1 Monod law is insufficient to explain biomass growth 2 in nitrogen-limited yeast fermentation

3 David Henriques,^{a,*} Eva Balsa-Canto^a

4 Bioprocess engineering group, IIM-CSIC, Vigo, Spain ^a;

5 Address correspondence to David Henriques, davidh@iim.csic.es

6 **ABSTRACT** The yeast *Saccharomyces cerevisiae* is an essential microorganism in food biotechnology; par-
7 ticularly, in wine and beer making. During wine fermentation, yeasts transform sugars present in the grape
8 juice into ethanol and carbon dioxide. The process occurs in batch conditions and is, for the most part, an
9 anaerobic process. Previous studies linked limited-nitrogen conditions with problematic fermentations, with
10 negative consequences for the performance of the process and the quality of the final product. It is, there-
11 fore, of the highest interest to anticipate such problems through mathematical models. Here we propose a
12 model to explain problematic fermentations under nitrogen-limited anaerobic conditions. We separated the
13 biomass formation into two phases: growth and carbohydrate accumulation. Growth was modelled using the
14 well-known Monod law while carbohydrate accumulation was modelled by an empirical function, analogous to
15 a proportional controller activated by the limitation of available nitrogen. We also proposed to formulate the
16 fermentation rate as a function of the total protein content when relevant data are available. The final model
17 was used to successfully explain a series of experiments taken from the literature, performed under normal
18 and nitrogen-limited conditions. Our results revealed that Monod law is insufficient to explain biomass forma-
19 tion kinetics in nitrogen-limited fermentations of *S. cerevisiae*. The goodness-of-fit of the herewith proposed
20 model is superior to that of previously published models, offering the means to predict, and thus control,
21 problematic fermentations.

22 **IMPORTANCE** Problematic fermentations still occur in the winemaking industrial practise. Problems include
23 sluggish rates of fermentation, which have been linked to insufficient levels of assimilable nitrogen. Data and
24 relevant models can help anticipate poor fermentation performance. In this work, we proposed a model to
25 predict biomass growth and fermentation rate under nitrogen-limited conditions and tested its performance
26 with previously published experimental data. Our results show that the well-known Monod law does not suffice
27 to explain biomass formation. A second term accounting for carbohydrate accumulation is required to predict
28 successfully, and thus control, problematic fermentations.

29 **KEYWORDS:** problematic fermentations, nitrogen-limited, kinetic model, *S. cerevisiae*, biomass, wine,
30 sluggish

31 INTRODUCTION

32 The yeast species *Saccharomyces cerevisiae* is the best-studied eukaryote. *S. cerevisiae* presents unique charac-
33 teristics such as its fermentation capacity and its ability to withstand adverse conditions of osmolarity and low
34 pH. These desirable properties have made of *S. cerevisiae* one of the workhorses of bio-industries including
35 food, beverage -especially wine- and biofuel production industries (1).

36 In winemaking, the fermentation turns grape must into an alcoholic beverage. Wine production occurs in
37 a closed system (i.e. in batch conditions) and is, for the most part, an anaerobic process. During fermentation,
38 yeasts transform sugars present in the must into ethanol and carbon dioxide. A complete alcoholic fermenta-
39 tion is achieved when the residual fermentable sugar is less than 2 g/L. Despite improvements in fermentation
40 control, problematic fermentations such as, stuck - with a higher than desired sugar residual- and sluggish
41 -unusually long- fermentations still occur in real practice.

Various studies have shown that insufficient levels of assimilable nitrogen contribute to stuck or sluggish fermentations (2). A minimum of 120–140 mg/L of assimilable nitrogen is required to achieve a standard fermentation rate, while nitrogen content in grape juice may be as low as 60 mg/L (3). Therefore, it has become common practice to supplement nitrogen-deficient musts with diammonium phosphate (4). Nevertheless, both excessive or insufficient nitrogen can lead to the production of undesired metabolites affecting the organoleptic properties of wine (5, 4).

The need to predict and control wine fermentation has motivated a quest for mechanistic models of alcoholic fermentation (see the reviews by Marín (6) and Miller and Block (7). Several approaches linking nitrogen consumption to growth and fermentation rate have been proposed in the literature. Alternatives differ on the way they explain biomass formation and the relation of cell mass (or sometimes cell numbers) to fermentation rate.

Dynamic kinetic models incorporate the Monod equation to describe biomass growth (8, 9, 10). The estimation of the Monod parameters requires nitrogen and sugars uptake data throughout the fermentation. Otherwise, lack of identifiability of the parameters will limit the predictive capability of the model, and thus logistic functions may be more suitable (11, 12). Both formulations, Monod or logistic, achieve a maximum biomass value which can not be exceeded.

However, Schulze et al. (13) showed that, for nitrogen-limited anaerobic fermentations of *S. cerevisiae*, there is a substantial increment in biomass content after the depletion of ammonia. Besides, measurements of protein and messenger RNA (mRNA) contents revealed that their concentrations stay stable while the concentration of carbohydrates increases. Similar effects have been detected in baker's yeast exposed to nitrogen starvation but not in the carbon starved cells; see (14) for a fed-batch aerobic example, or (15) for a chemostat anaerobic example. These data would imply that, at least for ammonium deficient scenarios, the use of the Monod model would be insufficient.

Similarly, Varela et al. (16) showed that protein/carbohydrate fractions in the biomass of *S. cerevisiae* vary substantially throughout the fermentation of wine fermentations. To account for this dynamics, Pizarro et al. (17) and Vargas et al. (18) included, in their flux balance analysis models, an empirical function that controls the concentration of carbohydrates in the newly formed biomass as a function of extracellular sugar.

In what regards to the modelling of fermentation rate, previous studies proposed two alternatives. The first implies that the fermentation rate is proportional to biomass or cell numbers (8, 19, 9, 10, 12). The second includes the role of hexose transporters as a separate entity that is dependent on the transport of ammonia (20, 11, 21, 22, 23).

The advantage of the second over the first is that it enables the simulation of the effect of nitrogen additions. Still, further experimental analysis of the underlying hypothesis is required. Also, a good adjustment to the data comes at the cost of a large number of estimated parameters.

Pizarro et al. (17) and Vargas et al. (18) proposed a third alternative. These authors assumed that cells use different hexose transporters at different fermentation stages based on previously published experimental data (24). However, this approach also requires the adjustment of the maximum transport rate for each hexose carrier. As the maximum transport rate is deemed to be a function of temperature and the initial nitrogen concentration, the modelling also involves the estimation of a substantial number of parameters.

In this work, we propose a dynamic kinetic model capable of successfully explaining nitrogen-limited fermentations. The model describes growth and fermentation rate while distinguishing protein from carbohydrate fractions of biomass. The Monod equation describes the growth, and the carbohydrate accumulation observed after the depletion of ammonia is described by an empirical function, analogous to a proportional controller.

We tested the model properties by fitting the data provided by Schulze et al. (13) and Coleman et al. (9) who explored fermentations under various nitrogen regimes. Our results show why standard Monod models often produce poor results in modelling nitrogen-limited fermentations of *S. cerevisiae*.

Fermentation rate can be a function of the biomass or, as proposed here, a function of the protein fraction since the glycolytic enzymes are an important part of the protein pool. Remarkably, even if protein content is not measured, the model is identifiable provided the uptake of hexoses and nitrogen, as well as the production

Monod law and nitrogen-limited yeast fermentations

of ethanol, are measured over time.

We compared the performance of the proposed model with the Monod model and the model proposed by Cramer et al. (8) who used it to explain nitrogen-limited fermentations. Our results showed that the proposed model captures the data more accurately than previously published alternatives (up to 6.8 times improvement in normalised root-mean-square error in the fitting of biomass dynamics, and up to 3.8 times improvement in the fitting of the overall process dynamics). In consequence, our model can be used to predict, and thus, control, fermentation problems.

RESULTS

Modeling biomass formation during nitrogen-limited fermentation. The microbial growth rate was described as the sum of two different terms. The first, μ_N , corresponds to the healthy cell division, mainly when assimilable nitrogen sources (YAN) are abundant. The second, μ_C , corresponds to a secondary increase in biomass after the depletion of nitrogen sources- which we assumed corresponds, mostly to carbohydrate accumulation. The biomass growth equation reads:

$$\frac{dX}{dt} = \mu_N(X) + \mu_C(X) \quad (1)$$

where X is the biomass (g/L).

To model growth associated with cell division we used a Monod type kinetics (25) where nitrogen sources were considered the limiting nutrient:

$$\mu_N = \text{lag}(t) \cdot X \cdot \mu_{\max N} \cdot \frac{N}{N + k_{sN}} \quad (2)$$

where $\mu_{\max N}$ is the maximum specific growth rate, N is extracellular YAN (g/L), k_{sN} is the Monod equation parameter and lag is a function of time (t) representing the lag phase. The lag phase is typically a period of adjustment in which a given microbial population adapts to a new medium before it starts growing exponentially. To model the lag we used the model proposed by (26):

$$\text{lag} = \frac{a_0}{a_0 + (1 - a_0) \cdot e^{-\mu_{\max N} \cdot t}} \quad (3)$$

where a_0 is an estimated parameter and $\mu_{\max N}$ is the maximum growth rate associated with μ_N .

The secondary increase in biomass concentration (μ_C) is activated by the decrease the concentration of YAN:

$$\phi_N = \left(1 - \frac{N}{N + k_{sC}}\right) \quad (4)$$

where k_{sN} is a parameter controlling the half-maximal inactivation of μ_N which is modeled with an empirical expression analogous to a proportional controller:

$$\mu_N = \text{lag}(t) \cdot \mu_{\max C} \cdot \phi_N \cdot \left(\theta_C - \frac{X_C}{X}\right) \quad (5)$$

where θ_C is the set-point (target value) for the ratio of carbohydrates in the biomass content (X_C/X) and $\mu_{\max C}$ controls the velocity of the convergence towards that reference point.

The generation of carbohydrates is affected both by μ_N and μ_C and its dynamics is described as:

$$\frac{dX_C}{dt} = \lambda_C \cdot \mu_N + \mu_C \quad (6)$$

where λ_C is the carbohydrate content in the biomass during exponential growth.

On the contrary the formation of protein is only affected by μ_N and its dynamics is described by:

$$\frac{dX_P}{dt} = \lambda_P \cdot \mu_N \quad (7)$$

where λ_P is the protein content in the biomass during exponential growth.

YAN is the limiting nutrient in the primary growth phase and its consumption is proportional to μ_N :

$$\frac{dN}{dt} = -lag(t) \frac{\mu_N}{Y_{X/N}} \quad (8)$$

where $Y_{X/N}$ is the YAN to biomass yield.

For the sake of comparison we used two additional models: the first, regarded as M_M , corresponds to that for which the biomass follows the Monod description:

$$\frac{dX}{dt} = \mu_N(X) \quad (9)$$

and the second, regarded as M_C , corresponds to the model proposed by (8) and subsequently used by (9) to explain nitrogen-limited fermentations. In this model, the biomass follows a Monod equation 9 depending on the active biomass and the dynamics of cell mass includes a death rate proportional to the extracellular ethanol concentration.

Modeling fermentation rate and production of extracellular metabolites. Regarding the fermentation rate, we proposed two models with different underlying hypotheses:

- M_P : assuming that fermentation rate is proportional to the total protein content (X_P),
- M_X : assuming that fermentation rate is proportional to the total biomass content.

The uptake of glucose reads as follows:

$$M_P : \quad \frac{dGl_x}{dt} = -lag(t) \cdot X_P \cdot v_{Gl_x} \quad (10)$$

$$M_X : \quad \frac{dGl_x}{dt} = -lag(t) \cdot X \cdot v_{Gl_x} \quad (11)$$

where v_{Gl_x} is the expression governing glucose transport.

Similarly to (27), we used Micheaelis-Menten type kinetics kinetics coupled to ethanol inhibition to model glucose transport:

$$v_{Gl_x} = v_{max,Gl_x} \cdot \frac{Gl_x}{Gl_x + k_{Gl_x}} \phi_{Eth} \quad (12)$$

where v_{max,Gl_x} is the maximum rate of glucose transport, k_{Gl_x} is the Michaelis-Menten constant and ϕ_{Eth} is competitive inhibition induced by ethanol (28) :

$$\phi_{Eth} = \frac{1}{1 + Eth/k_{E_i}} \quad (13)$$

where Eth is the extracellular ethanol concentration and k_{E_i} defines the strength of the inhibitory effect.

The excretion of extracellular metabolites was assumed to be proportional to the consumption of glucose. In the case of ethanol, this is described by:

$$\frac{dEth}{dt} = -k_{Eth} \cdot \frac{dGl_x}{dt} \quad (14)$$

where k_{Eth} controls the quantity of consumed glucose which is converted into ethanol.

Monod law and nitrogen-limited yeast fermentations

Model identifier	Model characteristics
M_P	Biomass is split into two terms, due to nitrogen and carbohydrates Lag-phase Fermentation rate is proportional to the total protein content 12 unknown parameters
M_X	Biomass is split into two terms, due to nitrogen and carbohydrates Lag-phase Fermentation rate is proportional to the biomass 12 unknown parameters
M_M	Biomass follows Monod, nitrogen the limiting substrate Lag-phase Fermentation rate is proportional to the total protein content 8 unknown parameters
M_C	Biomass is split into total and active Lag-phase (not included in the original formulation) Fermentation rate is proportional to the active biomass 8 unknown parameters

TABLE 1 Main characteristics of candidate models.

Calibration of candidate models. The Table 1 summarizes the main characteristics of the four candidate models.

The structural identifiability analysis of the models revealed that it is possible to uniquely estimate parameter values provided the dynamics of biomass, glucose uptake, nitrogen sources uptake and ethanol are measured throughout the fermentation. Remarkably, it is not necessary to measure protein for the identification of the model M_P .

Candidate models were calibrated by data fitting using data-sets available in the literature. In particular we used data from Schulze et al. (13) and Coleman et al. (9). The protein and carbohydrate contents were estimated from the data provided by Schulze et al. (13) and kept fixed during parameter estimation.

We performed 10 runs of the parameter estimation problem for each case, so to guarantee convergence to the best possible solution. Since we used a metaheuristic for the optimisation of parameters, we obtained a distribution of values. We selected the best fit for ulterior analysis. Table 2 presents the value of the corrected Akaike information criterion for each experiment and model. Results show that, on average, the models proposed in this work, M_P and M_X , are significantly better than existing models based on the Monod equation. Only in experiments 4 and 5 models M_P , M_X and M_M can not really be distinguished, as their re-scaled Akaike Δ values are within 4 units difference.

Experiment identifier	AIC_c M_P	AIC_c M_X	AIC_c M_M	AIC_c M_C
Exp1	-1314	-1262	-828	-822
Exp2	-294	-296	-278	-264
Exp3	-363	-376	-331	-330
Exp4	-408	-409	-411	-405
Exp5	-227	-226	-231	-225
Mean	-521	-513	-416	-409

TABLE 2 Corrected Akaike information criterion as computed for the different models and experiments.

Since the value obtained for the corrected Akaike information does not provide a measure of goodness-of-fit, Figure 1 presents the quality of fit of the different models for the available experiments in terms of the normalised root mean squared error (NMRSE). Figure 1 shows that all models could explain the data within

178 a 10% normalised root mean squared error. However, the performance of the different models differed sub-
 179 stantially. Models proposed in this work, M_P and M_X , which explain biomass growth using two terms due to
 180 nitrogen and carbohydrates, were the best in all experiments; M_P outperforms M_X in 3 out of 5 experiments.
 181 NMSRE values for model M_P were within the range 2.42% and 5.72% (obtained for the first and second ex-
 182 periments, respectively); for the model M_X , the range was 2.69% to 5.86% (obtained for the first and second
 183 experiments, respectively); for model M_M the range was 3.21% and 8.53% (obtained for the fourth and the first
 184 experiments respectively); for model M_C the range was 3.49% and 9.32% (obtained for the fourth and the first
 185 experiments respectively).

FIG 1 Comparison of candidate models. A) Normalised root-mean-square error for candidate models and all experiments considered. The models proposed in this work, which consider biomass growth split in two phases, due to nitrogen and carbohydrates, are the best in all experiments.

186 **Biomass formation.** Our results suggest that Monod law is insufficient to explain biomass formation ki-
 187 netics in nitrogen-limited fermentations of *S. cerevisiae*.

188 Figure 2.A) presents the normalised mean squared error obtained for the biomass for all data sets and
 189 models M_P (proposed here), M_M (Monod) and M_C ((9)). Results illustrate the proposed model M_P repre-
 190 sents better the data in all cases, even in normal nitrogen conditions. We found the most important differ-
 191 ence in the quality of fitting for the (13) data. Remarkably the normalised root-mean-square error obtained
 192 for the proposed model ($NMSRE(M_P) : 3.04\%$) is 7.1 times lower than the one obtained with the Monod
 193 model ($NMSRE(M_M) : 21.49\%$) and 6.6 times lower than that obtained with the model by Coleman et al.
 194 ($NMSRE(M_C) : 20.11\%$). For the experiments in (9), differences are lower; still, the model proposed in this
 195 work was always superior in the quality of fit to the biomass data. The improvement was more notorious in
 196 those cases using low nitrogen media. Figure 2.B) presents the fit for three out of five experiments. In all three
 197 cases, the proposed model could recover the increase of biomass produced after YAN depletion. However,
 198 in experiment 5, performed at 35°C and with high-sugar and low-ethanol, a slight decrease of the biomass is
 199 observed in the final data; the model tries to fit the tendency up but also that final data; thus it performs only
 200 slightly better than the Monod model.

201 Figure 2.C) shows how after nitrogen depletion, at around 20 min (see Figure 3.A), carbohydrates accumu-
 202 late and would explain the role of μ_2 in models M_P and M_X in recovering the secondary growth observed in
 203 the biomass dynamics.

204 **Fermentation rate.** Both models, M_P and M_X , were able to recover the overall dynamics of YAN and
 205 glucose consumption, and ethanol formation. Remarkably the performance of the model M_X , which uses the

Monod law and nitrogen-limited yeast fermentations

FIG 2 Comparison of candidate models in the fitting of biomass data. A) Normalised root-mean-square error for candidate models and all experiments considered. B) Illustrative examples of the various fit to the data. c) Prediction of the total protein, mRNA and carbohydrates percentage for experiment 1 (in which data are available). Model M_P , proposed here, performs better in explaining the biomass as measured by the NMSRE. The model can recover the increase of biomass observed in several experiments after YAN depletion. This increase is significant in experiment 1; while in experiment 3, the increase is followed by a final decrease which is not recovered by the model. Note, however, that the quality of fit is still better even for that particular experiment.

biomass instead of the protein to explain the fermentation rate is slightly worse than the performance of the model M_P but for the experiment 3 in which normal sugar and low nitrogen were used and experiment 4 in which sugar and nitrogen were normal. Remarkably the values of the NMRSE in experiment 4 were practically the same for both models.

Figure 3 shows the trajectories for all four candidate models and the corresponding experimental data for

external compounds: YAN and glucose consumption plus ethanol production, as measured in three out of five experiments. The figure concentrates in those three experiments for which the differences between the best and the worst models are higher (namely Exp1, Exp2 and Exp5).

Results show that all models can explain the data with $R^2 > 0.9$ for most cases. Different models presented different performance; in general, M_P and M_X could successfully recover all data with $R^2 > 0.96$. M_M and M_C showed higher discrepancies with data. Remarkably M_C was the best to recover the final glucose consumption and ethanol production in experiment 5, which corresponds to a stuck fermentation.

FIG 3 Comparison of candidate models in the fitting of external compounds. A) Experiment 1; B) Experiment 2 and C) Experiment 5. R^2 values are reported for each observable. Note that all models were quite successful in fitting the experimental data. Models M_M and M_C experienced some problems in fitting YAN data; while all models, but M_C , found some problems in fitting the data corresponding to a stuck fermentation.

218 DISCUSSION

219 In this study, we proposed two candidate models to explain nitrogen-limited fermentations of *S. cerevisiae*.
220 We proposed two modifications to the standard modelling approaches: 1) the biomass growth accounts for
221 protein and carbohydrates, and 2) fermentation rate is proportional to the total protein content. We reconciled
222 candidate models to previously published data.

223 Our results revealed that Monod law is insufficient to explain biomass formation kinetics in nitrogen-
224 limited fermentations of *S. cerevisiae*. This is so because, during exponential growth, the biomass composition
225 seems to remain unaltered with mRNA and protein comprising an important percentage of the biomass. How-
226 ever, this is not the case during secondary growth phase where μ_1 fades and μ_2 increases, as shown in (13)
227 data.

228 Several recent studies on the modelling of alcoholic fermentation of *S. cerevisiae* relied on logistic growth
229 (20, 12), Monod or the use of kinetic constraints coupled to a stoichiometric model (which in practice holds
230 similar results to Monod).

231 Logistic growth is well suited to explain cell numbers in nitrogen-limited fermentations. However, in light
232 of the current results, we argue the same might not be valid in the case of biomass formation. A logistic curve
233 model might not adequately describe the delay during the transition from exponential growth ($\mu_N \gg \mu_C$) to
234 carbohydrate accumulation ($\mu_C \gg \mu_N$) and the different dynamics between μ_N and μ_C observed in cell mass
235 data.

236 From a practical point of view, nitrogen-limited Monod models seem to explain biomass formation until
237 the nitrogen source is depleted but struggle to explain later stages. The models proposed by Cramer et al.
238 (8), Coleman et al. (9) distinguished active from inactive cell mass and used the methylene blue method to de-
239 tect metabolic activity. Their experimental data shows that total and active biomass are practically equivalent
240 until depletion of ammonia. After that, total biomass increases while active biomass stays stable or decreases.
241 Because the model included Michaelis-Menten type expressions without ethanol inhibition, the excessive bur-
242 den is put in the biomass decay - considered to be ethanol dependent. The authors observed that their model
243 showed some discrepancies in the fits for cell mass. These discrepancies were more relevant at lower temper-
244 atures, for which the model predicted that cell activity decreased more rapidly than shown by the data. The
245 models proposed in this work, outperformed Coleman et al. (9) results, but for the prediction of glucose and
246 ethanol consumption in stuck fermentations. Our model could be further refined by introducing the concept
247 of cellular decay to improve the accuracy in the prediction of stuck fermentations.

248 Malherbe et al. (20) developed a fermentation model for enological conditions that accounted for different
249 temperatures and initial nitrogen concentrations. These authors estimated the number of hexose transporters
250 present in the yeast cells as a function of time, temperature and nitrogen. The model represents glucose con-
251 sumption data extremely well. However, it should be noted that it comes at the cost of a heavily parameterised
252 and nonlinear hexose transport function that also includes inhibition by substrate and ethanol. Nevertheless,
253 the notion that hexose transporters vary with nitrogen content along the fermentation is likely to find some
254 overlap with our simpler model for glucose transport.

255 Pizarro et al. (17) included the effect of ethanol inhibition on sugar transport and the existence of different
256 hexose transporters with different maximum transport rates and substrate affinities. From a qualitative point
257 of view, this model corresponded to a significant improvement over past models. Noticeably, the hexose trans-
258 port functions considered the role of the initial nitrogen concentration and temperature. Additionally, Varela
259 et al. (16) showed that the protein content of biomass is heavily dependent on the initial nitrogen concentration
260 which supports the notion that hexose transport is likely to be better described by total protein content rather
261 than biomass *per se*. Our model would support this hypothesis. Results obtained by M_P model are superior to
262 those obtained with the M_X model, at least, for those cases in which initial nitrogen is restricted.

263 Here we proposed an alternative modelling strategy which using a minimal number of parameters, that can
264 be estimated from easily measured data, provides a considerable improvement in the description of nitrogen-
265 limited fermentations as compared to previously published models. Other mechanisms, such as the role of
266 ethanol in cellular decay, can be easily incorporated. The model brings a new possibility to anticipate, and thus
267 control, problematic fermentations.

268 MATERIALS AND METHODS

269 **Data** We considered data-sets from five different experiments as published in the literature. The first
270 experiment (Exp1) data set was obtained from Schulze et al. (13). The authors conducted a nitrogen-limited
271 (165 mg NH₃ and 50g of glucose) batch fermentation while measuring biomass, protein content of biomass
272 along with several extracellular metabolites (glycerol, ethanol and ammonia). We also considered four data-
273 sets from Coleman et al. (9). The authors performed various experiments at different conditions. In this work
274 we denote Exp2 as that performed at 11°C, with high-sugar and low-nitrogen; Exp3, the one performed at 15°C,
275 with normal sugar and low-nitrogen; Exp4, the one performed at 30°C with normal sugar and normal nitrogen;
276 and Exp5, the one performed at 35°C with high-sugar and low-nitrogen.

277 **Model building** The proposed model is composed of six ordinary differential equations (as presented in
278 the section Results). Solutions for the system are of type:

$$279 \quad y = \int_0^{t_F} y(\theta, y, t) dt \quad (15)$$

280 where y is the solution of the ODE system and corresponds to the dynamics of the relevant state variables -
281 biomass, carbohydrates, protein, glucose, YAN, ethanol, glycerol- between the beginning and the end of the
282 experiment (t_F); y represents the states time derivative.

283 **Structural identifiability analysis** The proposed model depends on fourteen unknown parameters θ -
284 namely $a_0, \mu_{maxN}, k_{sN}, \mu_{maxC}, k_{sC}, \theta_C, \lambda_C, \lambda_P, Y_{X/N}, v_{max,Gl}, k_{Gl}, k_{Ei}, k_{Eth}, k_{Gly}$ - to be estimated from the
285 experimental data. The proposed model introduces the description of the total protein content. However
286 for some experimental set-ups protein was not measured. Therefore we performed a structural identifiability
287 analysis ((29, 30)) to assess if all parameters of the model can be estimated even if protein measurements are
288 not available.

289 We performed the analysis using GenSSI2 toolbox (31). The toolbox uses symbolic manipulation to imple-
290 ment the so-called generating series of the model and to compute identifiability *tableaus*.

291 **Parameter estimation** The aim of parameter estimation is to compute the unknown parameters - growth
292 related constants and kinetic parameters - that minimize the weighted least squares function which provides
293 a measure of the distance between data and model predictions (?). In the absence of error estimates for
294 the experimental data a fixed value for each state, the maximum experimental data, was assumed in order to
295 normalize the least squares residuals:

$$296 \quad J_{mc}(\theta) = \sum_{k=1}^{n_{exp}} \sum_{j=1}^{n_{obs}} \sum_{i=1}^{n_{st}} \left(\frac{y_{k,j,i}(\theta) - y_{k,j,i}^m}{\max(y_{k,j})} \right)^2, \quad (16)$$

297 where n_{exp} , n_{obs} and n_{st} are, respectively, the number of experiments, observables, and sampling times. y_j^m
298 represents each of the measured quantities and $y_j(\theta)$ corresponds to model predicted values.

299 In our particular case, we used five different experiments: the first, taken from (13), and second to fifth,
300 taken from (9). All experiments provided time series data of the biomass (X), glucose uptake (G/x), ammonia
301 consumption (N) and ethanol production Eth .

302 Parameters were estimated by solving a nonlinear optimization problem to find the unknown parameter
303 values (θ) to minimize $J_{mc}(\theta)$, subject to the system dynamics - the model- and parameter bounds (32).

304 To avoid the risk of premature convergence by the optimization routine while searching for the optimal
305 parameter set, the parameter estimation procedure was repeated 10 times for each candidate model starting
306 from different initial guesses.

307 **Model selection** Models were compared in terms of the Akaike's information criterion (AIC). AIC compares
308 multiple competing models (or working hypotheses) in those cases where no single model stands out as being
309 the best. AIC is calculated using the number of data (n_d), the number fitted parameters (n_θ) and the sum of
310 squares of the weighted residuals (WRSS). For the cases in which the number of data is low relative to the
311 number of parameters (roughly $n_d/n_\theta < 40$ for the most complex model), the corrected Akaike's information

Monod law and nitrogen-limited yeast fermentations

312 criterion is computed as follows:

$$313 \quad AIC_c = n_d \log(WRSS) + 2n_\theta + 2n_\theta \frac{2n_\theta(n_\theta + 1)}{(n_d - 2 + n_\theta - 1)} \quad (17)$$

314 The models were ranked by AIC_c , with the best approximating model being the one with the lowest AIC_c
315 value. The minimum AIC value (AIC_{min}) was used to re-scale the Akaike's information criterion. The re-scaled
316 value $\Delta = |AIC_c - AIC_{min}|$ was used to assess the relative merit of each model. Models such as $\Delta \leq 2$ have
317 substantial support, models for which $4 \leq \Delta \leq 7$ have considerably less support and models with $\Delta > 10$ have
318 no support (33).

319 **Numerical methods and tools** The parameter estimation and model selection were implemented in the
320 AMIGO2 toolbox (34). The system of ODEs was compiled to speed up calculations and solved using the CVODES
321 solver (35), a variable-step, variable-order Adams-Bashforth-Moulton method. The parameter estimation prob-
322 lem was solved using a hybrid meta-heuristic, the enhanced scatter search method (eSS, (36)).

323 SUPPLEMENTAL MATERIAL

324 All scripts and data required to reproduce results and figures can be accessed in [https://sites.google.com/site/](https://sites.google.com/site/amigo2toolbox/examples)
325 [amigo2toolbox/examples](https://sites.google.com/site/amigo2toolbox/examples).

326 ACKNOWLEDGMENTS

327 This project has received funding from MCIU/AEI/FEDER, UE (grant reference RTI2018-093744-B-C33) and GAIN
328 Xunta de Galicia (Contrato Programa and grant reference IN607B 2020-03).

329 REFERENCES

1. Parapouli M, Vasileiadis A, Afendra AS, Hatziloukas E. 2020. *Saccharomyces cerevisiae* and its industrial applications. AIMS Microbiol 6 (1):1–31.
2. Bisson LF. 1999. Stuck and sluggish fermentations. Am. J. Enol. Viticult. 50 (1):107–119.
3. Alexandre H, Charpentier C. 1998. Biochemical aspects of stuck and sluggish fermentation in grape must. J Ind Microbiol Biotech 20:20–27.
4. Tesnière C, Brice C, Blondin B. 2015. Responses of *Saccharomyces cerevisiae* to nitrogen starvation in wine alcoholic fermentation. Appl. Microbiol. Biotechnol. 99 (17):7025–7034.
5. Bell SJ, Henschke PA. 2005. Implications of nitrogen nutrition for grapes, fermentation and wine. Aust. J. Grape Wine Res. 11 (3):242–295.
6. Marín MR. 1999. Alcoholic fermentation modelling: Current state and perspectives. Am. J. Enol. Viticult. 50 (2):166–178.
7. Miller KV, Block DE. 2020. A review of wine fermentation process modeling. J. Food Eng. 273:109783.
8. Cramer AC, Vlassides S, Block DE. 2002. Kinetic model for nitrogen-limited wine fermentations. Biotechnol. Bioeng. 77 (1):49–60.
9. Coleman MC, Fish R, Block DE. 2007. Temperature-dependent kinetic model for nitrogen-limited wine fermentations. Appl. Environ. Microbiol. 73 (18):5875–5884.
10. Kelkar S, Dolan K. 2012. Modeling the effects of initial nitrogen content and temperature on fermentation kinetics of hard cider. J. Food Eng. 109 (3):588–596.
11. David R, Dochain D, Mouret JR, Wouwer AV, Sablayrolles JM. 2010. Dynamical modeling of alcoholic fermentation and its link with nitrogen consumption. IFAC Proc Vol 43 (6):496–501.
12. Henriques D, Alonso-del Real J, Querol A, Balsa-Canto E. 2018. *Saccharomyces cerevisiae* and *S. kudriavzevii* synthetic wine fermentation performance dissected by predictive modeling. Front. Microbiol. 9:88.
13. Schulze U, Liden G, Villadsen J. 1996. Dynamics of ammonia uptake in nitrogen limited anaerobic cultures of *Saccharomyces cerevisiae*. J Biotechnol. 46 (1):33–42.
14. Jørgensen H, Olsson L, Rønnow B, Palmqvist E. 2002. Fed-batch cultivation of baker's yeast followed by nitrogen or carbon starvation: effects on fermentative capacity and content of trehalose and glycogen. Appl. Microbiol. Biotechnol. 59 (2-3):310–317.
15. Thomsson E, Gustafsson L, Larsson C. 2005. Starvation response of *Saccharomyces cerevisiae* grown in anaerobic nitrogen-or carbon-limited chemostat cultures. Appl. Environ. Microbiol. 71 (6):3007–3013.
16. Varela C, Pizarro F, Agosin E. 2004. Biomass content governs fermentation rate in nitrogen-deficient wine musts. Appl. Environ. Microbiol. 70 (6):3392–3400.
17. Pizarro F, Varela C, Martabit C, Bruno C, Pérez-Correa JR, Agosin E. 2007. Coupling kinetic expressions and metabolic networks for predicting wine fermentations. Biotechnol. Bioeng. 98 (5):986–998.
18. Vargas FA, Pizarro F, Pérez-Correa JR, Agosin E. 2011. Expanding a dynamic flux balance model of yeast fermentation to genome-scale. BMC Syst. Biol. 5 (1):75.
19. Sainz J, Pizarro F, Pérez-Correa JR, Agosin E. 2003. Modeling of yeast metabolism and process dynamics in batch fermentation. Biotechnol. Bioeng. 81 (7):818–828.
20. Malherbe S, Fromion V, Hilgert N, Sablayrolles JM. 2004. Modeling the effects of assimilable nitrogen and temperature on fermentation kinetics in enological conditions. Biotechnol. Bioeng. 86 (3):261–272.

21. **Charnomordic B, David R, Dochain D, Hilgert N, Mouret JR, Sablayrolles JM, Vande Wouwer A.** 2010. Two modelling approaches of winemaking: first principle and Metab. Eng. Math. Comput Model Dyn. Syst. 16 (6):535–553.
22. **David R, Dochain D, Mouret JR, Wouwer AV, Sablayrolles JM.** 2014. Nitrogen-backed modeling of wine-making in standard and nitrogen-added fermentations. Bioproc. Biosyst. Eng. 37 (1):5–16.
23. **Mouret JR, Farines V, Sablayrolles JM, Trelea IC.** 2015. Prediction of the production kinetics of the main fermentative aromas in winemaking fermentations. Biochem Eng. J. 103:211–218.
24. **Perez M, Luyten K, Michel R, Riou C, Blondin B.** 2005. Analysis of *Saccharomyces cerevisiae* hexose carrier expression during wine fermentation: both low-and high-affinity Hxt transporters are expressed. FEMS Yeast Res. 5 (4-5):351–361.
25. **Monod J.** 1949. The growth of bacterial cultures. Annu Rev. Microbiol. 3 (1):371–394.
26. **Baranyi J, Roberts TA.** 1994. A dynamic approach to predicting bacterial growth in food. Int. J. Food. Microbiol. 23 (3-4):277–294.
27. **Hjersted JL, Henson MA, Mahadevan R.** 2007. Genome-scale analysis of *Saccharomyces cerevisiae* metabolism and ethanol production in fed-batch culture. Biotechnol. Bioeng. 97 (5):1190–1204.
28. **Leão C, Van Uden N.** 1984. Effects of ethanol and other alkanols on passive proton influx in the yeast *Saccharomyces cerevisiae*. Biochim. Biophys. Acta Biomembr. 774 (1):43–48.
29. **Chis OT, Banga JR, Balsa-Canto E.** 2011. Structural identifiability of systems biology models: A critical comparison of methods. PLoS ONE 6 (11).
30. **Chis OT, Villaverde AF, Banga JR, Balsa-Canto E.** 2016. On the relationship between sloppiness and identifiability. Math Biosci 282:147–161.
31. **Ligon TS, Fröhlich F, Chiş OT, Banga JR, Balsa-Canto E, Hasenauer J.** 2018. GenSSI 2.0: Multi-experiment structural identifiability analysis of SBML models. Bioinformatics 34 (8):1421–1423.
32. **Vilas C, Arias-Méndez A, García MR, Alonso AA, Balsa-Canto E.** 2018. Toward predictive food process models: a protocol for parameter estimation. Crit. Rev. Food Sci. Nutr. 58 (3):436–449.
33. **Burnham KP, Anderson R.** 2004. Multimodel Inference: Understanding AIC and BIC in Model Selection. Sociol. Meth. & Res. 33:261–304.
34. **Balsa-Canto E, Henriques D, Gábor A, Banga JR.** 2016. AMIGO2, a toolbox for dynamic modeling, optimization and control in systems biology. Bioinformatics 32 (21):3357–3359.
35. **Hindmarsh AC, Brown PN, Grant KE, Lee SL, Serban R, Shumaker DE, Woodward CS.** 2005. SUNDIALS: Suite of nonlinear and differential/algebraic equation solvers. ACM Trans Math Softw.(TOMS) 31 (3):363–396.
36. **Egea JA, Balsa-Canto E, García MSG, Banga JR.** 2009. Dynamic optimization of nonlinear processes with an enhanced scatter search method. Ind. Eng. Chem. Res. 48 (9):4388–4401.