



HAL
open science

A universal dynamical metabolic model representing mixotrophic growth of *Chlorella* sp. on wastes

Bruno Assis Pessi, Caroline Baroukh, Anais Bacquet, Olivier Bernard

► To cite this version:

Bruno Assis Pessi, Caroline Baroukh, Anais Bacquet, Olivier Bernard. A universal dynamical metabolic model representing mixotrophic growth of *Chlorella* sp. on wastes. *Water Research*, 2023, 229, pp.119388. 10.1016/j.watres.2022.119388 . hal-03920752

HAL Id: hal-03920752

<https://hal.science/hal-03920752>

Submitted on 3 Jan 2023

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

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

Copyright

A universal dynamical metabolic model representing mixotrophic growth of *Chlorella* sp. on wastes

Bruno Assis Pessi^a, Caroline Baroukh^b, Anais Bacquet^c, Olivier Bernard^{a,c}

^a Biocore, INRIA, Université Côte d'Azur, Sophia Antipolis, France

^b LIPME, Université de Toulouse, INRAE, CNRS, Castanet Tolosan, France

^c LOV, UMR 7093, Sorbonne university, CNRS, Villefranche-sur-mer

Abstract

An emerging idea is to couple wastewater treatment and biofuel production using microalgae to achieve higher productivities and lower costs. This paper proposes a metabolic modelling of *Chlorella* sp. growing on fermentation wastes (blend of acetate, butyrate and other acids) in mixotrophic conditions, accounting also for the possible inhibitory substrates. This model extends previous works by modifying the metabolic network to include the consumption of glycerol and glucose by *Chlorella* sp., with the goal to test the addition of these substrates in order to overcome butyrate inhibition. The metabolic model was built using the DRUM framework and consists of 188 reactions and 173 metabolites. After a calibration phase, the model was successfully challenged with data from 122 experiments collected from scientific literature in autotrophic, heterotrophic and mixotrophic conditions. The optimal feeding strategy estimated with the model reduces the time to consume the volatile fatty acids from 16 days to 2 days. The high prediction capability of this model opens new routes for enhancing design and operation in waste valorisation using microalgae.

Keywords: Chlorella, metabolic modelling, heterotrophy, mixotrophy, diauxic growth, dynamical modelling

1. Introduction

Microalgae have been extensively studied during the past decade. Some species are capable of producing lipids or carbohydrates that can in turn be converted into biofuel (Sajjadi et al., 2018). Microalgae use light energy, via photosynthesis, to fix carbon dioxide. Not only, their growth rate is much faster than that of higher plants, but they can also be cultivated in wastewater, reducing the environmental impact of bioproducts (Morales et al., 2019, Arashiro et al., 2022). An emerging idea suggests using wastewater treatment to obtain biofuels as a co-product (Barsanti and Gualtieri, 2018). Even though the production efficiency appears to be attractive,

19 many optimization steps still need to be carried out for this process to become
20 sufficiently cost-effective and environmentally-friendly (Tan et al., 2018).

21 There is already an extensive list of works about wastewater treatment with
22 microalgae. Nevertheless there is no unified framework for the modeling of this
23 process (Shoener et al., 2019). Most of these models are variants of active sludge
24 models (ASM) and anaerobic digestion models (ADM), which were originally de-
25 signed for bacterial systems (Casagli et al., 2021b, Wágner et al., 2016). In addition
26 mixotrophy is rarely considered in these models and microalgae are assumed to grow
27 autotrophically, neglecting the simultaneous use of organic compounds. In contrast
28 to these models, metabolic models can acquire the reaction yields, including pho-
29 tosynthesis and organic carbon uptake by using the knowledge of the biochemical
30 reactions taking place in the organism which is reconstructed from genomic data.
31 Despite this, less than 5% of models of water recover facilities have metabolic re-
32 constructions and usually rely on empirical yield coefficients (Shoener et al., 2019).
33 Metabolic models also have the advantage of estimating internal metabolic fluxes,
34 providing valuable information for future strain improvement via metabolic engi-
35 neering. Nevertheless, the applicability of metabolic models to predict dynamical
36 systems is constrained by the size of the metabolic network - which in genome scale
37 models usually consists of thousands of reactions and metabolites. Consequently, to
38 embed mixotrophic microalgal models in larger frameworks representing the mass
39 fluxes within the wastewater process, it is necessary to use techniques to reduce
40 the metabolic models to a reasonable size, while still keeping the phenotype of the
41 original network.

42 The development of techniques to reduce the size of metabolic networks started
43 more than 20 years ago (Singh and Lercher, 2020). Since then new methods have
44 been developed relying on a variety of different approaches, such as linear program-
45 ming, graph-based search and elementary flux modes (Singh and Lercher, 2020).
46 Our increasing understanding of genomic information led to the construction of
47 more complex genome scale networks. This has motivated the field of bioinformat-
48 ics to research new ways to reduce and analyze metabolic models, consequently new
49 methods are still being developed (Küken et al., 2021, Hameri et al., 2021). The
50 choice of the reduction method will depend on the intentions of the modeler, the
51 wished degree of flexibility of the phenotype prediction and the final size of the
52 model. In this work, we will rely on the DRUM framework (Baroukh et al., 2014,
53 2016), which although requires input from the modeler and prior knowledge about
54 the network, it is able to reproduce behaviour outside the steady state, while greatly
55 reducing the number of metabolites and reactions in the final model. On top of this,
56 such dynamic model can support a control strategy to enhance process efficiency.

57 During dark fermentation of organic wastes, anaerobic bacteria and archaea con-
58 vert complex and non-assimilable compounds into Volatile Fatty Acids (VFAs) us-

59 able by microalgae. Indeed, VFAs can support heterotrophic growth of microalgae,
60 while they use ammonium and phosphate in the wastewater as source for nitrogen
61 and phosphorus (Baroukh et al., 2017, Turon et al., 2015b,a, Gao et al., 2022). In
62 this perspective, *Chlorella sp.* was selected for its potential in associating biofuel
63 production with effluent treatment (Casagli et al., 2021b, Gao et al., 2022, Wagner
64 et al., 2016). Indeed, this species can accumulate up to 50% of its dry weight in
65 lipids, essentially in the form of triacylglycerol (TAG). In addition, its capacity to
66 grow in the main dark fermentation effluents has been demonstrated, under het-
67 erotrophic or mixotrophic conditions (Turon et al., 2015a, Lacroux et al., 2022).
68 The VFA mixture resulting from dark fermentation is typically made of a blend
69 composed of about 30% acetate and 70% butyrate. Other organic acids can also
70 be found (Rafrafi et al., 2013, Turon et al., 2015a), among which lactate, which is
71 not consumed and does not affect the growth of *Chlorella* (Turon et al., 2015a).
72 Turon et al. (2015a) first proposed a kinetic model of the consumption of butyrate
73 and acetate by *Chlorella*. Later, an extended metabolic model was proposed using
74 the DRUM framework (Baroukh et al., 2017). Here, we extend further this model
75 by including the consumption of glucose and glycerol by *Chlorella* and propose a
76 universal multi-substrate dynamic reduced metabolic model.

77 This model is the cornerstone for tackling the major issue due to the high concen-
78 tration of butyrate which is slowly consumed by microalgae and inhibited by acetate
79 (Hu et al., 2012, Turon et al., 2015b, Lacroux et al., 2022). To address this problem,
80 another organic substrate can be added to lever the inhibition effect of butyrate and
81 eventually accelerate growth in dark fermentation effluents. Microalgae will first
82 use a more efficient carbon substrate to reach a higher biomass concentration. It
83 follows that acetate will be faster consumed, and in the end butyrate. Glucose could
84 ideally play this role, but its cost is not compatible with process economics. Acetate
85 addition is more reasonable but would generate large variations in pH. Glycerol is
86 a by-product of biodiesel synthesis by transesterification. Its low cost is likely to be
87 compensated by the enhanced microalgal productivity. Developing a mixotrophic
88 multi-substrate metabolic model is the main objective of this paper, and such ap-
89 proach can be used to identify strategies to more efficiently use dark fermentation
90 products and optimize their conversion into algal biomass.

91 The model developed here represents, in detail, the growth under different au-
92 totrophic, heterotrophic and mixotrophic conditions and for four organic substrates.
93 To our knowledge, this is the first model including such a large range of potential
94 substrates in mixotrophic regimes. The model was validated in different cultivation
95 conditions using the abundant literature available on autotrophic, heterotrophic
96 or mixotrophic growth of *Chlorella*. To this end, data from 122 experiments was
97 extracted from 15 publications, amounting to more than 2600 concentration data
98 points (see Table 1), thus reaching an unprecedented level of validation. This model

Table 1: Considered experiments for each substrate. In parenthesis, the number of experiments in mixotrophic conditions out of the total experiments.

Substrate	# exp.	Species	References
Glycerol	12 (9)	<i>C. sorokiniana</i>	León-Vaz et al. (2019)
		<i>C. sp.</i>	Sen and Martin (2018)
		<i>C. protothecoides</i>	Chen and Walker (2011), O’Grady and Morgan (2011)
		<i>C. vulgaris</i>	Ma et al. (2016)
Glucose	22 (11)	<i>C. sorokiniana</i>	León-Vaz et al. (2019), Li et al. (2013, 2014)
		<i>C. protothecoides</i>	Espinosa-Gonzalez et al. (2014), Chen and Walker (2011)
		<i>C. pyrenoidosa</i>	Ogbonna et al. (1997)
Glucose/Glycerol	2 (2)	<i>C. protothecoides</i>	O’Grady and Morgan (2011)
Acetate	40 (25)	<i>C. sorokiniana</i>	León-Vaz et al. (2019), Turon et al. (2015a,b), Chen et al. (2017a,b), Xie et al. (2020a)
		<i>C. saccharophila</i>	Xie et al. (2020b)
Butyrate	10 (4)	<i>C. sorokiniana</i>	Turon et al. (2015b,a)
Acetate/Butyrate	23 (6)	<i>C. sorokiniana</i>	Turon et al. (2015b,a)
Autotrophic	13	<i>C. sorokiniana</i>	Li et al. (2014), León-Vaz et al. (2019), Turon et al. (2015b,a)
		<i>C. sp.</i>	Sen and Martin (2018)

99 is shown to support a strategy to enhance the bioconversion of VFA into microalgal
100 biomass by managing the way the different substrates are supplied.

101 2. Materials and methods

102 2.1. General Principles of the DRUM approach

103 The dynamic model development follows the DRUM (Dynamic Reduction of
104 Unbalanced Metabolism) approach. The full description and complete explanation
105 of the approach is available in Baroukh et al. (2014).

106 Briefly, the metabolism of a microorganism can be described by its metabolic
107 network composed of a set of n_r biochemical reactions (here $n_r = 188$) involving

108 n_m metabolites (here $n_m = 173$) and represented by the stoichiometric matrix $K \in$
 109 $\mathcal{R}^{n_m \times n_r}$ (see Appendix for the full list of reactions and metabolites). The biomass B
 110 is produced from a set of substrates S and excretes a set of products P. In a perfectly
 111 mixed reactor with a constant volume, the system can be described by the following
 112 set of ordinary differential equations:

$$\frac{dM}{dt} = \frac{d}{dt} \begin{pmatrix} S \\ C \\ P \\ B \end{pmatrix} = \begin{pmatrix} K_S \\ K_C \\ K_P \\ K_B \end{pmatrix} \cdot v(M) \cdot B - DM + DM_{in} = K \cdot v(M) \cdot B - DM + DM_{in} \quad (1)$$

113 where M represents the vector of the concentrations of metabolites composed of
 114 substrate (S), intracellular metabolites (C), excreted products (P) and biomass (B).
 115 M_{in} is the influent concentration of these quantities. The dilution rate of the reactor
 116 (ratio of influent flow rate over the reactor volume) is D ($D = 0$ for a batch process).
 117 All the concentrations are expressed as total concentrations in the solution. $v \in \mathcal{R}^{n_r}$
 118 is the reaction kinetic vector, while the matrices $K_S \in \mathcal{R}^{n_S \times n_r}$, $K_C \in \mathcal{R}^{n_C \times n_r}$,
 119 $K_P \in \mathcal{R}^{n_P \times n_r}$ and $K_B \in \mathcal{R}^{1 \times n_r}$ correspond, respectively, to the stoichiometric
 120 matrices of substrates S, products P, intracellular metabolites C and biomass B
 121 ($n_S + n_C + n_P + 1 = n_m$).

122 In metabolic models, intracellular metabolites are generally assumed to be quasi-
 123 stationary ($\frac{dC}{dt} = K \cdot v = 0$), i.e. they are assumed to be consumed as soon as they
 124 have been synthesised. However, in the case of microalgae, this hypothesis has
 125 proven to be false for certain of its metabolites (denoted A) during mixotrophic or
 126 autotrophic growth (Baroukh et al., 2014, 2017). The DRUM method (Baroukh
 127 et al., 2014), consists in dividing the metabolic network into n quasi-stationary
 128 subnetworks ($K = (K_{SN_1}, \dots, K_{SN_n})$, $K_{SN_i} \cdot v_{SN_i} = 0$ for $i \in 1, \dots, n$). These are
 129 linked by the A metabolites that are, in contrast, non-stationary, can accumulate
 130 and be later consumed. This division into subnetworks is justified by the presence of
 131 metabolic pathways that correspond to metabolic functions, to reaction groups that
 132 are regulated simultaneously and to the presence of compartments within the cell.
 133 Cellular mechanisms are therefore employed for assessing the subnetwork. Hence,
 134 the A metabolites can either be found at the junction of several metabolic pathways,
 135 or they can be transported from one compartment to another, or they can be final
 136 products that accumulate in the cell. The system of ordinary differential equations
 137 (1) therefore becomes:

$$\frac{dM'}{dt} = \frac{d}{dt} \begin{pmatrix} S \\ A \\ P \\ B \end{pmatrix} = \begin{pmatrix} K_S \\ K_A \\ K_P \\ K_B \end{pmatrix} \cdot \alpha \cdot B = K' \cdot \alpha \cdot B - DM' + DM'_{in} \quad (2)$$

138 With K' the stoichiometric matrix of macroscopic reactions obtained through the
139 analysis of elementary modes (Schuster et al., 1999) on the subnetworks, and α
140 the kinetic vector associated to these macroscopic reactions. B now represents the
141 structural biomass, i.e. the fraction of biomass that does not contain the inert
142 compartments of reserve A . The total biomass can be deduced using a mass balance
143 of the elemental compounds (C, N, P, O, ...).

144 *2.2. Construction of the model*

145 The core of the metabolic network from Baroukh et al. (2017) has been used and
146 modified in order to add the glucose and glycerol consumption pathways (Figure
147 1). This network contains the central autotrophic, mixotrophic and heterotrophic
148 metabolic pathways including photosynthesis, glycolysis, the pentose phosphate
149 pathway, the Krebs cycle, oxidative phosphorylation and the synthesis of chloro-
150 phyll, carbohydrates (e.g. starch), amino-acids and nucleotides. The synthesis
151 pathways of macromolecules such as proteins, lipids, starch, DNA, RNA as well
152 as the functional biomass are represented through macroscopic reactions.

153 The DRUM method requires the partitioning of the metabolism into subnet-
154 works as well as the identification of the metabolites, in between the subnetworks,
155 which can accumulate. The subnetworks are defined by their metabolic function
156 and/or their affiliation to a cellular compartment. Different partitions among the
157 188 reactions have been tested, with a view to select the one which minimized the
158 number of parameters to assess while providing a correct representation of the ex-
159 perimental data. The best result was obtained when the network is divided into four
160 subnetworks (Figure 1), corresponding to, 1) the glyoxysome, 2) the chloroplast, 3)
161 the absorption of glycerol and 4) the synthesis of biomass. The glyoxysome and
162 chloroplast subnetworks remain unchanged in comparison with the initial Baroukh
163 et al. (2017) model.

164 The macroscopic reactions associated to each subnetwork are deduced from the
165 elementary mode analysis (Klamt and Stelling, 2003). The Matlab “efmtool” was
166 run to calculate the Elementary Flux Modes (EFMs) (Terzer and Stelling, 2008). In
167 total, 86, 142 EFMs including 3, 310 associated to futile cycles (dissipation of carbon
168 substrate in the form of CO_2) have been achieved. These macroscopic reactions are
169 further used to determine the mass fluxes in the different parts of the network by
170 assembling reactions belonging to the same kinetics.

171 *2.3. Analysis of the sub-networks*

172 *2.3.1. Motivations*

173 In this section, we present the macroscopic reactions which result from the re-
174 duction of each subnetwork by the computation of the EFMs. As recommended
175 by Baroukh et al. (2014), the reaction kinetics must be mathematically represented

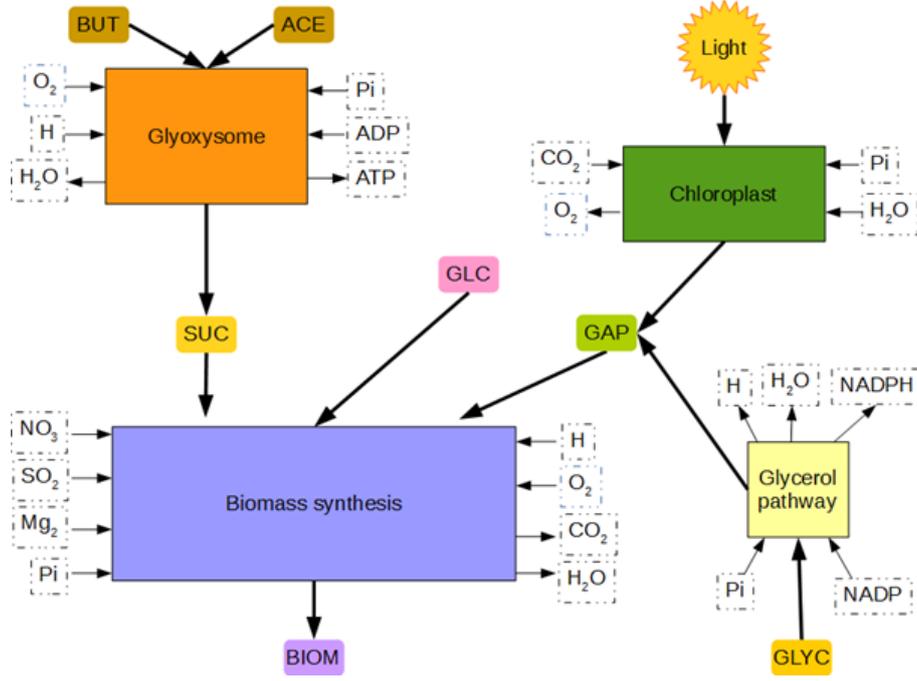


Figure 1: Considered metabolic subnetworks to represent growth of *Chlorella* on a mixture of glycerol, glucose, acetate and butyrate

	Subnetwork	Macroscopic reaction
MR_1	Glyoxysome	$2 \text{ ACE} + 3.5 \text{ H} + 0.5 \text{ O}_2 \rightarrow \text{SUC} + 0.5 \text{ H}_2\text{O}$
MR_2		$\text{BUT} + 7\text{H} + 1.5\text{O}_2 \rightarrow \text{SUC} + 5\text{H}_2\text{O}$
MR_3	Chloroplast	$\text{Light} + 3\text{CO}_2 + 2\text{H}_2\text{O} + \text{Pi} \rightarrow \text{GAP} + 3\text{O}_2$
MR_4	Glycerol pathway	$\text{GLY} + \text{Pi} \rightarrow \text{GAP} + \text{H}_2\text{O}$
$MR_5(\text{NH}_4)$	Biomass synthesis	$4.15 \text{ GAP} + 2.54\text{O}_2 + 0.99\text{NH}_4 + 0.02\text{SO}_4 + 0.01\text{Mg}_2 \rightarrow \text{B} + 0.99\text{H} + 2.90\text{H}_2\text{O} + 3.92\text{CO}_2 + 4.02\text{Pi}$
$MR_5(\text{NO}_3)$		$4.64\text{GAP} + 2.04\text{O}_2 + 0.99\text{NO}_3 + 0.98\text{H} + 0.02\text{SO}_4 + 0.01\text{Mg}_2 \rightarrow \text{B} + 5.39\text{CO}_2 + 2.90\text{H}_2\text{O} + 4.51\text{Pi}$
$MR_6(\text{NH}_4)$	Biomass synthesis	$4.15 \text{ SUC} + 7.30\text{H} + 4.61\text{O}_2 + 0.99\text{NH}_4 + 0.12\text{Pi} + 0.02\text{SO}_4 + 0.01\text{Mg}_2 \rightarrow \text{B} + 7.04\text{H}_2\text{O} + 8.06\text{CO}_2$
$MR_6(\text{NO}_3)$		$4.90\text{SUC} + 5.28\text{O}_2 + 0.99\text{NO}_3 + 0.12\text{Pi} + 10.78\text{H} + 0.02\text{SO}_4 + 0.01\text{Mg}_2 \rightarrow \text{B} + 11.07\text{CO}_2 + 8.31\text{H}_2\text{O}$
$MR_7(\text{NH}_4)$	Biomass synthesis	$2.07 \text{ GLC} + 2.54\text{O}_2 + 0.99\text{NH}_4 + 0.12\text{Pi} + 0.02\text{SO}_4 + 0.01\text{Mg}_2 \rightarrow \text{B} + 3.91\text{CO}_2 + 7.04\text{H}_2\text{O} + 0.99\text{H}$
$MR_7(\text{NO}_3)$		$2.34\text{GLC} + 2.14\text{O}_2 + 0.99\text{NO}_3 + 0.12\text{Pi} + 0.98\text{H} + 0.02\text{SO}_4 + 0.01\text{Mg}_2 \rightarrow \text{B} + 5.49\text{CO}_2 + 7.63\text{H}_2\text{O}$

Table 2: List of the macroscopic reactions in each respective subnetwork. For biomass production, the stoichiometric values differ if the nitrogen source is nitrate or ammonium.

176 using minimal hypotheses, and when possible applying a mass action law. A list of
 177 all sub-networks and the macroscopic reactions can be found at Table 2.

178 2.3.2. Glyoxysome subnetwork

179 The glyoxysome pathway consists of 26 reactions, from which 8 are exchange
 180 reactions. The glyoxysome is the peroxysome compartment where the glyoxylate
 181 cycle occurs. Here carbon compounds are converted to succinate, also allowing the
 182 production of glucose from lipids. In this compartment, fatty acids can be used as
 183 a source of energy and carbon for growth is produced when no photosynthesis takes
 184 place. Two EFMs have been achieved for this subnetwork (MR1 and MR2). In the
 185 glyoxysome, butyrate and acetate are converted into acetyl-CoA, which is in turn
 186 converted, via the glyoxylate cycle, into succinate. The succinate then enters the
 187 cytosol and is injected into the Krebs cycle, thus producing the different metabolites
 188 necessary for the synthesis of biomass.

189 Butyrate is known for inhibiting algal growth under heterotrophic and mixotrophic
 190 conditions (Turon et al., 2015a). Furthermore, acetate inhibits the absorption of
 191 butyrate, thus leading to diauxic growth (Turon et al., 2015a). Thereby, Michaelis-
 192 Menten kinetics have been proposed to describe the absorption of acetate (α_{MR1}).

$$\alpha_{MR1} = \frac{k_{MR1} \cdot ACE}{K S_{MR1} + ACE} \quad (3)$$

193 For butyrate (α_{MR2}) Haldane kinetics have been chosen with an inhibition by
 194 acetate term.

$$\alpha_{MR2} = \frac{k_{MR2} \cdot BUT}{BUT + \frac{k_{MR2}}{\beta_{MR2}} \cdot \left(\frac{BUT}{S_{optMR2}} - 1\right)^2} \frac{k_d}{(ACE + k_d)} \quad (4)$$

195 2.3.3. Chloroplast subnetwork

196 The chloroplast subnetwork is composed of 21 reactions, from which 7 are ex-
 197 change reactions. The glycerate-3-phosphate produced by photosynthesis is assumed
 198 to be transferred from the chloroplast towards the cytosol where it can be converted
 199 by glycolysis into glucose-6-phosphate or pyruvate. These metabolites are essential
 200 for the synthesis of functional biomass.

201 For autotrophic growth, light drives the reaction rate. When algae are grow-
 202 ing on a turbid medium like wastewater, the average light intensity stays low and
 203 the local photoinhibition impact can be neglected (Martínez et al., 2018). Also un-
 204 der mixotrophic conditions with elevated concentration of carbon substrates, large
 205 biomass densities can be reached. Dense microalgal cultures strongly attenuate
 206 light. In line with Baroukh et al. (2017), photosynthesis rate is assumed to be lin-
 207 early depending upon the average light intensity I_μ in the culture (see Equation
 208 6).

209 Moreover, light attenuation in the culture medium is expected to follow the Beer-
 210 Lambert law. The light intensity at depth z depends on the incident light I_0 and
 211 the extinction coefficient α due to the biomass (the turbidity of the medium without
 212 algae is negligible):

$$I(z) = I_0 \cdot e^{-\alpha \cdot z \cdot B} \quad (5)$$

213 The average light intensity of the culture medium in the reactor of depth L is
 214 given as follows (with $\beta_{MR3} = \alpha \cdot L$):

$$I_\mu = \frac{I_0}{L} \int_0^L e^{-\alpha \cdot B \cdot z} dz = \frac{I_0(1 - e^{-\beta_{MR3} \cdot B})}{\beta_{MR3} \cdot B} \quad (6)$$

215 The kinetics in the chloroplast subnetwork is finally given by:

$$\alpha_{MR3} = k_{MR3} \cdot I_\mu \quad (7)$$

216 2.3.4. Glycerol absorption subnetwork

217 The glycerol pathway subnetwork consists of 5 core reactions, plus the exchange
 218 reactions. Only one EFM was found for the glycerol absorption subnetwork (MR4).

219 The glycerol in the medium is transferred to the cytosol. Within three steps,
 220 it is transformed into glycerate-3-phosphate. During glycolysis, this glycerate-3-
 221 phosphate is then used for the synthesis of precursor metabolites that are in turn
 222 required for the synthesis of functional biomass. Since inhibition has been observed
 223 for glycerol assimilation (Chen and Walker, 2011, Ma et al., 2016, Liang et al., 2009),
 224 a Haldane reaction kinetics with inhibition was chosen (α_{MR4}):

$$\alpha_{MR4} = \frac{k_{MR4} \cdot GLY}{GLY + \frac{k_{MR4}}{\beta_{MR4}} \cdot \left(\frac{GLY}{Sopt_{MR4}} - 1\right)^2} \quad (8)$$

225 2.3.5. Functional biomass synthesis subnetwork

226 The reactions for the synthesis of lipids, proteins, DNA, RNA, chlorophyll and
 227 carbohydrates are all lumped together in the functional biomass synthesis subnet-
 228 work. This subnetwork includes glycolysis, the Krebs cycle, oxidative phosphoryla-
 229 tion, the pentose phosphate pathway, carbohydrate, lipid, amino-acid and nucleotide
 230 synthesis, as well as the assimilation of nitrogen, sulphur and glucose. In total there
 231 are 141 reactions in the functional biomass subnetwork.

232 This subnetwork generated 86,167 EFMs, including 3310 that did not produce
 233 biomass. Nearly all of the calculated EFMs are part of the biomass synthesis net-
 234 work. They can be sorted by using a similar method to the FBA (Flux Balance
 235 Analysis). The standard hypothesis supporting FBA is that evolution has selected
 236 metabolisms maximising biomass growth on each substrate (Orth et al., 2010), or

237 equivalently, minimising the loss of carbon as CO_2 . Therefore, for each substrate, the
 238 EFM presenting the highest GAP/BIOM, SUC/BIOM and GLC/BIOM yields were
 239 selected. In this way, the use of GAP, SUC or GLC for the synthesis of biomass
 240 resulted from three macroscopic reactions (MR5, MR6 and MR7). The yield of
 241 biomass on the carbon substrate depends on the nitrogen source (NH_4 or NO_3).
 242 Table 1 shows the resulting macroscopic reactions for both cases.

243 Glycerate-3-phosphate originates from the chloroplast and from the assimila-
 244 tion of glycerol. It is injected into the glycolysis so as to produce the necessary
 245 metabolites for growth (α_{MR5}). The kinetics is supposed to be linear with respect
 246 to glycerate-3-phosphate (GAP):

$$\alpha_{ME5} = k_{MR5} \cdot GAP \quad (9)$$

247 Succinate originates from the glyoxysome. It enters the Krebs cycle, thus also
 248 leading to the production of metabolites required for growth (α_{MR6}). Assuming a
 249 linear kinetics we get:

$$\alpha_{MR6} = k_{MR6} \cdot SUC \quad (10)$$

250 Glucose in the medium is transferred to the cytosol where it enables the produc-
 251 tion of biomass. Glucose is inhibiting at high concentration (osmotic stress), and
 252 its consumption is assumed to follow a Haldane kinetics (α_{MR7}) (Azma et al., 2011,
 253 Wu and Shi, 2007, Liang et al., 2009). MR7:

$$\alpha_{MR7} = \frac{k_{MR7} \cdot GLC}{GLC + \frac{k_{MR7}}{\beta_{MR7}} \cdot \left(\frac{GLC}{S_{optMR7}} - 1\right)^2} \quad (11)$$

254 2.4. Global dynamics of the network

255 Finally, the dynamical evolution of the metabolic fluxes associated to the 188
 256 considered metabolic core reactions can be derived from a system with 17 ordinary
 257 differential equations comprising 17 metabolites and 7 macroscopic reactions:

$$\frac{dM''}{dt} = \frac{d \begin{pmatrix} S \\ A \\ B \end{pmatrix}}{dt} = K'' \cdot \alpha \cdot B - DM'' + DM''_{in} \quad (12)$$

258 Where M'' is the metabolite vector (17x1) comprising the substrates S, the
 259 metabolites that can accumulate A (SUC and GAP) and the functional biomass
 260 B. K' is the stoichiometric matrix (17x7) of the macroscopic reactions and α the as-
 261 sociated kinetics vector (7x1). It is worth noting that, even if the model results from
 262 a reduction process through the DRUM approach, it can still predict the evolution
 263 of the 188 metabolic fluxes in the cell:

264 Moreover, the total biomass comprising the functional biomass and the metabo-
 265 lites A can be described as follows:

$$X_z(t) = \sum_A Z_A \cdot A(t) + Z_B \cdot B(t) \quad (13)$$

266 Where Z is a chemical element ($Z \in \{C; N; O; H; P; \dots\}$), Z_A and Z_B are the
 267 number of chemical elements Z per mole of metabolites A and biomass B, $A(t)$ and
 268 $B(t)$ are the concentrations of A and B at time t $X_Z(t)$ is the concentration of the
 269 chemical element in the total biomass X at time t.

270 Finally, the metabolic fluxes within the whole network can be derived from the
 271 α kinetics and the elementary modes associated to the $E_{SN_i}, i \in 1, 2, 3$ subnetworks:

$$v = \begin{pmatrix} v_{SN_1} \\ \dots \\ v_{SN_k} \end{pmatrix} = \begin{pmatrix} E_{SN_1} \cdot \alpha_{SN_1} \\ \dots \\ E_{SN_k} \cdot \alpha_{SN_k} \end{pmatrix} \quad (14)$$

272 2.5. Sensitivity analysis

273 We calculate the sensitivity coefficient for the model parameters as defined in
 274 Bernard et al. (2001):

$$\sigma_y^{\Delta p} = \frac{1}{t_f} \int_0^{t_f} \frac{y(p + \Delta p, x_0, u, \tau) - y(p, x_0, u, \tau)}{y(p, x_0, u, \tau)} d\tau \quad (15)$$

275 where y is the simulated output at time τ with parameter set p , initial condition
 276 x_0 and input variables u (e.g. light intensity and dilution rate). We calculate
 277 the global sensitivity using Morris's sampling method implemented in the SALib
 278 Python toolbox (Morris, 1991, Herman and Usher, 2017), replacing the standard
 279 elementary effect by the sensitivity coefficient defined above. We analyse the region
 280 between $\pm 20\%$ of the calibrated values of the parameters. The analysis is conducted
 281 separately for each of the four carbon substrates in mixotrophic conditions at a light
 282 intensity of $500 \mu\text{mol}/(\text{m}^2 \cdot \text{s})$. $\sigma_y^{\Delta p}$ is calculated as the average of the simulations
 283 considering a cultivation of 6 hours and three different initial concentration of the
 284 considered carbon substrate (0.1, 1 and 10 g/L). Table 3 shows the result of the
 285 sensibility analysis for all the parameters.

286 2.6. Reduced model calibration

287 In order to calibrate, and then validate the model, a large set of experiments
 288 from the literature have been used. In total 122 selected experiments (see Table 1)
 289 gather data on growth i) under autotrophic conditions, without any organic carbon
 290 input and submitted to light intensities ranging from 30 to $540 \mu\text{E} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$ ii)
 291 under heterotrophic conditions, without any light, and with varying concentrations

Table 3: Kinetic parameters obtained after model calibration. Results of sensibility analysis ($\sigma_y^{\Delta p}$ and standard deviation (SD) of $\sigma_y^{\Delta p}$). Parameters calibrated in this work ^A and in Baroukh et al. (2017) ^B. MB: mole of biomass

Parameter	Value	Unit	$\sigma_y^{\Delta p}$	SD $\sigma_y^{\Delta p}$	Ref.
k_{MR1}	$3.79 \cdot 10^{-1}$	$M.h^{-1}.MB^{-1}$	4.5910^{-03}	2.5210^{-03}	<i>B</i>
KS_{MR1}	$5.50 \cdot 10^{-5}$	M	9.2010^{-05}	5.2410^{-03}	<i>B</i>
k_{MR2}	$3.60 \cdot 10^{-2}$	$Mh^{-1}.MB^{-1}$	9.9610^{-05}	1.4010^{-03}	<i>B</i>
$Sopt_{MR2}$	$1.90 \cdot 10^{-5}$	M	2.4810^{-04}	1.4110^{-03}	<i>B</i>
β_{MR2}	$2.58 \cdot 10^5$	$h^{-1}.MB^{-1}$	1.8910^{-04}	1.4010^{-03}	<i>B</i>
k_{MR3}	$1.90 \cdot 10^{-3}$	$Mh^{-1}.MB^{-1}.(m^2.s.\mu mol^{-1})$	9.2110^{-04}	5.7010^{-04}	<i>B</i>
β_{MR3}	$2.48 \cdot 10^3$	MB^{-1}	-5.3810^{-04}	8.8210^{-05}	<i>B</i>
k_{MR4}	1.01	$M.h^{-1}.MB^{-1}$	5.5410^{-04}	1.8610^{-03}	<i>A</i>
$Sopt_{MR4}$	$7.00 \cdot 10^{-2}$	M	9.2710^{-04}	2.2510^{-03}	<i>A</i>
β_{MR4}	9.05	$h^{-1}.MB^{-1}$	1.1010^{-04}	2.4310^{-03}	<i>A</i>
k_{MR5}	$2.82 \cdot 10^1$	$h^{-1}.MB^{-1}$	9.7310^{-04}	1.1610^{-03}	<i>B</i>
k_D	$5.39 \cdot 10^{-10}$	M	-1.1010^{-04}	1.4310^{-03}	<i>B</i>
k_{MR6}	$2.37 \cdot 10^5$	$h^{-1}.MB^{-1}$	3.0710^{-06}	3.3310^{-03}	<i>B</i>
k_{MR7}	$6.81 \cdot 10^{-2}$	$M.h^{-1}.MB^{-1}$	2.3910^{-03}	2.1610^{-03}	<i>A</i>
$Sopt_{MR7}$	$5.07 \cdot 10^{-2}$	M	9.5410^{-05}	3.2210^{-03}	<i>A</i>
β_{MR7}	$2.95 \cdot 10^{+1}$	$h^{-1}.MB^{-1}$	4.2410^{-04}	3.2410^{-03}	<i>A</i>

292 in acetate, butyrate, glucose and glycerol, pure or combined iii) under mixotrophic
 293 conditions, with light and varying concentrations in acetate, butyrate, glucose and
 294 glycerol. Depending on the studies, different combinations of these substrates were
 295 tested.

296 Only the parameters for glucose and glycerol consumption were calibrated. All
 297 the other kinetic parameters for the macro reactions are taken from Baroukh et al.
 298 (2017). The calibration was done following a two-step process. First a stochastic
 299 global optimizer, Differential Evolution algorithm (Storn and Price, 1997), calculates
 300 the set of parameters minimizing the relative error between model and experimental
 301 data of biomass and substrate concentration over time. This parameter set is, then,
 302 used as initial point in a Markov Chain Monte Carlo sampler, which returns the
 303 parameters set inside a confidence interval (Foreman-Mackey et al., 2013). Glucose
 304 kinetics parameters were calibrated using concentration data from 5 experiments of
 305 Li et al. (2013), while glycerol parameters were fitted using data from 6 experiments
 306 of Ma et al. (2016). Table 3 shows the value of calibrated parameters and Figure 2
 307 shows the simulation of the calibrated model together with the experimental data
 308 used for calibration.

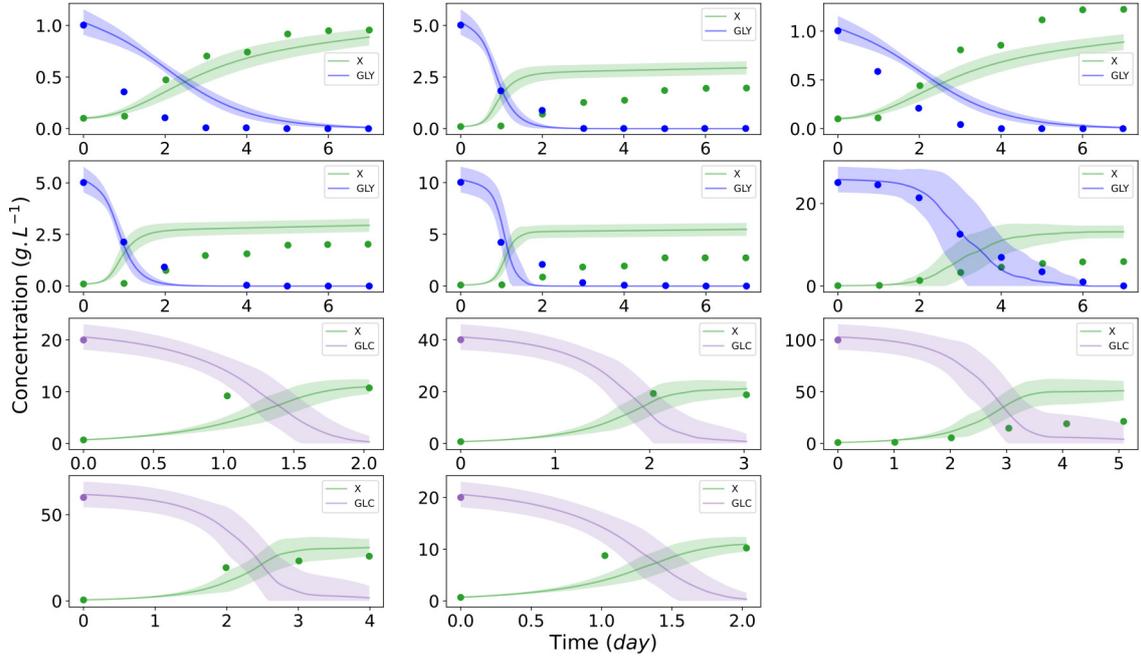


Figure 2: Comparison of the model against experimental data used for the calibration of glucose and glycerol kinetic parameters. Lines represent the average of 100 simulations and the colored region represents ± 1 standard deviation of estimated concentrations using Markov Chain Monte Carlo method.

309 2.7. Optimization of butyrate consumption

310 We consider the optimal control problem (Harmand et al., 2019), whose objective
 311 is to minimize the time t_f , where the chemical oxygen demand of the remaining
 312 waste substrates falls below the regulation threshold. The control variable is the
 313 concentration of the substrate to be added to the dark fermentation effluent (glucose
 314 or glycerol).

$$\min_{S_0} t_f : S(t_f) \leq \bar{S} \quad (16)$$

315 To solve the minimization problem we use a Nelder-Mead algorithm (Gao and
 316 Han, 2012). The output function simulates the metabolic model for a given S_0
 317 returning t_f , the time required to reach the regulation threshold.

318 3. Results and discussion

319 3.1. Validation of the model

320 The experimental data not used during the calibration stage were used to vali-
 321 date the model. A coherent set of experiments, representing various experimental

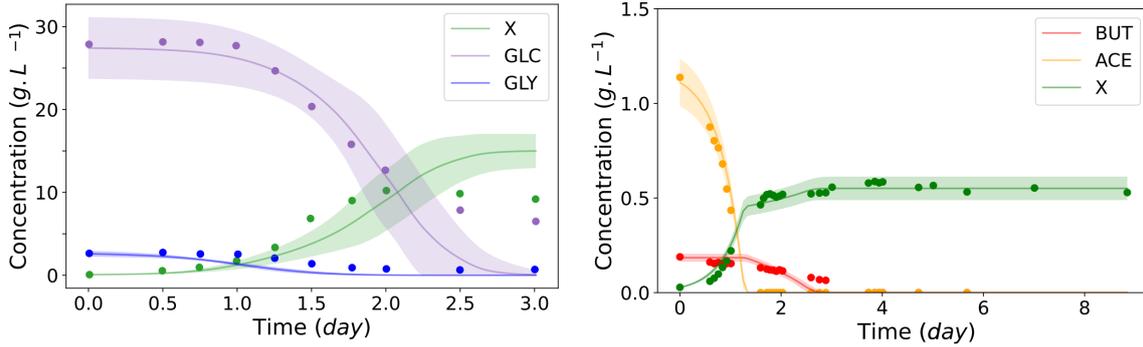


Figure 3: Model simulation and experimental data with two substrates. From left to right: glycerol and glucose; butyrate and acetate. The line represents the average of 100 simulations and the colored region represents ± 1 standard deviation of estimated concentrations using Markov Chain Monte Carlo method.

322 conditions, was kept for this validation stage (see Table 1). We used a Markov
 323 Chain Monte Carlo method to select parameters (Foreman-Mackey et al., 2013).
 324 We considered a $\pm 20\%$ uncertainty in the initial concentration of substrates. In
 325 Figure 3, the results of the model simulations are compared with the experimental
 326 data with more than one substrate.

327 As illustrated in Figure 2 for single substrates, the model efficiently predicts the
 328 production of biomass and the consumption of substrates, whether in autotrophic,
 329 heterotrophic or in mixotrophic conditions. Furthermore, the model is still accurate
 330 when there are two substrates (Figure 3). More generally, the predictive performance
 331 of the model is summarized in the Taylor diagram in Figure 4 for the whole data
 332 set. This diagram represents at the same time the standard deviation of the biomass
 333 prediction error and the Pearson correlation coefficient (Taylor, 2001). It illustrates
 334 both the centered and reduced quadratic errors between the experimental data sets
 335 and the associated simulations, as well as the correlation between the model and
 336 the data. It thus summarizes the degree of resemblance between the data and the
 337 simulations for the vast range of considered data. Indeed, the closer a data point
 338 is to (0;1), the better the model reproduces the experimental data (Taylor, 2001).
 339 Figure 5 represents all the data points versus model prediction, also demonstrating
 340 the goodness of fit of the model.

341 The results of the sensibility analysis is also show in Table 3. Most parameters
 342 have the same order of sensibility (10^{-4}) showing that all have an importance in
 343 the model. The high standard deviation of the sensibility coefficient demonstrates
 344 the intertwined influence of parameters, for example maximum uptake rate and the
 345 optimal concentration of the substrate, also and the dependence of the actual sub-
 346 strate and biomass concentration in the dynamics of the system. This demonstrates
 347 the necessity of calibrating the model in a wide range of conditions.

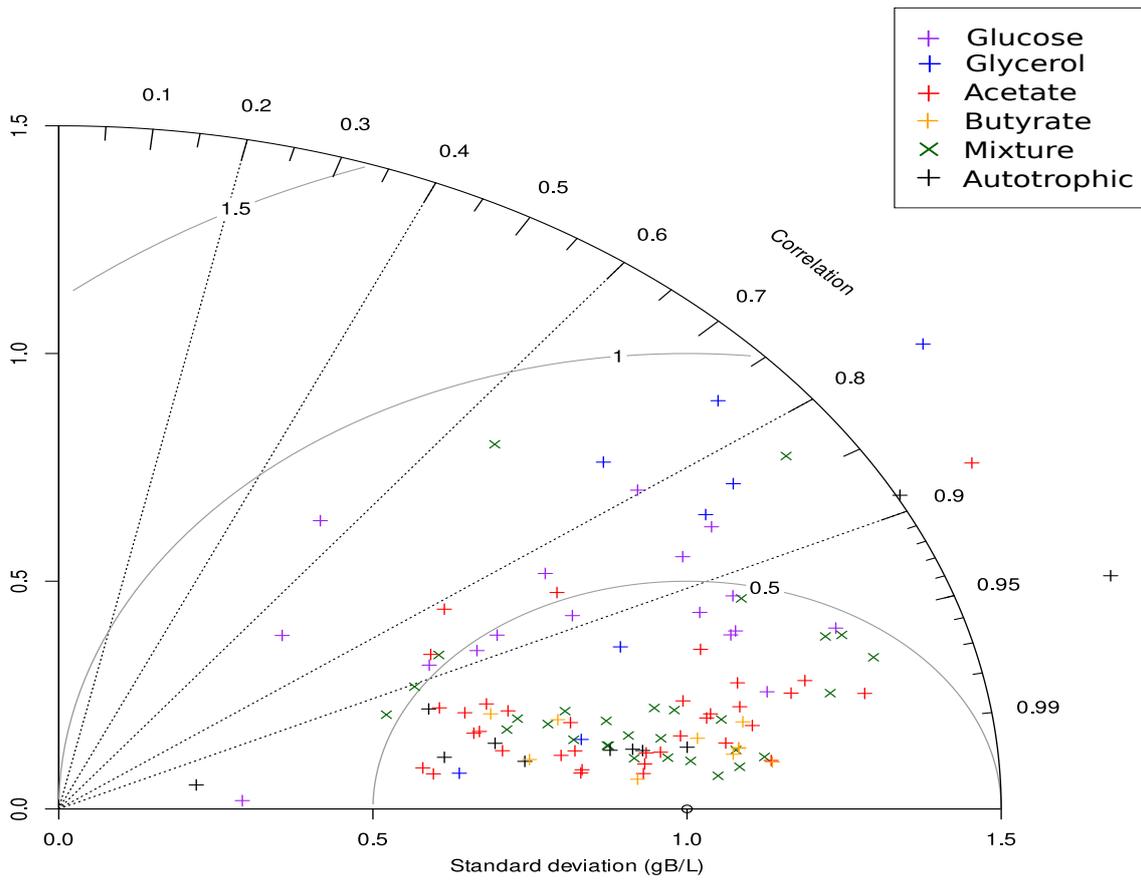


Figure 4: Taylor Diagram where each point represents the Pearson correlation coefficient and a normalized standard deviation of one experiment and model simulation. The semi circles centered at standard deviation 1.0 show the root-mean-square error.

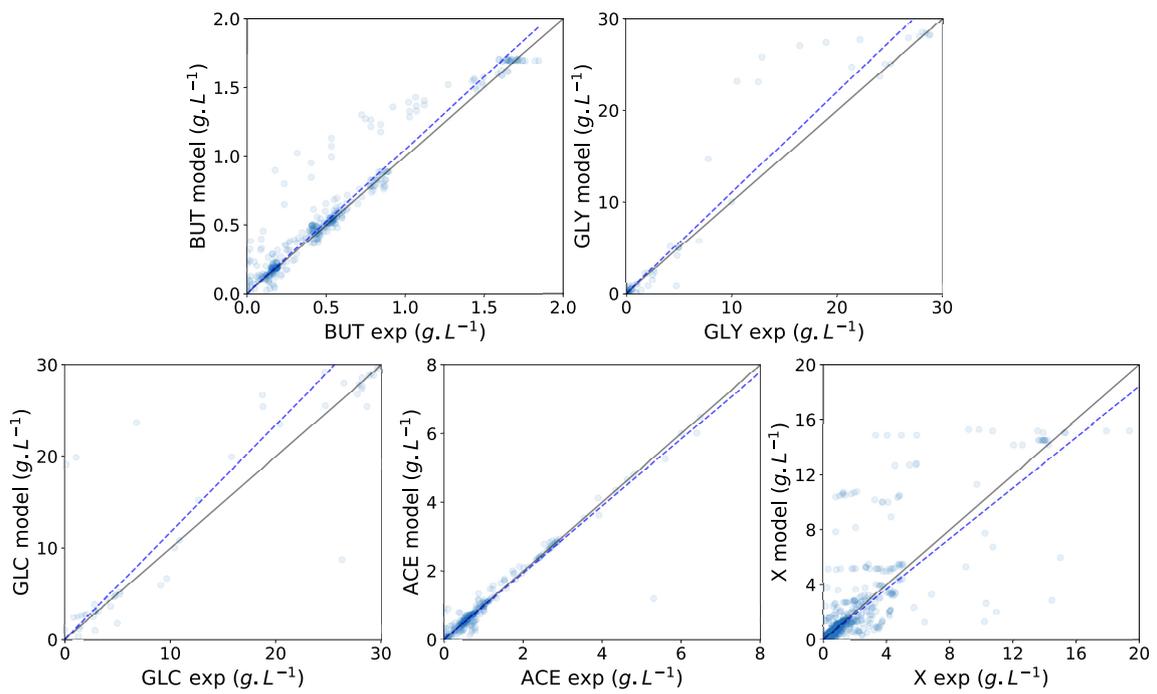


Figure 5: Validation of model predictions based on experimental data of butyrate, glycerol, glucose, acetate and biomass. All p-values for the regression are below 10^{-3} . R^2 for the lines are, respectively, 0.97, 0.71, 0.95, 0.97 and 0.64. Darker colors represent a concentration of data points.

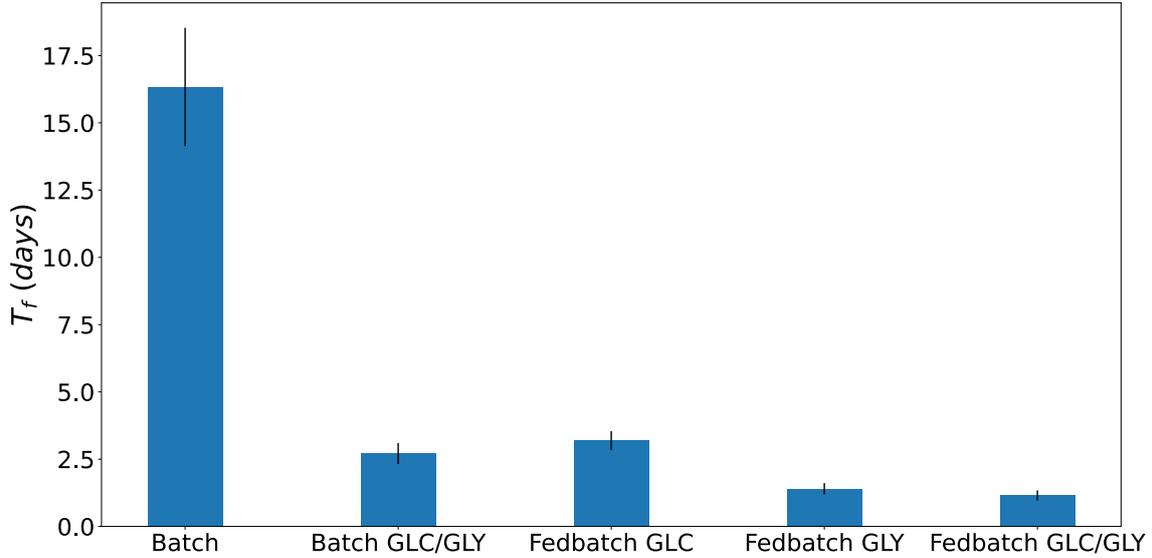


Figure 6: Time to reach the regulation threshold of a typical dark fermentation effluent using different conditions of cultivation - Batch (typical waste effluent), Batch GLC/GLY (typical waste with addition of an optimal concentration of glucose and glycerol), Fed-batch GLC, GLY, GLC/GLY (typical waste feeding, respectively, glucose, glycerol and a mixture of both.)

3.2. Optimization of microalgae growing on a mixture of dark fermentation products

Acetate and butyrate are the main volatile fatty acids (VFA) products of dark fermentation. Butyrate will inherently lead to growth inhibition, and a strategy must be found to unblock this inhibition. Here, we propose a strategy to enhance batch and fed-batch cultivation of microalgae from dark fermentation wastes, by adding glucose and glycerol. We first study mixotrophic conditions, considering a typical effluent from dark fermentation with $3.5g/L$ of butyrate and $1.7g/L$ of acetate (Lacroux et al., 2020, Ghimire et al., 2015) and a continuous light intensity of $500 \mu mol/(m^2.s)$. The objective was that the chemical oxygen demand (COD) of the effluent at the end of the effluent treatment must be below \bar{S} (here $\bar{S} = 125mg/L$) so that it can satisfy the state policies for discharge in the environment. We call this objective the regulation threshold. We solve the resulting optimal control problem in minimal time, with the objective that the COD of the remaining waste substrates falls below the regulation threshold. The control variable is the concentration of the substrate to be added to the dark fermentation effluent (glucose or glycerol). In the initial situation, without addition of organic carbon substrates, 16 days are necessary to reach the the regulation threshold.

Considering that glucose can be added, it turns out that the addition $98g/L$ of glucose reduces the time to reach the COD threshold to 4.0 days (see Figures 6 and 7). Using only glycerol takes longer: 5.6 days with an addition of $39.5g/L$ glycerol. Using a mixture of both substrates, the optimal starting concentration for the batch

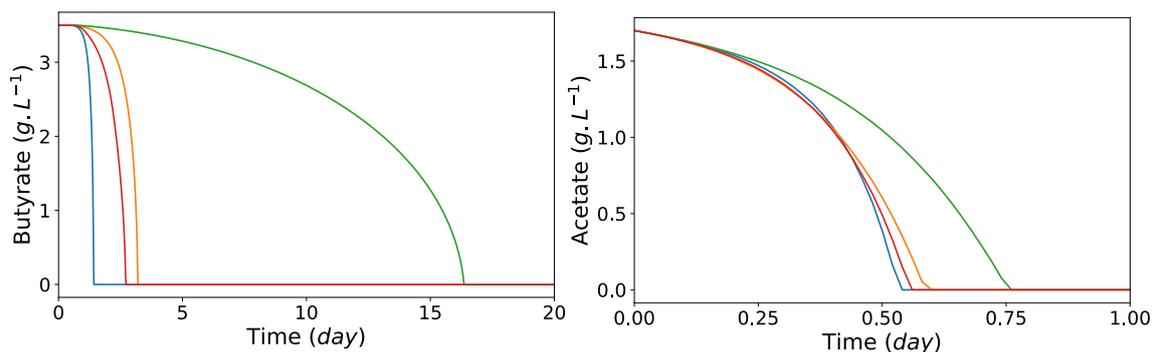


Figure 7: Concentrations of butyrate and acetate over time in different cultivation conditions. Optimal fed-batch with glycerol (blue), optimal fed-batch with glucose (red), batch with optimal addition of glycerol and glucose (orange), batch with only butyrate and acetate as substrates (green).

369 is $115g/L$ of glucose and $19g/L$ of glycerol, reducing the time to consume the VFAs
 370 to only 2.7 days.

371 Considering now a fed-batch cultivation systems instead of a batch one, the
 372 minimisation problem can be rewritten as following:

$$\min_D t_f : S(t_f) \leq \bar{S} \quad (17)$$

373 where D is the dilution rate of the inflow containing only the additional substrates,
 374 at a high concentration, so that the volume of the reactor does not change. t_f is the
 375 time to reach the regulation threshold. The optimal strategy can be approximated
 376 into a sub-optimal strategy, which would maximize the reaction rates for glucose
 377 and glycerol consumption. The strategy thus consists in computing the dilution
 378 rate such that glucose and glycerol concentrations stay constant close to the values
 379 for which the consumption of glucose and glycerol is maximum. The control problem
 380 is then reduced to finding the optimal final time of the inlet flux in the cases of a
 381 glucose, glycerol, and mixture inlet. In the case of the mixture, a fraction of 0.21 of
 382 glucose and 0.79 of glycerol is obtained, keeping glycerol at an optimal concentration.
 383 Using this control strategy, the final times for fed-batch cultivation are: 3.2 days for
 384 glucose, 1.4 days for glycerol and 1.2 days for a mixture of glucose and glycerol (see
 385 Figure 6), to be compared to the 16 days without inorganic carbon addition.

386 3.3. Analysis of metabolic maps

387 One of the advantages of metabolic model is that the main metabolic fluxes can
 388 be estimated. They are represented in Figure 8 considering the exponential growth
 389 phase. During autotrophic growth, a strong activity takes place in the chloroplast
 390 subnetwork where photosynthesis occurs. The greatest fluxes are associated to the
 391 fixation of CO_2 by RUBISCO, the conversion to 3PG (3-phosphoglycerate) and the

392 conversion of 3PG to glycerate-3-phosphate (GAP). GAP is then mainly transported
393 toward the cytoplasm, where it is injected into the glycolysis. From then on, it
394 enables the synthesis of precursor metabolites that are necessary for the synthesis of
395 functional biomass composed of proteins, DNA, RNA, chlorophyll, carbohydrates,
396 and lipids. During pure heterotrophic growth, in the dark, no reaction occurs in
397 the chloroplast subnetwork. Considering a mix of acetate or butyrate, the largest
398 fluxes are concentrated in the glyoxysome subnetwork where carbon substrates are
399 converted into succinate. The synthesised succinate is exported from the glyoxysome
400 and injected into the Krebs cycle. The latter then enables the synthesis of precursor
401 metabolites for the production of functional biomass. To ascend the glycolysis, the
402 anaplerotic pathways are in an upward direction.

403 No reaction takes place in the glyoxysome, neither in the glycerol utilization
404 pathway for growth on glucose only. Glucose carried from the medium into the
405 cytoplasm is directly injected into the upper glycolysis for the production of pre-
406 cursor metabolites necessary for the synthesis of functional biomass. Glycolysis is
407 therefore entirely in a downward direction, as the anaplerotic pathways that enable
408 the synthesis of oxaloacetate (OXA) fueling the Krebs cycle.

409 For growth on glycerol only, the greatest fluxes are located in the glycerol utiliza-
410 tion pathway subnetwork, where the uptaken glycerol is converted to glycerate-3-
411 phosphate. The glycerate-3-phosphate is then injected into the middle of glycolysis,
412 this time in an upward direction for the high glycolysis and in a downward direction
413 for the lower glycolysis. The remnants of the fluxes are similar to those observed
414 for glucose.

415 In mixotrophic conditions, when all carbon substrates are present in the medium,
416 all metabolic reactions are activated. First, diauxic growth occurs, and acetate is
417 consumed instead of butyrate. Glycerol and glucose are also used, but the flux
418 remains after the acetate is depleted. Secondly, once all acetate has disappeared,
419 butyrate is in turn consumed.

420 Figure 8 shows the metabolic fluxes after 30 hours of batch culture with opti-
421 mal initial conditions (110 *g/l* of glucose and 20 *g/L* of glycerol). At this point,
422 the acetate in the culture is practically depleted, but the concentration of butyrate
423 remains similar to the initial condition. The rate of consumption of butyrate then
424 remains stable for 2 days, and the rate rapidly increases until the concentration
425 threshold is reached. Since the optimal concentration is about 40 times lower than
426 the regulation threshold, the butyrate consumption rate does not reach the max-
427 imum value. Glycerol has a much lower concentration, while the flux of glucose
428 concentration is still high. As a result of the higher concentration of biomass and
429 the lower availability of light per cell, the metabolic fluxes in the chloroplast are
430 reduced.

431 3.4. Model limitations and perspectives

432 The reduced metabolic model efficiently represents the microalgal growth un-
433 der various substrates in heterotrophic or mixotrophic conditions. More accurate
434 predictions could probably be obtained by expanding the model to include other
435 factors such as pH, CO_2 , O_2 and temperature which were differing among the large
436 experimental data set considered in our study. According to Lacroux et al. (2020),
437 pH fluctuations when algae consume VFA can strongly impact growth and should
438 now be included in the model. Associating a pH model that accounts for the various
439 chemical species and their speciation, as proposed by Casagli et al. (2021a), would
440 allow the calculation of the pH and the concentration of the undissociated form of
441 the acids, which is the one actually taken up by the microalgae.

442 The effect of mixotrophic growth, here considered as the sum of autotrophic and
443 heterotrophic conditions, can be more subtle in some cases. There is no consen-
444 sus in the literature for a general model that fits all cases. For example, according
445 to Martínez and Orús (1991), mixotrophic growth is greater than the sum of au-
446 totrophic and heterotrophic conditions when the concentration of CO_2 is limiting,
447 since CO_2 produced by respiration can be recycled for the photosynthesis pathway.
448 Recently, circular use of CO_2 and O_2 and preferential consumption of inorganic car-
449 bon has also been evidenced for *Chlorella vulgaris*, nonetheless heterotrophic growth
450 was promoted at high O_2 concentrations (Manhaeghe et al., 2020). Furthermore,
451 under certain conditions an inverse relationship between light intensity and glucose
452 consumption has been recorded (Patel et al., 2019, Wan et al., 2011).

453 Temperature is one of the most important factors affecting microalgae growth,
454 even if only a minority of the models account for it (Shoener et al., 2019). The
455 experimental data considered here were carried out at various temperatures and
456 accounting for it would increase model accuracy. In the case of metabolic models,
457 including temperature is challenging, and will involve a large number of parameters
458 to characterize the influence of temperature on each individual reaction. The influ-
459 ence of short variations of temperature could be well estimated by the Arrhenius
460 equation, but large variations of temperature typically occurring in outdoor cultiva-
461 tion require more advanced models (Casagli and Bernard, 2022, Pessi et al., 2022).
462 For these reasons, calibrating the temperature effect would require a large amount
463 of dedicated experiments.

464 It is likely that other factors were influencing the experimental outcomes in the
465 considered data base, especially for the experiments carried out with high concen-
466 trations of glucose and glycerol. It seems that another substrate was sometimes
467 limiting growth, most probably nitrogen according to the mass balance from the
468 medium initial composition.

469 Extending the metabolic model to account for all these mechanisms is beyond
470 the scope of this paper. It will require a large number of experiments to further

471 calibrate and validate the model. Our objective was primarily to validate the model
472 across a wide range of conditions, demonstrating a strong foundation for future
473 improvements in the model, knowing that most models of water resource recovery
474 facilities are calibrated, but not validated (Shoener et al., 2019). Overall, despite
475 the simplicity of the model in its present form, it is already very efficient. Especially
476 when accounting for the diversity of strain and experimental conditions considered
477 through the 15 studied papers. The model can then, already in its present form,
478 can be used as a tool for optimizing microalgae growth on a mix of substrates.

479 Better predictions will be achieved if biochemical and also cellular level processes
480 (e.g. metabolic reactions) are considered in the next generation of water treatment
481 models (Batstone et al., 2019). Metabolic models are able to accurately predict
482 the specific VFAs produced by a mixed-culture depending on the components of
483 the input medium (Regueira et al., 2020). We could envisage the coupling of these
484 models - VFA production by mixed-culture and treatment by microalgae - to opti-
485 mize production of hydrogen constrained by the capacities of the waste treatment
486 system, or to predict in advance the necessity of adding another carbon substrate
487 such as glycerol. Metabolic knowledge could also give a strong foundation for esti-
488 mating possible medium conditions take could lead to undesirable emissions, such
489 as N_2O , and which are not well modeled through the current approaches (Casagli
490 et al., 2021a).

491 **4. Conclusion**

492 The metabolic model developed in this work for *Chlorella* accurately predicts
493 growth under autotrophic, heterotrophic and mixotrophic conditions with acetate,
494 butyrate, glucose and glycerol for more than a hundred experiments from the liter-
495 ature. Covering a large range of conditions, strains and substrates, the predictive
496 capacity of this reduced metabolic model remains remarkable. Moreover, to lift the
497 inhibition exerted by the presence of butyrate, the optimal addition of different sub-
498 strates in the medium has been predicted by the model. Optimising the cultivation
499 conditions reduce the time to reach the regulation threshold from 16 to less than 2
500 days.

501 Thanks to this approach, it will become possible to streamline two-stage wa-
502 ter treatment strategies, and to recycle, carbon, nitrogen and phosphorus into the
503 microalgal biomass. In particular, this approach highlights noteworthy synergies be-
504 tween waste molecules, and the important role of of the metabolic network for future
505 models. The anticipated management of these molecules could improve productivity
506 significantly.

507 **Acknowledgment**

508 The authors acknowledge the support of the ControlAB ANR project (ANR-20-
509 CE45-0014) and of the BIOMSA Ademe project. The UCAJEDI and EUR DS4H in-
510 vestments in the Future projects managed by the National Research Agency (ANR)
511 with the reference numbers ANR-15-IDEX-0001.

512 **References**

- 513 Arashiro, L.T., Josa, I., Ferrer, I., Van Hulle, S.W.H., Rousseau, D.P.L., Garfí,
514 M., 2022. Life cycle assessment of microalgae systems for wastewater treatment
515 and bioproducts recovery: Natural pigments, biofertilizer and biogas. *Sci. Total*
516 *Environ.* 847, 157615. doi:10.1016/j.scitotenv.2022.157615.
- 517 Azma, M., Mohamed, M.S., Mohamad, R., Rahim, R.A., Ariff, A.B., 2011. Improve-
518 ment of medium composition for heterotrophic cultivation of green microalgae,
519 *Tetraselmis suecica*, using response surface methodology. *Biochem. Eng. J.* 53,
520 187–195. doi:10.1016/j.bej.2010.10.010.
- 521 Baroukh, C., Muñoz-Tamayo, R., Bernard, O., Steyer, J.P., 2016. Reply
522 to the Comment on “Mathematical modeling of unicellular microalgae and
523 cyanobacteria metabolism for biofuel production” by Baroukh et al. [*Curr.*
524 *Opin. Biotechnol.* 2015, 33:198–205]. *Curr. Opin. Biotechnol.* 38, 200–202.
525 doi:10.1016/j.copbio.2016.02.018.
- 526 Baroukh, C., Muñoz-Tamayo, R., Steyer, J.P., Bernard, O., 2014. DRUM: A
527 new framework for metabolic modeling under non-balanced growth. Applica-
528 tion to the carbon metabolism of unicellular microalgae. *PLOS ONE* 9, 1–15.
529 doi:10.1371/journal.pone.0104499.
- 530 Baroukh, C., Turon, V., Bernard, O., 2017. Dynamic metabolic modeling of het-
531 erotrophic and mixotrophic microalgal growth on fermentative wastes. *PLoS Com-*
532 *put. Biol.* 13, 1–18. doi:10.1371/journal.pcbi.1005590.
- 533 Barsanti, L., Gualtieri, P., 2018. Is exploitation of microalgae economically and
534 energetically sustainable? *Algal Res.* 31, 107–115. doi:10.1016/j.algal.2018.02.001.
- 535 Batstone, D., Hülsen, T., Oehmen, A., 2019. Metabolic modelling of mixed
536 culture anaerobic microbial processes. *Curr. Opin. Biotechnol.* 57, 137–144.
537 doi:10.1016/j.copbio.2019.03.014.
- 538 Bernard, O., Hadj-Sadok, Z., Dochain, D., Genovesi, A., Steyer, J.P., 2001. Dynam-
539 ical model development and parameter identification for an anaerobic wastewater
540 treatment process. *Biotechnol. Bioeng.* 75, 424–438. doi:10.1002/bit.10036.

- 541 Casagli, F., Bernard, O., 2022. Simulating biotechnological processes affected by
542 meteorology: Application to algae–bacteria systems. *J. Cleaner Prod.* 377, 134190.
543 doi:10.1016/j.jclepro.2022.134190.
- 544 Casagli, F., Rossi, S., Steyer, J.P., Bernard, O., Ficara, E., 2021a. Balancing Mi-
545 croalgae and Nitrifiers for Wastewater Treatment: Can Inorganic Carbon Lim-
546 itation Cause an Environmental Threat? *Environ Sci Technol* 55, 3940–3955.
547 doi:10.1021/acs.est.0c05264.
- 548 Casagli, F., Zuccaro, G., Bernard, O., Steyer, J.P., Ficara, E., 2021b. ALBA: A
549 comprehensive growth model to optimize algae-bacteria wastewater treatment in
550 raceway ponds. *Water Res.* 190, 116734. doi:10.1016/J.WATRES.2020.116734.
- 551 Chen, C.Y., Ho, S.H., Liu, C.C., Chang, J.S., 2017a. Enhancing lutein produc-
552 tion with *Chlorella sorokiniana* Mb-1 by optimizing acetate and nitrate con-
553 centrations under mixotrophic growth. *J. Taiwan Inst. Chem. Eng.* 79, 88–96.
554 doi:10.1016/j.jtice.2017.04.020.
- 555 Chen, J.H., Chen, C.Y., Chang, J.S., 2017b. Lutein production with wild-type
556 and mutant strains of *Chlorella sorokiniana* MB-1 under mixotrophic growth. *J.*
557 *Taiwan Inst. Chem. Eng.* 79, 66–73. doi:10.1016/j.jtice.2017.04.022.
- 558 Chen, Y.H., Walker, T.H., 2011. Biomass and lipid production of heterotrophic
559 microalgae *Chlorella protothecoides* by using biodiesel-derived crude glycerol.
560 *Biotechnol. Lett.* 2011 3310 33, 1973–1983. doi:10.1007/S10529-011-0672-Y.
- 561 Espinosa-Gonzalez, I., Parashar, A., Bressler, D.C., 2014. Heterotrophic growth
562 and lipid accumulation of *Chlorella protothecoides* in whey permeate, a dairy
563 by-product stream, for biofuel production. *Bioresour. Technol.* 155, 170–176.
564 doi:10.1016/J.BIORTECH.2013.12.028.
- 565 Foreman-Mackey, D., Hogg, D.W., Lang, D., Goodman, J., 2013. Emcee : The
566 MCMC Hammer. *Publ. Astron. Soc. Pac.* 125, 306–312. doi:10.1086/670067.
- 567 Gao, F., Han, L., 2012. Implementing the Nelder-Mead simplex algorithm with adap-
568 tive parameters. *Comput Optim Appl* 51, 259–277. doi:10.1007/s10589-010-9329-
569 3.
- 570 Gao, Y., Guo, L., Jin, C., Zhao, Y., Gao, M., She, Z., Wang, G., 2022. Metagenomics
571 and network analysis elucidating the coordination between fermentative bacteria
572 and microalgae in a novel bacterial-algal coupling reactor (BACR) for mariculture
573 wastewater treatment. *Water Res.* 215, 118256. doi:10.1016/j.watres.2022.118256.

- 574 Ghimire, A., Frunzo, L., Pirozzi, F., Trably, E., Escudie, R., Lens, P.N., Esposito,
575 G., 2015. A review on dark fermentative biohydrogen production from organic
576 biomass: Process parameters and use of by-products. *Appl. Energy* 144, 73–95.
577 doi:10.1016/j.apenergy.2015.01.045.
- 578 Hameri, T., Fengos, G., Hatzimanikatis, V., 2021. The effects of model complexity
579 and size on metabolic flux distribution and control: Case study in *Escherichia*
580 *coli*. *BMC Bioinf.* 22, 1–25. doi:10.1186/s12859-021-04066-y.
- 581 Harmand, J., Lobry, C., Rapaport, A., Sari, T., 2019. Optimal Control in Bio-
582 processes: Pontryagin’s Maximum Principle in Practice. volume 3 of *Chemical*
583 *Engineering Series - Chemostat and Bioprocesses*. Wiley.
- 584 Herman, J., Usher, W., 2017. SALib: An open-source Python library for Sensitivity
585 Analysis. *J Open Source Softw* 2. doi:10.21105/joss.00097.
- 586 Hu, B., Min, M., Zhou, W., Du, Z., Mohr, M., Chen, P., Zhu, J., Cheng, Y., Liu, Y.,
587 Ruan, R., 2012. Enhanced mixotrophic growth of microalga *Chlorella* sp. on pre-
588 treated swine manure for simultaneous biofuel feedstock production and nutrient
589 removal. *Bioresour. Technol.* 126, 71–79. doi:10.1016/j.biortech.2012.09.031.
- 590 Klamt, S., Stelling, J., 2003. Two approaches for metabolic pathway analysis?
591 *Trends Biotechnol.* 21, 64–69. doi:10.1016/S0167-7799(02)00034-3.
- 592 Küken, A., Wendering, P., Langary, D., Nikoloski, Z., 2021. A structural property
593 for reduction of biochemical networks. *Sci Rep* 11, 17415. doi:10.1038/s41598-
594 021-96835-1.
- 595 Lacroux, J., Jouannais, P., Atteia, A., Bonnafous, A., Trably, E., Steyer, J.P., van
596 Lis, R., 2022. Microalgae screening for heterotrophic and mixotrophic growth on
597 butyrate. *Algal Res* 67, 102843. doi:10.1016/j.algal.2022.102843.
- 598 Lacroux, J., Trably, E., Bernet, N., Steyer, J.P., van Lis, R., 2020. Mixotrophic
599 growth of microalgae on volatile fatty acids is determined by their undissociated
600 form. *Algal Res.* 47, 101870. doi:10.1016/j.algal.2020.101870.
- 601 León-Vaz, A., León, R., Díaz-Santos, E., Vígara, J., Raposo, S., 2019.
602 Using agro-industrial wastes for mixotrophic growth and lipids production
603 by the green microalga *Chlorella sorokiniana*. *N. Biotechnol.* 51, 31–38.
604 doi:10.1016/j.nbt.2019.02.001.
- 605 Li, T., Zheng, Y., Yu, L., Chen, S., 2013. High productivity cultivation of a heat-
606 resistant microalga *Chlorella sorokiniana* for biofuel production. *Bioresour. Tech-*
607 *nol.* 131, 60–67. doi:10.1016/j.biortech.2012.11.121.

- 608 Li, T., Zheng, Y., Yu, L., Chen, S., 2014. Mixotrophic cultivation of a *Chlorella*
609 *sorokiniana* strain for enhanced biomass and lipid production. *Biomass Bioenergy*
610 66, 204–213. doi:10.1016/j.biombioe.2014.04.010.
- 611 Liang, Y., Sarkany, N., Cui, Y., 2009. Biomass and lipid productivities of *Chlorella*
612 *vulgaris* under autotrophic, heterotrophic and mixotrophic growth conditions.
613 *Biotechnol. Lett.* 2009 317 31, 1043–1049. doi:10.1007/S10529-009-9975-7.
- 614 Ma, X., Zheng, H., Addy, M., Anderson, E., Liu, Y., Chen, P., Ruan, R., 2016.
615 Cultivation of *Chlorella vulgaris* in wastewater with waste glycerol: Strategies for
616 improving nutrients removal and enhancing lipid production. *Bioresour. Technol.*
617 207, 252–261. doi:10.1016/j.biortech.2016.02.013.
- 618 Manhaeghe, D., Blomme, T., Van Hulle, S.W., Rousseau, D.P., 2020. Experimental
619 assessment and mathematical modelling of the growth of *Chlorella vulgaris* under
620 photoautotrophic, heterotrophic and mixotrophic conditions. *Water Res.* 184,
621 116152. doi:10.1016/J.WATRES.2020.116152.
- 622 Martínez, C., Mairet, F., Bernard, O., 2018. Theory of turbid microalgae cultures.
623 *J. Theor. Biol.* 456, 190–200. doi:10.1016/j.jtbi.2018.07.016.
- 624 Martínez, F., Orús, M.I., 1991. Interactions between Glucose and Inorganic Carbon
625 Metabolism in *Chlorella vulgaris* Strain UAM 101. *Plant Physiol.* 95, 1150–1155.
626 doi:10.1104/PP.95.4.1150.
- 627 Morales, M., Collet, P., Lardon, L., Hélias, A., Steyer, J.P., Bernard, O., 2019. Chap-
628 ter 20 - Life-cycle assessment of microalgal-based biofuel, in: Pandey, A., Chang,
629 J.S., Soccol, C.R., Lee, D.J., Chisti, Y. (Eds.), *Biofuels from Algae* (Second Edi-
630 tion). Elsevier. *Biomass, Biofuels, Biochemicals*, pp. 507–550. doi:10.1016/B978-
631 0-444-64192-2.00020-2.
- 632 Morris, M.D., 1991. Factorial Sampling Plans for Preliminary Computational Ex-
633 periments. *Technometrics* 33, 161–174. doi:10.2307/1269043.
- 634 Ogbonna, J.C., Masui, H., Tanaka, H., 1997. Sequential heterotrophic/autotrophic
635 cultivation—an efficient method of producing *Chlorella* biomass for health food
636 and animal feed. *J. Appl. Phycol.* 9, 359–366. doi:10.1023/A:1007981930676.
- 637 O’Grady, J., Morgan, J.A., 2011. Heterotrophic growth and lipid production
638 of *Chlorella protothecoides* on glycerol. *Bioprocess Biosyst. Eng.* 34, 121–125.
639 doi:10.1007/s00449-010-0474-y.
- 640 Orth, J.D., Thiele, I., Palsson, B.Ø., 2010. What is flux balance analysis? *Nat.*
641 *Biotechnol.* 28, 245–248. doi:10.1038/nbt.1614.

- 642 Patel, A.K., Joun, J.M., Hong, M.E., Sim, S.J., 2019. Effect of light conditions on
643 mixotrophic cultivation of green microalgae. *Bioresour. Technol.* 282, 245–253.
644 doi:10.1016/J.BIORTECH.2019.03.024.
- 645 Pessi, B.A., Pruvost, E., Talec, A., Sciandra, A., Bernard, O., 2022. Does tempera-
646 ture shift justify microalgae production under greenhouse? *Algal Res.* 61, 102579.
647 doi:10.1016/j.algal.2021.102579.
- 648 Rafrafi, Y., Trably, E., Hamelin, J., Latrille, E., Meynial-Salles, I., Benomar, S.,
649 Giudici-Orticoni, M.T., Steyer, J.P., 2013. Sub-dominant bacteria as keystone
650 species in microbial communities producing bio-hydrogen. *Int. J. Hydrogen En-
651 ergy* 38, 4975–4985. doi:10.1016/j.ijhydene.2013.02.008.
- 652 Regueira, A., Bevilacqua, R., Lema, J.M., Carballa, M., Mauricio-Iglesias, M.,
653 2020. A metabolic model for targeted volatile fatty acids production by cofer-
654 mentation of carbohydrates and proteins. *Bioresour. Technol.* 298, 122535.
655 doi:10.1016/j.biortech.2019.122535.
- 656 Sajjadi, B., Chen, W.Y., Raman, A.A.A., Ibrahim, S., 2018. Microalgae lipid and
657 biomass for biofuel production: A comprehensive review on lipid enhancement
658 strategies and their effects on fatty acid composition. *Renewable Sustainable
659 Energy Rev.* 97, 200–232. doi:10.1016/j.rser.2018.07.050.
- 660 Schuster, S., Dandekar, T., Fell, D.A., 1999. Detection of elementary flux modes
661 in biochemical networks: A promising tool for pathway analysis and metabolic
662 engineering. *Trends Biotechnol.* 17, 53–60. doi:10.1016/S0167-7799(98)01290-6.
- 663 Sen, R., Martin, G.J.O., 2018. Glycerol and nitrate utilisation by marine
664 microalgae *Nannochloropsis salina* and *Chlorella* sp. and associated bacte-
665 ria during mixotrophic and heterotrophic growth. *Algal Res.* 33, 298–309.
666 doi:10.1016/j.algal.2018.06.002.
- 667 Shoener, B.D., Schramm, S.M., Béline, F., Bernard, O., Martínez, C., Plósz, B.G.,
668 Snowling, S., Steyer, J.P., Valverde-Pérez, B., Wágner, D., Guest, J.S., 2019.
669 Microalgae and cyanobacteria modeling in water resource recovery facilities: A
670 critical review. *Water Res. X* 2, 100024. doi:10.1016/j.wroa.2018.100024.
- 671 Singh, D., Lercher, M.J., 2020. Network reduction methods for genome-scale
672 metabolic models. *Cell. Mol. Life Sci.* 77, 481–488. doi:10.1007/s00018-019-03383-
673 Z.
- 674 Storn, R., Price, K., 1997. Differential evolution—a simple and efficient heuristic
675 for global optimization over continuous spaces. *J. Global Optim.* 11, 341–359.
676 doi:10.1023/A:1008202821328.

- 677 Tan, X.B., Lam, M.K., Uemura, Y., Lim, J.W., Wong, C.Y., Lee, K.T.,
678 2018. Cultivation of microalgae for biodiesel production: A review on up-
679 stream and downstream processing. *Chinese J. Chem. Eng.* 26, 17–30.
680 doi:10.1016/j.cjche.2017.08.010.
- 681 Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single
682 diagram. *J. Geophys. Res. Atmos.* 106, 7183–7192. doi:10.1029/2000JD900719.
- 683 Terzer, M., Stelling, J., 2008. Large-scale computation of elemen-
684 tary flux modes with bit pattern trees. *Bioinformatics* 24, 2229–2235.
685 doi:10.1093/bioinformatics/btn401.
- 686 Turon, V., Baroukh, C., Trably, E., Latrille, E., Fouilland, E., Steyer, J.P., 2015a.
687 Use of fermentative metabolites for heterotrophic microalgae growth: Yields and
688 kinetics. *Bioresour. Technol.* 175, 342–349. doi:10.1016/j.biortech.2014.10.114.
- 689 Turon, V., Trably, E., Fouilland, E., Steyer, J.P., 2015b. Growth of *Chlorella*
690 *sorokiniana* on a mixture of volatile fatty acids: The effects of light and tem-
691 perature. *Bioresour. Technol.* 198, 852–860. doi:10.1016/j.biortech.2015.10.001.
- 692 Wágner, D.S., Valverde-Pérez, B., Sæbø, M., Bregua de la Sotilla, M., Van Wa-
693 genen, J., Smets, B.F., Plósz, B.G., 2016. Towards a consensus-based bioki-
694 netic model for green microalgae – The ASM-A. *Water Res.* 103, 485–499.
695 doi:10.1016/j.watres.2016.07.026.
- 696 Wan, M., Liu, P., Xia, J., Rosenberg, J.N., Oyler, G.A., Betenbaugh, M.J., Nie, Z.,
697 Qiu, G., 2011. The effect of mixotrophy on microalgal growth, lipid content, and
698 expression levels of three pathway genes in *Chlorella sorokiniana*. *Appl. Microbiol.*
699 *Biotechnol.* 91, 835–844. doi:10.1007/s00253-011-3399-8.
- 700 Wu, Z., Shi, X., 2007. Optimization for high-density cultivation of heterotrophic
701 *Chlorella* based on a hybrid neural network model. *Lett. Appl. Microbiol.* 44,
702 13–18. doi:10.1111/j.1472-765X.2006.02038.x.
- 703 Xie, Y., Li, J., Ho, S.H., Ma, R., Shi, X., Liu, L., Chen, J., 2020a. Pilot-scale
704 cultivation of *Chlorella sorokiniana* FZU60 with a mixotrophy/photoautotrophy
705 two-stage strategy for efficient lutein production. *Bioresour. Technol.* 314, 123767.
706 doi:10.1016/j.biortech.2020.123767.
- 707 Xie, Z., Lin, W., Liu, J., Luo, J., 2020b. Mixotrophic cultivation
708 of *Chlorella* for biomass production by using pH-stat culture medium:
709 Glucose-Acetate-Phosphorus (GAP). *Bioresour. Technol.* 313, 123506.
710 doi:10.1016/j.biortech.2020.123506.