Systems & Control Transactions

Research Article - Peer Reviewed Conference Proceeding
ESCAPE 35 - European Symposium on Computer Aided Process Engineering
Ghent, Belgium. 6-9 July 2025
Jan F.M. Van Impe, Grégoire Léonard, Satyajeet S. Bhonsale,
Monika E. Polańska, Filip Logist (Eds.)

PSE PRESS

Parameter Estimation and Model Comparison for Mixed Substrate Biomass Fermentation

Tom Vinestock[a], Miao Guo[a]*

[a] King's College London, Department of Engineering, London, United Kingdom
* Corresponding Author: miao.guo@kcl.ac.uk

ABSTRACT

Most industrial fermentations in food and drink use a single, high purity sugar as a substrate. These pure substrates are more expensive and less sustainable than mixed substrates, that can be derived from agricultural byproducts such as straw. However, use of mixed substrates in fermentation leads to challenging modelling and parameter estimation problems, particularly when much academic research, intended to inform industrial applications, uses batch fermentations, while large-scale fermentation is usually continuous, thanks to its cost and productivity advantages. Our findings highlight key challenges in using batch-derived experimental data to inform models of the continuous fermentation processes used at industrial scale. Extrapolating from data obtained in batch to continuous fermentation is risky, as models with near-equivalent data-fit and predictions in a batch context give very different predictions for continuous culture. For continuous fermentations to switch to mixed substrates, we need to improve understanding of dual substrate growth in continuous fermentation, to allow for optimised process design and operation.

Keywords: Fermentation, Modelling and Simulations, Biosystems, Design Under Uncertainty, Food & Agricultural Processes, Dual Substrate Growth, Lignocellulosic Hydrolysates, Continuous Fermentation

BACKGROUND

Single Cell Protein & Agricultural Residues

Currently, animal agriculture accounts for at least 16.5% of global greenhouse gas emissions [1]. Given Net Zero targets, this highlights the need for sustainable alternatives to livestock for dietary protein production. Single cell protein (SCP) fermentation is one such alternative, offering an effective means of transforming carbohydrate-rich substrates into high-protein foodstuffs, and is significantly more sustainable than conventional animal-based protein production, with lower greenhouse gas emissions, water use and land use [2-4]. Therefore, there is a strong environmental interest in increasing consumption and production of SCP to reduce dependence on animal-derived proteins [5].

However, whereas cows and other ruminants can be fed agricultural residues, such as corn stover, wheat and rice straw, SCP fermentations generally depend on high-purity single-substrate feedstocks [6], which are expensive, and directly compete for food crops. This raises the question of how to adapt SCP fermentations to more sustainable feedstocks derived from cheap residues of agriculture and forestry.

Globally, 11 100 megatons of these residues are produced annually [7], which if used as a feedstock for SCP fermentation would be more than enough to provide for all human dietary protein needs [7]. These residues are mainly composed of lignocellulose, a mixture of fibrous lignin and the carbohydrates cellulose and hemicellulose. While hydrolysis can be used to break down this mixture into simpler components, yielding 360g fermentable sugar per kg lignocellulose [7], only about 60% of the resulting fermentable sugar is glucose, the favoured 6-carbon sugar, with the other 40% coming from xylose, a 5-carbon sugar, along with trace sugars such as mannose [7].

Batch & Continuous Fermentation

Modelling biomass growth on such a mixture of glucose and xylose is complicated and is one of the barriers to widespread adoption of lignocellulosic feedstock in SCP fermentation [8]. To design and operate a large-scale SCP fermentation using a lignocellulosic feedstock, accurate growth models that allow the growth behaviour

on mixed substrates and the resulting yields to be understood are needed. However, academic work to obtain these models often uses batch fermentation experiments [8,9], with all the nutrients available to the system added at the start, while industrial scale SCP production is usually operated as a continuous process, with biomass constantly harvested, and substrate added as it is consumed, as this affords a much higher biomass productivity for a given reactor volume [10].

Structured & Unstructured Kinetic Models

In bioprocess kinetic modelling, a distinction is made between structured and unstructured approaches. Structured approaches incorporate detailed information about metabolic pathways, and so are useful for metabolic engineering [11] but this complexity can lead to challenges in parameter and state estimation, impeding practical applications [9]. Unstructured approaches, such as the Monod model, use simpler mathematical formulations based on empirical data [9]. The benefit of these models is their greater ease of use, with many fewer state variables and parameters to estimate. As this work focuses on bioprocesses modelling for process optimisation and control, the unstructured approach is adopted.

RESEARCH OBJECTIVES

This work has two main aims. The first is to compare different unstructured models for dual substrate growth, using previously published batch fermentation data. The second aim is to evaluate these models' ability to predict continuous fermentation dynamics using parameters estimated from batch fermentation data.

The study focuses on the filamentous fungi strain *Fusarium Venenatum* A3/5, which is commercially significant as the most widely used SCP strain for food-grade protein production [12]. Unlike previous research that often examines a single model in the batch context [8,9], this study evaluates five different dual substrate growth models and extends the analysis to continuous fermentation systems. The findings are expected to enhance modelling of lignocellulose-derived mixed substrate fermentations, and illuminate the differences between batch and continuous fermentation in this context.

METHODOLOGY

Experimental Data

The data used in this work is from Banks *et al*"s [8] recent publication, in which a microplate reader and high-performance liquid chromatography were combined to measure biomass and (extracellular) glucose and xylose concentrations at regular intervals over a 120-hour batch fermentation on two glucose-xylose mixtures (15 gL^{-1} glucose: 15 gL^{-1} xylose and 20 gL^{-1} glucose:10 gL^{-1} xylose). 3 biological replicates were used, resulting in 3 sets of measurements for each timepoint. 5 different versions of each dataset were generated using leave-one-out validation, to give an indication of the uncertainty of each model's outputs.

Parameter Estimation

Optimal parameters for each model were found using SciPy's differential evolution (DE) function for parameter optimisation in combination with ODEInt, a SciPy wrapper for LSODA, for numerical integration of the dynamical batch fermentation models defined below for each set of parameters [13, 14]**.** Stochastic methods such as DE can better identify global minima in non-convex problems such as parameter estimation than traditional gradient methods [8]. 50 generations were used for the DE, with initial populations generated using Latin hypercube sampling. At the end of this process the best population member was 'polished' using L-BFGS, a gradient descent method [15]. Default values were used for all the other algorithmic parameters. Mean absolute error (MAE) was used as the cost function, as shown in Equation 1.0. As in Banks *et al.* [8], the feedstock with equal concentrations of glucose and xylose was used as a calibration dataset, while validation was conducted using the 20 gL^{-1} glucose:10 gL^{-1} xylose data.

$$MAE = \frac{1}{3N}\sum_{k=0}^{N-1}\sum_{r=1}^{3}|X_{Pred}[k,r] - X_{Meas}[k,r]| + \left|S_{1,Pred}[k,r] - S_{1,Meas}[k,r]\right| + \left|S_{2,Pred}[k,r] - S_{2,Meas}[k,r]\right| \quad (1.0)$$

In the validation dataset, unmeasurable initial states, such as initial biomass, and where applicable, initial enzyme concentrations/intensities were back-calculated from measured values, using the DE approach described previously.

Each model, parameterised based on the calibration dataset, was then used to predict steady-state continuous fermentation behaviour, assuming a single, mixed feed containing 15 gL^{-1} glucose and 15 gL^{-1} xylose. Optimal dilution rates, defined as the dilution rates leading to the maximum steady-state biomass output rate, were found using exhaustive search for each model, with ODEInt used to numerically integrate the continuous fermentation models. Substrate efficiencies were then calculated based on this operating point.

Growth Models

Five growth models are compared. Their parameters and states variables are detailed in Tables 1-3.

Single Substrate Growth Models

While there are many single substrate growth models, the Monod and Contois models are particularly widely used. The Monod model (EQ 1.1) [16] assumes that

the growth rate depends only on the substrate concentration, approaching a maximum rate of μ_{max} when the substrate concentration is much higher than the fixed threshold concentration given by K_s. The Contois model (EQ 1.2) [17] is similar to the Monod model but assumes that the substrate threshold concentration increases linearly with the biomass concentration.

$$\mu = \mu_{max}\frac{S}{S+K_s}, \qquad \mu = \mu_{max}\frac{S}{S+K_{SX}X} \qquad (1.1, 1.2)$$

Multi-Substrate Growth Models

In the multi-substrate case, the specific biomass growth rate ($\frac{\dot{X}}{X}$) can be considered the sum of the growth rates (μ_i) on each substrate, which may in general depend on the substrate concentration itself (S_i), as well as the concentrations of other substrates (S_j), and key enzymes required for the metabolism of the substrate itself (Z_i) and other substrates(Z_j), with each growth rate's upper limit given by a substrate-specific parameter (μ_{mi}).

$$\frac{\dot{X}}{X} = \sum_{i\in\{1,..N\}, j\in\{1,...,N\}, j\neq i} \mu_i\left(S_i, S_j, Z_i, Z_j\right) \qquad (1.3)$$

$$\mu_i \in [0, \mu_{m,i}]$$

As in [8,9], the consumption rate of each substrate ($\dot{S}_i$) is assumed to be specified by:

$$\dot{S}_i = \frac{-\mu_i}{Y_{xi}}X \qquad (1.4)$$

Table 1: Parameters common to all growth models considered.

Symbol	Name	Description	Unit
$\mu_{m,i}$	Maximum growth rate on substate *i*	Fixed parameter	h^{-1}
$Y_{x,i}$	Yield coefficient	Mass of biomass per mass of substrate *i* consumed	$g_{biomass}$ $g_{substrate}^{-1}$
μ_i	Current growth rate of substrate *i*	Function depending on substrate and enzyme concentrations.	h^{-1}

In a continuous system with a dilution rate of D per hour, we have, in the simpler case of two substrates:

$$\frac{\dot{X}}{X} = -D + (\mu_1 + \mu_2) \qquad (1.5)$$

$$\dot{S}_1 = \frac{-\mu_1}{Y_{x1}}X + D(S_{in1} - S_1) \qquad (1.6)$$

$$\dot{S}_2 = \frac{-\mu_2}{Y_{x2}}X + D(S_{in2} - S_2) \qquad (1.7)$$

$$\dot{y} = D * X \qquad (1.8)$$

In a batch system, $D = 0$, resulting in a further simplified system. We consider five different possible functions for μ_i. We assume this function can be split into two parts: the first dealing with single substrate dynamics, the second dealing with interaction between substrates.

Table 2: State variables employed in the dual substrate growth models

Symbol	Name	Applicability	Unit
X	Biomass Concentration	All models	gL^{-1}
S_i	Concentration of Substrate *i*	All models	gL^{-1}
y	Cumulative Biomass Output	Continuous Models	g
Z	Concentration of key enzyme required for growth on xylose	Enzyme Inhibition Model	gL^{-1}
z_i	Relative intensity of key enzyme required for growth on substrate *i*	Optimised Enzyme Production Model	%

Inhibition Models

When more than one substrate is present, there may be an interaction between the substrates. Usually, the growth rate on a mixture of substrates is less than or equal to the maximum of the single substrate maximum growth rates, because of substrate cross-inhibition [18].

Direct Inhibition Model

Vega-Ramon *et al.* [9] introduce a direct inhibition model, with an inhibition term of the form:

$$\mu_i = \mu_i' * \frac{1}{1+\frac{S_j}{K_{ij}}} \; for \; i,j \in \{1,2\}, i \neq j, \qquad (2.1)$$

where μ_i' is the growth rate on substrate *i* in the absence of substrate *j*. Banks *et al.* [8] applied this model to the fermentation of *F. Venenatum* A3/5 on a glucose-xylose mixture, with Contois growth. Glucose growth was assumed to be unaffected by xylose, giving the following equations, with glucose as the preferred substrate (S_1) and xylose the less-favoured substrate (S_2).

$$\mu_1 = \mu_{m1}\frac{S_1}{S_1 + K_{X1}X}, \qquad (2.2)$$

$$\mu_2 = \mu_{m2}\frac{S_2}{S_2 + K_{X2}X}\frac{1}{1+\frac{S_1}{K_I}} \qquad (2.3)$$

Enzyme Inhibition Model

Table 3: Model Specific Parameters

Symbol	Name	Description	Model(s)	Unit
K_s	Monod Half-Saturation	Substrate Concentration at which growth rate is half full single substrate rate.	Monod	gL-1
K_{SX}	Contois Half-Saturation	Ratio of substrate to biomass concentration giving a growth rate that is half the full rate.	Contois	$g_{substrate}$ $g_{biomass}^{-1}$
K_I	Inhibition Constant	Value of S_1 at which growth rate on S_2 is halved relative to the growth rate on S_2 in the absence of S_1.	Direct Inhibition	gL-1
K_{Z_S}	Enzyme Half Saturation	Enzyme concentration at which growth rate on S_2 is halved relative to case of excess enzyme.	Enzyme Inhibition	%
K_{Z_d}	Enzyme Degradation Rate		Enzyme Inhibition, Optimal Enzyme Production	h-1
K_{Z_c}	Enzyme Creation Rate		Enzyme Inhibition, Optimal Enzyme Production	%

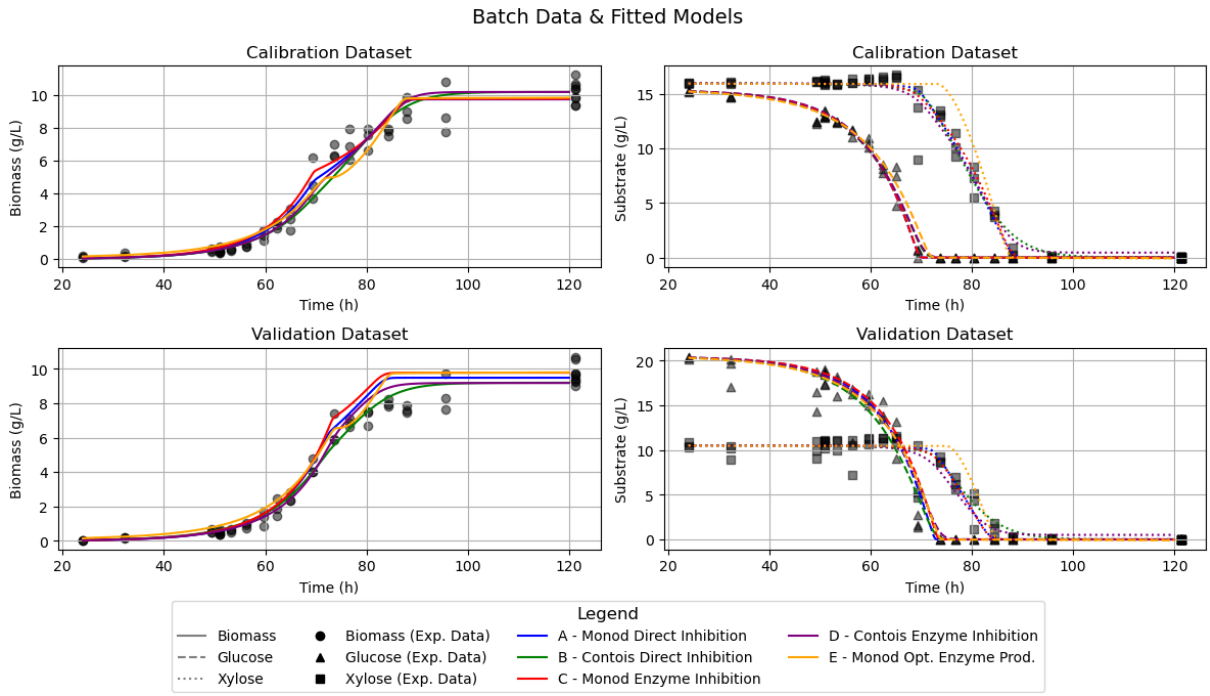

Figure 1: Fitted models, and the batch fermentation data from [8] used to fit the models.

Considering potential mechanisms for substrate cross-inhibition, instead of S_1 inhibiting S_2 growth directly, it can be modelled as reducing the production rate of an enzyme Z that is key for S_2 growth and is assumed to be rate limiting at low concentrations. The Monod-type term with a constant K_{Z_S} reflects that at a sufficient value of Z, it is no longer meaningfully rate limiting.

$$\mu_i = \mu_i * \frac{Z}{Z + K_{Z_S}} \tag{3.1}$$

$$\dot{Z} = -(K_{Z_d} + D) * Z + (K_{Z_c} * \mu * X * \frac{S_i}{S_i + K_{Si}} \frac{1}{1 + \frac{S_i}{K_{Z_{S_i}}}}) \tag{3.2}$$

This model is a simplification of one proposed by Nakamura *et al.* [19]. Using the Monod model, and assuming as previously that glucose (S_1) growth is unaffected by xylose (S_2) gives the following:

$$\mu_1 = \mu_{m1} \frac{S_1}{S_1 + K_{S1}}, \quad \mu_2 = \mu_{m2} \frac{S_2}{S_2 + K_{S2}} \frac{Z}{Z + K_{Z_S}} \tag{3.3, 3.4}$$

Optimal Enzyme Production Model

A third option is to model the concentration of the key enzymes for both S_1 and S_2 separately. This requires some modelling of cellular decision-making. In Kompala *et al.* [18], it is argued that cells optimise the allocation of resources to maximise growth rate. This leads to:

$$r_i = \mu_{mi} * z_i * S_i \tag{4.1}$$

$$v_i = \frac{r_i}{\max_j(r_j)}, \quad u_i = \frac{r_i}{\sum_j r_j} \tag{4.2, 4.3}$$

$$r_{Z_i} = K_{z_c} * \frac{S_i}{K_{si} + S_i}, \quad \mu_i = r_i * v_i \tag{4.4, 4.5}$$

$$Z_i = r_{Z_i} * u_i - K_{Z_d} * z_i - \sum_j \frac{\mu_j}{X} * z_i \tag{4.6}$$

Here, z_i is a relative enzyme intensity, defined as the enzyme mass per unit biomass, normalised by the maximum possible enzyme intensity.

RESULTS & DISCUSSION

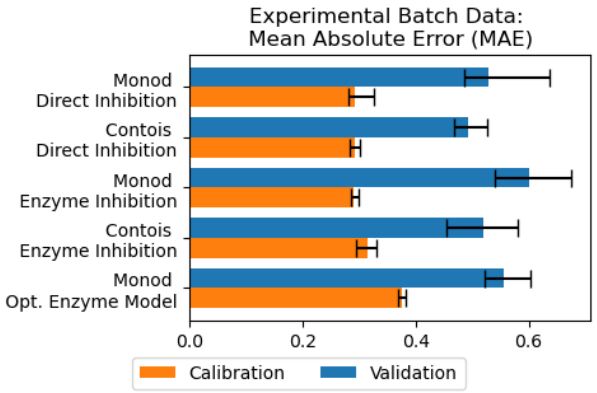

Figure 2: MAE in model predictions for calibration and validation batch data. Error bars show minimum and maximum MAE from 5 variant datasets created using leave-one-out method; the main bar shows the mean.

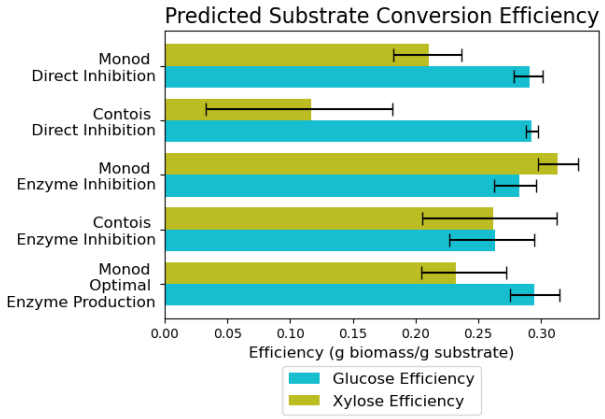

Figure 3: Predicted substrate conversion efficiency in continuous fermentation for glucose and xylose for each of the 5 dual substrate models evaluated.

Figure 1 shows that all 5 models give highly similar fits to the both the calibration and validation datasets (for legibility, only a single fit for each model is shown). This can also be seen from Figure 2, which shows the different models have similar errors. The similar values for validation error between the models suggests that the more complex models aren't overfitting.

One interpretation of the results presented in Figures 1-2 would be that the different models aren't meaningfully different. However, applying these models to continuous fermentation challenges this. Looking at Figures 3 and 4, the models make significantly different predictions for substrate yields, biomass production rates and optimal dilution rates in continuous fermentation.

Figure 3 shows the two Direct Inhibition Models predict much lower substrate yields for xylose compared to glucose, whereas the enzyme-based models predict more similar yields for the two substrates. This corresponds to the Direct Inhibition Models predicting that the glucose preference results in xylose being washed through the system in the continuous case, as glucose concentrations don't drop low enough to encourage the switch to xylose when it is diluted at the rate that maximises biomass output. The difficulty in discerning whether co-utilisation of glucose and xylose occurs in batch is the cause of much of the uncertainty in xylose efficiency, production rate and optimal dilution rate in Figures 3-4.

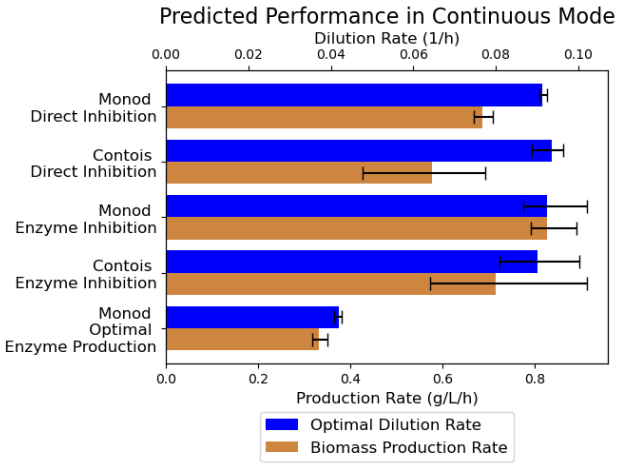

Figure 4: Maximum biomass production rates and the corresponding optimal dilution rates calculated for each model, based on the batch data parameterisations.

Figure 4 shows that the different models also predict very different biomass production rates and optimal dilution rates, although the biomass production rates and the optimal dilution rate are closely related, suggesting the steady-state biomass concentration is similar in all cases. There is a factor of about two between the largest predicted production rate, for Monod Enzyme Inhibition, and the smallest, for Monod Optimal Enzyme Production, which demonstrates the difficulty of extrapolating to continuous culture from batch fermentation results.

Additionally, differences, or uncertainties, between models exists alongside the parametric uncertainty within a model. For biomass production rate, optimal dilution rate and xylose efficiency, the inter-model uncertainty, characterised by the difference between model

mean values, exceeded the intra-model uncertainty, characterised by the range of values for one model, by factors of 3.2, 5.6 and 2.4, respectively.

Limitations

While this work focuses on the uncertainty introduced from extrapolating from batch fermentation data to continuous fermentation performance, the change in reactor volume, and the consequent effects on mixing and gas-transfer, are another major source of uncertainty in the transition from lab scale to industry scale [20]. Additionally, the higher shear stresses from the more vigorous mixing required in large reactor can reduce cell viability and growth [21]. These effects are not addressed in this work, and the uncertainties they introduce would be in addition to the model uncertainty discussed here.

CONCLUSIONS & FUTURE WORK

This research demonstrates the significant uncertainty generated by using batch fermentation data to assess and parameterise models intended for application to large-scale continuous fermentations. As a result, it suggests a need for greater focus on continuous fermentation within academia, in order to support industrial adoption of lignocellulosic feedstocks. It is relevant to academics and professionals working on design and optimisation of continuous fermentation, as it shows the risks of extrapolating from batch to continuous fermentation.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge funding from Quorn and the BBSRC, under project reference 2725958.

REFERENCES

1. Twine R. Emissions from Animal Agriculture—16.5% Is the New Minimum Figure. *Sustainability* (2021)
2. Good Food Institute, Fermentation: State of the industry report. (2021)
3. Nijdam D, Rood T*et al.* The price of protein: Review of land use and carbon footprints from life cycle assessments of animal food products and their substitutes. *Food Policy* (2012)
4. Carbon Trust. Quorn Footprint Comparison Report (2023)
5. Li YP, Ahmadi F. Recent advances and challenges in single cell protein (SCP) technologies for food and feed production. *npj Sci Food* (2024)
6. Vlaeminck E, Uitterhaegen E *et al.* Single-cell protein production from industrial off-gas through acetate: techno-economic analysis for a coupled fermentation approach. *Fermentation* (2023)
7. Piercy E, Verstraete W *et al.* A sustainable waste-to-protein system to maximise waste resource utilisation for developing food- and feed-grade protein solutions. *Green Chemistry* (2023)
8. Banks M, Taylor M, and Guo M. High throughput parameter estimation and uncertainty analysis applied to the production of mycoprotein from synthetic lignocellulosic hydrolysates. *Curr. Res. Food Sci.* (2024)
9. Vega-Ramon F, Zhu X, *et al.* Kinetic and hybrid modeling for yeast astaxanthin production under uncertainty. *Biotechnol. Bioeng.* (2021)
10. Olszewska-Widdrat A, Alexandri M, *et al.* Batch and continuous lactic acid fermentation based on a multi-substrate approach. *Microorganisms* (2020)
11. Ramkrishna D, Song HS. Dynamic models of metabolism: Review of the cybernetic approach. *AIChE Journal* (2012)
12. Ritala A, Häkkinen ST *et al.* Single cell protein—state-of-the-art, industrial landscape and patents 2001–2016. *Front. Microbiol.* (2017)
13. Virtanen P, Gommers R *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods* (2020)
14. Hindmarsh AC, Petzold, LR. LSODA, Ordinary Differential Equation Solver for Stiff or Non-Stiff System. OECD/NEA (2005).
15. Liu DC, Nocedal J. On the limited memory BFGS method for large scale optimization. *Math. Program* (1989)
16. Monod, J. The growth of bacterial cultures. *Annu. Rev. Microbiol.* (1949)
17. Contois DE. Kinetics of bacterial growth: relationship between population density and specific growth rate of continuous cultures. *J. Gen. Microbiol.* (1959)
18. Kompala DS, Ramkrishna D *et al.* Investigation of bacterial growth on mixed substrates: Experimental evaluation of cybernetic models. *Biotechnol. Bioeng.* (1986)
19. Nakamura Y, Sawada T *et al.* Stability analysis of continuous culture in diauxic growth. *J. Ferment. Bioeng.* (1996)
20. Garcia-Ochoa F, Gomez E. Bioreactor scale-up and oxygen transfer rate in microbial processes: An overview. *Biotechnol. Adv.* (2009)
21. Wang C, Lan CQ. Effects of shear stress on microalgae – A review. *Biotechnol. Adv.* (2018)