8), results that are statistically significant, as their p-values are less than $1 \times 10^{-6}\%$. The model also provides an excellent prediction of glycerol production. These results validate the quality, applicability, and accuracy of the model, even under different operating conditions.

Fig. 9 depicts the individual metabolic pathway contributions during glucose consumption and ethanol production without gas injection.

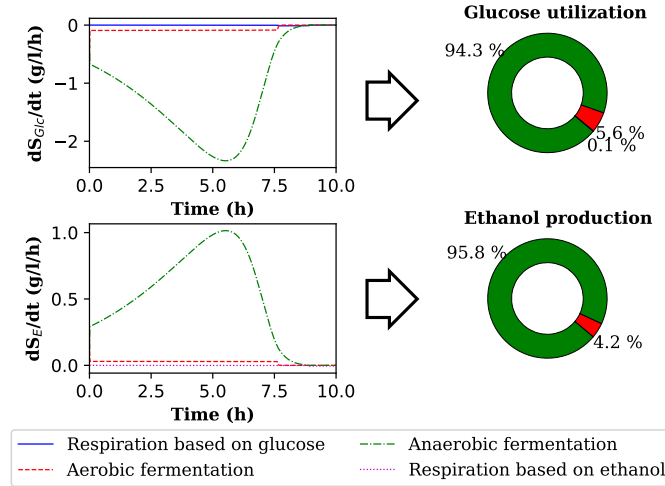


Figure 9: Model validation: individual metabolic pathways contribution to glucose and ethanol in the absence of gas injection (Line graph). Stoichiometry of glucose utilization and ethanol production during glucose degradation process (Pie graph) (Experiment G).

Even without gas injection, the aerobic metabolic pathway was partly activated thanks to the oxygen transfer from the gas headspace. This limited mass transfer explains the low activation of the aerobic metabolic pathway compared to the dominant metabolism of the anaerobic pathway. Most of the glucose consumption and ethanol production were associated with the anaerobic metabolic pathway (94.3 % and 95.8 %, respectively). No ethanol consumption was observed during the experiment, as Fig. 9 confirms.

Indeed, this particular case of yeast culture without gas injection more closely resembles real world scenarios, where alcoholic fermentation processes are partially carried out under anaerobic conditions. Certain designs of bioreactors can allow, the oxygen in the gas headspace to be continuously renewed if atmospheric air is allowed to enter, for example, through filters. The pie graphs in Fig. 9 prove that the model is able to activate/deactivate the corresponding metabolic pathways according to the medium conditions and, more importantly, under different operating conditions than those used for model calibration. The results obtained allow a better understanding of the phenomena occurring in partial anaerobic systems in order to improve process control and optimize operating conditions according to the conditions of the medium.

4.4. Potential application of this model

Producers and scientists currently consider modeling to be a promising tool for enhancing bio-production. To this end, databases, mechanistic models, and machine learning need to work in synergy for online process monitoring. This approach is known as hybrid modeling and offers a promising route in the general quest of the digital twin in bio-production (Muldbak et al., 2022; Sokolov et al., 2021; Babi et al., 2022). For this approach to be efficient, the mechanistic model needs to be as predictive as possible. The mechanistic model could be improved either inline by real-time tuning some key parameters or online, using a dynamic learning database.

Even though the mechanistic model was applied to a specific commercial strain of *Saccharomyces cerevisiae*, the general modeling approach proposed in this study is a perfect brick

for hybrid modeling for any application of bio-production. Our model is constructed for this commercial strain by including the main metabolic pathways of mutant and wild-type yeast strains used for ethanol or yeast production reported in the literature. The model was successfully calibrated and validated for the commercial yeast strain provided by the Institut Oenologique de Champagne under the name IOC Fizz+, thus demonstrating its applicability and universality in ethanol or yeast production systems. It is important to note that, for predicting the metabolism of other yeast strains, model calibration alone may be sufficient if the mutated yeast strain exhibits the same metabolic pathways described in the model. The structural basis of the model can allow other metabolic pathways to be easily included or removed when describing the metabolism of the microorganisms that do not exhibit the same metabolic pathways more accurately.

The predictive model proposed in this paper is valuable not only for alcoholic fermentation but also for other processes, such as the production of chemicals, fuels, foods, and pharmaceuticals, as yeast is one of the most widely used hosts for synthetic biology (Vieira Gomes et al., 2018). One of the disadvantages of the Crabtree effect is the carbon loss due to the ethanol production under aerobic conditions, which leads to a lower biomass formation and consequently a lower production of recombinant proteins (Mattanovich et al., 2012). Therefore, the structural base of the developed model could be adapted or serve as a basis for the modeling of other Crabtree-positive yeasts used for the production of therapeutic proteins. The model could then be used as a tool for achieving a better understanding, control, and optimization of the production process.

Beyond the huge domain of engineered yeast strains, the formulation proposed in this work can also be applied to other strains. In particular, the proposed functions that account for activation or inhibition of pathways and sudden shifts of pathways are universal. For example, our team is currently using this framework to model the Chinese Hamster Ovary (CHO) cell metabolism to produce antibodies. CHO cells are analogous with the Crabtree effect in that they exhibit a phenomenon known as the Warburg effect, where glucose is fermented to produce lactate even in the presence of oxygen (Potter et al., 2016). Even in the presence of oxygen, this first stage of lactic fermentation corresponds to a peak of exponential cell growth, followed by a metabolic shift from net production to net consumption of lactate (known as stationary phase) during which proteins are produced. The similarity between the systems mentioned above suggests that our mechanistic models has the potential to predict the metabolic shift observed for CHO cells, which has the possibility of considerably improving the current state of CHO cells modeling.

Finally, bringing together mechanistic modeling and machine learning can better explain system phenomena that are traditionally difficult to describe. For instance, well-established theoretical knowledge can be formulated as explicit equations, while parameters which cannot be derived from first principles or space-time-varying (latent) states are estimated via a machine learning approach (Reichstein et al., 2022). The development of online sensors using the Raman spectroscopy mechanistic and machine learning models and their hybridization variants have considerably increased their application in bioprocess retro-control, allowing maximum productivity with lower resource consumption.

5. Conclusions

In this study, a robust and predictive yeast model was developed and successfully validated with experimental data from experiments with a commercial yeast strain used for

wine production. The model includes a comprehensive set of metabolic pathways that are always present in the model but are more or less activated depending on the growth conditions. A general framework is proposed for the formulation, including functions that account for activation, inhibition, and shift of metabolic pathways. The model parameters were determined by a blend of literature data and experimental data. The inverse analysis method was used to determine parameters that could not be directly obtained from literature and experiments. Once the model was calibrated, its quality and robustness were again confirmed with an additional experiment performed without gas injection, and the model describes the main process variables with an overall mean relative error of 7%. The complete formulation and set of parameters are provided in the document so that the reader can implement them for their own needs.

The model is able to switch between aerobic fermentation and glucose-based respiration when the glucose reaches values below 0.10-0.15 g/l. It is also able to activate the anaerobic fermentation metabolic pathway even under aerobic conditions when the rate of oxygen uptake is higher than the rate of dissolved oxygen, as supplied by aeration. Once glucose is depleted under aerobic conditions, the model automatically switches to ethanol degradation. The results provided in this work give new insights towards the behavior of *S. cerevisiae*. For example the model better informs the emerging nature of the global stoichiometry and differences in ethanol tolerances, which depend on the evolution of yeast growth conditions and the active metabolic pathway.

Beyond the application to the yeast strain studied here, this work gives a general framework of mechanistic modeling able to predict the coexistence of several metabolic pathways and their shift along the growth conditions. This framework can be used as a building block of a digital twin of any bio-production.

6. Acknowledgements

This study was carried out in the Centre Européen de Biotechnologie et de Bioéconomie (CEBB), and supported by Région Grand Est, Département de la Marne, Grand Reims and the European Union. In particular, the authors would like to thank Département de la Marne, Grand Reims, Région Grand Est and European Union with European Regional Development Fund (ERDF Champagne Ardenne 2014-2020) for their financial support to the Chair of Biotechnology of CentraleSupélec.

CRedit author statement

YGH: Conceptualization, Methodology, Formal analysis, Modeling, Visualization, Writing-Original draft preparation, Writing-Reviewing and Editing. EM: Investigation, Methodology. PP: Conceptualization, Methodology, Formal analysis, Modeling, Writing-Reviewing and Editing, Supervision, Funding acquisition.

References

Amenaghawon, N., Okieimen, C., Ogbeide, S., 2012. Kinetic modelling of ethanol inhibition during alcohol fermentation of corn stover using *saccharomyces cerevisiae*. Int. J. Eng. Res. Appl. 2, 798–803.

- Arroyo-López, F.N., Salvadó, Z., Tronchoni, J., Guillamón, J.M., Barrio, E., Querol, A., 2010. Susceptibility and resistance to ethanol in *saccharomyces* strains isolated from wild and fermentative environments. *Yeast* 27, 1005–1015. doi:<https://doi.org/10.1002/yea.1809>.
- Babi, D.K., Griesbach, J., Hunt, S., Insaído, F., Roush, D., Todd, R., Staby, A., Welsh, J., Wittkopp, F., 2022. Opportunities and challenges for model utilization in the biopharmaceutical industry: current versus future state. *Current Opinion in Chemical Engineering* 36, 100813.
- Bakker, B.M., Overkamp, K.M., van Maris, A.J., Kötter, P., Luttik, M.A., van Dijken, J.P., Pronk, J.T., 2001. Stoichiometry and compartmentation of nadh metabolism in *saccharomyces cerevisiae*. *FEMS microbiology reviews* 25, 15–37. doi:<https://doi.org/10.1111/j.1574-6976.2001.tb00570.x>.
- Carteni, F., Occhicone, A., Giannino, F., Vincenot, C.E., de Alteriis, E., Palomba, E., Mazzoleni, S., 2020. A general process-based model for describing the metabolic shift in microbial cell cultures. *Front. Microbiol.* 11, 2332. doi:<https://doi.org/10.3389/fmicb.2020.521368>.
- Ding, J., Huang, X., Zhang, L., Zhao, N., Yang, D., Zhang, K., 2009. Tolerance and stress response to ethanol in the yeast *saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* 85, 253–263. doi:<https://doi.org/10.1007/s00253-009-2223-1>.
- García-Ochoa, F., Gómez, E., 2009. Bioreactor scale-up and oxygen transfer rate in microbial processes: an overview. *Biotechnol. Adv.* 27, 153–176. doi:<https://doi.org/10.1016/j.biotechadv.2008.10.006>.
- Ghose, T., Tyagi, R., 1979. Rapid ethanol fermentation of cellulose hydrolysate. ii. product and substrate inhibition and optimization of fermentor design. *Biotechnol. Bioeng.* 21, 1401–1420. doi:<https://doi.org/10.1002/bit.260210808>.
- Green, D.W., Perry, R.H., 2008. Perry's chemical engineers' handbook. eighth ed., The McGraw-Hill Companies, Inc, New York. doi:<https://doi.org/10.1036/0071511245>.
- Hagman, A., Säll, T., Piškur, J., 2014. Analysis of the yeast short-term crabtree effect and its origin. *FEBS J.* 281, 4805–4814. doi:<https://doi.org/10.1111/febs.13019>.
- Hanly, T.J., Henson, M.A., 2013. Dynamic metabolic modeling of a microaerobic yeast co-culture: predicting and optimizing ethanol production from glucose/xylose mixtures. *Biotechnol. Biofuels* 6, 1–16. doi:<https://doi.org/10.1186/1754-6834-6-44>.
- Hindmarsh, A.C., 1982. *RS Stepleman et al. (eds.), North-Holland, Amsterdam, 1983 (vol. 1 of IMACS transactions on scientific computation), pp. 55-64.*
- Johnston, M., Kim, J.H., 2005. Glucose as a hormone: receptor-mediated glucose sensing in the yeast *saccharomyces cerevisiae*. doi:<https://doi.org/10.1042/bst0330247>.
- La, A., Du, H., Taidi, B., Perré, P., 2020. A predictive dynamic yeast model based on component, energy, and electron carrier balances. *Biotechnol. Bioeng.* 117, 2728–2740. doi:<https://doi.org/10.1002/bit.27442>.
- La, A., Perré, P., Taidi, B., 2019. Process for symbiotic culture of *Saccharomyces cerevisiae* and *Chlorella vulgaris* for *in situ* CO₂ mitigation. *Appl. Microbiol. Biotechnol.* 103, 731–745. doi:<https://doi.org/10.1007/s00253-018-9506-3>.
- Lambert, R., Pearson, J., 2000. Susceptibility testing: accurate and reproducible minimum inhibitory concentration (MIC) and non-inhibitory concentration (NIC) values. *J. Appl. Microbiol.* 88, 784–790. doi:<https://doi.org/10.1046/j.1365-2672.2000.01017.x>.
- Lencastre Fernandes, R., Bodla, V.K., Carlquist, M., Heins, A.L., Eliasson Lantz, A., Sin, G., Gernaey, K.V., 2012. Applying mechanistic models in bioprocess development. *Adv.*

- Biochem. Engin./Biotechnol. , 137–166doi:https://doi.org/10.1007/10_2012_166.
- Luong, J., 1985. Kinetics of ethanol inhibition in alcohol fermentation. Biotechnol. Bioeng. 27, 280–285. doi:<https://doi.org/10.1002/bit.260270311>.
- Ma, M., Liu, Z.L., 2010. Mechanisms of ethanol tolerance in *saccharomyces cerevisiae*. Appl. Microbiol. Biotechnol. 87, 829–845. doi:<https://doi.org/10.1007/s00253-010-2594-3>.
- Mattanovich, D., Branduardi, P., Dato, L., Gasser, B., Sauer, M., Porro, D., 2012. Recombinant protein production in yeasts., in: Lorence, A. (Ed.), Recombinant Gene Expression. Methods in Molecular Biology (Methods and Protocols). Humana Press, Totowa, NJ, pp. 329–358. doi:https://doi.org/10.1007/978-1-61779-433-9_17.
- Medina, V.G., Almering, M.J., van Maris, A.J., Pronk, J.T., 2010. Elimination of glycerol production in anaerobic cultures of a *saccharomyces cerevisiae* strain engineered to use acetic acid as an electron acceptor. Appl. Environ. Microbiol. 76, 190–195. doi:<https://doi.org/10.1128/aem.01772-09>.
- Muldbak, M., Gargalo, C., Krühne, U., Udugama, I., Gernaey, K.V., 2022. Digital twin of a pilot-scale bio-production setup, in: Computer Aided Chemical Engineering. Elsevier. volume 49, pp. 1417–1422.
- Nissen, T.L., Hamann, C.W., Kielland-Brandt, M.C., Nielsen, J., Villadsen, J., 2000. Anaerobic and aerobic batch cultivations of *saccharomyces cerevisiae* mutants impaired in glycerol synthesis. Yeast 16, 463–474. doi:[https://doi.org/10.1002/\(SICI\)1097-0061\(20000330\)16:5<3C463::AID-YEA535>3E3.0.CO;2-3](https://doi.org/10.1002/(SICI)1097-0061(20000330)16:5<3C463::AID-YEA535>3E3.0.CO;2-3).
- Ochoa-Estopier, A., Lesage, J., Gorret, N., Guillouet, S., 2011. Kinetic analysis of a *saccharomyces cerevisiae* strain adapted for improved growth on glycerol: Implications for the development of yeast bioprocesses on glycerol. Bioresour. Technol. 102, 1521–1527. doi:<https://doi.org/10.1016/j.biortech.2010.08.003>.
- Overkamp, K.M., Bakker, B.M., Kötter, P., van Tuijl, A., de Vries, S., van Dijken, J.P., Pronk, J.T., 2000. In vivo analysis of the mechanisms for oxidation of cytosolic nadh by *saccharomyces cerevisiae* mitochondria. J. Bacteriol. 182, 2823–2830. doi:<https://doi.org/10.1128/jb.182.10.2823-2830.2000>.
- Parapouli, M., Vasileiadis, A., Afendra, A.S., Hatziloukas, E., 2020. *Saccharomyces cerevisiae* and its industrial applications. AIMS Microbiol. 6, 1. doi:<https://doi.org/10.3934/microbiol.2020001>.
- Petzold, L., 1983. Automatic selection of methods for solving stiff and nonstiff systems of ordinary differential equations. SIAM J. Sci. Comput. 4, 136–148. doi:<https://doi.org/10.1137/0904010>.
- Pham, H.T.B., Larsson, G., Enfors, S.O., 1998. Growth and energy metabolism in aerobic fed-batch cultures of *saccharomyces cerevisiae*: Simulation and model verification. Biotechnol. Bioeng. 60, 474–482. doi:[https://doi.org/10.1002/\(sici\)1097-0290\(19981120\)60:4<3C474::aid-bit9>3E3.0.co;2-j](https://doi.org/10.1002/(sici)1097-0290(19981120)60:4<3C474::aid-bit9>3E3.0.co;2-j).
- Potter, M., Newport, E., Morten, K.J., 2016. The warburg effect: 80 years on. Biochem. Soc. Trans. 44, 1499–1505. doi:<https://doi.org/10.1042/BST20160094>.
- Reichstein, M., Ahrens, B., Kraft, B., Camps-Valls, G., Carvalhais, N., Gans, F., Gentine, P., Winkler, A.J., 2022. Combining system modeling and machine learning into hybrid ecosystem modeling, in: Knowledge-Guided Machine Learning. Chapman and Hall/CRC, pp. 327–352.
- Rieger, M., Käppeli, O., Fiechter, A., 1983. The role of limited respiration in the incomplete oxidation of glucose by *saccharomyces cerevisiae*. Microbiology 129, 653–661. doi:<https://doi.org/10.1093/micro/129.4.653>.

- [//doi.org/10.1099/00221287-129-3-653](https://doi.org/10.1099/00221287-129-3-653).
- Rosenfeld, E., Beauvoit, B., Blondin, B., Salmon, J.M., 2003. Oxygen consumption by anaerobic *saccharomyces cerevisiae* under enological conditions: effect on fermentation kinetics. *Appl. Environ. Microbiol.* 69, 113–121. doi:<https://doi.org/10.1128/aem.69.1.113-121.2003>.
- Sablayrolles, J.M., Barre, P., 1986. Evaluation des besoins en oxygène de fermentations alcooliques en conditions oenologiques simulées. *Sciences des aliments* 6, 373–383.
- Scheiblaue, J., Scheiner, S., Joks, M., Kavsek, B., 2018. Fermentation of *saccharomyces cerevisiae*—combining kinetic modeling and optimization techniques points out avenues to effective process design. *J. Theor. Biol.* 453, 125–135. doi:<https://doi.org/10.1016/j.jtbi.2018.05.016>.
- Sokolov, M., von Stosch, M., Narayanan, H., Feidl, F., Butté, A., 2021. Hybrid modeling—a key enabler towards realizing digital twins in biopharma? *Current Opinion in Chemical Engineering* 34, 100715.
- Sonnleitner, B., Käppli, O., 1986. Growth of *saccharomyces cerevisiae* is controlled by its limited respiratory capacity: formulation and verification of a hypothesis. *Biotechnol. Bioeng.* 28, 927–937. doi:<https://doi.org/10.1002/bit.260280620>.
- Sostarić, N., Arslan, A., Carvalho, B., Plech, M., Voordeckers, K., Verstrepen, K.J., van Noort, V., 2021. Integrated multi-omics analysis of mechanisms underlying yeast ethanol tolerance. *J. Proteome Res.* 20, 3840–3852. doi:<https://doi.org/10.1021/acs.jproteome.1c00139>.
- Suliman, A.K., Putra, M.D., Abasaeed, A.E., Gaily, M.H., Al-Zahrani, S.M., Zeinelabdeen, M.A., 2018. Kinetic modeling of the simultaneous production of ethanol and fructose by *saccharomyces cerevisiae*. *Electron. J. Biotechnol.* 34, 1–8. doi:<https://doi.org/10.1016/j.ejbt.2018.04.006>.
- Thierie, J., 2004. Modeling threshold phenomena, metabolic pathways switches and signals in chemostat-cultivated cells: the crabtree effect in *saccharomyces cerevisiae*. *J. Theor. Biol.* 226, 483–501. doi:<https://doi.org/10.1016/j.jtbi.2003.10.017>.
- Thierie, J., 2019. Two way crabtree-effect model enhancement by maintenance considerations addition. *J. Theor. Comput. Sci.* 5, 2. doi:<https://doi.org/10.4172/2376-130X.1000164>.
- Veloso, I.I., Rodrigues, K.C., Sonogo, J.L., Cruz, A.J., Badino, A.C., 2019. Fed-batch ethanol fermentation at low temperature as a way to obtain highly concentrated alcoholic wines: Modeling and optimization. *Biochem. Eng. J.* 141, 60–70. doi:<https://doi.org/10.1016/j.bej.2018.10.005>.
- Venter, G., Sobieszczanski-Sobieski, J., 2003. Particle swarm optimization. *AIAA J.* 41, 1583–1589. doi:<https://doi.org/10.2514/2.2111>.
- Verduyn, C., Postma, E., Scheffers, W.A., van Dijken, J.P., 1990a. Physiology of *saccharomyces cerevisiae* in anaerobic glucose-limited chemostat cultures. *Microbiology* 136, 395–403. doi:<https://doi.org/10.1099/00221287-136-3-395>.
- Verduyn, C., Postma, E., Scheffers, W.A., Van Dijken, J.P., 1990b. Energetics of *saccharomyces cerevisiae* in anaerobic glucose-limited chemostat cultures. *Microbiology* 136, 405–412. doi:<https://doi.org/10.1099/00221287-136-3-405>.
- Verduyn, C., Zomerdijk, T.P., van Dijken, J.P., Scheffers, W.A., 1984. Continuous measurement of ethanol production by aerobic yeast suspensions with an enzyme electrode. *Appl. Microbiol. Biotechnol.* 19, 181–185. doi:<https://doi.org/10.1007/bf00256451>.
- Vieira Gomes, A.M., Souza Carmo, T., Silva Carvalho, L., Mendonça Bahia, F., Parachin,

N.S., 2018. Comparison of yeasts as hosts for recombinant protein production. *Microorganisms* 6, 38. doi:<https://doi.org/10.3390/microorganisms6020038>.

Appendix A. Biological rate expressions

Anaerobic fermentation based on glucose

$$R_1 = \mu_{max1} \frac{S_{Glc}}{K_{Glc,Ax} + S_{Glc}} \frac{K_{O_2}}{K_{O_2} + S_{O_2}} \frac{K_{E,Ax}^{inh}}{K_{E,Ax}^{inh} + S_E} X_Y \quad (A.1)$$

Aerobic fermentation based on glucose

$$R_2 = \mu_{max2} \lambda_r \frac{S_{Glc}}{K_{Glc,Ox} + S_{Glc}} \frac{S_{O_2}}{K_{O_2} + S_{O_2}} \frac{K_{E,Ox}^{inh}}{K_{E,Ox}^{inh} + S_E} X_Y \quad (A.2)$$

Respiration based on glucose

$$R_3 = \mu_{max3} (1 - \lambda_r) \frac{S_{Glc}}{K_{Glc,Ox} + S_{Glc}} \frac{S_{O_2}}{K_{O_2} + S_{O_2}} \frac{K_{E,Ox}^{inh}}{K_{E,Ox}^{inh} + S_E} X_Y \quad (A.3)$$

Respiration based on ethanol

$$R_4 = \mu_{max4} \frac{\delta K_{Glc,Ox}}{\delta K_{Glc,Ox} + S_{Glc}} \frac{S_E}{K_E + S_E} \frac{S_{O_2}}{K_{O_2} + S_{O_2}} \frac{K_{E,Ox}^{inh}}{K_{E,Ox}^{inh} + S_E} X_Y, \quad (A.4)$$

where δ is a safety coefficient to guarantee low values of the mean saturation coefficient in the switching function for the activation of ethanol respiration process without conditioning the value of the glucose saturation coefficient, with a value of 1×10^{-3} .

Aeration

$$R_5 = k_L a (M_{O_2} H_{(T)} P_{O_2} - S_{O_2}) \quad (A.5)$$

The constant Henry's value as function of temperature is estimated using Eq. A.6:

$$H_{(T)} = X_{(T)} \rho_{w(T)}, \quad (A.6)$$

where X is the mole fraction of the oxygen dissolved in water when the oxygen partial pressure is 1 atm and where X was estimated using a correlation (Eq. A.7) for the temperature range of 273 to 333 K taken from [Green and Perry \(2008\)](#):

$$\ln X_{(T)} = -171.2542 + 8391.24/T + 23.24323 \ln T \quad (A.7)$$

The water density $\rho_{w(T)}$ in (mol/l) was estimated using a correlation (Eq. A.8) for water over the entire temperature range of 273.16 to 647.096 K taken from [Green and Perry \(2008\)](#):

$$\rho_{w(T)} = 17.863 + 58.606\tau^{0.35} - 95.396\tau^{2/3} + 213.89\tau - 141.26\tau^{4/3} \quad (\text{A.8})$$

$$\tau = 1 - T/647.096 \quad (\text{A.9})$$