frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 31 October 2023
DOI 10.3389/fpls.2023.1183254

Check for updates

OPEN ACCESS

EDITED BY
Friedrich Altmann,
University of Natural Resources and Life Sciences Vienna, Austria

REVIEWED BY
Smita Srivastava,
Indian Institute of Technology Madras, India
Steffen Waldherr,
University of Vienna, Austria

*CORRESPONDENCE
Johannes Felix Buyel
johannes.buyel@rwth-aachen.de

†These authors have contributed equally to this work

RECEIVED 09 March 2023
ACCEPTED 27 September 2023
PUBLISHED 31 October 2023

CITATION
Nausch H, Baldan M, Teichert K, Lutz J, Claussen C, Bortz M and Buyel JF (2023) Simulation and optimization of nutrient uptake and biomass formation using a multi-parameter Monod-type model of tobacco BY-2 cell suspension cultures in a stirred-tank bioreactor.
Front. Plant Sci. 14:1183254.
doi: 10.3389/fpls.2023.1183254

COPYRIGHT
© 2023 Nausch, Baldan, Teichert, Lutz, Claussen, Bortz and Buyel. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Simulation and optimization of nutrient uptake and biomass formation using a multi-parameter Monod-type model of tobacco BY-2 cell suspension cultures in a stirred-tank bioreactor

Henrik Nausch [1†], Marco Baldan [2†], Katrin Teichert [2], Jannik Lutz [1], Carsten Claussen [3], Michael Bortz [2] and Johannes Felix Buyel [1,4,5*]

[1]Department Bioprocess Engineering, Fraunhofer Institute for Molecular Biology and Applied Ecology IME, Aachen, Germany, [2]Division Optimization, Fraunhofer Institute for Industrial Mathematics ITWM, Kaiserslautern, Germany, [3]Fraunhofer Institute for Translational Medicine and Pharmacology ITMP, Hamburg, Germany, [4]Institute for Molecular Biotechnology, RWTH Aachen University, Aachen, Germany, [5]Institute of Bioprocess Science and Engineering (IBSE), University of Natural Resources and Life Sciences, Vienna (BOKU), Vienna, Austria

Introduction: Tobacco (*Nicotiana tabacum*) cv Bright Yellow-2 (BY-2) cell suspension cultures enable the rapid production of complex protein-based biopharmaceuticals but currently achieve low volumetric productivity due to slow biomass formation. The biomass yield can be improved with tailored media, which can be designed either by laborious trial-and-error experiments or systematic, rational design using mechanistic models, linking nutrient consumption and biomass formation.

Methods: Here we developed an iterative experiment-modeling-optimization workflow to gradually refine such a model and its predictions, based on collected data concerning BY-2 cell macronutrient consumption (sucrose, ammonium, nitrate and phosphate) and biomass formation.

Results and discussion: The biomass formation was well predicted by an unstructured segregated mechanistic Monod-type model as long as the nutrient concentrations did not approach zero (we omitted phosphate, which was completely depleted). Multi-criteria optimization for sucrose and biomass formation indicated the best tradeoff (in a Paretian sense) between maximum biomass yield and minimum process time by reducing the initial sucrose concentration, whereas the inoculation biomass could be increased to maximize the biomass yield or minimize the process time, which we

confirmed in calibration experiments. The model became inaccurate at biomass densities > 8 g L^{-1} dry mass when sucrose was almost depleted. We compensated for this limitation by including glucose and fructose as sucrose hydrolysis products in the model. The remaining offset between the simulation and experimental data might be resolved by including intracellular pools of sucrose, ammonium, nitrate and phosphate. Overall, we demonstrated that iterative models can be used to systematically optimize conditions for bioreactor-based processes.

KEYWORDS

biopharmaceuticals, cultivation medium, mechanistic model, multi-criteria optimization, upstream production

Introduction

Plant cell cultures (PCCs) can be used to produce complex protein-based biopharmaceuticals such as growth factors, cytokines or antibodies, especially proteins that are unsuitable for expression in microbial or mammalian cells (Xu and Zhang, 2014; Santos et al., 2016). PCCs from carrot (*Daucus carrota*), rice (*Oryza sativa*) and tobacco (*Nicotiana tabacum*), in particular tobacco cell lines NT-1 and BY-2, are often used for this purpose (Xu et al., 2011; Xu and Zhang, 2014; Santos et al., 2016; Moon et al., 2020). Tobacco cells have short doubling times of 25–30 h compared to 50–500 h for other PCCs and can be grown to high cell densities of up to 300–600 g L^{-1} fresh mass (FM) in large-scale stirred-tank reactors (STRs) with working volumes of up to 100,000 L (Fischer et al., 1999). However, PCCs generally have a low productivity of up to 1 g L^{-1} of protein, compared to 5–20 g L^{-1} for microbial and mammalian cells (Schillberg and Spiegel, 2022).

One option to increase the productivity of PCCs is to optimize the growth medium (Terashima et al., 2001; Holland et al., 2010; Holland, 2013; Vasilev et al., 2013; Häkkinen et al., 2018; Sadoch et al., 2020). PCCs are typically cultivated in chemically defined media containing sucrose ($C_{12}H_{22}O_{11}$) as a carbon source, ammonium nitrate (NH_4NO_3) and/or potassium nitrate (KNO_3) as nitrogen sources, and potassium dihydrogen phosphate (KH_2PO_4) as a phosphate source. The feeding strategy also affects the volumetric productivity. For example, varying nitrogen sources in trial-and-error screening experiments led to recipes such as Murashige and Skoog (MS) medium (Murashige and Skoog, 1962), Gamborg B5 (Gamborg et al., 1968), Chu N6 (Chu et al., 1975), Schenk and Hildebrandt (SH) medium (Schenk and Hildebrandt, 1972), and amino acid (AA) medium (Thompson et al., 1986). More recently, carbon, nitrogen and phosphate sources have been optimized more systematically using a Design of Experiments (DoE) approach, confirming the potential to increase biomass formation and volumetric productivity of protein-based biopharmaceuticals. These studies also demonstrated that mineral salts, vitamins and plant growth regulators, including substances referred to as phytohormones, are required for the growth of PCCs, but have only a minor impact on biomass formation and biopharmaceutical production (Holland, 2013; Vasilev et al., 2013). Importantly, the experiments were conducted in shake flasks and the transferability of such conditions to STRs has yet to be demonstrated.

Model-based optimization is another systematic option to improve media composition (Schinn et al., 2021; Yeo et al., 2022). Model-based DoE methods attempt to find new sets of experiments that improve the model based on prior knowledge about the system and a small number of initial tests (Seufert et al., 2021). In general, there are three modeling strategies: data-driven (black box), mechanistic (white box), and hybrid (gray box). The mechanistic modeling are based on physical, chemical and biological principles that describe the interdependency of parameters such as nutrient uptake, biomass formation and target protein production, thus enabling robust bioprocess engineering (Tsopanoglou and Del Jiménez Val, 2021). Additionally, mechanistic models can be described as structured or unstructured. Whereas unstructured models consider nutrient (substrate) uptake from the medium, biomass formation and protein production, structured models also include the intracellular metabolism of nutrients. Compared to structured models, unstructured models require a limited number of equations and parameters, making them less prone to overfitting when presented with limited input data (Hawkins, 2004). Both unstructured and structured models can be either unsegregated, which assumes the cell population is homogeneous, or segregated, which differentiates cells into viable and dead cells or differentiates by nutrient uptake, biomass formation, and protein production.

A mechanistic, kinetic modeling approach is preferred for PCCs, where all metabolic effects are lumped into unstructured kinetic functions for simplification (Resat et al., 2009), especially for model-based DoE where the number of experiments is small and hence only few model parameters can be fitted. In particular, Monod-type models are preferred because they simulate nutrient uptake, biomass formation and protein production while differentiating between viable and dead cells (Prakash and Srivastava, 2006; Prakash and Srivastava, 2008; Jiménez-Hornero et al., 2009a; Jiménez-Hornero et al., 2009b; Puad et al., 2017; Villegas et al., 2017; Puad and Abdullah, 2018; Jacob et al., 2020).

Model-based optimization requires an iterative workflow for process characterization and model setup followed by process and

model improvement in three steps: experiment, modeling, and optimization (Figure 1). A similar workflow has been applied in chemical engineering (Höller et al., 2019; Asprion et al., 2022). In the first iteration, initial experiments are carried out to collect starting data, e.g., a few experiments to characterize nutrient consumption, cell growth and biomass formation. Then a model is set up to identify relevant parameters (e.g., concentration of specific nutrients), and the model is used to predict improved/optimal process conditions, e.g., altered initial nutrient concentrations (Table S1). In the second iteration, these optimized process conditions are experimentally verified and used to update the model calibration and/or process for further model improvement (Table S1). Both, the process and model, can then be refined in subsequent iterations until the desired model fidelity or process improvement is achieved.

Here we applied an unstructured segregated Monod-type model to describe the nutrient uptake and biomass formation of tobacco BY-2 cells as model PCC in 2-L and 5-L STRs as part of a semi-continuous fermentation process (Patent WO2015165583A1). We used the model to maximize the biomass yield and minimize the process time by medium optimization. The purpose of the study was to demonstrate that relevant bioprocess optimization is possible using a very small number of experiments in combination with corresponding models.

Materials and methods

Cultivation of tobacco cells

Tobacco (*N. tabacum* BY-2) cells were grown in either 100-mL or 1000-mL Erlenmeyer shake flasks containing 20 mL or 200 mL of modified MS medium as previously described (Murashige and Skoog, 1962; Rademacher et al., 2019). Flasks were placed on a Climo-Shaker ISF1-X orbital shaker (Kuhner Shaker, Herzogenrath, Germany) at 26°C, shaking at 160 rpm with a displacement of 25 mm. The BY-2 cells were transferred to fresh MS medium every 7 days using an inoculation cell density of 20 g L^{-1} FM.

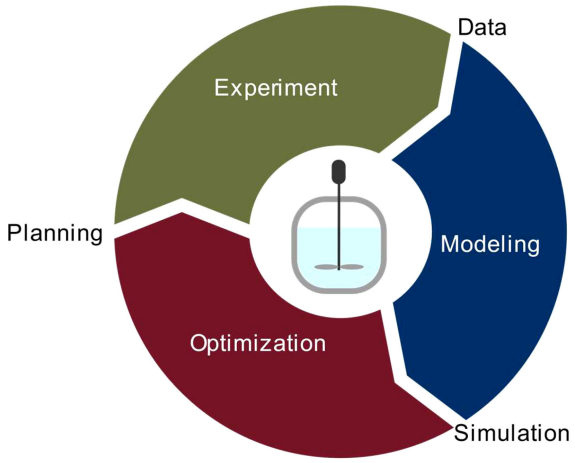


FIGURE 1
Iterative experiment-model-optimization workflow. In the first iteration, initial experiments were carried out to collect starting data (e.g., for nutrient consumption, cell growth and biomass formation), a model was set up in order to identify relevant parameters (e.g., nutrients), and the model, once embedded in an optimization frame, was used to predict optimal process conditions (e.g., initial nutrient concentrations). In the second iteration, these optimized process conditions were experimentally verified and used to update the model calibration and/or process and for model improvement.

The 7-day-old cells were used to inoculate MS medium in 2-L or 5-L STRs (Getinge Deutschland, Rastatt, Germany) (Tables S2, S3). The double-walled glass STRs were equipped with dissolved oxygen (dO_2), capacitance, and pH probes, and contained a porous sparger for aeration. Two axial marine and two radial flat-blade impellers were used for mixing. The bottom marine impeller was installed at a height corresponding to the mid-level of the minimum culture volume, whereas the second marine impeller was placed at the corresponding height of the maximum culture volume (Figure S1). The bottom flat-blade impeller was installed equidistant in between the two marine impellers. The second flat-blade impeller was positioned at the same distance above the top marine impeller. The dO_2 level was kept constant by controlling the stirrer speed. The pH were monitored but not controlled. Online data were recorded using the BioExpert process information management system (Getinge Deutschland). The FM data were used to calibrate the capacitance probe that controlled the feeding rate. The concentration of sucrose, glucose and fructose in the medium was adjusted for each fermentation run as described below (Table S3). The feed phase was initiated at a FM of 100 g L^{-1} and this FM concentration was maintained by adding fresh medium. In addition, interspersed draining of cultivation broth was used every 22–26 h to restore the starting volume. The feed medium was of the same composition as the batch medium but contained 59 mM (or 20 g L^{-1}) instead of 88 mM (or 30 g L^{-1}) sucrose. The feed rate was dynamically adjusted in dependence of the growth rate that was monitored via a capacitance probe. Accordingly, the drain volumes differed over time. Process samples were taken for subsequent analysis every 22–26 h (simultaneously with draining during the semi-continuous phase). A 5-L bioreactor was used in the initial experiments (#1-3) for model setup, whereas 2-L reactors were used for validation experiments (#4-9). The feeding and control strategy as well as the impeller configuration were the same in both cases.

Cell biomass and macronutrient analysis

To determine the BY-2 cell concentration, each 100-µL sample was diluted 1:10 by mixing with 1 mL 0.9% (m v^{-1}) sodium chloride containing 0.025% (m v^{-1}) Evans blue, and the cells were counted in a Fuchs-Rosenthal counting chamber. To determine the FM, DM and macronutrient levels, 5.0-mL samples were applied to cellulose filter paper in a Büchner funnel, and the medium removed by vacuum filtration at 0.08 MPa for 5 s. For the FM, the BY-2 pellet was transferred to a weighing dish and the FM was determined using a fine balance. For the dry mass (DM), the weighing dish with the BY-2 pellet was dried at 60°C for 3 days and weighed again. For the macronutrients, the flow-through, obtained from the vacuum filtration, was collected in a 15-mL Falcon tube and stored at –20°C

before analysis using commercial assay kits for sucrose (cat. MAK013; Merck, Darmstadt, Germany), glucose (cat. ABIN5067615; antibodies-online, Aachen, Germany), fructose (cat. K619-100; BioVision, Ilmenau, Germany), ammonium (cat. MAK310; Merck), nitrate (cat. Cay780001; Biomol, Hamburg, Germany), and phosphate (cat. KA0815; Abnova Germany, Heidelberg, Germany).

Model fitting and multi-criteria optimization

Calculations were performed using the dynamic programming language Julia[1]. The optimization solver was Ipopt (Wächter and Biegler, 2006). We used the default values of the termination criteria, i.e., the (scaled) non-linear problem error ($<10^{-8}$) and the (absolute) criteria according to "dual_inf_tol" (1), "constr_viol_tol" (10^{-4}), and "compl_inf_tol" (10^{-4}) [2]. The performance of different models was compared by K-fold cross-validation (Stone, 1974). Confidence intervals (CIs) were calculated based on bootstrapping (Dogan, 2007). Inequality constraints were used in the parameter identification problem to ensure that values of inhibition constants were at least equal to or more extreme than the corresponding saturation constants.

Results and discussion

Phosphate is rapidly depleted from the PCC medium whereas sucrose, ammonium and nitrate are not completely consumed

In the standard setting for the BY-2 semi-continuous fermentation (experiments #1, #2 and #3) (Figure 2 and S2, Table S3), BY-2 cells were cultivated in MS medium in 5-L STRs, starting with a cell density of 20 ± 5 g L^{-1} FM (corresponding to 0.75 ± 0.22 g L^{-1} DM) at the beginning of the batch phase, and were kept at 100 ± 20 g L^{-1} FM during the semi-continuous phase. Under these conditions, the initial sucrose concentration of 88 mM decreased to 25–30 mM at the end of the batch phase and remained at 15–30 mM during the semi-continuous phase, which correlated with the slight variation in the FM/DM ratios (Figure S2). In line with the declining concentration of sucrose, the concentration of its hydrolysis product glucose increased from 0 mM at t_0 to 15 mM at the end of the batch phase, and remained at 12–18 mM in the semi-continuous phase. This is consistent with the fact that sucrose is not only taken up by plant cells via sucrose transporters, but is also hydrolyzed into glucose and fructose externally by invertases bound to the cell wall, and these products are then taken up by hexose transporters (Lemoine et al., 2013).

Like sucrose, ammonium levels decreased from 21 to 10–12 mM at the end of the batch phase and stayed at 5–10 mM during the semi-continuous phase, correlating with the FM/DM. The decrease in ammonium may reflect its uptake by ammonium transporters and the intracellular detoxification of ammonium via the glutamine synthase/glutamine oxoglutarate aminotransferase (GS/GOGAT) cycle that produces glutamate (Bittsánszky et al., 2015). To ensure ammonium levels remain sub-toxic, other PCC media including Gamborg B5, Chu N6, SH medium and AA medium contain 2, 7, 3 and 0 mM ammonium, respectively, compared to the 21 mM present in MS medium. However, Ullisch (2012) observed that increasing the ammonium concentration in MS medium from 21 to 41 mM caused NT-1 to accumulate 20% more biomass, possibly due to the formation of more glutamate (Ullisch, 2012). Even though not all the ammonium is consumed, it may therefore be a suitable target for medium optimization when aiming to accelerate biomass formation.

The decrease in nitrate levels was much less pronounced than the other components, falling from an initial 39 mM to 25–30 mM at the end of the batch phase, which was maintained during the semi-continuous phase. This small change in nitrate levels may reflect the conversion of nitrate into ammonium by nitrate reductase and nitrite reductase, which are subject to feedback inhibition to ensure ammonium stays at sub-toxic levels (Behrend and Mateles, 1975; Behrend and Mateles, 1976). Notably, increasing nitrate levels from 39 to 139 mM reduced growth by 30%, and when nitrate was the sole nitrogen source the plant cells did not grow at all (Holland et al., 2010; Ullisch et al., 2012).

In contrast to sucrose, ammonium and nitrate, the initial 2.7 mM phosphate present in the MS medium was almost completely consumed within 3 days of batch fermentation and remained at 0.1–0.3 mM during the semi-continuous phase. This agrees with previous studies reporting that phosphate is a growth-limiting factor during the fermentation of BY-2 cells, even if the phosphate concentration is increased by 400% (Holland, 2013; Vasilev et al., 2013).

Therefore, the initial set of experiments yielded plausible data that we used for an initial model calibration.

The model for the standard cultivation conditions in STRs suggests that sucrose levels can be reduced without affecting biomass formation and volumetric biomass yield

For the standard BY-2 semi-continuous fermentation in STRs (experiments #1, #2 and #3) (Figure 2 and S2), we applied a Monod-type unstructured segregated model to describe the extracellular dynamics of cell growth/biomass formation (DM increase over time) and nutrient uptake. As proposed by Jacobs et al. (2020), we modeled two physiological states of cell DM: viable (active) (X_a) and dead (X_d). This yields the total DM concentration X [g L^{-1}] (Eq. 1).

$$X = X_a + X_d \quad (1)$$

1 https://julialang.org/

2 https://coin-or.github.io/Ipopt/OPTIONS.html

The model assumes an irreversible progression in which active cells (active DM) are converted into dead cells (dead DM) (Jacob et al., 2020) (Eq. 2).

$$V\frac{dX_a}{dt} + X_a\frac{dV}{dt} = \mu X_a V - k_d X_a V \quad (2)$$

where μ [h^{-1}] is the specific growth rate, k_d [h^{-1}] is the death rate constant of active DM, and V [L] is the volume of the cultivation medium. Based on the volumetric change of the cultivation medium, we calculated the feed rate F_t [L h^{-1}] for each time point t (Jiménez-Hornero et al., 2009a) (Eq. 3).

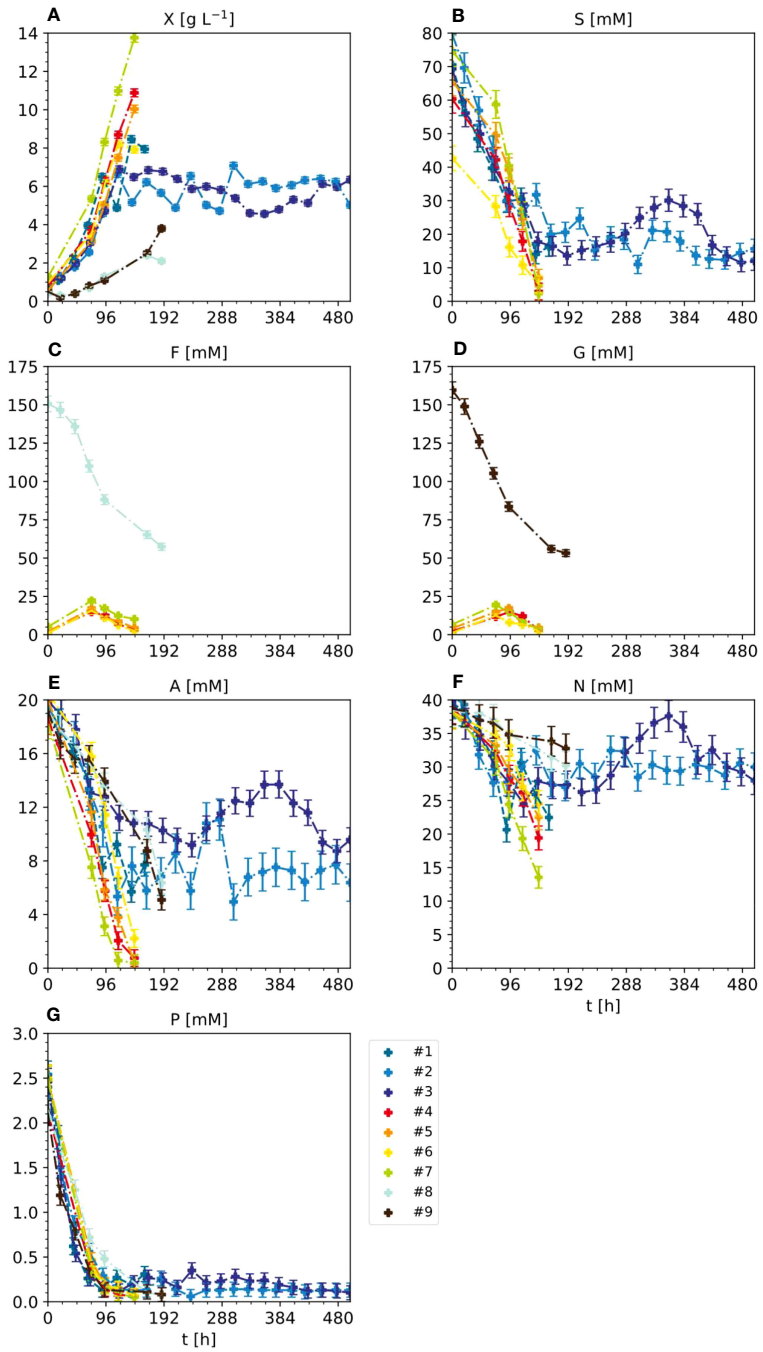


FIGURE 2
Nutrient consumption and cell growth/biomass formation under standard, optimal and non-optimal cultivation conditions. Experiments 1–3: standard cultivation conditions used for model setup (iteration 1). Experiments 4 and 5: optimal cultivation conditions according to the model-based optimization for model validation (iteration 2). Experiments 6 and 7: non-optimal cultivation conditions used for model validation (iteration 2). Experiments 8 and 9: non-optimal cultivation conditions but different carbon sources (either fructose or glucose instead of sucrose) used for model improvement (iteration 2). Measurement uncertainty made of a constant and proportional component (Table S4). Numbers – individual experiments- **(A)** Dry mass X. **(B)** Sucrose S. **(C)** Fructose **(F, D)** Glucose **(G, E)** Ammonium **(A, F)** Nitrate N. **(G)** phosphate P.

10.3389/fpls.2023.1183254

$$\frac{dV}{dt} = F_t \tag{3}$$

The dependency of the specific growth rate (biomass formation) of active DM (μ) on sucrose (S), ammonium (A) and nitrate (N) concentrations can be expressed as a Monod kinetic (Eq. 4).

$$\mu = \mu_m \left(\frac{S}{S+K_S}\frac{K_{IS}}{S+K_{IS}}\right)\left(\frac{A}{A+K_A}\frac{K_{IA}}{A+K_{IA}}\right)\left(\frac{N}{N+K_N}\frac{K_{IN}}{N+K_{IN}}\right) \tag{4}$$

where K_r is the saturation constant and K_{Ir} the inhibition constant of corresponding nutrient r, i.e. S, A and N in [g L^{-1}], and μ_m is the maximum specific growth rate [h^{-1}] (Prakash and Srivastava, 2006; Prakash and Srivastava, 2008). For each nutrient, we used inequality constraints to force the inhibition constant to be equal to or greater than the saturation constant during parameter identification to enforce actual plateaus, i.e., saturation.

The phosphate concentration was modeled but its effect was not included in the specific growth rate because it dropped below the limit of quantification after 3 days of batch cultivation. Accordingly, adding a term for the phosphate concentration in the specific growth rate (Eq. 4) would have zeroed out the DM growth after 3 days, which did not agree with the observed DM increase. This limitation of the current model might be circumvented in the future by including intracellular phosphate levels in a structured model but the determination of such pools was beyond the scope of this study.

The sucrose, ammonium, nitrate and phosphate concentrations during BY-2 cultivation were modeled in Equations 5–8.

$$V\frac{dS}{dt} + S\frac{dV}{dt} = F_t S_f - \mu_S\left(\frac{S}{S+K_S}\frac{K_{IS}}{S+K_{IS}}\right)X_a V \tag{5}$$

$$V\frac{dN}{dt} + N\frac{dV}{dt} = F_t N_f - \mu_N\left(\frac{N}{N+K_N}\frac{K_{IN}}{N+K_{IN}}\right)X_a V \tag{6}$$

$$V\frac{dA}{dt} + A\frac{dV}{dt} = F_t A_f - \mu_A\left(\frac{A}{A+K_A}\frac{K_{IA}}{A+K_{IA}}\right)X_a V \tag{7}$$

$$V\frac{dP}{dt} + P\frac{dV}{dt} = F_t P_f - \mu_P P \;\; V \tag{8}$$

where S_f, N_f, A_f and P_f are the feed medium concentrations of sucrose, nitrate, ammonium and phosphate respectively, and μ_S, μ_N, μ_A and μ_P denote the corresponding sucrose, nitrate, ammonium and phosphate consumption rates [h^{-1}]. As for equation 4, the influence of the active DM (X_a) on the phosphate uptake in (Eq. 8) was neglected because i) the phosphate concentration dropped close to zero within ~90 h (Figure 2) and the "0" values would compromise numeric stability of the resulting models and ii) the necessary additional parameter would increase the likelihood of overfitting given the small calibration data set (n=3). As above, potential intracellular phosphate pools might be accounted for in future studies to refine our model, the measurement of such pools was beyond the scope of this study where we intended to demonstrate rapid optimization potentials for bioprocesses operation with a minimal number of experiments.

Parameter identification focused on calibrating the 12 parameters (Table 1) of the model (1)-(8), which is described hereafter as the "initial model".

Model parameters can be identified by minimizing an objective function dependent on a norm for the error made in measuring the process outputs (Jiménez-Hornero et al., 2009b). Here, we used the weighted least-squares function as an objective function, which for normally distributed errors constitutes a maximum likelihood estimator (Eq. 9 and 10).

$$\theta^* = \min_{\theta} Q(\theta;\tilde{x},\tilde{t},\tilde{y}) = \sum_{h=1}^{H}\sum_{t=1}^{T(h)} q_{h,t}(\theta;\tilde{x}_h,\tilde{t}_{h,t},\tilde{y}_{h,t}) \tag{9}$$

$$q_{h,t}(\theta;\tilde{x}_h,\tilde{t}_{h,t},\tilde{y}_{h,t}) = \sum_{r=1}^{R(h,t)} w_r\left[1 - \frac{g_r(\theta;\tilde{x}_h,\tilde{t}_{h,t})}{\tilde{y}_{h,t,r}}\right]^2 \tag{10}$$

TABLE 1 Parameter ranges and identified optimal parameter values.

Parameter	Description	Unit	Min	Max	Optimal	90% CI
k_d	Death rate constant	h^{-1}	0	0.1	6.62×10^{-4}	5.8×10^{-5}
μ_m	Maximum specific growth rate	h^{-1}	0	10	9.47	0.06
K_S	Sucrose saturation constant	g L^{-1}	0	100	17.64	3.5
K_{IS}	Sucrose inhibition constant	g L^{-1}	0	100	17.64	3.4
K_A	Ammonium saturation constant	g L^{-1}	0	1	0.796	0.040
K_{IA}	Ammonium inhibition constant	g L^{-1}	0	1	0.797	0.024
K_N	Nitrate saturation constant	g L^{-1}	0	5	0.133	0.057
K_{IN}	Nitrate inhibition constant	g L^{-1}	0	5	0.135	0.097
μ_s	Sucrose consumption rate	h^{-1}	0	1	2.235×10^{-1}	4.65×10^{-2}
μ_A	Ammonium consumption rate	h^{-1}	0	1	4.863×10^{-3}	6.34×10^{-4}
μ_N	Nitrate consumption rate	h^{-1}	0	1	3.013×10^{-2}	2.49×10^{-3}
..	Phosphate consumption rate	h^{-1}	0	1	2.883×10^{-2}	$2.36\ 10^{-3}$

where index $t = 1, \ldots, T$ denotes the time points at which the individual model factors $r = 1, \ldots, R$ (i.e. ammonium, …, sucrose as well as the DM) were measured for each experiment $h = 1, \ldots, H$ which created $\tilde{t}, \tilde{y}$ pairs ($\tilde{t}$ is the time of measurement and $\tilde{y}$ denotes the measured value). The model solution was synthetically described by $y = g(\theta; t)$ and had R responses (in this case represented by nutrients and DM). The term $w \in \mathbb{R}^R$ is a vector of weights for these responses to balance their priorities. Here, we used equal values, i.e., the unit vector, giving equal weights for each response. θ represents the model parameters (Table 1 for the initial model).

The CI for parameter estimates was processed using a residual-based non-parametric bootstrap resampling approach (Dogan, 2007). The error matrix Σ was obtained by the measured data $\tilde{y}_{h,t,r}$. Specifically, the entries of the matrix were the absolute and proportional measurement uncertainties u_r^0 and $u_r^{\%}$, respectively, that were determined for each model factor r (Table S4).

$$\sum = \begin{bmatrix} u_1^0 + u_1^{\%}\tilde{y}_{1,1,1} & \ldots & u_1^0 + u_1^{\%}\tilde{y}_{t,h,1} & \ldots \\ u_2^0 + u_2^{\%}\tilde{y}_{1,1,2} & \ldots & u_2^0 + u_2^{\%}\tilde{y}_{h,t,2,} & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ u_R^0 + u_R^{\%}\tilde{y}_{1,1,R} & \ldots & u_R^0 + u_R^{\%}\tilde{y}_{h,t,R} & \ldots \end{bmatrix} \quad (11)$$

This approach for error estimation is preferred over the more widely used model-data mismatch because the measurement error is assumed known (Table S4) (Joshi et al., 2006). The error values for model factors were resampled with replacement to create a large pool of error matrix sets (~1,000). Subsequently, each of the resampled error matrices was added back to the original modeled data (i.e., model prediction) to create a pool of synthetic data sets. Each of the generated data sets was independently processed through least-squares estimation (Eq. 9 and 10) to identify a set of model parameter values that formed the basis for CI calculation as described elsewhere (Dekking et al., 2005).

The quality measure of the model prediction for the r-*th* output (i.e., model factor) was quantified by the mean average error MAE (Borchani et al., 2015), which measures the average error between model prediction g and data $\tilde{y}$.

$$MAE_r^h(\theta^*) = \frac{1}{T(h)} \sum_{t=1}^{T(h)} \left| \tilde{y}_{h,t,r} - g_r(\theta^*; \tilde{x}_h, \tilde{t}_{h,t}) \right| \quad (12)$$

Moreover, we introduced a second quality measure (nMAE) corresponding to the MAE normalized with respect to the available data in order to normalize the data from experiments #1 to #7 (i.e., $\bar{H} = 7$).

$$nMAE_r^h(\theta^*) = \frac{MAE_r^h(\theta^*)}{\frac{1}{\sum_{k=1}^{H} T(k)} \sum_{h=1}^{\bar{H}} \sum_{t=1}^{T(j)} \tilde{y}_{h,t,r}} = \frac{MAE_r^h(\theta^*)}{\bar{y}_r} \quad (13)$$

The indices h and k indicate subsets of the dataset when more than one dataset $\mathcal{D} = \{1, 2, .\} \subseteq \{1, \ldots, 9\}$ was considered at the same time:

$$MAE_r^{\mathcal{D}}(\theta^*) = \frac{1}{\sum_{k \in \mathcal{D}} T(k)} \sum_{h \in \mathcal{D}} \sum_{t=1}^{T(h)} \left| \tilde{y}_{h,t,r} - g_r(\theta^*; \tilde{x}_h, \tilde{t}_{h,t}) \right| \quad (14)$$

The initial mass concentration of carbohydrates was 30 g L^{-1} in all experiments. In experiments #1-7 carbohydrates were provided as 88 mM sucrose whereas it was 151 mM fructose in #8 and 151 mM glucose in #9. Accordingly, experiments #1 - #7 did not contain fructose or glucose at the start of the fermentation. Instead, the monosaccharides formed in these experiments due to enzymatic cleavage of sucrose during fermentation. Because glucose was rapidly taken up by the BY-2 cells, the glucose concentrations in the culture medium of #1-7 was only in the 15–21 mM range. This was approximately 10-fold lower than the 151 mM fructose and glucose present at the start of experiments #8 and #9 respectively. Therefore, experiments #8 and #9 were not considered during this analysis.

Parameter identification in equations (9) and (10) for the model (1)-(8) considered all available initial experiments, namely #1, #2 and #3. Therefore, the following quality measures refer to the "training" set, because the same data adopted to establish the model were used to evaluate model performance. For the DM, the MAE was 0.64 [g L^{-1}] (Figure 3). This corresponds to a nMAE of 0.13. For sucrose, the MAE was 6.58 [mM] (nMAE = 0.22), for ammonium it was 2.21 [mM] (nMAE = 0.22), for nitrate it was 3.48 [mM] (nMAE = 0.11), and for phosphate it was 0.497 [mM] (nMAE = 0.79) (Figure 3). Later, the quality measures will be evaluated based on a "test" set, namely on unseen data that has not been used for parameter identification. A higher MAE in the test set than the training set would indicate model overfitting.

Interestingly, in case of all three nutrients (*S, N, A*) the inhibition constants were close the respective saturation constants ($K_r \approx K_{Ir}$) (Table 1). We had defined the constraint $K_r \geq K_{Ir}$ during parameter identification to reflect a typical activity plateau of enzymes and to exclude an actual reduction in growth at high substrate concentrations. However our observation that parameter optimization favors inhibition constant close to the value of the saturation constants indicates that an actual substrate inhibition, i.e., reduced growth at high concentrations, can occur and therefore allowing for "free-floating" values of K_{Ir} might be a better option in future models.

For sucrose and nitrate, the inhibition and saturation constants were lower than the initial standard concentrations, whereas the constants were higher for ammonium. This suggests that the initial concentrations of sucrose and nitrate can be reduced without significantly changing the specific DM growth (Eq. 4). The predicted phosphate trajectory was satisfactory only during the batch stage (Figure 3), i.e., the MAE was ≤0.03 [mM].

Model-based multi-criteria optimization suggests that reducing the amount of sucrose maximizes the yield and minimizes the process time for most of the Pareto front

Based on the model fitted to the standard BY-2 semi-continuous fermentation setting, we solved a multi-criteria (short

10.3389/fpls.2023.1183254

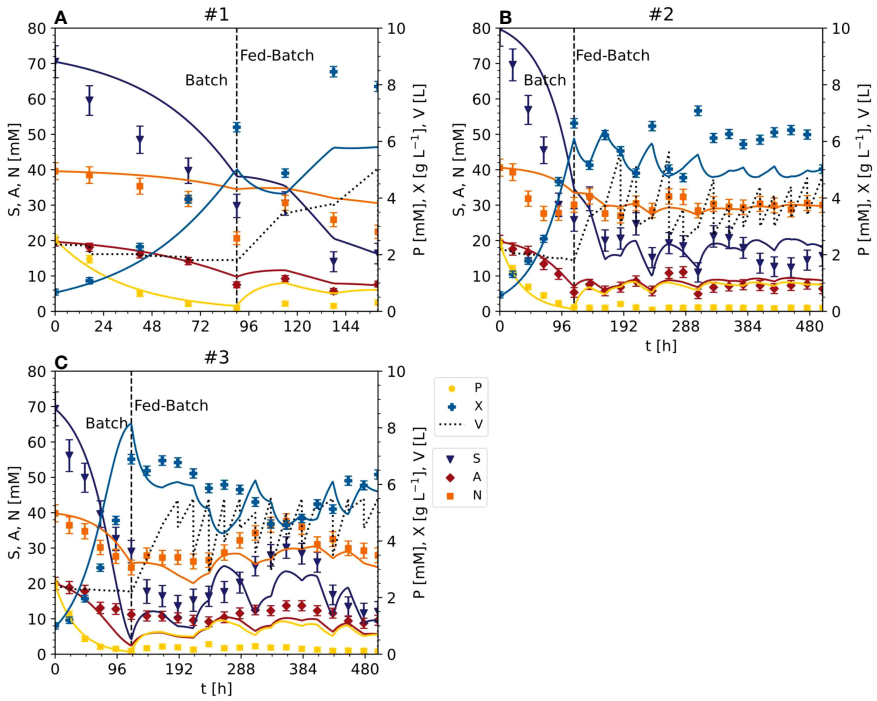


FIGURE 3
Time trajectories of measured and predicted data (continuous line) for the model set representing the cultivation of BY-2 cells under standard conditions (experiments #1, #2 and #3) **(A–C)** Iteration 1. Measurement uncertainty made of a constant and proportional component (Table S4). A, ammonium; F, fructose; G, glucose; N, nitrate; P, phosphate; S, sucrose; V, volume; X, cell dry mass. See Table S3 for measured initial values.

time, high yield) optimization problem in which multiple objective functions, were "simultaneously" minimized (Eq. 15):

$$\min_{x \in X} \; (f_1(x), f_2(x), \ldots, f_r(x)) \tag{15}$$

where is the independent variable (e.g., process time, nutrient concentration or starting DM). In the single objective case (i.e., r = 1), this results in one optimal solution. For r > 1, the result is given by the so-called Pareto optimal set in which one objective can only be improved if the values of the other objectives are worsened (Miettinen, 1998). To solve the multi-objective problem (15), it must be transformed into a single objective problem by mathematical scalarization (Finlayson et al., 2015).

Here, we used two objective functions, i.e., maximized the dimensionless yield of the active DM (Eq. 16), i.e., the gram DM at fermentation end ($X_a(t)$) divided by the inoculum (i.e., starting) cell DM (X_0):

$$Y_a = X_a(t)/X_0 \overset{t \to min}{\Rightarrow} Y_a^* = X_a^*(t^*)/X_0 \tag{16}$$

evaluated at time t^* (i.e., $Y_a{}^*$), and, simultaneously, minimized the time (t*) at which the maximum active DM was achieved (i.e., the process time), whereby the active DM presented a maximum that was due to both limited cell growth and increasing cell death (Figure 4A). To maximize the active DM, we used the sucrose concentration S_0 and the inoculum DM concentration (X_0) as variables (i.e., degrees of freedom). The first was varied between 17 and 88 mM while the second was varied between 0.3 and 1.5 g L^{-1} in order to investigate whether lowering the sucrose concentration would be beneficial. The range of the inoculum concentration included the values of the existing experiments. Based on the previous experiments (#1, #2 and #3), we assumed that the initial concentration of the dead DM was zero. The scarcity of data did not allow us to test the model before optimization. Therefore, we limited the degrees of freedom to the inoculum and sucrose, leaving out nitrate and ammonium. Hence, the bi-criteria optimization problem was formulated as follows (Eq. 17):

$$\min_{X_0,\, S_0} \; -Y_a^*(X_0,\; S_0), \quad t^*(X_0,\; S_0) \tag{17}$$

The ϵ-constraint or weighted sum methods can be used for scalarization (Finlayson et al., 2015). The weighted sum method identified the convex portion of the Pareto front whereas the ϵ-constraint method identified the remaining (non-convex) section of the front (Figure 4B). To ease the solution of the optimization problem, we implemented an iterative approach. First, we identified the maximal $Y_a{}^*$ (~28) and recorded the associated t^*_{max} (the minimal time at which $Y_a{}^*$ is maximized; ~180 h; top right orange star in Figure 4B). Then, we minimized t* under the side condition that the first derivative of $Y_a{}^*$ is ~0. This corresponded to the minimal time t^*_{min} at which $Y_a{}^*$ peaks, i.e., the earliest maximum of $Y_a{}^*$ (bottom left orange start in Figure 4B; ~96 h). We continued by handling t^* as a degree of freedom too (in addition to X_0 and S_0). Specifically, in the ϵ-constraint method, the process time t^* was allowed to adopt a set of pre-defined values that were uniformly distributed across the range spanned up by t^*_{min} and t^*_{max} (96–180 h) with a step-width of 3 h (i.e., 28 steps) to cover

the entire Pareto front (yellow stars in Figure 4B). According to the Pareto front, the shortest t^* (i.e., 102 h) is obtained with the highest initial DM concentration (1.5 g L^{-1}) (Figure 4C). However, the highest DM yield Y_a^* of 27.9 (Eq. 16 with X_a evaluated at t^*) would be achieved with the maximum sucrose concentration (88 mM) and the minimum possible initial DM concentration (0.3 g L^{-1}) at a time t^* of 185 h (Figure 4C). This is to be expected because under such conditions the specific concentration of carbohydrate per cell is highest at fermentation start, facilitating the highest relative biomass increase as expressed with Y_a^*. Importantly, alternative definitions of the objective function can be used. For example, the dimensionless biomass yield Y_a^* can be replaced by the biomass concentration at fermentation end $X_a(t^*)$.

Moreover, moving along the Pareto front, reducing the sucrose concentration by 25% from 88 mM to 67 mM would increase the active yield and reduce the t^* compared to standard conditions (Figure 4C). Accordingly, compared to 88 mM sucrose and 0.68, 0.58 and 1.00 g L^{-1} DM in experiments #1, #2 and #3, respectively, we proposed two experiments taken from the Pareto front (#4 and #5) with values of S_0 = 67 mM and DM = 0.80 and 0.56 g L^{-1}, respectively. For comparison and model validation, we also set up two further experiments that were not Pareto optimal (#6 and #7), with 47 mM sucrose/0.92 g L^{-1} DM and 87 mM sucrose/1.25 g L^{-1} DM, respectively, the latter representing the standard BY-2 cell cultivation conditions.

Reducing the sucrose concentration does not affect biomass formation or volumetric biomass yield when using Pareto optimal conditions

In the optimized setting for the semi-continuous fermentation according to the Pareto front (experiments #4 and #5), the BY-2 cells were cultivated with a starting biomass of 17 g L^{-1} FM (0.80 g L^{-1} DM) and a sucrose concentration of 67 mM (#4), or a starting biomass of 14 g L^{-1} FM (0.56 g L^{-1} DM) and a sucrose concentration of 67 mM (#5), at the beginning of the batch phase (Figures 2, 4; S3) in 2-L STRs. We choose a smaller bioreactor volume to accommodate the number of verification runs, i.e., 6 vs the initial 3 runs for model calibration (see also next section). In our hands, both reactor settings (2-L and 5-L) had performed equally in the last 8 years in terms of, for example, oxygen supply, biomass build-up etc. (data not shown).

The validation experiments that were not Pareto optimal started with a biomass of 17 g L^{-1} FM (0.92 g L^{-1} DM) and 47 mM g L^{-1} sucrose (#6) or 22 g L^{-1} FM (1.25 g L^{-1} DM) and 87 mM sucrose (#7), the latter as the standard BY-2 cell cultivation conditions (Figures 2, 4 and S3). Moreover, because the maximum predicted biomass was not achieved under standard cultivation settings (experiments #1, #2 and #3) because the semi-continuous stage started at 100 g L^{-1} FM, in these experiments the batch stage lasted until the sucrose in the medium was depleted.

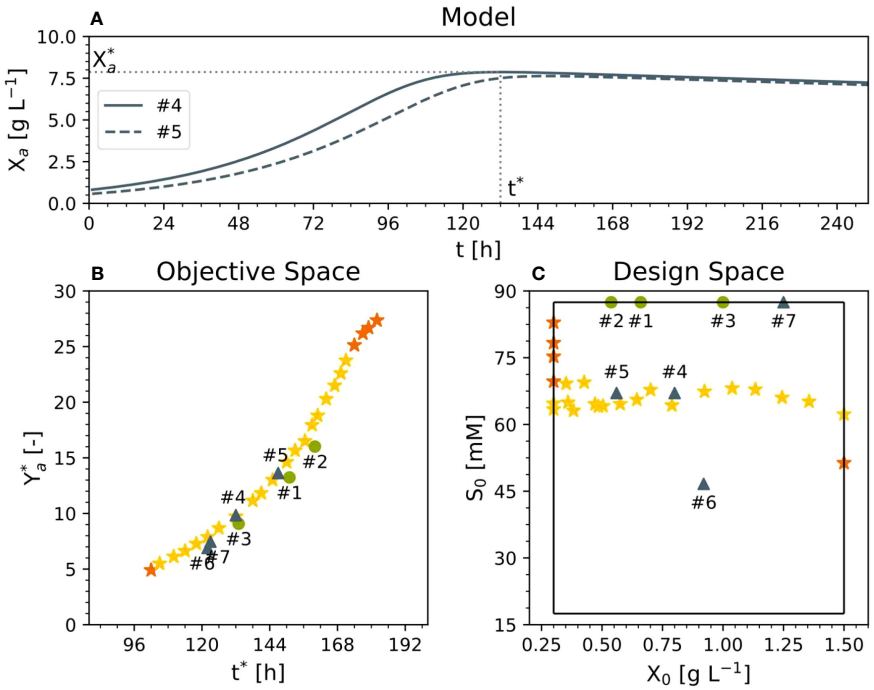


FIGURE 4
Results of multi-criteria optimization. **(A)** Representative time trajectory of two Pareto solutions relative to active DM. **(B)** Pareto front in the objective and **(C)** design space. Experiments #1–3 were used for model setup, experiments 4 and 5 had an optimal cultivation medium, and experiments 6 and 7 were performed for comparison and model validation purposes (dry mass X, sucrose S, process time t*, active yield Y_a* at time t*, initial dry mass X_0, initial sucrose S_0).

In experiments #4 and #5, the FM reached 207.90 g L^{-1} (10.88 g L^{-1} DM) and 181.70 g L^{-1} (10.04 g L^{-1} DM) by the time the sucrose was depleted (within 6 days). The sucrose was hydrolyzed into glucose and fructose, which resulted in a transient peak of 15–17 mM on days 3–4, which declined to <5 mM on day 6. The ammonium was depleted along with the sucrose, but nitrate levels fell by only ~50% from 39 to 19 and 23 mM in experiments #4 and #5, respectively. As observed in experiments #1, #2 and #3, phosphate was almost completely consumed with 2 days during batch fermentation.

In experiment #6, the biomass increased to 172.20 g L^{-1} FM (7.92 g L^{-1} DM) and the starting concentration of sucrose was also depleted in 6 days. Glucose and fructose peaked at 13–16 mM on day 4. Ammonium was depleted by day 6, but nitrate levels fell by only 38%, from 39 to 25 mM.

In experiment #7, a control repeating the standard BY-2 cultivation settings that was not included in the initial model setup, the biomass increased to 270.80 g L^{-1} FM (13.76 g L^{-1} DM) before the sucrose was depleted within 6 days. The transient peak of glucose and fructose (19–22 mM) was observed on day 3 but also declined by day 6. In contrast to sucrose, the ammonium was completely consumed by day 5, and nitrate levels declined by 66% from 39 to 14 mM on day 6. Phosphate was depleted within 3 days.

These data indicate that (1) once the FM exceeded a threshold of 50–60 g L^{-1} FM (3.0–3.5 g L^{-1} DM), the glucose and fructose resulting from sucrose hydrolysis were immediately consumed, (2) the conversion of nitrate to ammonium appeared less efficient than ammonium uptake into cells and its conversion to glutamate and glutamine, and (3) the depletion of sucrose, ammonium and phosphate did not appear to limit the cell growth. To understand the discrepancy between nutrient depletion and cell growth, it may be necessary to include the intracellular metabolism of nutrients in future models. Nevertheless, the relative cell growth was apparently not affected by the lower sucrose concentration and inoculation cell density. Importantly, whereas there was some variability in the fresh-to-dry mass ratio of the inoculum across the set of 9 experiments (21.8 ±4.0, ± standard deviation), that ratio was consistent at the end of the batch phase (19.7 ±1.5, ± standard deviation) indicating that there was no substantial difference in the biomass obtained from the different experiments (i.e., conditions) in terms of water content as it can arise from water uptake into the vacuole.

Model calibration based on optimized cultivation conditions confirms the biomass yield increase caused by reducing the initial sucrose concentration

The data from the Pareto optimal experiments (#4 and #5) and non-optimal comparators (#6 and #7) (Figures 2 and S3) allowed us to test the Monod-type model against unknown data (test set). The MAE was 1.02 [g L^{-1}] for DM and 8.18 [mM], 1.38 [mM] and 1.86 [mM] for sucrose, ammonium and nitrate, respectively (Figures 3, 5), confirming in principle that the model could fit the observed nutrient consumption and biomass formation. Therefore, the MAE for nutrients in the test set was lower than or equal to that in the training set, but the MAE for DM was higher in the test set.

The model was able to fit the cell growth until the DM reached 8 g L^{-1} but underestimated the real values above this threshold (MAE = 0.30 [g L^{-1}] for DM <8 g L^{-1}) (Figure 6). There are two possible explanations. First, the data used for model fitting (experiments #1, #2 and #3) did not include DM values > 8 g L^{-1}. Second, specific cell growth (Eq. 4) was directly dependent on nutrients such as sucrose and ammonium, and the model was not able to calculate DM values when at least one nutrient was close to zero (the case when DM > 8 g L^{-1}). Even fitting the model to the data (#4, #5, #6 and #7) did not improve the prediction quality. These considerations, together with the absence of model overfitting (i.e., the MAE of the nutrients in the test set never exceeded that in the training set when DM < 8 g L^{-1}), suggested the need for a more complex model, for example, one that includes the hydrolysis of sucrose and/or intracellular metabolism.

To compare DM growth in different experiments with distinct initial inoculum DM values (0.56–1.25 g L^{-1}), all measured and predicted DM trajectories were normalized against their initial DM values (Figure 7). However, there was a mismatch that caused the model either to underestimate (e.g., experiment #1) or overestimate (e.g., experiment #3) the biomass formation. Due to this mismatch, despite experiments #1, #2, and #3 using the standard initial sucrose concentration (88 mM) and differing only in DM, the model predicted a growing biomass yield for decreasing initial DM values, whereas the measured biomass yields were maximum and minimum at intermediate values of the initial DM (Figure 7A). Independently, measured and predicted trajectories confirmed that the model did not fit cell growth accurately when DM >8 g L^{-1} (Figure 7B). The assessment of the model's prediction errors is therefore an important task for future investigations. Nevertheless, we confirmed that a 25% reduction in the concentration of sucrose can increase the volumetric biomass yield by 13%, confirming that model-based optimization of the cultivation medium can improve the performance of BY-2 cell suspension cultures.

An improved model including glucose and fructose enables the prediction of cell growth at low sucrose concentrations

Given that the model could not predict DM concentrations >8 g L^{-1} when any of the nutrients reached zero, we used a second unstructured model (Eq. 2-3, 5-8, 18-20), accommodating the hydrolysis of sucrose into fructose and glucose (Eq. 18-20). Therefore, the specific growth term depended on glucose and fructose but not on sucrose. Moreover, the influence of ammonium and phosphate were included in the maximum specific growth as additive terms because including them as factors would have zeroed out the growth rate for times >90 h as the phosphate concentration approaches zero, which did not agree with experimental observations as discussed above.

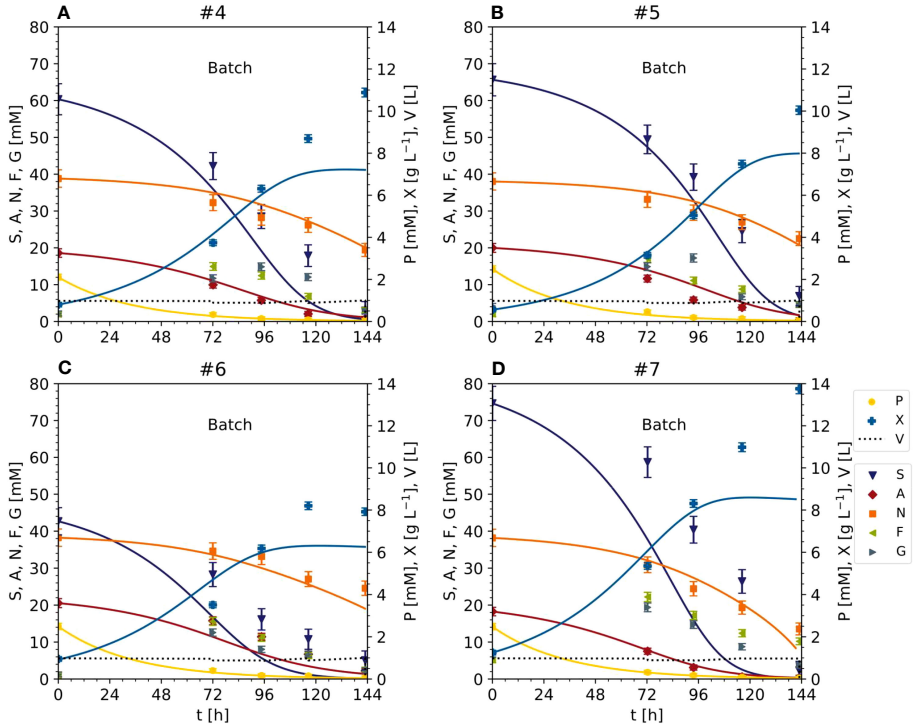


FIGURE 5
Time trajectories of measured and predicted data (continuous line) for the initial model validation representing the cultivation of BY-2 cells under optimal (experiments #4 and #5) and non-optimal (experiments #6 and #7) conditions **(A–D)** Iteration 2. Measurement uncertainty made of a constant and proportional component (Table S4). A, ammonium; F, fructose; G, glucose; N, nitrate; P, phosphate; S, sucrose; V, volume; X, cell dry mass. See Table S3 for measured initial values.

$$\mu = \left(\mu_m + \mu_m^A \frac{A}{A+\delta_A} + \mu_m^P \frac{P}{P+\delta_P}\right) \left(\frac{G}{G+K_G}\frac{K_{IG}}{G+K_{IG}} + \frac{F}{F+K_F}\frac{K_{IF}}{F+KI_F}\right)\left(\frac{N}{N+K_N}\frac{K_{IN}}{N+K_{IN}}\right) \quad (18)$$

$$V\frac{dF}{dt} + F\frac{dV}{dt} = m_{SF}\alpha\left(\frac{S}{S+K_S}\frac{K_{IS}}{S+K_{IS}}\right)X_a V - \mu_F\left(\frac{F}{F+K_F}\frac{K_{IF}}{F+KI_F}\right)X_a V \quad (19)$$

$$V\frac{dG}{dt} + G\frac{dV}{dt} = m_{SG}\alpha\left(\frac{S}{S+K_S}\frac{K_{IS}}{S+K_{IS}}\right)X_a V - \mu_G\left(\frac{G}{G+K_G}\frac{K_{IG}}{G+K_{IG}}\right)X_a V \quad (20)$$

The sucrose to glucose or fructose DM conversion parameters (m_{SG} and m_{SF}) were set at 0.526 because one mole of sucrose always produces one mole each of glucose and fructose (Puad et al., 2017). This model is defined hereafter as the "improved model". It has 20 parameters to identify, including the death rate constant k_d , maximum specific growth rate μ_m, maximum specific growth rate for ammonium and phosphate $\mu^A{}_m$, $\mu^P{}_m$, saturation and inhibition constants K_h and K_{Ih} , hydrolysis constant α, and consumption rates μ_h. Coefficients δ_A, δ_P can be interpreted as the minimal ammonium and phosphate concentration in the medium at which cell growth occurs. They were assumed known in the context of this study to limit the model complexity and avoid overfitting and their values were set to 10% of the starting concentrations. This allowed a nearly constant growth contribution when they are not depleted as observed in the experiments. As the number of experiments available for model calibration iteratively increase, the two parameters can be fitted too.

We used K-fold cross-validation to show the better prediction capability of the improved model compared to the initial one (Stone, 1974). With K = 4 full data sets available, i.e., all experiments performed so far in which fructose and glucose were measured: #4-7. Experiments #1-3 were not considered since the improved model would need the initial concentrations of fructose and glucose. The *k-th* fold consists of splitting the data so that the *k-th* set is used for testing, while the remaining K–1 are used for model setup (i.e., model training). It follows that for each model (i.e., initial and improved), k parameter identifications were performed (each relying on 3 experiments) with the performance measures estimated on the corresponding test set. The improved model was supposed to have a lower average validation (i.e., test) error than the initial one:

$$aMAE_i = \frac{1}{K}\sum_{k=1}^{K} = 1K\ MAE_i^k(\theta_k^*) \quad (21)$$

Analogously:

$$anMAE_i = aMAE_i/\bar{y}_i \quad (22)$$

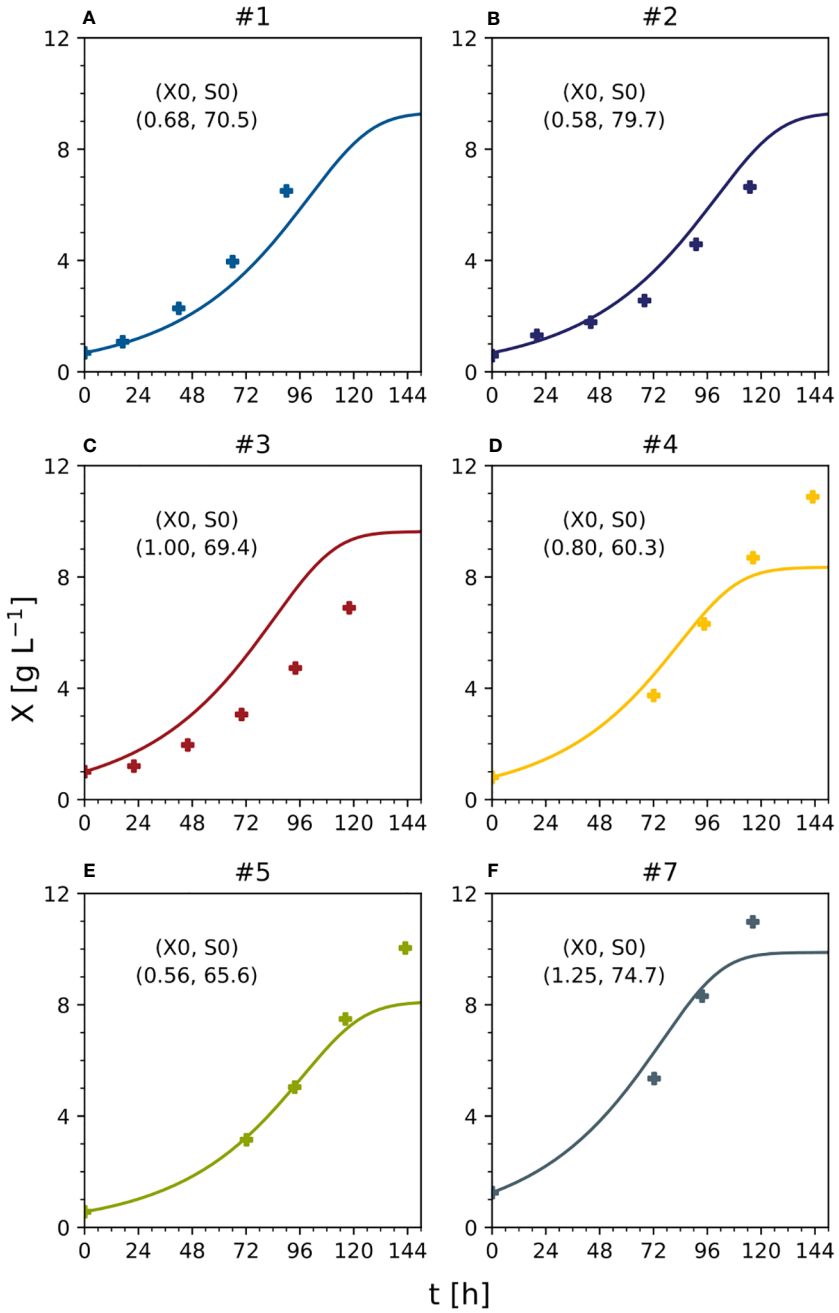


FIGURE 6
Predicted and measured time trajectories of dry mass under standard, optimal and non-optimal cultivation conditions (batch stage). Text indicates initial DM [g L^{-1}] and sucrose [mM] in each experiment. **(A)** Experiment #1. **(B)** experiment #2. **(C)** experiment #3. **(D)** experiment #4. **(E)** experiment #5. **(F)** experiment #7. S_0 – starting sucrose concentration; t – process time; X – cell dry mass; X_0 – starting dry mass. See Table S3 for measured initial conditions of the process parameters.

Whereas the MAE is a quality measure that refers to a single set of model parameters θ*, the aMAE is the average MAE over different models (i.e., different parameter sets). The following analysis allowed a fair comparison between models and the aMAE provided an estimation of the modeling prediction error on unseen data.

Compared to the initial model, the improved model had a lower or equal $aMAE_i$ for all outputs (Table 2). The most significant improvements occurred for the DM (the $anMAE_X$ dropped from 0.214 to 0.149) and sucrose (the $aMAE_S$ was 30% lower in the improved model). The average error of the prediction for ammonium and nitrate remained the same. Such results indicated the better modeling capability of the improved model, in particular the ability to capture high DM values (Figure 8). For DM greater than 8 [g L^{-1}], $anMAE_X$ dropped from 0.371 to 0.097 (Table 2).

The improved model was then retrained with all available full data sets (i.e., experiments #4-7) resulting in the proposed identified parameters (Table 3). The MAE_i estimated on the training data

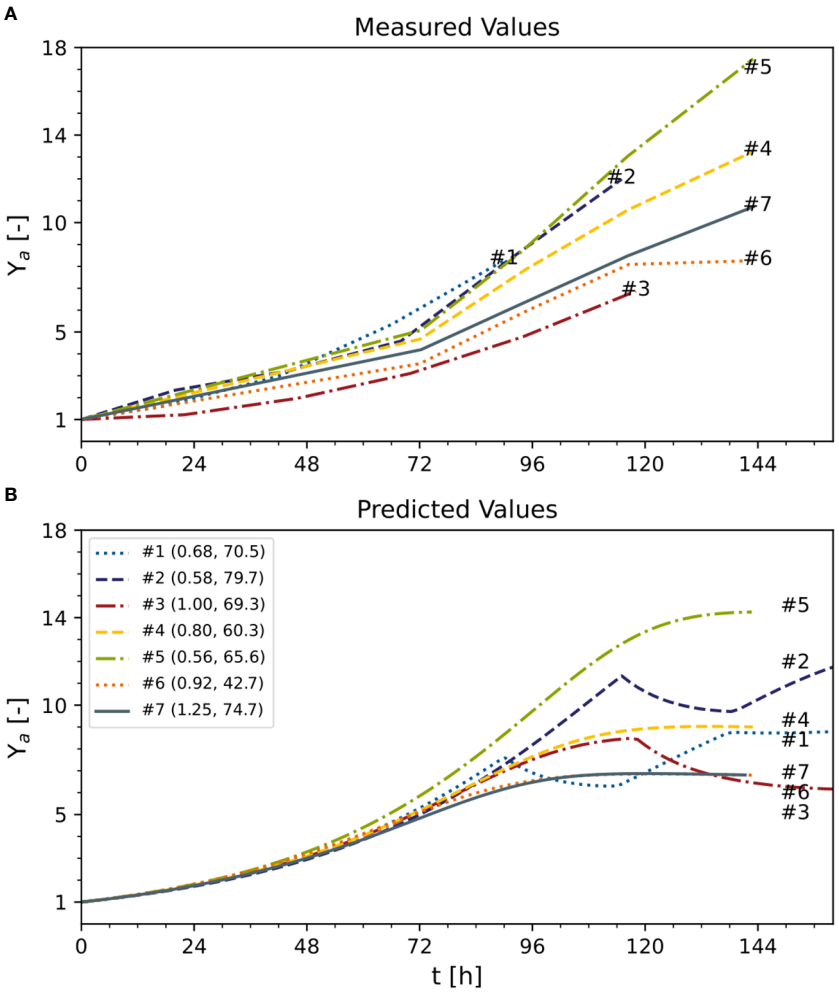


FIGURE 7
Predicted and measured normalized time trajectories of dry mass under standard optimal and non-optimal cultivation conditions. Text following the experiment number indicates initial DM [g L^{-1}] and sucrose [mM] of each experiment. **(A)** Measured values. **(B)** predicted values. Y_a – active yield (Eq. 16). Numbers represent the individual experiments. See Table S3 for measured initial conditions.

(Figure 9) was 0.58 [g L^{-1}] for DM and 4.31, 1.35, 3.03 and 0.02 [mM] for sucrose, ammonium, nitrate, and phosphate respectively. The MAE was equal or lower (Figure 9) compared to the initial model (Figure 3) in all model outputs. Nevertheless, this did not necessarily imply a better performance because different data sets were used and the MAE relied on training data. The MAEs for fructose and glucose, not included in the initial model, were 3.81 and 2.29 [mM] with nMAE values of 0.409 and 0.261, respectively. Among the nutrients, fructose showed the highest nMAE and its concentration was thus most difficult to predict (Figure 9). The reason for the low predictability is unknown but may reflect a staged metabolism that is not solely dependent on fructose alone and would require a modification of the kinetic model (Eq. 5). CIs were in the same order of magnitude as the values of the identified parameters (Table 3). Such a result is not desirable because it indicates a moderate predictive power of the underlying model and can hamper the identification of relevant differences between samples/conditions (O'Brien and Yi, 2016). However, wide CIs were expected because of the large number of parameters and the small data set. The effect of the wide CIs on the predictive ability of the models was assessed by generating 1,000 random sets of model parameter values θ, uniformly sampled from inside the CIs. These

TABLE 2 Average and average normalized test MAE values for initial and improved models.

Model	$aMAE_X$ [g L^{-1}] ($anMAE_X$ [-])	$aMAE_{X>8}$ [g L^{-1}] ($anMAE_{X>8}$ [-])	$aMAE_S$ [mM] ($anMAE_S$ [-])	$aMAE_A$ [mM] ($anMAE_A$ [-])	$aMAE_N$ [mM] ($anMAE_N$ [-])
Initial	1.069 (0.214)	1.850 (0.371)	5.841 (0.198)	1.983 (0.194)	3.082 (0.101)
Improved	0.741 (0.149)	0.488 (0.097)	4.541 (0.153)	1.808 (0.177)	3.232 (0.106)

10.3389/fpls.2023.1183254

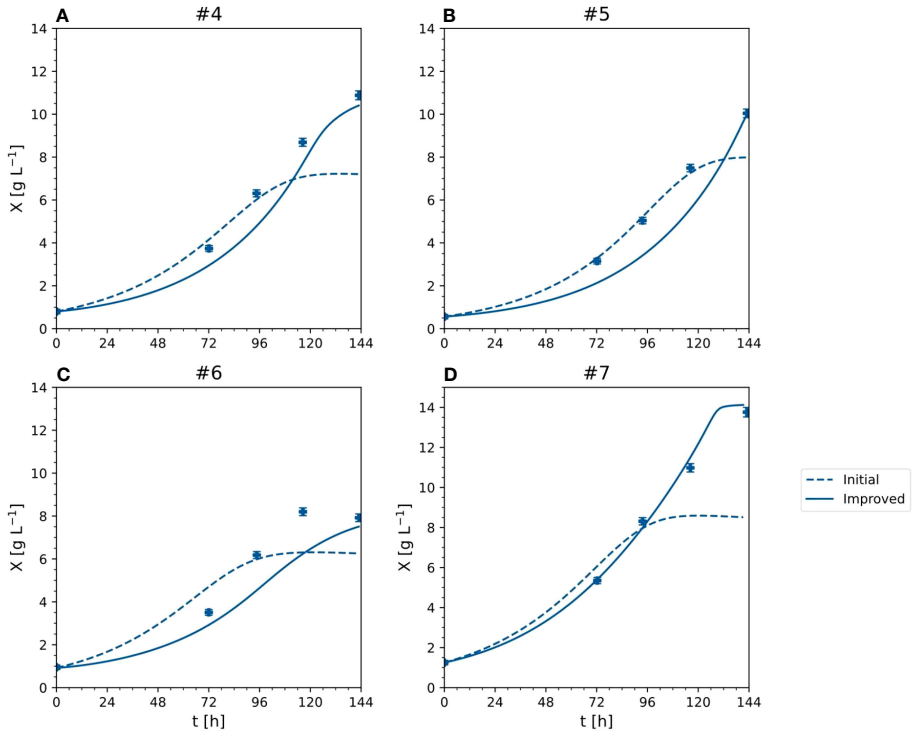


FIGURE 8
X - cell dry mass time trajectories of measured and predicted data for the initial and improved models representing the cultivation of BY-2 cells under optimal and non-optimal conditions (experiment #4, ... #7). **(A–D)** Iteration 2.

sets were then used to create model predictions and estimate the model-data mismatch based on MAE (Eq. 12), aMAE (Eq. 21) and anMAE (Eq. 22) for each model factor using experiments #4-7 as references (Figure 10). In case of the initial model, the anMAE was equal to 26%, 17%, 8%, 5% for DM, ammonium, nitrate, and phosphate, respectively, which was less than the sum of the model independent process parameter uncertainties (Table S4), which was approximately 27%. For this sum, constant and proportional measurement uncertainties of each process parameter were unified under a single relative uncertainty,

TABLE 3 Parameter ranges and identified optimal parameter values for the improved model (experiments #1 to #7 were used for model setup).

Parameter	Description	Unit	Min	Max	Optimal	90% CI
k_d	death rate constant	h^{-1}	0	0.01	5.44×10^{-4}	6.2×10^{-5}
μ_m	maximum specific growth rate	h^{-1}	0	10	0.162	0.17
μ_m^A	maximum specific growth rate ammonium	h^{-1}	0	10	0.578	0.48
μ_m^P	maximum specific growth rate phosphate	h^{-1}	0	10	2.1×10^{-5}	0.37
K_S	Sucrose saturation constant	g L^{-1}	0	100	72.53	73.4
K_{IS}	Sucrose inhibition constant	g L^{-1}	0	100	100.0	73.3
K_A	Ammonium saturation constant	g L^{-1}	0	1	0.221	0.22
K_{IA}	Ammonium inhibition constant	g L^{-1}	0	1	0.221	0.065
K_N	Nitrate saturation constant	g L^{-1}	0	5	0.835	0.25
K_{IN}	Nitrate inhibition constant	g L^{-1}	0	5	1.0	0.06
K_F	Fructose saturation constant	g L^{-1}	0	30	1.108	0.71
K_{IF}	Fructose inhibition constant	g L^{-1}	0	30	1.404	0.51
..	Glucose saturation constant	g L^{-1}	0	30	18.0	17.1

(Continued)

TABLE 3 Continued

Parameter	Description	Unit	Min	Max	Optimal	90% CI
K_{IG}	Glucose inhibition constant	g L^{-1}	0	30	18.0	12.0
α	Sucrose hydrolysis constant	h^{-1}	0	1	0.396	0.24
μ_A	Ammonium consumption rate	h^{-1}	0	1	6.9×10^{-3}	1.9×10^{-3}
μ_N	Nitrate consumption rate	h^{-1}	0	1	1.71×10^{-2}	2.7×10^{-2}
μ_F	Fructose consumption rate	h^{-1}	0	1	8.34×10^{-2}	2.9×10^{-2}
μ_G	Glucose consumption rate	h^{-1}	0	1	0.273	0.26
μ_P	Phosphate consumption rate	h^{-1}	0	1	2.68×10^{-2}	3.2×10^{-3}

namely 3% for DM, 6% for each nutrient. This observation suggested that the model still retained predictive validity when model parameters θ were inside the CIs because the variability of the model predictions were in the same range as the uncertainty of the independent process variables. The sole exception in this context was sucrose, which had an anMAE of 34%, which was higher than the sum of process parameter uncertainties.

In the improved model, the anMAE values were 53%, 35%, 21%, 21%, 4%, 42%, 24%, for DM, sucrose, ammonium, nitrate, phosphate, fructose, and glucose, respectively (Figure 10), whereas the sum of the modeled independent process parameter uncertainties was 39%. Like the initial model, the improved model suggested that, with respect to the standard cultivation conditions, a 30% reduction of sucrose in the medium is beneficial for the DM growth (Figure 11A, e.g., experiment #5). Also, glucose seemed to be preferred over fructose as a carbon source (Figures 11B, C) which was qualitatively confirmed in experiments #8 and #9. Therefore, including the individual carbon sources in the model allowed an improvement of the prediction compared to the model accounting only for the "precursor" sucrose, that is not taken up by the cells directly, but only after breakdown to the monosaccharides.

Finally, the model was adopted to predict DM and nutrients for experiments #8 and #9 (unseen data), in which sucrose was replaced

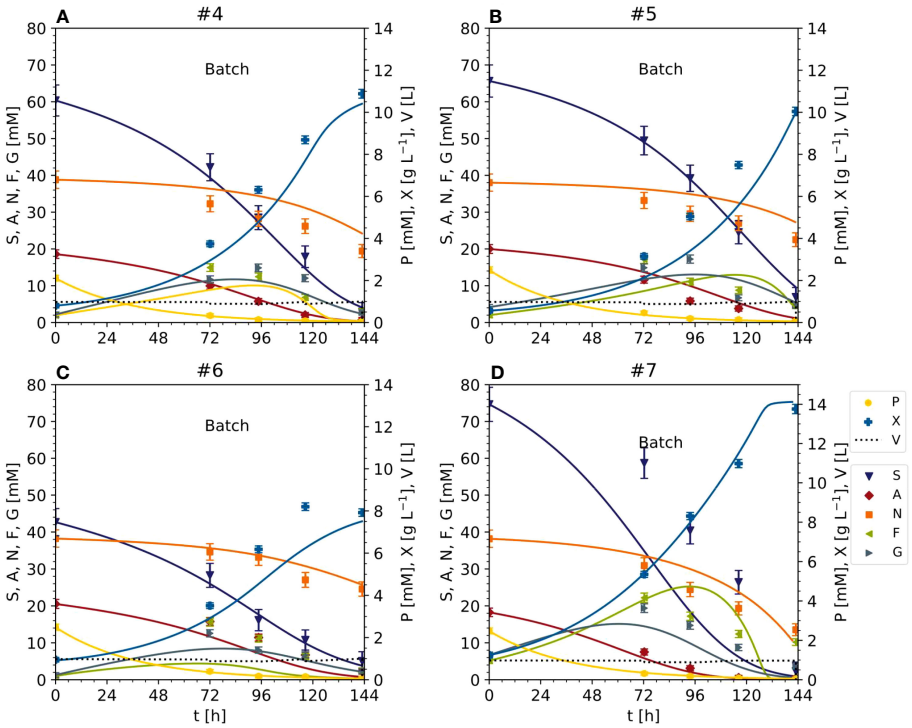


FIGURE 9
Time trajectories of measured and predicted data (continuous line) for the improved model representing the cultivation of BY-2 cells under optimal and non-optimal conditions (experiment #4, ... #7). **(A–D)** Iteration 2. Measurement uncertainty made of a constant and proportional component (Table S4). A, ammonium; F, fructose; G, glucose; N, nitrate; P, phosphate; S, sucrose; V, volume; X, cell dry mass. See Table S3 for measured initial values.

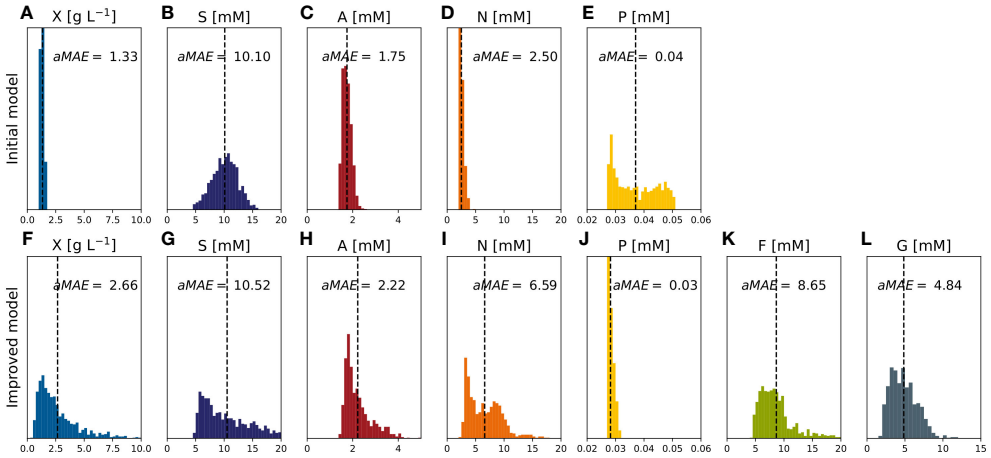


FIGURE 10
MAE of model factors describing the model-data mismatch obtained with random parameter set values (~1000) uniformly sampled within the confidence intervals (experiments #4,...,#7). **(A–E)** Initial model. **(F–L)** Improved model. A, ammonium; F, fructose; G, glucose; N, nitrate; P, phosphate; S, sucrose; X, cell dry mass.

with either fructose (#8) or glucose (#9) as the sole carbon source (Figures 2, 12, S4). The DM prediction was mediocre (MAE_X = 0.37 [g L^{-1}]), and were very good for ammonium (MAE_A = 0.88 [mM]) and nitrate (MAE_N = 0.92[mM]), but poor for fructose or glucose (MAE_F = 25.06[mM], MAE_G = 19.60 [mM]) (Figure 12). Indeed, MAE_F and MAE_G were one order of magnitude larger than in the model setup (Figure 12). The decline in the concentrations of fructose and glucose was underestimated by the model (50 mM measured compared to 100 mM predicted), possibly because the model was fitted against data sets in which the fructose and glucose concentrations were 0–25 mM (Figure S3), whereas in experiments #8 and #9 the concentrations were 150 mM (Figure S4). Nevertheless, the inclusion of fructose and glucose in the model enabled the prediction of cell growth and biomass formation at high cell densities and low sucrose concentrations, and improved the modeling capability with respect to the initial model (Table 3). To increase the model accuracy further, the unstructured model might be converted to a structured one that includes the intracellular metabolism of nutrients.

Conclusions

We have shown that adopting the iterative "experiment-modeling-optimization" workflow achieved a 13% increase in the growth of tobacco BY-2 cell suspension cultures while reducing the

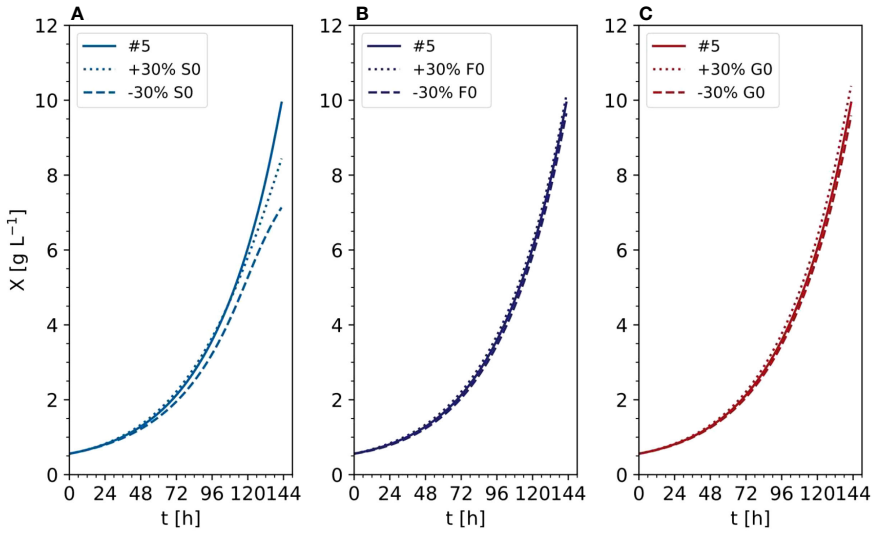


FIGURE 11
Time trajectories of predicted data with the improved model comparing the influence of initial carbon sources **(A)** Sucrose, **(B)** Fructose, **(C)** Glucose) in the DM (experiment #5). X, cell dry mass.

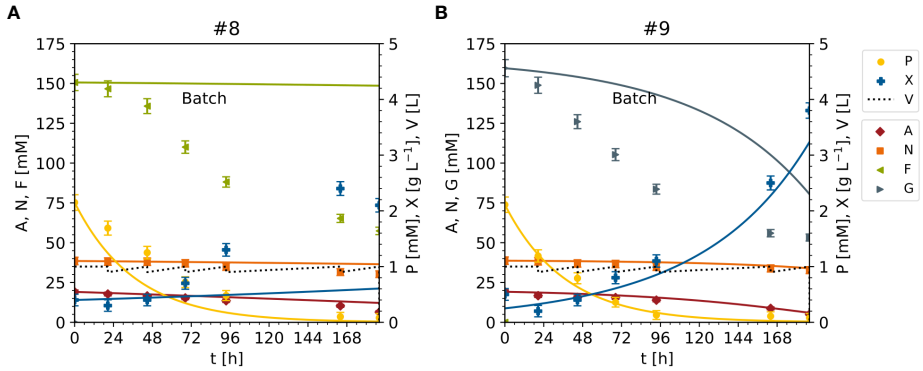


FIGURE 12
Time trajectories of measured and predicted data (continuous line) for the model set up representing the cultivation of BY-2 cells under non-optimal culture conditions with either fructose (experiment #8) or glucose (experiment #9) as the sole carbon source for improved model validation. **(A, B)** Iteration 2. A – ammonium. F – fructose. G – glucose. N – nitrate. P – phosphate. S – sucrose. V – volume. X – cell dry mass. See Table S3 for measured initial values. Measurement uncertainty made of a constant and proportional component (Table S4).

sucrose concentration in the cultivation medium by 25%. This was based on a mechanistic unstructured segregated Monod-type model using the sucrose concentration as a relevant parameter. However, the initial model could not predict cell growth at high cell densities, which was resolved by an improved model that included glucose and fructose as sucrose hydrolysis products. Moreover, cell growth could only be fitted if the nutrients did not reach zero, so we removed phosphate from the model because this was completely depleted. This shortcoming might be addressed by a structured model that includes intracellular nutrient metabolism. Nevertheless, the model suggested that sucrose concentration and inoculation cell density can be reduced to maximize the yield, which we confirmed in validation experiments. Mechanistic models can therefore be used to maximize the productivity of cell suspension cultures while reducing upstream production costs. This may be applicable not only to plant cells, but also to microbial and mammalian cell cultures used for the production of biopharmaceuticals.

Data availability statement

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Author contributions

HN and JB designed the experiments. HN and JL conducted the experiments and pre-processed the data. MaB, KT and MiB analyzed the data and built the models. HN and MB wrote the manuscript. CC, MiB, and JB revised the manuscript and secured the funding. All authors contributed to the article and approved the submitted version.

Funding

This work was funded in part by the Fraunhofer-Gesellschaft Internal Programs under grant no. Attract 125-600164 and the state of North-Rhine-Westphalia under the Leistungszentrum grant no. 423 “Networked, adaptive production” as well as the BMBF-funded Fraunhofer innovation program project number 800 093 (digitalization of bioprocesses – dibi).

Acknowledgments

We thank Dr. Richard M Twyman for editorial assistance.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1183254/full#supplementary-material

References

Asprion, N., Böttcher, R., Schwientek, J., Höller, J., Schwartz, P., Vanaret, C., et al. (2022). Decision support for the development, simulation and optimization of dynamic process models. *Front. Chem. Sci. Eng.* 16, 210–220. doi: 10.1007/s11705-021-2046-x

Behrend, J., and Mateles, R. I. (1975). Nitrogen metabolism in plant cell suspension cultures: I. Effect of amino acids on growth. *Plant Physiol.* 56, 584–589. doi: 10.1104/pp.56.5.584

Behrend, J., and Mateles, R. I. (1976). Nitrogen metabolism in plant cell suspension cultures: II. Role of organic acids during growth on ammonia. *Plant Physiol.* 58, 510–512. doi: 10.1104/pp.58.4.510

Bittsánszky, A., Pilinszky, K., Gyulai, G., and Komives, T. (2015). Overcoming ammonium toxicity. *Plant Sci.* 231, 184–190. doi: 10.1016/j.plantsci.2014.12.005

Borchani, H., Varando, G., Bielza, C., and Larrañaga, P. (2015). A survey on multi-output regression. WIREs Data Mining Knowl. *Discov.* 5, 216–233. doi: 10.1002/widm.1157

Chu, C.-C., Wang, C.-C., Sun, C.-S., Hsu, C., Yin, K.-C., Chi, Y.-C., et al. (1975). Establishment of an efficient medium for anther culture of rice through comparative experiments on the nitrogen sources. *Sci. Sin.* 18, 659–668. doi: 10.1360/ya1975-18-5-659

Dekking, F., Kraaikamp, C., Lopuhaä, H. P., and Meester, L. E. (2005). *A modern introduction to probability and statistics: understanding why and how* (Springer London: London UK). doi: 10.1007/1-84628-168-7

Dogan, G. (2007). Bootstrapping for confidence interval estimation and hypothesis testing for parameters of system dynamics models. *Syst. Dyn. Rev.* 23, 415–436. doi: 10.1002/sdr.362

Finlayson, B. A., Biegler, L. T., Grossmann, I. E., Küfer, K.-H., and Bortz, M. (2015). "Mathematics in chemical engineering," in *Ullmann's encyclopedia of industrial chemistry* (Weinheim, Germany: Wiley-VCH), 1–161.

Fischer, R., Liao, Y. C., and Drossard, J. (1999). Affinity-purification of a TMV-specific recombinant full-size antibody from a transgenic tobacco suspension culture. *J. Immunol. Methods* 226, 1–10. doi: 10.1016/s0022-1759(99)00058-7

Gamborg, O. L., Miller, R. A., and Ojima, K. (1968). Nutrient requirements of suspension cultures of soybean root cells. *Exp. Cell Res.* 50, 151–158. doi: 10.1016/0014-4827(68)90403-5

Häkkinen, S. T., Reuter, L., Nuorti, N., Joensuu, J. J., Rischer, H., and Ritala, A. (2018). Tobacco BY-2 media component optimization for a cost-efficient recombinant protein production. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.00045

Hawkins, D. M. (2004). The problem of overfitting. *J. Chem. Inf. Comput. Sci.* 44, 1–12. doi: 10.1021/ci0342472

Holland, T. (2013). "*PhD thesis: Development of plant cell cultures with respect to the industrial production of biopharmaceutical ingredients*" *[Entwicklung von Pflanzenzellkulturen im Hinblick auf die industrielle Produktion biopharmazeutischer Wirkstoffe]* (Aachen, Germany: RWTH Aachen University).

Holland, T., Sack, M., Rademacher, T., Schmale, K., Altmann, F., Stadlmann, J., et al. (2010). Optimal nitrogen supply as a key to increased and sustained production of a monoclonal full-size antibody in BY-2 suspension culture. *Biotechnol. Bioeng.* 107, 278–289. doi: 10.1002/bit.22800

Höller, J., Bickert, P., Schwartz, P., Kurnatowski, M., Kerber, J., Künzle, N., et al. (2019). *Parameter estimation strategies in thermodynamics* (ChemEngineering. 3). doi: 10.3390/chemengineering3020056

Jacob, A., Mahanty, B., and Thomas, J. (2020). Dynamic modelling of growth and flavonoid production from Ocimum tenuiflorum suspension culture. *Bioprocess Biosyst. Eng.* 43, 2053–2064. doi: 10.1007/s00449-020-02394-6

Jiménez-Hornero, J. E., Santos-Dueñas, I. M., and García-García, I. (2009a). Optimization of biotechnological processes. The acetic acid fermentation. Part I: The proposed model. *Biochem. Eng. J.* 45, 1–6. doi: 10.1016/j.bej.2009.01.009

Jiménez-Hornero, J. E., Santos-Dueñas, I. M., and García-García, I. (2009b). Optimization of biotechnological processes. The acetic acid fermentation. Part II: Practical identifiability analysis and parameter estimation. *Biochem. Eng. J.* 45, 7–21. doi: 10.1016/j.bej.2009.01.010

Joshi, M., Seidel-Morgenstern, A., and Kremling, A. (2006). Exploiting the bootstrap method for quantifying parameter confidence intervals in dynamical systems. *Metab. Eng.* 8 (5), 447–455. doi: 10.1016/j.ymben.2006.04.003

Lemoine, R., La Camera, S., Atanassova, R., Dédaldéchamp, F., Allario, T., Pourtau, N., et al. (2013). Source-to-sink transport of sugar and regulation by environmental factors. *Front. Plant Sci.* 4. doi: 10.3389/fpls.2013.00272

Miettinen, K. (1998). "Nonlinear multiobjective optimization," in *Part of the book series: International Series in Operations Research & Management Science (ISOR, volume 12)* (New York, USA: Springer US). doi: 10.1007/978-1-4615-5563-6

Moon, K.-B., Park, J.-S., Park, Y.-I., Song, I.-J., Lee, H.-J., Cho, H. S., et al. (2020). Development of systems for the production of plant-derived biopharmaceuticals. *Plants* 9. doi: 10.3390/plants9010030

Murashige, T., and Skoog, F. (1962). A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol. Plant* 15, 473–497. doi: 10.1111/j.1399-3054.1962.tb08052.x

O'Brien, S. F., and Yi, Q. L. (2016). How do I interpret a confidence interval? *Transfusion* 56, 1680–1683. doi: 10.1111/trf.13635

Prakash, G., and Srivastava, A. K. (2006). Modeling of azadirachtin production by Azadirachta indica and its use for feed forward optimization studies. *Biochem. Eng. J.* 29, 62–68. doi: 10.1016/j.bej.2005.02.027

Prakash, G., and Srivastava, A. K. (2008). Production of biopesticides in an in *situ* cell retention bioreactor. *Appl. Biochem. Biotechnol.* 151, 307–318. doi: 10.1007/s12010-008-8191-6

Puad, N., Abd-Karim, K., and Mavituna, F. (2017). A model for Arabidopsis thaliana cell suspension growth and sugar uptake kinetics. *J. Teknol* 79, 5–3. doi: 10.11113/jt.v79.11331

Puad, N., and Abdullah, T. A. (2018). "Monitoring the growth of plant cells in suspension culture," in *Multifaceted protocol in biotechnology* Eds. A. Amid, S. Sulaiman, D. N. Jimat and N. F. M. Azmin (Springer Singapore: Singapore), 203–214doi: 10.1007/978-981-13-2257-0_17

Rademacher, T., Sack, M., Blessing, D., Fischer, R., Holland, T., and Buyel, J. (2019). Plant cell packs: a scalable platform for recombinant protein production and metabolic engineering. *Plant Biotechnol. J.* 17, 1560–1566. doi: 10.1111/pbi.13081

Resat, H., Petzold, L., and Pettigrew, M. F. (2009). Kinetic modeling of biological systems. *Methods Mol. Biol.* 541, 311–335. doi: 10.1007/978-1-59745-243-4_14

Sadoch, J., Pyc, M., Urbanowicz, A., Iglewski, A., and Pilarski, R. (2020). High-throughput evolutionary optimization of the induction medium towards recombinant protein production in BY-2 tobacco. *Biotechnol. Bioeng.* 118, 676–689. doi: 10.1002/bit.27594

Santos, R. B., Abranches, R., Fischer, R., Sack, M., and Holland, T. (2016). Putting the spotlight back on plant suspension cultures. *Front. Plant Sci.* 7. doi: 10.3389/fpls.2016.00297

Schenk, R. U., and Hildebrandt, A. C. (1972). Medium and techniques for induction and growth of monocotyledonous and dicotyledonous plant cell cultures. *Can. J. Bot.* 50, 199–204. doi: 10.1139/b72-026

Schillberg, S., and Spiegel, H. (2022). Recombinant protein production in plants: A brief overview of strengths and challenges. *Methods Mol. Biol.* 2480, 1–13. doi: 10.1007/978-1-0716-2241-4_1

Schinn, S.-M., Morrison, C., Wei, W., Zhang, L., and Lewis, N. E. (2021). Systematic evaluation of parameters for genome-scale metabolic models of cultured mammalian cells. *Metab. Eng.* 66, 21–30. doi: 10.1016/j.ymben.2021.03.013

Seufert, P., Schwientek, J., and Bortz, M. (2021). An Adaptive Algorithm based on High-Dimensional Function Approximation to obtain Optimal Designs. *arXiv preprint. arXiv* 2101, 6214. doi: 10.48550/arXiv.2101.06214

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *J. R. Stat. Soc Ser. B.* 36, 111–133. doi: 10.1111/j.2517-6161.1974.tb00994.x

Terashima, M., Ejiri, Y., Hashikawa, N., and Yoshida, H. (2001). Utilization of an alternative carbon source for efficient production of human alpha(1)-antitrypsin by genetically engineered rice cell culture. *Biotechnol. Prog.* 17, 403–406. doi: 10.1021/bp010024p

Thompson, J. A., Abdullah, R., and Cocking, E. C. (1986). Protoplast culture of rice (Oryza sativa L.) using media solidified with agarose. *Plant Sci.* 47, 123–133. doi: 10.1016/0168-9452(86)90059-2

Tsopanoglou, A., and Del Jiménez Val, I. (2021). Moving towards an era of hybrid modelling: advantages and challenges of coupling mechanistic and data-driven models for upstream pharmaceutical bioprocesses. *Curr. Opin. Chem. Eng.* 32, 100691. doi: 10.1016/j.coche.2021.100691

Ullisch, D. (2012). *"PhD thesis: A fundamental research of growth,metabolismand product formation of tobacco suspension cells at different scales" [Eine grundlegende Untersuchung von Wachstum, Metabolismus und Produktbildung von Tabaksuspensionszellen in unterschiedlichen Maßstäben]* (Aachen, Germany: RWTH Aachen University).

Ullisch, D. A., Müller, C. A., Maibaum, S., Kirchhoff, J., Schiermeyer, A., Schillberg, S., et al. (2012). Comprehensive characterization of two different Nicotiana tabacum cell lines leads to doubled GFP and HA protein production by media optimization. *J. Biosci. Bioeng.* 113, 242–248. doi: 10.1016/j.jbiosc.2011.09.022

Vasilev, N., Grömping, U., Lipperts, A., Raven, N., Fischer, R., and Schillberg, S. (2013). Optimization of BY-2 cell suspension culture medium for the production of a human antibody using a combination of fractional factorial designs and the response surface method. *Plant Biotechnol. J.* 11, 867–874. doi: 10.1111/pbi.12079

Villegas, A., Arias, J. P., Aragón, D., Ochoa, S., and Arias, M. (2017). First principle-based models in plant suspension cell cultures: a review. *Crit. Rev. Biotechnol.* 37, 1077–1089. doi: 10.1080/07388551.2017.1304891

Wächter, A., and Biegler, L. T. (2006). On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math. Program.* 106, 25–57. doi: 10.1007/s10107-004-0559-y

Xu, J., Ge, X., and Dolan, M. C. (2011). Towards high-yield production of pharmaceutical proteins with plant cell suspension cultures. *Biotechnol. Adv.* 29, 278–299. doi: 10.1016/j.bioteChadv.2011.01.002

Xu, J., and Zhang, N. (2014). On the way to commercializing plant cell culture platform for biopharmaceuticals: present status and prospect. *Pharm. Bioprocess.* 2, 499–518. doi: 10.4155/pbp.14.32

Yeo, H. C., Park, S.-Y., Tan, T., Ng, S. K., Lakshmanan, M., and Lee, D.-Y. (2022). Combined multivariate statistical and flux balance analyses uncover media bottlenecks to the growth and productivity of Chinese hamster ovary cell cultures. *Biotechnol. Bioeng.* 119, 1740–1754. doi: 10.1002/bit.28104