



**HAL**  
open science

## Reconciling competing models: a case study of wine fermentation kinetics

Rodrigo Assar, Felipe A. Vargas, David James Sherman

► **To cite this version:**

Rodrigo Assar, Felipe A. Vargas, David James Sherman. Reconciling competing models: a case study of wine fermentation kinetics. Algebraic and Numeric Biology 2010, Research Institute for Symbolic Computation, Johannes Kepler University of Linz, Jul 2010, Hagenberg, Austria. pp.68–83, 10.1007/978-3-642-28067-2\_6 . inria-00541215

**HAL Id: inria-00541215**

**<https://inria.hal.science/inria-00541215>**

Submitted on 30 Nov 2010

**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.

# Reconciling competing models: a case study of wine fermentation kinetics

Rodrigo Assar<sup>1</sup>, Felipe A. Vargas<sup>2</sup>, David J. Sherman<sup>1</sup>

<sup>1</sup>INRIA Team MAGNOME and CNRS UMR 5800 LaBRI. Université Bordeaux I.  
33405 Talence Cedex, France.

<sup>2</sup>Department of Chemical and Bioprocess Engineering. Pontificia Universidad Católica de Chile.  
Casilla 306, Correo 22, Santiago, Chile.

rodrigo.assar@labri.fr, ftvargas@ing.puc.cl, david.sherman@labri.fr

## Abstract

Mathematical models of wine fermentation kinetics promise early diagnosis of stuck or sluggish winemaking processes as well as better matching of industrial yeast strains to specific vineyards. The economic impact of these challenges is significant: worldwide losses from stuck or sluggish fermentations are estimated at 7 billion € annually, and yeast starter production is a highly competitive market estimated at 40 million € annually. Additionally, mathematical models are an important tool for studying the biology of wine yeast fermentation through functional genomics, and contribute to our understanding of the link between genotype and phenotype for these important cell factories.

Here, we have developed an accurate combined model that best matches experimental observations over a wide range of initial conditions. This model is based on mathematical analysis of three competing ODE models for wine fermentation kinetics and statistical comparison of their predictions with a large set of experimental data. By classifying initial conditions into qualitative intervals and by systematically evaluating the competing models, we provide insight into the strengths and weaknesses of the existing models, and identify the key elements of their symbolic representation that most influence the accuracy of their predictions. In particular, we can make a distinction between main effects that are linear in the modeled variable, and secondary quadratic effects that model interactions between cellular processes.

We generalize our methodology to the common case where one wishes to combine existing, competing models and refine them to better agree with experimental data. The first step is symbolic, and rewrites each model into a polynomial form in which main and secondary effects can be conveniently expressed. The second step is statistical, classifying the match of each model's predictions with experimental data, and identifying the key terms in its equations. Finally, we use a combination of those terms and their coefficients to instantiate the combined model expressed in polynomial form. We show that this procedure is feasible for the case of wine fermentation kinetics, allowing predictions which closely match experimental observations in normal and problematic fermentation.

## 1 Introduction

Obtaining ways to combine models has become a necessity. When a development area has promoted interest of many investigation teams, it is common that competing models are developed, each one constructed to meet particular needs and that fit specific data. The same problem is approached of different ways, doing different estimations or working in different complexity levels. Reconciling models into a combined one it is possible to answer particular needs and obtaining general models. Different models may use different mathematical approaches and description formalisms, so we must conserve a degree of independence between the models.

In this particular study, we are interesting in the wine industry that annually produces 5 millions of tonnes only in France and whose worldwide losses from stuck or sluggish fermentations are estimated at 7 billion € annually. Yeast starter production is highly competitive (estimated at 40 million € annually). While fermentation has been used from antiquity for the production of wine and other alcoholic drinks, advances in the understanding of this process have allowed the modernization of this industry throughout

time. As mechanisms that participate in fermentation are understood better it is possible to control the process. Fermentation is carried out by the action of yeasts, that convert the two grape sugars, glucose and fructose, to ethanol and by-products like aromas, flavors, carbon dioxide gas, and heat. Fermentation is the anaerobic alternative to respiration for generating energy molecules (ATP), but it is remarkably less efficient. While respiration generates ATP with yield of 30 per glucose molecule, fermentation produces only 2 ATP molecules. In spite of the efficiency of respiration, some types of yeasts, in particular *Saccharomyces cerevisiae*, prefer to ferment in the presence of oxygen given a high enough concentration of sugar. Because it is more rapid to get ATP from glycolysis and fermentation than from respiration, *S.cerevisiae* has developed biological mechanisms that, for high sugar concentrations, repress the synthesis of mRNA from genes that are involved in respiration and oxidizing metabolism, privileging fermentation ([5]). There exist several factors that affect the process of fermentation: the composition of the flora that is used to ferment, and those related with yeast nutrition and maintenance of viability of cells are fundamental. Although yeasts carry out the fermentation, some types of bacteria like *lactic*, *acetic bacteria* and *streptomyces* can survive fermentation ([10], [9]), affecting the results. Different strains and species of yeasts also influence the development of the process. Among the nutrients that are important are carbon sources (glucose, fructose) and nitrogen sources (amino acids, ammonia, nucleotide base peptides); temperature, presence of oxygen, low ethanol levels, and controlled pH are also important for the viability of cells. The pH index affect flora composition: low levels ( $< 3.5$ ) inhibit many bacteria but yeasts are resistant to wide pH ranges. On the other hand, yeasts are tolerant to ethanol only in limited ranges, up to 17%v/v approximately. Low temperatures favor non-yeast flora while high temperatures work against the diversity. Carbon and nitrogen sources are essential for the growth of yeast cells but their excess can generate problems too. One considers two types of fermentation problems in industrial applications: stuck and sluggish fermentation ([3]). Stuck fermentation happens when not all the sugar is consumed by the yeasts, and sluggish fermentations are those where the process is very slow. The risk of stuck fermentation is increased when the levels of temperature and sugar concentrations are high. Lower temperatures promote slower fermentation rates and consequently the risk of sluggish fermentation. High levels of nitrogen diminish the risk of sluggish fermentation but affects the flavor and quality of the wine. Strains of yeast that are highly tolerant to ethanol also have reduced risk of sluggish fermentation.

Different approaches have been used to explain the influence of fermentation factors on the production of ethanol. Even though the qualitative effect of the temperature or the inhibiting effect of the ethanol is known, it is not easy to quantify these relations in a mathematical model. A difficulty is the complexity of the biological mechanisms and the great amount of factors, making it necessary to estimate relations and reducing the number of variables. The variables that have been considered most important in previous studies are temperature, sugar, carbon dioxide, nitrogen, biomass, glycerol and ethanol. The influence of the type of yeast and mixtures have been studied little ([31], [23]). There exist two groups of models for the wine fermentation process; the first type focuses on the predictability and the second type on interpretability. By means of data mining techniques like decision trees, machine learning, support vector machines or neural networks ([29], [30], [33] and [32]), the first type of models exploit advances in computing technologies and large databases to predict fermentation profiles. These types of models have the advantage of including a large quantity of factors of the process, but they lack biological, physical or mechanical foundations and generally they are complex and difficult to interpret. The more predictable a the model, the more it is complex. The second type of models correspond to mechanistic kinetic models that are based on physical and biological principles. One of the first mechanistic models of oenological fermentation was developed by Boulton in 1980 ([4]), with the possibility of obtaining accurate measures of fermentation factors concentrations once the parameters of the models have been adjusted. This last approach allows the inclusion of black box models to estimate some behaviors, which also facilitates the generalization of the models to extend them to different environmental conditions,

strains or cultures.

Our goal is to build a general method to build a combined model that reconciles existing models. We select models validated by their authors and look for the conditions in that they fit better reality, without modifying the internal coefficients of the models. For doing this, first we homogenize the notation and evaluate the models in different conditions and then we select the model that best represents reality in function of these conditions. Finally we build a combined model whose terms and coefficients are obtained of the original models. We applied our method to three interpretable models of the fermentation process: Coleman ([6]), Scaglia ([26]) and Pizarro ([22], [24]).

## Methods

We divided our work in three steps: the symbolic, the statistical and the constructive one. First we rewrite each model to homogenize the notations and to separate the different effects that are included in the models. In particular for ODE models, we rewrite the models into polynomial form (or other appropriate base) to separate main and interaction effects. In the statistical step we classify each model according to how they agree with experimental data. For doing this, we identify the *factors* and select the *independent variables* that represent the process. To classify the results in function of configurations of factors, based in the availability of experimental data and the considerations of other studies, we construct a discreet set of intervals or levels of their domains. We compare the model's predictions of independent variables with experimental data for each configuration of factors. To decide the quality of the fitting we used two statistical criteria: confidence intervals and shape analysis. The first one is local and the second one is global. Given a configuration, one says that a simulation is *locally right* if it belongs to the confidence interval of experimental results for each time point (with experimental measures). This criterion allows us to decide if the model agrees with the experimental results at that time. For each time we can observe that there exist significant differences between simulations and experimental results, but the global behavior (the shape) can be the same. To decide if the shape of the simulated profile is similar to that of experimental results, we used linear regression and techniques of linearization to fit the experimental and simulated curves and compare them. We consider that a model is better than other one when the adjustment with experimental data is statistically better. Finally, we define a criterion to select the best model in function of configuration of factors and the independent variable to study. We build a combined model that optimizes the results to obtain the estimated profiles of the variables. This model uses coefficients of the original models and considers main and interaction effects.

In this case we coded the methods in *Matlab* (version 7.0; The MathWorks, Inc., Natick, MA) for doing independent simulations. The models that we considered have associated resolution methods to compute the coefficients. To simplify we include in the term *model* the resolution method.

## Description of analyzed models

We analyzed three mechanistic approaches of fermentation process. The models have the common characteristic of being composed of a set of first degree differential equations, whose coefficients relate one variable with the others. Table 1 lists the number of equation associated to each variable and model in section of Results.

The Coleman model ([6]) was built on the model previously presented by [7] to include the effects of temperature. Its main goal is introducing temperature dependency to predict difficult fermentations. It consists of a 5 coupled ODEs (equations 1- 5) that are combined with 4 one-dimensional regression models to estimate parameters. The variables that are represented in differential equations are concentrations of: biomass ( $X$ ), active biomass ( $X_A$ ), nitrogen ( $N$ ), ethanol ( $EtOH$ ) and sugar ( $S$ ). In basic terms,

	Model			
	Coleman ([6])	Pizarro ([22]).	Scaglia ([26])	Logistic ([20])
$X$	1	6	11	15
$X_A$	2			
$N$	3	7		
$EtOH$	4	8	12	
$S$	5	9	13	
$Gly$		10		
$CO_2$			14	

Table 1: Number of equation associated to each variable and model.

the Coleman model considers biomass concentration controlled by the growth rate and death rate ( $\mu$  and  $\tau$ ). The growth rate is computed using Monod's equation with nitrogen nutrition and without considering competition ([19]; see Table 2); it considers that the lower is the remaining nitrogen the lower the growth rate. The death rate is considered proportional to ethanol concentration. The other fermentation variables are obtained by estimating production rates (for  $EtOH$ ) or consumption rates (for  $N$  and  $S$ ) per biomass unit. The effect of temperature is included to estimate parameters of the system with regression techniques. The Pizarro model ([22], [24], equations 6- 10) uses essentially the same differential equations as Coleman model, it adds the fermentation variable glycerol and does not consider active biomass concentration. The Pizarro method uses a different way for estimating uptakes and consumption rates per biomass unit. In this case intracellular behavior, studied by flux balance analysis, gives the specific production and consumption rates for the environmental conditions that are modeled by dynamic mass balance. It is built through an iterative process where intracellular network fluxes are bounded according to extracellular conditions, and for each iteration a maximization (of growth or glucose consumption rate) is performed to obtain uptakes and consumption rates that are used to predict extracellular concentrations of metabolites. The main goal is to better introduce the influence of environmental conditions in wine fermentation.

The Scaglia model includes other types of relations between variables (equations 11- 13). It considers only 4 fermentation variables:  $X$  (biomass concentration),  $S$  (sugar concentration),  $CO_2$  (carbonic dioxide gas concentration) and  $EtOH$  (ethanol concentration). The cell growth expression (equation 11) is based on Verlhust's logistic equation (15), where  $\mu$  represents the growth rate and the quadratic coefficient of population,  $\beta$ , models the competition for available resources. The growth rate ( $\mu$  in Table 2 for Scaglia method) of equation 15 is obtained by adaptations of Monod's model ([19]) with sugar nutrition: the lower is the remaining sugar the lower the growth rate. To model the death rate  $\tau$  they observe that the faster the decrease of substrate concentration, the larger the increase in the cellular death rate. Scaglia included coefficients associated to proportion of yeast cells in growth cellular step,  $F_\mu$ , and for those in death step,  $F_\tau$  (see Table 2), to avoid the discontinuity of Blackman's equation to model growth rate ([28]). These factors estimate the transition according to the proximity of carbon dioxide emission to the maximum expected for a normal fermentation progress ( $max(CO_2)$ ). It is considered Carbon dioxide concentration ( $CO_2$ , equation 14), being estimated with a emission rate coefficient per biomass unity, and the rate of an additional coefficient that we called  $CO_2Form$  (see Table 2). The ethanol production rate is obtained by estimating the conversion factor (yield) of carbon dioxide emitted to ethanol produced (equation 12). Nitrogen consumption is not considered, and the sugar consumption rate (equation 13) is composed by a term acting on biomass and a quadratic term (logistic equation type, 15).

The three teams worked on different strains of *Saccharomyces cerevisiae*. The Coleman team worked with the yeast *Premier Cuvee* (Red Star, Milwaukee, WI). It is a commercial strain of *Saccharomyces cerevisiae*. The fermentations were prepared at pH 3.35, total nitrogen was determined by measuring

ammonia concentration and alpha amino acid concentration. The Pizarro team worked with *Prise de Mousse* EC1118 (Lalvin, Zug, Switzerland), which is another commercial strain. The pH was kept constant at 3.5, nitrogen measures considered ammonia and free aminoacidic nitrogen. The Scaglia team used experimental results of other two studies: [31] and [11]. The *Saccharomyces cerevisiae Bsc411* of [31] was identified according to *Kurtzman & Fell* protocols and was taken from Argentina. The Fleet team studies are widely recognized. In both cultures acidity was controlled, the reducing sugars were determined colorimetrically using the *3,5-dinitrosalicylic acid (DNS) method* ([18]). Coleman and Pizarro models reviewed an ample range of temperatures and initial conditions of sugar and nitrogen. The Scaglia model was adjusted only on moderate levels of temperature, sugar and nitrogen concentrations. For more details about cultures and fermentation conditions which were used, refer directly to the papers.

## Experimental data

In our study we considered experimental data of three papers: [22], [15] and [16]. The experimental measures of the Pizarro team correspond to a wide range of data. We used laboratory results that were obtained with the strain *Prise de Mousse EC1118*, and industrial results for Industrial Cabernet Sauvignon, wine fermentations that were monitored during the 2003 vintage at a commercial winery in Chile. Sugar profiles for six batch fermentations at 28 °C with high/low nitrogen and other two at 12 °C and 17 °C with high conditions of nitrogen were used to calibrate the model, and consequently we expect a better adjustment of this model for sugar in this conditions. By direct communication with [15], we obtained data of biomass profiles in two particular conditions: moderate temperature (24 °C), high level of sugar (280 g/l) and moderate/high levels of nitrogen (approximately 220 mg/l and 551 mg/l respectively). The data set available in Mendes-Ferreira studie ([16]) used the strain *Saccharomyces cerevisiae PYCC4072* that was supplied by the Portuguese Yeast Culture Collection. The paper describes experimental biomass, ethanol and sugar (and other indexes) results for two experimental conditions, fermentation maintained at 20 °C with moderate initial sugar concentration (200 g/l) and initial nitrogen concentration high (267 mg/l) or low (66 mg/l). The acidity conditions were adjusted to pH 3.7, nitrogen is supplied by ammonium phosphate and sugar corresponds to glucose.

## 2 Results

### Exploratory analysis

The observation of experimental data gave us some ideas about the profiles of fermentation variables. In the three variables (biomass, ethanol and sugar concentrations) we observed two phases, transient and stable. Before a particular time, that we call *stabilization time*, fermentation variables change exponentially over time. After stabilization these are statistically constant. We verified the exponential behavior of biomass profiles statistically by means of linearization and linear regression (Figure 1), the growth rate can be assumed constant over time. In the case of biomass, in the first phase it increases exponentially until the cells stop their growth. Ethanol concentration increases while the yeast cells are active, after the stabilization time the production stops. Sugar concentration decays in an exponential way until it is completely consumed. Different samples show different uptake (for biomass and ethanol) and consumption rates (for sugar). For the Coleman ([6]) and Pizarro ([22]) models, fermentation variables evolve in time according to uptake (biomass, ethanol and glycerol) or consumption (sugar, nitrogen) factors per concentration unity of yeast cell (biomass). They assume that these coefficients change in time depending of the fermentation or environmental variables and do not depend only on initial conditions. When solving these models one obtains exponential behaviors whose rates change over time, according to the value of the fermentation variables or environmental conditions, finishing in a stable phase. According

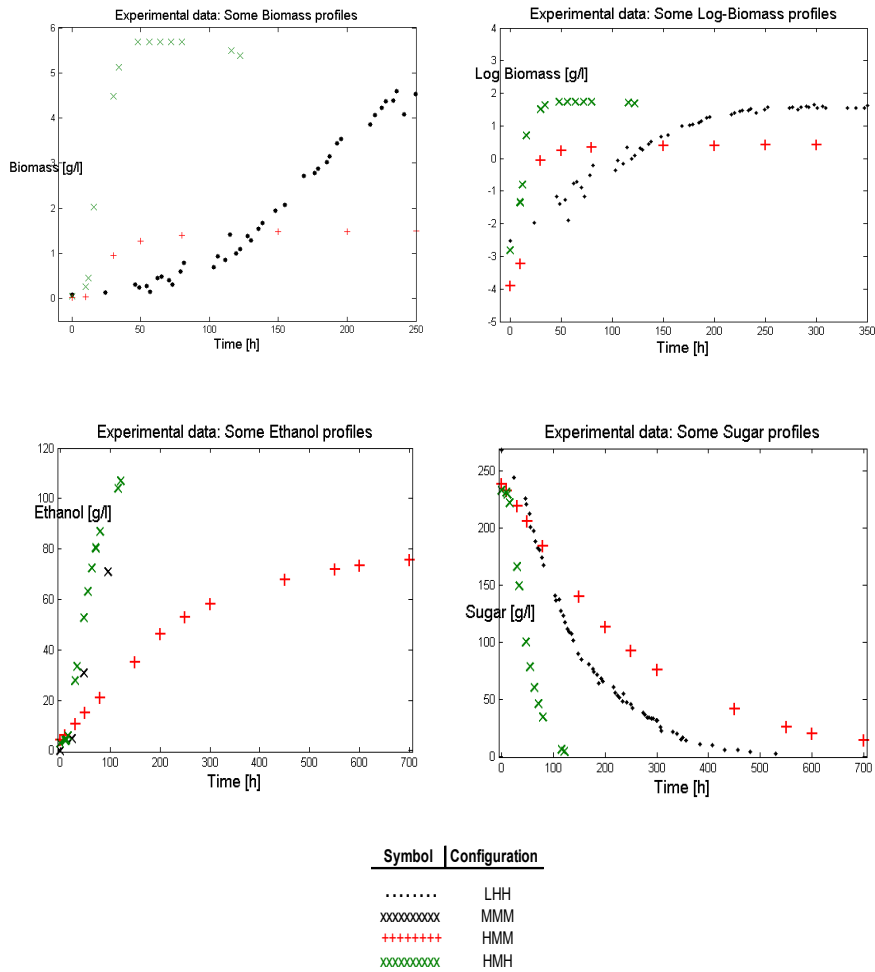


Figure 1: Fermentation variables profiles for some initial conditions configurations. For Biomass and Sugar one shows experimental results in *LHH* (low temperature: 12 °C, high initial sugar: 268 g/l, high initial nitrogen: 300 mg/l), *HMM* (28°C, sugar: 238 g/l, nitrogen: 50mg/l) and *HMH* (28 °C, sugar: 233 g/l, nitrogen: 300 mg/l) configurations. For Ethanol we show data for *MMM* configuration (20 °C, sugar: 200 g/l, nitrogen: 66 mg/l) instead of *LHH*. Log-Biomass profiles are shown too, we obtained linear correlations in transient phase of 0.98, 0.99 and 0.97 respectively.

to the sign of the factors we have exponential growth (positive sign) or decay (negative sign) followed by stabilization. Biomass profiles resemble the solutions of logistic differential equation 15, that are well known in ecology to model population growth. These types of differential equations were derived by Verhulst in 1838 to describe the self-limiting growth of a biological population (Verlhurst’s model; see [20]). Population starts to grow in an exponential phase, as it gets closer to the carrying capacity the growth slows down and reaches a stable level. The equations (11-13) that define the Scaglia model ([26]) include logistic components but they are more complex, and one observes relations between one-order differential expressions of variables.

### Symbolic step: rewriting the models

We rewrote the models into a polynomial form that in this case allowed to separate main effects of interaction effects that are represented by quadratic coefficients. The Coleman model is expressed by equations 1-5, the Pizarro model by equations 6-10 and the Scaglia model by equations 11-14. Table 2 describes the meaning of the variables and our notation as compared with the nomenclature of the original papers.

$$\frac{dX}{dt} = \mu \cdot X_A \quad (1)$$

$$\frac{dX_A}{dt} = (\mu - \tau) \cdot X_A \quad (2)$$

$$\frac{dN}{dt} = -v_N \cdot X_A \quad (3)$$

$$\frac{d[EtOH]}{dt} = v_{EtOH} \cdot X_A \quad (4)$$

$$\frac{dS}{dt} = -v_S \cdot X_A \quad (5)$$

$$\frac{dX}{dt} = \mu \cdot X \quad (6)$$

$$\frac{dN}{dt} = -v_N \cdot X \quad (7)$$

$$\frac{d[EtOH]}{dt} = v_{EtOH} \cdot X \quad (8)$$

$$\frac{dS}{dt} = -v_S \cdot X \quad (9)$$

$$\frac{d[Gly]}{dt} = v_{Gly} \cdot X \quad (10)$$

$$\frac{dX}{dt} = (F_\mu \cdot \mu - F_\tau \cdot \tau) \cdot X - F_\mu \cdot \beta \cdot X^2 \quad (11)$$

$$\frac{d[EtOH]}{dt} = \frac{1}{Y_{CO_2/EtOH}} \cdot \frac{dCO_2}{dt} \quad (12)$$

$$\frac{dS}{dt} = -((v_S + v_{S_0}) \cdot X - \frac{0.00002}{Y_{X/S}} \cdot X^2) \quad (13)$$

$$\frac{dCO_2}{dt} = v_{CO_2} \cdot X + \frac{d(CO_2Form)}{dt} \quad (14)$$

$$\frac{dX}{dt} = \mu \cdot X \cdot \left(1 - \frac{\beta}{\mu} \cdot X\right) \quad (15)$$

### Statistical step: Classifying the results

For fermentation process the considered factors are initial conditions and time. We studied the fermentation variables:  $X$  (biomass concentration),  $EtOH$  (ethanol concentration) and  $S$  (sugar concentration).



Notation	Meaning	Original notation and computation		
		Coleman ([6])	Pizarro ([22])	Scaglia ([26])
$X$	viable biomass concentration, $[g \cdot l^{-1}]$	$X$ total biomass concentration	$X_V$	$X$
$X_A$	active biomass concentration, $[g \cdot l^{-1}]$	$X_A$		
$N$	nitrogen concentration, $[g \cdot l^{-1}]$	$N$	$NH_4$	$N$
$EtOH$	ethanol concentration, $[g \cdot l^{-1}]$	$E$	$EtOH$	$P$
$S$	sugar concentration, $[g \cdot l^{-1}]$	$S$	$glu$	$S$
$Gly$	glycerol concentration, $[g \cdot l^{-1}]$	$Gly$		
$CO_2$	carbon dioxide concentration, $[g \cdot l^{-1}]$			$CO_2$
$\mu$	growth rate of yeast cells, $[h^{-1}]$	Monod's model nitrogen nutrition: $\frac{\max(\mu) \cdot N}{K_N + N}$	$\mu$ , Flux balance analysis	$12.072 \cdot \frac{\max(\mu) \cdot S}{S + K_S \cdot 93.0231 \cdot 508}$
$\tau$	death rate of yeast cells, $[h^{-1}]$	$k_d$ , temperature and ethanol dependent		$(0.0001 - 0.047 \cdot \frac{dS}{dt})$
$v_N$	nitrogen consumption rate per yeast mass, $[g \cdot g^{-1} \cdot h^{-1}]$	$\frac{\mu}{Y_{X/N}}$	$v_{NH_4}$ , Flux balance analysis	
$v_{EtOH}$	ethanol production rate per yeast mass, $[g \cdot g^{-1} \cdot h^{-1}]$	$\beta = \frac{\max(v_{EtOH}) \cdot S}{K_S + S}$	$v_{EtOH}$ , Flux balance analysis	
$v_S$	sugar consumption rate per yeast mass, $[g \cdot g^{-1} \cdot h^{-1}]$	$\frac{v_{EtOH}}{Y_{EtOH/S}}$	$v_{glu}$ , Flux balance analysis	$\frac{1}{Y_{X/S}} \cdot \frac{\max(\mu) \cdot S}{S + K_S \cdot 93.0231 \cdot 508}$
$v_{S_0}$	initial sugar consumption rate per yeast mass, $[g \cdot g^{-1} \cdot h^{-1}]$			$F = 0.008$ , initial sugar dependent
$v_{gly}$	glycerol production rate per yeast mass, $[g \cdot g^{-1} \cdot h^{-1}]$		$v_{gly}$ , Flux balance analysis	
$v_{CO_2}$	carbon dioxide production rate per yeast mass, $[g \cdot g^{-1} \cdot h^{-1}]$			$G \cdot \frac{\max(\mu) \cdot S}{S + K_S \cdot 93.0231 \cdot 445}$
$K_N$	constant for nitrogen-limited growth, $[g \cdot l^{-1}]$	0.009 Monod constant		
$K_S$	constant for sugar utilization in growth, $[g \cdot l^{-1}]$	10.278 Michaelis-Menten-type constant		2.15 Monod constant for sugar-limited growth
$Y_{X/N}$	yield coefficient for cell mass grown per mass of nitrogen used, $[g \cdot g^{-1}]$	$Y_{X/N}$ , temperature dependent		
$Y_{EtOH/S}$	yield coefficient for ethanol produced per sugar consumed, $[g \cdot g^{-1}]$	$Y_{E/S} = 0.55$		
$Y_{X/S}$	yield coefficient for cells formed per sugar consumed, $[g \cdot g^{-1}]$			0.029
$Y_{CO_2/EtOH}$	yield coefficient for $CO_2$ formed per ethanol produced, $[g \cdot g^{-1}]$			$Y_{CO_2/P}$ , initial conditions dependent
$\max(\mu)$	maximum growth rate of yeast cells, $[h^{-1}]$	$\mu_{max}$ , temperature dependent		$\mu_m$ , initial conditions dependent
$\max(v_{EtOH})$	maximum ethanol production rate, $[g \cdot g^{-1} \cdot h^{-1}]$	$\beta_{max}$ , temperature dependent		
$\max(CO_2)$	maximum $CO_2$ for normal fermentation progress, $[g \cdot l^{-1}]$			$CO_{2,95} = 78$ , corresponds to 95% of maximum emissions for initial conditions.
$F_\mu$	correction factor of growth rate			$\frac{\exp(-(CO_2 - \max(CO_2)))}{\exp(CO_2 - \max(CO_2)) + \exp(-(CO_2 - \max(CO_2)))}$
$F_\tau$	correction factor of death rate			$1 - F_g$
$\beta$	$\beta$ competition coefficient in logistic equation 15, $l \cdot g^{-1} \cdot h^{-1}$			$\beta$ , initial conditions dependent
$CO_2Form$	additional Scaglia concentration coefficient for carbon dioxide formation, $[g \cdot l^{-1}]$			$241.44 \cdot \max(\mu) \frac{S^2}{(S + K_S \cdot 93.0231 \cdot 0)(S + K_S \cdot 93.0231 \cdot 445)} \cdot X + 0.01 \cdot X$

Table 2: Notation, comparison with original papers nomenclature.

Temperature	Sugar	Nitrogen	Biomass	Ethanol	Sugar
Low (0-19 °C)	Moderate (160-240 [g/l])	Moderate (50-240 [mg/l])			
		High (240-551 [mg/l])			1
	High (240-308 [g/l])	Moderate			
		High	2		3
Moderate (20-27 °C)	Moderate	Moderate	1	1	1
		High	1	1	6 1
	High	Moderate	1		
		High	1		2
High (28-35 °C)	Moderate	Moderate	1	1	1
		High	2	1	2
	High	Moderate			
		High			1

	Pizarro data ([22])
	Malherbe data ([15])
	Mendes-Ferreira data ([16])

Table 3: Origin of experimental data. For each Initial Temperature-Sugar-Nitrogen configuration, and fermentation variable it is showed the origin of available data.

Temperature	Sugar	Nitrogen	Biomass		Ethanol		Sugar	
			Transient	Stable	Transient	Stable	Transient	Stable
Low (0-19 °C)	Moderate (180-240 [g/l])	Moderate (50-240 [mg/l])					Pizarro	Pizarro
		High (240-551 [mg/l])						
	High (240-308 [g/l])	Moderate					Scaqlia/Pizarro	Scaqlia/Pizarro
		High	Indiferent	Pizarro				
Moderate (20-27 °C)	Moderate	Moderate	Coleman/Pizarro	Coleman	Coleman	Indiferent	Coleman/Pizarro	Coleman
		High	Scaglia/Pizarro	Pizarro	Indiferent	Pizarro	Scaglia/Pizarro	Pizarro
	High	Moderate	Scaglia	Pizarro				
		High	Scaglia	Scaglia			Scaglia	Coleman
High (28-35 °C)	Moderate	Moderate	Coleman	Coleman	Pizarro	Coleman	Pizarro	Pizarro
		High	Coleman	Pizarro	Indiferent	Pizarro	Coleman	Pizarro
	High	Moderate						
		High					Pizarro	Scaqlia/Pizarro

Quality	Worst						Best	

Table 4: Criterion of selection of best models in function of initial conditions. For each combination variable-phase is written the best model, colors represent the quality of the adjustment when comparing between all the initial conditions.

Initial conditions were determined by ranges of initial temperature, sugar and nitrogen concentration, and we divided in transient and stable phase. The levels of initial conditions were ranges of initial temperature, sugar and nitrogen concentrations. We considered that the temperature is *low* when it is lower than 19 °C, is *moderate* for values between 20 °C and 27 °C, and *high* for larger values. Initial sugar concentration was called *moderate* for values less than 240 g/l and *high* for superior values. Initial nitrogen concentration was *moderate* for values less than 240 mg/l and *high* for those superior values. For each initial conditions configuration, we separated the profiles in *transient* and *stable phase* by analyzing the experimental results.

In Table 3 we show the origin of experimental data for different initial condition levels and fermentation variables. This classification allowed us to cover a wide range of configurations. In spite of this, for

some combinations we do not have experimental data because conditions of fermentation are difficult. This is the case for *low* initial temperature, sugar and nitrogen concentrations. For *high* temperature and sugar with insufficient levels of nitrogen source, we have the same situation. Part of the data have superior statistical quality. While the number of samples that describe an initial condition configuration is larger, the quantity of information that validates our assertions about the adjustment of each model in this configuration is also larger. In particular the Pizarro sample for *HMM* (high temperature, moderate initial sugar level and moderate initial nitrogen level) and *HMH* configurations give us standard deviations for variable profiles. Mendes-Ferreira ([16]) samples for *MHM* (moderate temperature, high initial sugar level and moderate initial nitrogen level) and *MHH* supply means and standard deviations of measures too.

We evaluated the three studied models according to how well they agree with the experimental results (Figure 2). For each sample we reviewed the adjustment in the transient and the stable phase. For the local criterion, at each point we built confidence intervals of experimental results by using measures of means and standard deviations, and computed the *p-values* associated to the decision of considering simulated value equal to experimental result. Because we observe exponential behavior, for the global criterion we have computed the correlation between the logarithm of simulations and the data over the time (Table 5). In general, an adjustment was considered *Very good* if the local criterion and the global one are very favorable ( $p\text{-value} \geq 0.1$ ,  $correlation \geq 0.98$ ); *Good* if a criterion is very favorable and the other one is only favorable ( $0.05 \leq p\text{-value} < 0.1$  or  $0.95 \leq correlation < 0.98$ ); *Little wrong* if a criterion is unfavorable ( $p\text{-value} < 0.05$  or  $correlation < 0.95$ ) and the other one is favorable or superior; and *Wrong* if both criteria are unfavorable. The cases near to the limits were checked especially. In case the local criterion is absolutely unfavorable ( $p\text{-value} = 0$ ) we qualified in *Wrong*, if local criterion is unfavorable (but not absolutely) and global criterion is optimum ( $correlation = 1$ ) we considered it *Good*.

Temperature	Sugar	Nitrogen	Biomass						Ethanol						Sugar									
			Coleman		Scaglia		Pizarro		Coleman		Scaglia		Pizarro		Coleman		Scaglia		Pizarro					
			Transient	Stable	Transient	Stable	Transient	Stable	Transient	Stable	Transient	Stable	Transient	Stable	Transient	Stable	Transient	Stable	Transient	Stable				
Low	Moderate	Moderate																						
		High																						
	High	Moderate																						
		High																						
Moderate	Moderate	Moderate																						
		High																						
	High	Moderate																						
		High																						
High	Moderate	Moderate																						
		High																						
	High	Moderate																						
		High																						

Very good

Good

Little wrong

Wrong

Figure 2: Summary of results of adjustment according to initial conditions. Quality of adjustment of each model for each initial conditions configuration and phase (transient and stable).

**Constructive step: Building the combined model**

As a direct consequence this criterion allows us to obtain better predictions of fermentation variable profiles than those obtained by each individual model. In Table 4 we summarize our criteria, and according to initial conditions and variable to consider (between Biomass, Ethanol and Sugar) we say which approach to use. We used the criteria to build a combined model where we separate the different effects in polynomial way to obtain a *combined model* (equations 16-18). These equations capture the three models

Config.	Variable	Coleman model				Pizarro model				Scaglia model				Combined model			
		Transient		Stable		Transient		Stable		Transient		Stable		Transient		Stable	
		C.1	C.2	C.1	C.2	C.1	C.2	C.1	C.2	C.1	C.2	C.1	C.2	C.1	C.2	C.1	C.2
MMM	X	0.006	1	0.006	0.92	0.109	1	0	1	0.012	0.96	0	0.77	0.033	1	0.006	0.92
	EtOH	0.001	1	0	0.97	0	1	0	1	0	1	0	0.93	0.001	1	0	0.97
	S	0.07	0.97	0.001	0.93	0.022	1	0	0.98	0.017	0.95	0	0.98	0.063	1	0.001	0.93
MMH	X	0.045	0.98	0.001	0.90	0.006	1	0.483	1	0.113	0.98	0.06	0.99	0.256	0.99	0.483	1
	EtOH	0.047	1	0	0.97	0.001	1	0.053	1	0.048	1	0.149	0.92	0.049	1	0.053	1
	S	0.129	0.93	0.152	0.98	0.140	0.98	0.27	0.98	0.064	0.99	0.015	0.97	0.129	0.99	0.27	0.98
HMM	X	0.388	0.99	0.476	0.97	0.102	0.83	0.027	1	0.082	0.99	0	0.87	0.388	0.99	0.476	0.97
	EtOH	0.049	0.95	0.155	0.97	0.129	0.95	0	1	0.08	0.92	0	-0.97	0.129	0.95	0.155	0.97
	S	0.171	0.79	0	-0.7	0.238	1	0.272	0.97	0.107	0.91	0.032	0.73	0.238	1	0.272	0.97
HMH	X	0.203	0.97	0	0.28	0.156	0.98	0.048	0.97	0.197	0.96	0	0.91	0.203	0.97	0.048	0.97
	EtOH	0.275	1	0.089	0.80	0.162	0.99	0.214	0.99	0.264	0.98	0.001	1	0.339	0.99	0.214	0.99
	S	0.327	1	0	0.59	0.135	0.98	0.167	0.99	0.197	0.95	0.001	1	0.327	1	0.167	0.99

Table 5: Statistical analysis of models. For each configuration of factors and fermentation variable, we show the average  $p$ -value for the local criterion (C.1), and *correlation* for global criterion (C.2). The lower the  $p$ -value, the bigger the local error in simulations. The bigger the *correlation*, the bigger the global similarity between data and simulations. Results for other configurations can be asked.

and the variables  $X_A$  and  $CO_2$  are those computed by the Coleman and Scaglia models respectively. The coefficients  $\mu_A$ ,  $\varepsilon_A$  and  $\sigma_A$  correspond to coefficients of the Coleman model to represent linear effect of  $X_A$  on  $X$ . The linear coefficient of  $X$  on  $EtOH$ ,  $\varepsilon^{(1)}$ , is associated to the Pizarro model;  $\mu^{(1)}$  is composed by contributions of the Pizarro and Scaglia models, and quadratic effects (coefficients  $\mu^{(2)}$  and  $\sigma^{(1)}$ ) are obtained from the Scaglia model. The coefficients are active or not in function of initial configuration and time (Table 6).

$$\frac{dX}{dt} = \mu_A \cdot X_A + \mu^{(1)} \cdot X - \mu^{(2)} \cdot X^2 \quad (16)$$

$$\frac{d[EtOH]}{dt} = \varepsilon_A \cdot X_A + \varepsilon^{(1)} \cdot X + \varepsilon_{CO_2} \cdot \frac{dCO_2}{dt} \quad (17)$$

$$\frac{dS}{dt} = -(\sigma_A \cdot X_A + \sigma^{(1)} \cdot X - \sigma^{(2)} \cdot X^2) \quad (18)$$

Coefficient	Meaning	Configuration of initial conditions								
		LMH	LHH	MMM	MMH	MHM	MHH	HMM	HMH	HHH
$\mu_A = \frac{\max(\mu) \cdot N}{K_N + N}$	linear effect of $X_A$	-	$\emptyset$	$\forall t$	$\emptyset$	$\emptyset$	$\emptyset$	$\forall t$	$\forall t$	-
$\mu^{(1)} = FBA, (F_\mu \cdot \mu + F_\tau \cdot \tau)$	linear effect of $X$	-	$t > 110$	$t \leq 96$	$\forall t$	$\forall t$	$\forall t$	$\emptyset$	$t > 30$	-
$\mu^{(2)} = F_\mu \cdot \beta$	quadratic effect of $X$	-	$\emptyset$	$\emptyset$	$t \leq 51$	$t \leq 27$	$\forall t$	$\emptyset$	$\emptyset$	-
$\varepsilon_A = \frac{\max(v_{EtOH}) \cdot S}{K_S + S}$	linear effect of $X_A$	-	-	$t \leq 96$	$t \leq 51$	-	-	$t > 300$	$t \leq 30$	-
$\varepsilon^{(1)} = FBA$	linear effect of $X$	-	-	$\emptyset$	$\forall t$	-	-	$t \leq 300$	$\forall t$	-
$\varepsilon_{CO_2} = \frac{1}{\bar{v}_{CO_2/EtOH}}$	linear effect of $\frac{dCO_2}{dt}$	-	-	$\emptyset$	$t \leq 51$	-	-	$\emptyset$	$t \leq 30$	-
$\sigma_A = \frac{v_{EtOH}}{Y_{EtOH/S}}$	linear effect of $X_A$	$\emptyset$	$\emptyset$	$\forall t$	$\emptyset$	-	$t > 107$	$\emptyset$	$t \leq 30$	$\emptyset$
$\sigma^{(1)} = FBA, 0.008 + \frac{\max(\mu) \cdot S}{Y_{X/S} \cdot (S + K_S \cdot 93.0231 \cdot 508)}$	linear effect of $X$	$\forall t$	$\forall t$	$t \leq 96$	$\forall t$	-	$t \leq 107$	$\forall t$	$t > 30$	$\forall t$
$\sigma^{(2)} = \frac{2 \cdot 10^{-2}}{Y_{X/S}}$	quadratic effect of $X$	$\emptyset$	$\forall t$	$\emptyset$	$t \leq 51$	-	$t \leq 107$	$\emptyset$	$\emptyset$	$t > 103$

Table 6: Temporal intervals at which the coefficients of equations 16-18 are active for each configuration of initial conditions and the formulas. One writes – if there are not experimental data to validate, *FBA* denotes the result by using Flux balance Analysis.

For instance let us consider the configuration *MMH*. The equations 16-18 take the form from the

equations 19-21. One observes that the main factor is the linear effect of  $X$ . In the transient phase there exists a quadratic effect of  $X$ , a linear effect of  $CO_2$  rate on  $EtOH$ , and  $X_A$  only affects the ethanol modelling.

$$\frac{dX}{dt} = \mu^{(1)} \cdot X - 1_{t \leq 51} \cdot \mu^{(2)} \cdot X^2 \quad (19)$$

$$\frac{d[EtOH]}{dt} = 1_{t \leq 51} \cdot \varepsilon_A \cdot X_A + \varepsilon^{(1)} \cdot X + 1_{t \leq 51} \cdot \varepsilon_{CO_2} \cdot \frac{dCO_2}{dt} \quad (20)$$

$$\frac{dS}{dt} = -(\sigma^{(1)} \cdot X - 1_{t \leq 51} \cdot \sigma^{(2)} \cdot X^2) \quad (21)$$

For configuration *MMH* we observed an initial effect of competition to consume resources; but for *HHH* the sugar consumption (equation 22) initially, when the substrates are abundant, competition does not exist but it appears when the resources become scarce.

$$\frac{dS}{dt} = -(\sigma^{(1)} \cdot X - 1_{t > 103} \cdot \sigma^{(2)} \cdot X^2) \quad (22)$$

As a result, in function of factors, we go from one profile type to another one. Temporal phase and initial condition affects the results. For each fermentation there exists a time at which the profiles change from transient to stable phase: *stabilization time*. According to environmental conditions one obtains different profiles of the fermentation variables; for different level of initial temperatures, sugar and nitrogen concentration, one observes different growth/decrease rates and the change to the stable phase happens in different time (Figure 1). The bigger the initial temperature, sugar or nitrogen concentration; the bigger the growth rate of biomass. As the biomass changes to its stable phase, ethanol production and the consumption of sugar stop.

## Conclusions and Discussion

We built a general method to combine models in function of configurations of factors. The method was applied to fermentation process modelling to explain the profiles of fermentation variables: concentration of yeast biomass, ethanol and sugar; by considering four factors: initial temperature, sugar and nitrogen, and growth phase. Our method starts with a symbolic step to homogenize the notation, for ODE models by rewriting into polynomial form and by identifying main and interaction effects. It continues with a statistical step to evaluate the models, in function of experimental data ([22], [15] and [16]). We defined discrete levels for each factor, for each configuration of factors and fermentation variable we statistically compared the results of three kinetic fermentation models ([6], [22] and [26]) with the experimental results and we obtained quality indexes of each model (Figure 2). We finished with the construction of a *combined model*, where one selects the best resolution method for each fermentation variable and configuration of factors (Table 4). The equations 16-18 allows to interpret the combined model in function of initial configuration, for instance equations 19-21 for *MHH*.

Although generally for all variables there exist combinations of models of good adjustment to the experimental data, for each one of the fermentation variables and initial conditions the approaches showed different quality levels (Table 4, Figure 2). The best simulations were obtained for sugar consumption, in general terms Pizarro model showed the best adjustments especially for *high* levels of temperature. For low temperature and *high* levels of sugar and nitrogen the Scaglia model showed results similar to those of the Pizarro model. The configuration *MMH* (Figure 3) was covered by two different data sets for sugar simulations, the Pizarro and Mendes-Ferreira data ([16]). We observed similar measures between data

sets and that for the transient phase the Scaglia model agreed better with the Mendes-Ferreira data ([16]), and for the stable phase the predictions of Pizarro are the best. For the *HMM* configuration we observed the best results with the Pizarro simulations; this agreed with the calibration data used to estimate sugar uptake parameters. For *HMH* configuration Coleman model showed the best results in transient phase, Pizarro worked better in stable phase and for *HHH* configuration.

The worst quality levels were obtained for biomass: Coleman and Scaglia models best agreed with experimental data in the transient phase. The configurations *HMH* and *HMM* showed the best results for Coleman model; *MHM* and *MHH* for the Scaglia model. Pizarro model worked better in stable phase, best results in *LHH*, *MMH*, *MHM*, *MHH* and *HMM*. Pizarro and Coleman models showed the best results for Ethanol production simulations.

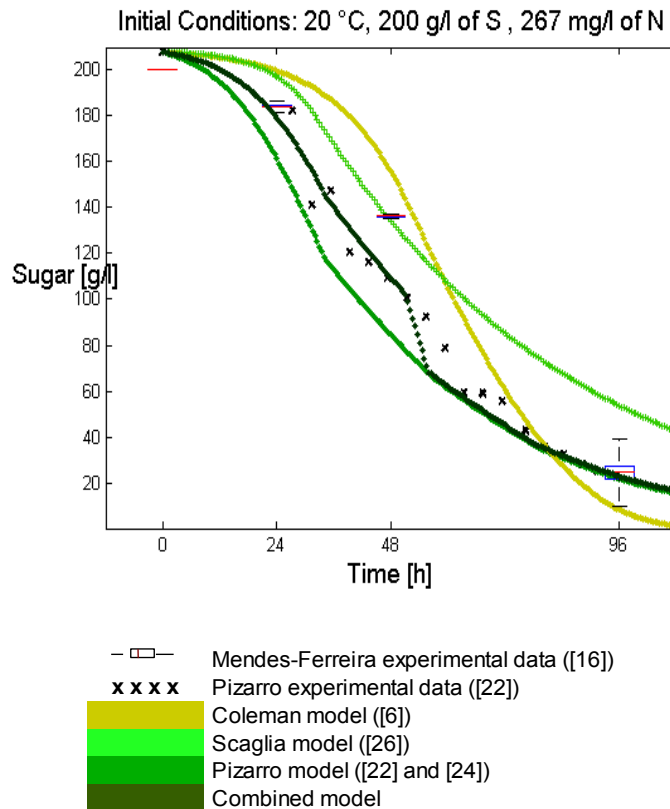


Figure 3: Comparison between models adjustments for sugar profiles in *MMH* initial configuration. Sugar profiles simulated by the three models and our combined model. Two experimental samples; with temperature: 20 °C, sugar: 200 and 207 g/l, nitrogen: 267 and 240 mg/l.

Our combined model, obtained good results for almost all the initial conditions configurations. As it is observed in Figure 2 there exist very few initial conditions in which no model obtains good results of adjustment. The only negative cases are *LHH* configuration in transient phase for biomass, and *MMM* configuration in stable phase of Ethanol. We observed the best results for sugar profiles simulations, for all the initial conditions one obtains that the transient or stable phase is adjusted with quality *good* (Table 4). For the transient phase the best configurations are *HHH*, *HMM* and *HMH*; for stable phase *HMM*, followed by *LMH* and *LHH*. For this fermentation variable, the configurations *LHH*, *HMM* and *HMH* were represented mainly by Pizarro model, in which they were used to calibrate sugar-uptake parameters. For *MMH* (Figure 3) we obtained *good* results in transient phase, by combining Scaglia and

Pizarro model, and *very good* results in stable phase with the Pizarro model. The best result for biomass adjustments were obtained for the configuration *HMM* that showed very good results for both transient and stable phase. We obtained good results for *MMH*, *MHM* and *MHH* in both temporal phases too. With respect to simulations of ethanol profiles, the best result was obtained for *HMH* configurations. For this configuration the transient phase is very good represented by the three models and the stable phase only by the Pizarro approach. Another important fermentation variable is nitrogen concentration. The experimental data of [15] and [16] give us nitrogen measurements for *MMM*, *MMH*, *MHM* and *MHH* configurations. We observed that for these initial conditions, the combined model is completely represented by the Pizarro model. It obtains the best results (good or very good quality), the Coleman approach showed poor results and the Scaglia model does not include this variable.

In this study we chose to build combined models without changing the individual models, since by design we assume that the original models are validated. The other alternative is to tune the internal parameters to refine the models and to obtain more generality. Several types of experimental results can be included to improve the estimation methods, looking for the correct inclusion of the relevant factors of the fermentation process. The effect of these factors can be directly computed by using devices to measure the number of cells (by using Hemacytometer or Neubauer for example) on fermentation samples or by biotechnological tools as *DNA Microarrays* ([14]) and *PCR* (Polymerase chain reaction, [13]). The estimations of profiles features as growth rate, death rate and yield coefficients developed in [6] can be extended to more strains and species. Microarrays and PCR can give us estimations of the profiles too. In [1], [16] and [17] were done studies of gene expression profiles of particular strains of *Saccharomyces cerevisiae* during fermentations with high level of sugar but different levels of nitrogen concentration. In [1] was observed that some genes involved in biosynthesis of macromolecular precursors have superior levels of expression in high nitrogen condition than low. Low levels of nitrogen showed expression levels superior for genes involved in translation and oxidative carbon metabolism. In [16] and [17] were observed early responses of yeast cells to low nitrogen. They identified 36 genes highly expressed under conditions of low or absent nitrogen in comparison with a nitrogen-replete condition for *Saccharomyces cerevisiae* *PYCC4072*, the behavior of four of these transcripts was confirmed by *RT-PCR* analysis in this and another wine yeast strain. The signature genes of both studies can be used to predict nitrogen deficiency and to prevent fermentation problems. These ideas can be extended to study the temporal transcriptional, responses of genes on different pH, initial temperature, sugar and nitrogen concentrations, strains and species of yeasts. The analysis can be oriented, for example, to genes associated with oxidation of glucose, glycolysis and anaerobic functions. Expression levels of enzymes allow to simulate phenotype by FBA on metabolic models to obtain uptakes and consumption rates on different conditions and yeast strains.

Another challenge to obtain better approaches of the reality is to construct and to calibrate fermentation models that consider interacting yeast populations competing by resources. Although it has been observed that *Saccharomyces cerevisiae* is dominant in the majority of spontaneous alcoholic fermentations ([12], [25]) and it is the most popular yeast in inoculated cultures, there exist many strains and other yeasts as *Candida cantarellii* that participate in the process ([21], [34]) and it influences the aroma ([23]). The intervention of *lactic acid and acetic bacteria* in fermentations is also documented ([10], [9]). One can consider competence between individuals of the same population, modelled by logistic-like models similar to equation 15, and interactions between different populations. A usual way to model the presence of two or more populations competing by resources is the *Lotka-Volterra-like models* ([27]). Different strains of *Saccharomyces cerevisiae* can present different levels of tolerance to ethanol, acidity, growth and death rate between other coefficients, Another fermenting yeast *Candida cantarellii* present different rates of growth, ethanol and glycerol yields [31]. In future versions we will introduce the dependency of these rates with respect to yeast strains or species, and pH conditions.

In *System Biology* the combination of models takes relevance to analyze hierarchical systems. The

notion *composition* ([8]) is used to build models by defining their components and the relations between them. A system is analyzed in a hierarchic way, defining it as being composed by subsystems. The components with different nature are well defined using different formalisms to generating sub-models, and *Base formalisms* capable of including the semantics of a wide variety of languages are used to define combined models ([2]). In this study we focused in the mathematical way of combining models, in future works we will approach the formalism for defining general combinations.

## References

- [1] Leilah E. Backhus, Joseph DeRisi, Patrick O. Brown, and Linda F. Bisson. Functional genomic analysis of a commercial wine strain of *saccharomyces cerevisiae* under differing nitrogen conditions. *FEMS Yeast Research*, 1:111–125, April 2001.
- [2] F.J. Barros, M.T. Mendes, and B.P. Zeigler. Variable DEVS-variable structure modeling formalism: an adaptive computer architecture application. In *Fifth Annual Conference on AI, and Planning in High Autonomy Systems*, pages 185–191, Gainesville, FL, USA, 1994.
- [3] Linda F. Bisson. Stuck and sluggish fermentations. *Am. J. Enol. Vitic.*, 50(1):107–119, March 1999.
- [4] Roger Boulton. The prediction of fermentation behavior by a kinetic model. *Am. J. Enol. Vitic.*, 31(1):40–45, March 1980.
- [5] Roger B. Boulton, Vernon L. Singleton, Linda F. Bisson, and Ralph E. Kunkee. *Principles and Practices of Winemaking*. Springer, 1st edition, 1996.
- [6] Matthew C. Coleman, Russell Fish, and David E. Block. Temperature-Dependent kinetic model for Nitrogen-Limited wine fermentations. *Applied and Environmental Microbiology*, 73(18):5875–5884, September 2007. PMID: 17616615 PMID: 2074923.
- [7] Amanda C. Cramer, Sophocles Vlassides, and David E. Block. Kinetic model for nitrogen-limited wine fermentations. *Biotechnology and Bioengineering*, 77(1):49–60, 2002.
- [8] SBML developers. SbmL composition workshop 2007. [http://sbml.info/Events/Other\\_Events/SBML\\_Composition\\_Workshop\\_2007](http://sbml.info/Events/Other_Events/SBML_Composition_Workshop_2007), September 2007.
- [9] G. S. Drysdale and G. H. Fleet. Acetic acid bacteria in winemaking: A review. *Am. J. Enol. Vitic.*, 39(2):143–154, June 1988.
- [10] G. H. Fleet, S. Lafon-Lafourcade, and P. Ribéreau-Gayon. Evolution of yeasts and lactic acid bacteria during fermentation and storage of bordeaux wines. *Applied and Environmental Microbiology*, 48(5):1034–1038, November 1984. PMID: 16346661 PMID: 241671.
- [11] Graham H. Fleet. *Wine Microbiology and Biotechnology*. CRC Press, 1 edition, 1993.
- [12] Valerie Frezier and Denis Dubourdieu. Ecology of yeast strain *saccharomyces cerevisiae* during spontaneous fermentation in a bordeaux winery. *Am. J. Enol. Vitic.*, 43(4):375–380, December 1992.
- [13] K. Kleppe, E. Ohtsuka, R. Kleppe, I. Molineux, and H. G. Khorana. Studies on polynucleotides , : XCVI. repair replication of short synthetic DNA's as catalyzed by DNA polymerases. *Journal of Molecular Biology*, 56(2):341–361, March 1971.
- [14] D A Kulesh, D R Clive, D S Zarlenga, and J J Greene. Identification of interferon-modulated proliferation-related cDNA sequences. *Proceedings of the National Academy of Sciences of the United States of America*, 84(23):8453–8457, December 1987.
- [15] S. Malherbe, V. Fromion, N. Hilgert, and J.-M. Sablayrolles. Modeling the effects of assimilable nitrogen and temperature on fermentation kinetics in enological conditions. *Biotechnology and Bioengineering*, 86(3):261–272, 2004.
- [16] A. Mendes-Ferreira, M. del Olmo, J. Garcia-Martinez, E. Jimenez-Marti, A. Mendes-Faia, J. E. Perez-Ortin, and C. Leao. Transcriptional response of *saccharomyces cerevisiae* to different nitrogen concentrations during alcoholic fermentation. *Appl. Environ. Microbiol.*, 73(9):3049–3060, May 2007.
- [17] A. Mendes-Ferreira, M. del Olmo, J. García-Martínez, E. Jiménez-Martí, C. Leão, A. Mendes-Faia, and J. E. Pérez-Ortín. *Saccharomyces cerevisiae* signature genes for predicting nitrogen deficiency during alcoholic



- fermentation. *Applied and Environmental Microbiology*, 73(16):5363–5369, August 2007. PMID: 17601813  
PMCID: 1950961.
- [18] G. L. Miller. Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Analytical Chemistry*, 31(3):426–428, March 1959.
- [19] Monod. La technique de culture continue; théorie et applications. *Ann Ist Pasteur Lille*, 79:390–410, 1950.
- [20] James D. Murray. *Mathematical Biology: I. An Introduction (Interdisciplinary Applied Mathematics)*. Springer, 3rd edition, December 2007.
- [21] Canan Nurgel, Huseyin Erten, Ahmet Canbas, Turgut Cabaroglu, and Serkan Selli. Yeast flora during the fermentation of wines made from vitis vinifera L. cv. emir and kalecik karasi grown in anatolia. *World Journal of Microbiology and Biotechnology*, 21(6):1187–1194, October 2005.
- [22] Francisco Pizarro, Cristian Varela, Cecilia Martabit, Claudio Bruno, J. Ricardo Pérez-Correa, and Eduardo Agosin. Coupling kinetic expressions and metabolic networks for predicting wine fermentations. *Biotechnology and Bioengineering*, 98(5):986–998, 2007.
- [23] María Eugenia Rodríguez, Christian A Lopes, Raúl J Barbagelata, Nora B Barda, and Adriana C Caballero. Influence of candida pulcherrima patagonian strain on alcoholic fermentation behaviour and wine aroma. *International Journal of Food Microbiology*, 138(1-2):19–25, March 2010. PMID: 20116878.
- [24] Javier Sainz, Francisco Pizarro, J. Ricardo Pérez-Correa, and Eduardo Agosin. Modeling of yeast metabolism and process dynamics in batch fermentation. *Biotechnology and Bioengineering*, 81(7):818–828, 2003.
- [25] Pilar Santamaría, Patrocinio Garijo, Rosa López, Carmen Tenorio, and Ana Rosa Gutiérrez. Analysis of yeast population during spontaneous alcoholic fermentation: Effect of the age of the cellar and the practice of inoculation. *International Journal of Food Microbiology*, 103(1):49–56, August 2005.
- [26] Gustavo J.E. Scaglia, Pablo M. Aballay, Carmen A. Mengual, Martha D. Vallejo, and Oscar A. Ortiz. Improved phenomenological model for an isothermal winemaking fermentation. *Food Control*, 20(10):887–895, 2009.
- [27] James F. Selgrade. Dynamical behavior of a competitive model with genetic variation. *Applied Mathematics Letters*, 2(1):49–52, 1989.
- [28] Michael L. Shuler and Fikret Kargi. *Bioprocess Engineering: Basic Concepts*. Prentice Hall College Div, 1st edition, November 1991.
- [29] Venkatanarayanan Subramanian, Kristan K.S. Buck, and David E. Block. Use of decision tree analysis for determination of critical enological and viticultural processing parameters in historical databases. *Am. J. Enol. Vitic.*, 52(3):175–184, September 2001.
- [30] P. Teissier, B. Perret, E. Latrille, J. M. Barillere, and G. Corrieu. A hybrid recurrent neural network model for yeast production monitoring and control in a wine base medium. *Journal of Biotechnology*, 55(3):157–169, July 1997.
- [31] M.E. Toro and F. Vazquez. Fermentation behaviour of controlled mixed and sequential cultures of candida cantarellii and saccharomyces cerevisiae wine yeasts. *World Journal of Microbiology and Biotechnology*, 18(4):347–354, January 2002.
- [32] Alejandra Urtubia, J. Ricardo Pérez-Correa, Alvaro Soto, and Philipppo Pszczółkowski. Using data mining techniques to predict industrial wine problem fermentations. *Food Control*, 18(12):1512–1517, December 2007.
- [33] S. Vlasides, J. Ferrier, and D. Block. Using historical data for bioprocess optimization: modeling wine characteristics using artificial neural networks and archives process information. *Biotechnology and Bioengineering*, 73(1):55–68, 2001.
- [34] Katharina Zott, Cecile Miot-Sertier, Olivier Claisse, Aline Lonvaud-Funel, and Isabelle Masneuf-Pomarede. Dynamics and diversity of non-Saccharomyces yeasts during the early stages in winemaking. *International Journal of Food Microbiology*, 125(2):197–203, July 2008. PMID: 18495281.