Optimization of microalgae efficiency using reduced metabolic models

Bruno Assis Pessi

UNIVERSITÉ CÔTE D'AZUR

ÉCOLE DOCTORALE
SCIENCES ET TECHNOLOGIES DE L'INFORMATION ET DE LA COMMUNICATION

THÈSE DE DOCTORAT

Optimisation de l'efficacité des microalgues par des modèles métaboliques réduits

Bruno ASSIS PESSI

INRIA, Biocore

Présentée en vue de l'obtention du grade de docteur en Automatique, Traitement du Signal et des Images (ATSI) **d'**Université Côte d'Azur

Dirigée par : Olivier BERNARD
Co-encadrée par : Laetitia GIRALDI
Soutenue le : 09/02/2023

Devant le jury, composé de :
Caroline BAROUKH, chargée de recherche, INRAE
Fabrizio BEZZO, professeur, University of Padova
Filipa LOPES, professeur, CentraleSupélec
Francis MAIRET, chercheur, IFREMER
Laetitia GIRALDI, chargée de recherche, INRIA
Olivier BERNARD, directeur de recherche, INRIA

Inria

Optimisation de l'efficacité des microalgues par des modèles métaboliques réduits

Jury:

Rapporteurs

Fabrizio BEZZO, professeur, University of Padova
Francis MAIRET, chercheur, IFREMER

Examinateurs

Caroline BAROUKH, chargée de recherche, INRAE
Filipa LOPES, professeur, CentraleSupélec
Olivier BERNARD, directeur de recherche, INRIA
Laetitia GIRALDI, chargée de recherche, INRIA

Optimisation de l'efficacité des microalgues par des modèles métaboliques réduits

Résumé:

Cette thèse s'intéresse à la modélisation des microalgues et à l'optimisation de leur productivité. Les microalgues sont des organismes unicellulaires capables de croître de manière autotrophe, via la photosynthèse, et également de manière hétérotrophe lorsque du carbone organique est disponible. Une grande variété de facteurs influencent la croissance des microalgues. La première partie de cette thèse se concentre sur les effets macroscopiques : température, intensité lumineuse et contamination par des prédateurs. Dans la deuxième partie, les effets au niveau métabolique sont étudiés : influence du substrat et des vitamines sur la croissance.

Premièrement, nous analysons comment la culture sous serre influence la productivité. Nous proposons un modèle pour *Tetraselmis suecica*, et l'adaptons à quatre autres espèces. Des expériences en raceways sous serre ont été réalisées, en décalant la température de deux raceways par rapport à une référence à température ambiante. Le modèle de productivité a ensuite été validé en tenant compte de l'évolution de la température et de la lumière. La production annuelle des raceways et le bénéfice de l'utilisation des serres ont été évalués selon différents scénarios pour les cinq espèces considérées.

Nous étudions également la contamination par le zooplancton qui représente une contrainte majeure dans les systèmes de culture à grande échelle. La contamination peut être contrôlée en régulant le taux de dilution. Cependant, il n'est pas simple de trouver la meilleure stratégie de contrôle du taux de dilution. Des taux de dilution faibles favorisent le développement des prédateurs. Des taux de dilution élevés empêchent l'établissement des prédateurs, mais au risque de réduire la productivité des microalgues. De plus, la présence de régimes périodiques issus de l'interaction prédateur-proie ne permet pas de savoir si la présence de prédateurs doit être complètement évitée. Nous proposons une stratégie de contrôle du taux de dilution pour limiter les populations de zooplancton et optimiser la productivité de biomasse. Nous montrons que, à long terme, le taux de dilution constant optimal doit assurer l'éradication de la population de zooplancton.

Dans la deuxième partie de la thèse, nous proposons un modèle métabolique de *Chlorella sp*. se développant dans des conditions mixotrophes. Nous considérons diverses molécules, déchets de fermentation, en tenant compte des effets inhibiteurs. Une idée émergente est de coupler le traitement des eaux usées et la production de biocarburants en utilisant des microalgues pour atteindre des productivités plus élevées et des coûts plus faibles. Ce modèle métabolique inclut la consommation de glycérol et de glucose. Nous étudions comment l'ajout de ces substrats permet de surmonter l'inhibition par le butyrate. Le modèle métabolique a été construit en utilisant la méthode DRUM et comprend 188 réactions et 173 métabolites. Après une phase de calibration, le modèle a été confronté avec des données de 122 expériences collectées dans la littérature scientifique. La stratégie d'alimentation optimale estimée par le modèle réduirait le temps nécessaire pour consommer les acides gras volatils de 16 jours à 2 jours.

Enfin, nous étendons ce cadre de modélisation métabolique aux bactéries. Tout d'abord, en modélisant une *E. coli* auxotrophe en thiamine pour maximiser la production de lactate. Ensuite, nous étendons le modèle métabolique développé précédemment pour décrire une co-culture de *Chlorella* avec *E. coli* surproduisant la vitamine biotine. Nous modélisons l'influence de la concentration de biotine dans l'accumulation de lipides. Les résultats du modèle confirment qu'une souche de *E. coli* surproductrice de biotine pourrait maintenir une production de vitamines stimulante pour l'accumulation de lipides de *Chlorella*. Ce cadre de modélisation des interactions symbiotiques ouvre de nouvelles voies pour améliorer la conception et le fonctionnement des bioprocédés.

Mots clés: Modélisation, microalgues, contrôle, métabolisme, bactéries

Optimization of microalgae efficiency using reduced metabolic models

Abstract:

This thesis deals with the modeling of microalgae and the optimization of their productivity. Microalgae are unicellular organisms capable of growing autotrophically, via photosynthesis, and also heterotrophically when organic carbon is available. A great variety of factors influence the growth of microalgae. The first part of this thesis focuses on macroscopic effects: temperature, light intensity and contamination by predators. The second part considers the effects at the metabolic level: substrate composition and vitamin influence on kinetics.

First, we analyze the advantages of using greenhouses for the cultivation of phytoplankton. For this we propose a model for the marine green algae *Tetraselmis suecica*, and adapt it to four other species. Experiments under a greenhouse were carried out for the marine green algae *Tetraselmis suecica*. The temperature of two raceways was shifted compared to a reference raceway with free evolving temperature. The productivity model was then parametrized and validated accounting for the recorded evolution of temperature and light. The yearly raceway pond production and the benefit of greenhouse usage were assessed under different scenarios for the five considered species.

Then, we study zooplankton contamination, which represents a major constraint in large-scale microalgal cultivation systems. While zooplankton contamination cannot be avoided, their development can be controlled by regulating the dilution rate. However, finding the best control strategy by playing on the dilution rate is not straightforward. Low dilution rates (or long retention times) favor grazer development while high dilution rates avoid their establishment but reduce microalgal productivity. Furthermore, the presence of periodic regimes arising from the interaction between predator–prey makes it unclear if the presence of grazers must be completely avoided. We study the role of the dilution rate in the control of zooplankton populations and in the optimization of biomass productivity. We show that in long-term operations, the optimal constant dilution rate must ensure the eradication of the zooplankton population.

In the second part of the thesis, we propose a metabolic model for *Chlorella sp.* growing in mixotrophic conditions on fermentation wastes, accounting for the possible inhibitory substrates. An emerging idea is to couple wastewater treatment and biofuel production using microalgae to achieve higher productivity and lower costs. We propose a metabolic network for *Chlorella sp.* also including the consumption of glycerol and glucose. Then we study the optimal 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. An optimal feeding strategy estimated with the model would reduce the time to consume the volatile fatty acids from 16 days to 2 days.

Finally, we extend our metabolic modeling framework to bacteria. First, by modeling a thiamine auxotroph *E. coli* to maximize the production of lactate. Then, we extend the previously developed metabolic model to describe a co-culture *of Chlorella* with bacteria *E. coli* overproducing the vitamin biotin. We represent the influence of biotin concentration in the accumulation of lipids. The model results support the idea that a biotin overproducer *E. coli* could support the needs for vitamin by *Chlorella* and favor lipid accumulation. This modeling framework for symbiotic interactions, accounting for a dynamic metabolic model, opens up new avenues for improving the design and operation of bioprocesses.

Keywords: Modeling, microalgae, control, metabolism, bacteria

Acknowledgements

I would not have completed this thesis all by myself. Many people were essential in helping me finish this work during these three years, which were especially difficult because of the Covid-19 pandemic.

I would like to thank my thesis director Olivier Bernard for his ideas, reviews, support, trust, and understanding throughout these years. My co-director Laetitia Giraldi for helping me with the optimization problems of this work.

I am thankful for the support and friendship from all my colleagues at INRIA, from teams BIOCORE and McTao, who were there during these three years, in particular: Ali Gharib, Juan Carlos Arceo, Ignacio Fierro, Clotilde Djuikem, Odile Burckard, Marielle Péré, Yan Gao, Francesca Casagli, Walid Djelma, Nicolas Augier, Agustín Yabo, Alésia Herasimenka. Carlos Martinez for the work we did about predation of microalgae.

I will be forever grateful to my parents for their love, support, and prayers.

Above all, I give thanks to God for the gift of faith, which gave me hope to continue and allowed me to be here and complete this work.

Funding

This work was supported by the UCAJEDI and EUR DS4H investments in the Future projects managed by the National Research Agency (ANR) with the reference numbers ANR-15-IDEX-0001

Contents

Chapter 1

Introduction

The current energy production sources are dominated by fossil fuels, around 80% of total primary energy in the world, which also contributes to 60% of greenhouse gases emissions in the form of CO_2 (Höök and Tang, 2013). The emissions from fossil sources contribute to radiative forcing leading, through the greenhouse effect, to an increase in Earth's temperature. To estimate the future increase in temperature due to these emissions, it is necessary to consider both climate models and economic development projections. The climate research community developed scenarios, such as the Representative Concentration Pathway (RCP) and more recently the Shared Socioeconomic Pathways (SSPs) to include a variety of different economic development scenarios and the resulting temperature increase. Although, the most pessimistic scenario predicting greenhouse gases emissions (RCP8.5 and SSP5) seems unlikely yet (Ritchie and Dowlatabadi, 2017, Burgess et al., 2020, Pielke and Ritchie, 2021), the most realistic projections of greenhouse gas emissions forecast an average temperature increase of about $2°C$ by the end of the century, compared to pre-industrial era temperature (Palmer et al., 2018).

It is expected that, at least until the year 2050, the sources of fossil fuel will fully support current and future projected demand (Bhagea et al., 2019). After the mid-century, primary energy from oil will decline and it will be necessary to increase energy production from sources such as coal and natural gas (Ritchie and Dowlatabadi, 2017). As a consequence, the replacement of fossil fuel by renewable sources - such as biofuel - if not done to mitigate climate change, will be anyway necessary due to the lack of fossil resources to support economic development by the end of the century.

The production of biofuel from land plants demonstrated its potential in the 20th century. The progress of land biofuel has been slowed down due to its limitations such as competition with food agricultural production. The evolution of biofuel is classified into four generations. The first (edible oil) and second (non-edible oil) generations of biofuels are already operational. Microalgae are considered the third generation of biofuel feedstocks, with various advantages compared to the previous generations, such as higher oil content and higher growth rate. The fourth generation of biofuel is still in its infancy and uses feedstock designed by synthetic biology (Singh et al., 2020).

Scientific research investigating the potential of microalgae as a source of biodiesel started in the 1970s, motivated by the Oil Crisis of 1973 (Williams and Laurens, 2010). The use of

microalgae for the production of biofuel was a promising innovation that pushed for research in the field of energy, because of their high growth rate and the elevated lipid storage leading to areal productivities an order of magnitude higher than the first biofuel generation (Shuba and Kifle, 2018, Bhagea et al., 2019).

Microalgae are unicellular organisms capable of growing autotrophically with solar energy, through photosynthesis. Some species can also grow heterotrophically by absorbing a source of organic carbon compounds, such as glucose and acetate. Microalgae are also a very important organism in the carbon cycle of the planet, being responsible for 40% of global fixation of carbon. They can be used to produce a variety of products such as proteins, vitamins, cosmetics, feedstock, and food. The commercialization of microalgae-based products has been successful, especially in low-volume/high-value markets, such as products targeted for personal care (Barsanti and Gualtieri, 2018).

The two most widespread processes for producing microalgae are closed photobioreactors and open raceways (Schade and Meier, 2019). Photobioreactors can lead to a higher production output with better resistance to biological contaminants, but the energy input necessary for mixing and cooling strongly penalizes the economic and environmental balances (Tan et al., 2018, Schade and Meier, 2019). The more rustic raceways are a simpler and cheaper way for producing microalgae outdoors. They need less energy input and the functional design is simpler. The drawback is the higher contamination in the culture by grazers, bacteria, viruses, or even other competitive microalgae species (Williams and Laurens, 2010, Mata et al., 2010).

Despite advances made in microalgae research during the past decades, the production of microalgae at industrial scale is still limited. Different bottlenecks explain the limited use of these processes in comparison with their potential, in particular the economic and environmental profitability must still be improved.

The limited productivity of many cultivated species comes from the fact they are wild species, being the result of millions of years of natural selection with the objective to survive in nature, in very different conditions than the industrial ones. Finding new species, more efficient for a desired target is an important objective. Beyond this, the domestication of wild species and the enhancement of their production rate is fundamental. Various strategies can be used, such as genetic improvements to enhance a given function(Barsanti and Gualtieri, 2018). The RuBisCO enzyme involved in CO_2 fixation through the Calvin cycle is a limitation step in photosynthesis and genetic engineering using tools such as CRISPR can improve the process efficiency (Flynn and Raven, 2017, Lee et al., 2023). A recent success was reached with *Chlamydomonas reinhardtii*, by suppressing the expression of enzymes regulating the flux of carbon into the tricarboxylic acid (TCA) cycle lipid productivity almost doubled compared to the wild strain (Kao and Ng, 2017).

The use of mixotrophic growth, provided that the substrate is cheap and with a low environmental impact, is a promising way to increase productivity. This is of key interest if the substrate is a waste, and the process is used at the same time to produce biomass and to process wastewater (Castillo et al., 2021). It results that algae-bacteria systems are seen as a promising way for processing wastewater at lower economic and environmental costs. More generally, using bacteria that will stimulate microalgal growth and eliminate some toxic compounds is a promising approach. Algae-bacteria systems are more and more considered, with the idea

that algae can produce the requested dioxygen for the bacteria, while the bacteria produce CO_2 which will be used by photosynthesis. See Ramanan et al. (2016) for a review of Algae-bacteria interactions. Progresses in downstream processes are also expected, and they will contribute to improving the overall process efficiency.

In this work, we focus on the theoretical approach, using modeling to support advanced control in order to optimize the system's efficiency. Such approaches have proven to be efficient in many different biotechnological applications. In the microalgae field, they are probably even more important to rationally manage the complexity of these nonlinear systems, which are exposed to weather fluctuations affecting light and temperature. The development of numerical models is thus a prerequisite for understanding and managing these dynamical systems, involving several time scales, and permanently submitted to different perturbations. There is a need to bridge the gap between the detailed metabolic knowledge in the cell, and the necessity for control to keep a limited model complexity. Reducing metabolic models is difficult in a framework of permanent environmental fluctuation, maintaining the cell far from the balanced growth conditions which are generally the rule in metabolic modeling. Going from a metabolic model to a mechanistic model that can support process control is, therefore, a challenging objective, and this thesis proposes different approaches to meet this goal.

The PhD thesis is structured as follows. We first present the state of the art in metabolic modeling, introducing the concept of genome-scale metabolic models (GEMs), then current techniques to reduce the scale of these models in order to analyze them in steady-state or dynamical conditions. Chapters 3 and 4 are concerned with macroscopic modeling. We focus on two major factors influencing outdoor production: weather variation and contamination. In Chapter 3 we analyze how the cultivation of microalgae under greenhouses might improve productivity. In Chapter 4 we develop a control strategy to improve biomass production in the presence of predators. In Chapters 5, 6, 7 we use the DRUM framework (Baroukh et al., 2014) to reduce a large-scale metabolic network and propose a dynamical model of the metabolism. In Chapter 5, a metabolic model of microalgae *Chlorella* is used to optimize the consumption of a waste containing various molecules, including volatile fatty acids (VFAs). Chapter 6 is a preliminary work of modeling for Chapter 7, we optimize lactate production from a bacteria *E. coli* by controlling the supply of thiamine in the medium. In Chapter 7, we develop a metabolic model for an association of two microorganisms. *Chlorella* is producing lipids depending on the biotin supplied by an overproducer *E. coli*.

Chapter 2

Current state of microalgae research

2.1 Genome-scale metabolic models and network reduction

Many advances have been made concerning the mapping of metabolic reactions through the analysis of genomic data (Kim et al., 2017b). Genome-scale metabolic models (GEMs) are stoichiometric representations of the entire metabolism of a given organism, providing also the information linking genes, proteins (enzymes), and reactions (Kim et al., 2017b). They are constructed from the whole-genome sequencing of an organism, but several steps and iterations are necessary to build a functional metabolic model. First, it is necessary to identify functional roles in the genome; then, to connect them to enzyme complexes and then to reactions (Cuevas et al., 2016).

New GEMs are released every year and more and more organisms have their GEM. GEMs are also permanently updated and refined for model organisms such as *E. coli* following the knowledge consolidation of their genome and expressed proteins. This demonstrates the necessity of an experimental validation of the metabolic models. One of the first GEMs constructed for cyanobacteria predicted, through Flux Balance Analysis (FBA), that photorespiration would allow for optimal growth rates (Knoop et al., 2010). Analysis of GEMs helps to gain insights into possible metabolic engineering interventions and substrate allocation.

Usually, the construction of these models is first focused on the carbon-core metabolic network. Later on, they are refined by accounting for more details, such as improved compartmentalization by including more organelles (Tibocha-Bonilla et al., 2018). Microalgae metabolic models require, at least, the reactions taking place in the chloroplast, cytosol, and mitochondrion. The main metabolic pathways are photosynthesis (Calvin Cycle and light-dependent reactions), TCA cycle, Glyoxylate shunt, glycolysis, pentose phosphate pathway, respiration, and also the synthesis pathway of carbohydrates, amino acids, lipids, and nucleotides (Baroukh et al., 2014, Fachet et al., 2020). Detailed information about main metabolic pathways can be found in David et al. (2000).

2.1.1 Modeling metabolic networks

Metabolic networks are chains of reactions happening inside the cell. The different metabolic pathways keep the cell functioning, for instance: The production of energy, via ATP, the synthesis of macromolecules such as DNA, lipids, and proteins. Metabolic network models can be constructed based on the knowledge of biochemical processes, such as photosynthesis or glycolysis, or on the genomic knowledge of the organism (through the use of GEMs), which in general produces more accurate, though more complex models. The level of detail in the model can also be constrained by the objective of its use, and many reactions can be omitted. In general, simplifications and assumptions to reduce the size of the system are necessary because the large number of states in standard models of metabolic networks makes optimization and control impracticable. In general, it is assumed that the system is in Quasy Steady State (QSS), known as the Quasy Steady State Approximation (QSSA). In the QSSA internal metabolites are assumed to be in steady state, i.e. the equilibrium is reached instantaneously, while only the concentration of external metabolites or substrates behave dynamically. Mathematically, the QSSA is written as:

$$\frac{dc}{dt} = 0 \tag{2.1}$$

Where c is the concentration of metabolites inside the cell written as a fraction, i.e. mass of metabolites per total mass of the cell. The ordinary differential equations (ODE) representing the system in a continuous perfectly mixed stirred tank can be written in the following general form:

$$\frac{dC}{dt} = \frac{cX}{dt} = N.vX - D.C \tag{2.2}$$

$$\frac{dX}{dt} = \mu X - DX \tag{2.3}$$

$$\frac{dP}{dt} = N_p.vX - DP \tag{2.4}$$

$$\frac{dS}{dt} = N_s.vX - DS + DS_{in} \tag{2.5}$$

Where $C \in \mathbb{R}^{n_c}$, $S \in \mathbb{R}^{n_s}$,$P \in \mathbb{R}^{n_p}$are respectively the metabolites, substrates and products concentration vectors. X is the biomass concentration. Substrates, products, and biomass are written as mass per volume of the reactor; $N \in \mathbb{R}^{n_m * n_r}$ is the matrix of stochiometric indices of the reactions in the metabolic network; $v \in \mathbb{R}^{n_r}$ is the vector of the reactions kinetics; μ is the growth rate; D is the dilution rate; $S_{in} \in \mathbb{R}^{n_r}$ is the concentration vector of incoming substrates.

When the metabolite concentrations are written per mass/volume of the cell, the ODE is written as:

$$\frac{dc}{dt} = N.v - c\mu \tag{2.6}$$

Assuming a QSSA:

$$N.v - c\mu = 0 \tag{2.7}$$

The term $c\mu$, which describes the dilution of metabolites due to cellular growth is generally ignored because the dynamics of the chemical reactions are considerably greater than the loss of concentration due to the change in cell's mass (Provost and Bastin, 2004). In the end, we have the following equation:

$$N.v = 0 \tag{2.8}$$

The QSSA is a necessary assumption for most frameworks and modeling of metabolic networks. The QSSA cannot always be applied, for example in cases where metabolites accumulate inside the organism, such as in microalgae. Due to dial variations of light intensity microalgae accumulates different metabolites depending on light availability. Therefore, the use of these classic frameworks is limited when applied to the modeling of microalgae systems.

Constrained-based modeling techniques considering the QSSA are the most widely used when dealing with metabolic networks, enabling the estimation of intracellular fluxes at different conditions (Tibocha-Bonilla et al., 2018). The two most important techniques (Lotz et al., 2014) are Elementary flux modes (EFM) and Flux Balance Analysis (FBA).

2.2 Tools for computing the Elementary Flux Mode (EFM)

2.2.1 Definition

Elementary flux modes (EFMs) are often described as a minimum set of pathways capable of representing the total of the network at the steady state. A flux mode is defined mathematically as a set M:

$$M = \{v \in \mathbb{R}^r | v = \lambda v^*, \lambda > 0, Nv^* = 0\} \tag{2.9}$$

where v^* is a vector respecting the steady-state condition $Nv^* = 0$, having a subset $v^{irr} \geq 0$, corresponding to the irreversible reactions, while the subset v^{rev} corresponding to the subset of reversible reactions has no sign restriction (Schuster et al., 1999).

A representative v^* of M is an elementary flux mode if and only if it fulfills the simplicity condition:

There is no couple of vectors v', v'' with the following properties:

1. v^* is a non-negative linear combination of v' and v"
2. v' and v'' satisfies the conditions to be a flux mode
3. v' and v'' contain at least the same number of zero elements as v^*, and at least one of them contains more zero elements than v^*.
4. The elements at boundary reactions of v' and v'' have the same sign or one element is a zero (e.g. $v'_i = -1$, $v''_i = 0$)

The vectors v satisfying the steady state equation are necessarily non-negative belonging to the kernel of the stoichiometric matrix N. Therefore, the space generating these vectors is a polyhedral cone in the intersection between the kernel of N and the positive orthant. The

vectors v can then be written as a non-negative linear combination of a set of vector e_k which form the unique convex base of the polyhedral cone.

$$v = \sum \lambda_k e_k \; ; \lambda_k \geq 0 \tag{2.10}$$

The vectors e_k forming the convex basis are the elementary flux modes, being the simplest pathway connecting substrates to products at a steady state condition. The EFMs are useful to deduce macroscopic or global reactions in the metabolic network. Because of the QSSA, the kinects of internal metabolites can be ignored, simplifying the dynamic equation of the macro reaction.

Since $e_k \in \mathbb{R}^{n_r}$, each position corresponds to a reaction participating in the elementary mode.

The macro reactions are easily deduced by multiplying by zero the components representing the internal reactions between metabolites, then keeping the substrates and products in the remaining reactions. For more complex metabolic networks, the determination of EFMs requires much more computational effort, with the number of EFMs increasing exponentially with the size of the network. An efficient algorithm to calculate all EFMs may be necessary to reduce computational time. Also, the existence of reversible reactions in the network might increase the difficult of determining the set of EFM.

2.2.2 Minimal generating sets (MGS) and EFM reduction

The presence or not of reversible reactions in the metabolic network change the algorithm necessary to compute the set of EFMs. The simplest case is when all reactions are irreversible. In this case, every component of all flux vectors v is non-negative and the cone representing the space of allowed flux vectors at steady-steady is a pointed-cone. A convex cone K is pointed if:

$$K \cap -K = \{0\} \tag{2.11}$$

If a polyhedral cone is pointed there exists a unique minimal set of generating vectors and the elements of this set are the extreme rays of the cone. They also are a complete set of representatives of elementary modes. In summary, when all reactions are irreversible, i.e. $v = v^{irr}$, there is a unique minimal generating set (MGS) which is equivalent to the set of EFMs.

There are 3 cases when reversible reactions are present in the network:

1. The system has only irreversible elementary modes. Despite the presence of reversible reactions, no EFM can work in the reversible direction, i.e. $\nexists e_k = -e_{k'}$

2. The system has irreversible and reversible elementary modes.

3. The system has only reversible elementary modes

In the first case, the polyhedral cone is still pointed, the minimal generating set is unique, and it corresponds to the set of EFM. This is not the case anymore for the two remaining cases, where the cone is not pointed. Also, the MGS will not be unique and the set of EFM might be

greater than the size of the MGS. Given this, the set of EFM will always be a superset of the MGS.

To understand the difference between the MGS and the EFMs getting deeper into convex analysis is necessary. In convex analysis it is shown that the space generating the solution of a linear homogenous system of equations is a convex polyhedral cone, C. Every point of such a cone is a non-negative combination of fundamental vectors, f, and basis vectors, b,

$$C = \{v : v = \sum n_k f^k + \sum \lambda_m b^m \; ; n_k, \lambda_m \geq 0\} \tag{2.12}$$

The fundamental and basis vectors are also called the generating vectors. There is a minimum necessary number of generating vectors to span the cone. The basis vectors are the extreme rays of C for which the negative vector is also contained in C (Schuster et al., 1999). The definition of C here is identical to the Minimum set of Elementary Modes (MEMO) in Röhl and Bockmayr (2019), where every v vector in the steady state cone is written as a non-negative linear combination.

$$v = \sum \lambda_e e + \sum \lambda_f f \tag{2.13}$$

where $f \in \{U \cap E_N^{irr}\}$, $e \in \{U \cap E_N^{rev}\}$, where U is an inclusion-minimal set which is a subset of E_N which is the set of all EFMs for the stoichiometric matrix N.

The basis vector b is then equivalent to the vector e corresponding to the reversible EFMs, while the fundamental vectors f are equivalent to the irreversible set of EFMs. This implies that when there are no reversible EFMs, i.e. the cone is pointed, C has no basis vectors. By contrast, when there are only reversible EFMs, C has only basis vectors.

Jevremović and Boley (2013) and Röhl and Bockmayr (2019) provide algorithms to compute the minimal generating set when there are reversible reactions in the metabolic network. Röhl and Bockmayr (2019) relies on a method of splitting reversible reactions, with a minimal number of splits, until no reversible EFM is left creating a pointed cone. While Jevremović and Boley (2013) divide the stoichiometric matrix based on the reversibility or not of the reaction, then they compute the null space of a modified matrix representing the reversible reactions and the minimal generating set of the irreversible subnetwork.

The set of EFMs is in general much larger than the MGS. For example, the carbon metabolism of *Escherichia coli* has 6421 EFMs while only 15 vectors are in the MGS/MEMO. The division of a network into subnetworks, as in the DRUM method (described below), also reduces the number of EFMs. The use of MGS may be another way to reduce the size of the system, though only the number of macro reactions is guaranteed to be reduced compared to the use of EFMs - the number of metabolites might still be the same. Furthermore, it is not guaranteed that the MGS is able to create meaningful macro reactions capable of accurately modeling the network, as it happens with the use of EFMs.

The calculation of EFM becomes prohibitive when the metabolic network is too large, but the enumeration of EFMs is still possible by computing only a subset of the EFMs (Maton et al., 2022). Many methods have been developed in recent years to facilitate the computation of EFMs. Kaleta et al. (2009) is an example of subsystem analysis. This paper introduces the concept of Elementary Flux Patters, where instead of giving a stoichiometrical proportion to a reaction, it only considers the index. It means that it only calculates the list of reactions participating in an elementary mode.

Oddsdóttir et al. (2015) use optimization in metabolic flux analysis to reduce the number of EFMs. The idea is to find the best fitting EFMs to some measured external flux. There is an algorithm minimizing the difference between the measured flux and the EFMs to reproduce those fluxes. Tabe-Bordbar and Marashi (2013) couple EFM with FBA. The proposed algorithm removes reactions by FBA, considering a random objective reaction. They select a list of reactions to remove, followed by FBA calculation, if the objective flux is non-zero, then they proceed with the deletion of the reactions, on the other hand if the flux of the objective reaction is zero, then the reactions are kept. The goal is to find, at least, a subset of the EFMs of a genome-scale metabolic network, by reducing the size of the total network. In a recent paper Maton et al. (2022) calculate a reduced set of EFMs, based on several steps, including geometrical criteria, optimization techniques, and also external observations to derive macroreactions for the system.

2.3 Flux balance analysis (FBA)

Flux balance analysis (FBA) is one of the most common tools to analyze metabolic networks (Orth et al., 2010b). Together with EFMs, they can be used to identify feasible routes in the metabolic network and estimate internal metabolic fluxes based on substrate uptake and excretion rates (Lotz et al., 2014). As in the case of EFMs, FBA also assumes the cell to be at steady state. But, instead of trying to determine the possible set of reactions constructing the steady state, the method consists of the maximization (or minimization) of an objective function:

$$\max \ Z = c^T v \tag{2.14}$$

where $c \in \mathbb{R}^{n_r}$ is a vector of weights, indicating how much a certain reaction influences the objective function, and v is the vector of fluxes of metabolic reactions. Besides the constraint of the steady state ($Nv = 0$), it considers boundaries for the vector of reaction fluxes v.

$$l_i \leq v_i \leq u_i \tag{2.15}$$

where l_i and u_i are the lower and upper boundary, respectively. One of the most common cases, is the maximization of biomass production, in this case c will be a vector containing zeros in every position, except the position for the reaction of biomass.

The system of equations and constraints of FBA leads to a linear programming problem, which can be computed with standard algorithms such as Interior-Point methods.

2.4 State of the art of metabolic networks reduction

The increasing size of metabolic networks makes it difficult to apply numerical analysis, especially when considering dynamical aspects. Even in the case of the steady state, computational power becomes limiting. For example, as discussed above, the number of EFMs grows exponentially as the metabolic network increases. As a consequence, calculation of EFMs for genome-scale models even for simple organisms such as *E. coli* may not be possible due to computational limitations. Methods to reduce the size of these genome-scale metabolic models become imperative to analyze

steady state and dynamical behavior. Here, we will briefly mention the current methods used to reduce metabolic networks, with emphasis on genome-scale networks.

Many methods have been released in the literature in recent years regarding the reduction of metabolic networks (Singh and Lercher, 2020). They differentiate on the assumptions regarding the network in order to proceed with the model reduction, methods, and goals. While some techniques focus on keeping the same phenotype of the full network, others have a more greedy approach to the reduction, focusing on minimizing the most possible of the network and only keeping some desired reactions or phenotypes.

One of the first techniques used was the consideration of "*Enzymes Subsets*" (Pfeiffer et al., 1999). An Enzyme Subset is a group of enzymes that work together in a metabolic pathway and can be considered as a unit structure catalyzing a series of reactions. Mathematically, a group of reactions (or enzymes) belongs to an Enzyme Subset if in all flux vectors v satisfying the steady state condition, the ratio between the fluxes of the reactions in the Enzyme Subset, e.g. $v_n/v_{n'}$, has the same non-zero value and the direction of the reactions are not contradictory. It is possible therefore to reduce the network without losing the original information and capabilities, because it is considered that the enzymes belonging to such a subset are expressed coherently, regulating metabolism in unit. Nevertheless, the drawback of this method is the limited capability in reducing the total size.

A method to further reduce the metabolic network was later implemented by Burgard et al. (2001) called Minimal Reaction Sets. The method consists in solving a mixed-integer linear programming problem, where the objective is to minimize the number of reactions of the network, while still keeping a minimal flux of biomass production.

A more recent method *NetworkReducer* has recently been published where the objective is to reduce the network while at the same time keeping certain protected phenotypes, metabolites, and reactions (Erdrich et al., 2015). The algorithm functions in two major steps. First the pruning phase where reactions are iteratively removed until no more reaction can be deleted without breaking protected parts. Second, the compression phase is a loss-free simplification by the lumping of coupled reactions. An improvement of this method was made by Röhl and Bockmayr (2017), by including the minimization of the number of reactions as in Burgard et al. (2001).

Küken et al. (2021) propose a method of reduction based on the use of complexes (combination of the species participating in one side of the reaction), where the stoichiometric matrix is written as the product of two matrices, $N = Y.A$, where Y is a matrix having as columns the complexes and metabolites in the rows, A a matrix having indices of -1, 0 or 1 with reactions represented on the columns and the complexes on the rows. Depending on the structural conditions and the balancing of the complexes, the network is reduced while keeping the phenotype of the original network.

2.4.1 The DRUM framework

DRUM (Dynamic Reduction of Unbalanced Metabolism) is a metabolic modeling framework (Baroukh et al., 2014) created in order to circumvent the problem of inappropriate use of the QSSA to the whole metabolic network. It was initially developed for organisms which dynamically accumulate and reuse some metabolites, such as microalgae under varying environments.

The idea of the DRUM framework is to divide the complete metabolic network into subnetworks, in which the QSSA is valid, also reducing the total number of state variables representing the system. After the division of the network, the EFMs are calculated for each subnetwork, generating macroscopic reactions, representing the result of all the internal reactions of the subnetwork with much simpler kinetics. After the application of the DRUM method, the dynamical equations of the system are reduced to the number of metabolites that are allowed to accumulate, external substrates and products. The form of the system of differential equations is the same as the original one, but the stoichiometric matrix is reduced and modified to represent the new macroscopic reactions deduced from the EFMs. The DRUM method is able to accurately represent empirical data, predicting for example the accumulation of carbohydrates and lipids during the day and as well its consumption during night. Despite this, a more objective method to divide the subnetworks still needs to be defined. Finding new ways to split the metabolic network might reduce even more the size of the system, while still being able to predict the accumulation of metabolites.

In the DRUM framework the metabolic network is represented by the following system of Ordinary Differential Equations:

$$\frac{dM}{dt} = \frac{d}{dt}\begin{pmatrix} S \\ C \\ P \\ B \end{pmatrix} = \begin{pmatrix} N_S \\ N_C \\ N_P \\ N_B \end{pmatrix} .v(M).B - DM + DM_{in} = N.v(M).B - DM + DM_{in} \tag{2.16}$$

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

The DRUM method (Baroukh et al., 2014), consists in dividing the metabolic network into k quasi-stationary subnetworks, so the matrix N is rewritten in the following form:

$$N = [N_{SN_1}, \cdots N_{SN_k}] \tag{2.17}$$

where $N_{SN_i} \in \mathcal{R}^{n_m \times n_{SN_i}}$ and $\sum_{i=1}^{k} n_{SN_i} = n_r$. Each sub-network is assumed to be at steady state:

$$\forall i = 1, ..., k : \; N_{SN_i}.v_{SN_i} = 0 \tag{2.18}$$

By considering the steady-state condition, it is possible to calculate the EFMs for each of these N_{SN_i} sub-networks, then construct macro reactions:

$$\forall i = 1, ..., k : v_{SN_i} = E_{SN_i}\alpha_{SN_i} \; , \; \alpha_{SN_i} \geq 0 \tag{2.19}$$

$$\forall i = 1, ..., k : \ (N_{S_{SN_i}}.E_{SN_i}).S_{SN_i} \longrightarrow (N_{PSN_i}.E_{SN_i}).P_{SN_i} \tag{2.20}$$

where E_{SN_i} is the matrix of EFMs of the sub-network SN_i, and α_{SN_i} the kinetics of the macroscopic reactions described by the reduced stoichiometric matrix. Following this step, we group all the sub-networks, and considering that only metabolites A are allowed to accumulate. Meaning that other metabolites have simple dynamics and their concentration is directly determined by the A metabolites. We obtain a reduced dynamic model, defined by the new metabolites vector $M' \in \mathcal{R}^{n_m}$,the new stoichiometric matrix $N' \in \mathcal{R}^{n_m \times n_E}$ and α the kinetic vector associated to these macroscopic reactions:

$$\frac{dM'}{dt} = \frac{d}{dt}\begin{pmatrix} S \\ A \\ P \\ B \end{pmatrix} = \begin{pmatrix} N_S \\ N_A \\ N_P \\ N_B \end{pmatrix}.\alpha.B = N'.\alpha.B - DM' + DM'_{in} \tag{2.21}$$

Chapter 3

Temperature modelling applied to microalgae cultivation under greenhouse

This chapter has been adapted from the published article
Pessi, B.A., Pruvost, E., Talec, A., Sciandra, A., Bernard, O., 2022. Does temperature shift justify microalgae production under greenhouse? Algal Research 61, 102579.

Abstract

We analyze the influence of greenhouses in the cultivation of phytoplankton. For this we propose a model for the marine green algae *Tetraselmis suecica*, and adapt it to four other species (*Spirulina platensis*, *Dunaliella salina*, *Phaeodactylum tricornutum* and *Chlorella vulgaris*). Experiments under a greenhouse were carried out for the marine green algae *Tetraselmis suecica*, shifting the temperature of two raceways compared to a reference raceway with free evolving temperature. The productivity model was then parametrized and validated accounting for the recorded evolution of temperature and light. The yearly raceway pond production and the benefit of greenhouse usage was assessed under different scenarios for the five considered species. At year scale, greenhouse efficiency is notable only for few species, e.g. *Spirulina platensis*, where productivity can be increased by 20 %. Based on these results, cultivation under greenhouse is beneficial mainly to protect the culture against contamination and to increase productivity in cold regions for species susceptible to photoinhibition with optimal growth in high temperatures. Rotation of the cultivated species is also a good strategy to improve annual productivity.

3.1 Introduction

Microalgae are a promising and environmental-friendly source of high valuable chemicals for green chemistry and, on a longer run, for biofuel (Mobin and Alam, 2017). Despite its potential, large scale production of algal biofuel is challenging, and to date not economically and energetically viable (Baudry et al., 2017, Morales et al., 2019). Maximizing productivity through a better understanding and optimization of the growth conditions is still necessary so that industrial production of microalgae becomes a reality (Barsanti and Gualtieri, 2018, Bernard, 2011a).

Large scale cultivation of microalgae involves many challenges that differ depending of the cultivation system. Here, we will focus on raceway ponds, which is an outdoor method of production, with a simpler design than photobioreactors - another common cultivation system. One of the major drawbacks of cultivation in raceway ponds is its susceptibility to contamination by external organisms present in the outside environment. Using a greenhouse is an interesting trade-off to reduce contamination while modulating climate. Indeed, outdoor production is susceptible to external weather changes, especially temperature and solar radiation fluctuations, leading to growth conditions significantly different from the constant and optimal ones often maintained in the laboratory.

Microalgae are autotrophic organisms that fix CO_2 from the atmosphere through the Calvin cycle, using the photon energy from photosynthetically active radiations (PAR). Photosynthesis rate responds differently according to light intensity, which can be separated in 3 regions: photolimited, saturated, and photo inhibited (Béchet et al., 2013, Williams and Laurens, 2010). In the first region growth is limited by the quantity of absorbed photons. In the light saturated region, the photosynthetic system functions at its full capacity so growth does not increase with more light intensity. Surpassing a given intensity, photo inhibition takes place, and growth starts to diminish because of deactivation of key proteins in the photosynthetic apparatus (García-Camacho et al., 2012)

Temperature has also a strong influence on the growth rate of microalgae, since it affects the rate of enzymatic reactions. Moreover, it also modulates the solubility of several key molecules, such as CO_2 and O_2 that impact the growth rate (Ketheesan and Nirmalakhandan, 2013). The direct effect on growth rate for moderated temperatures is well represented by the Arrhenius Equation, representing how the reaction rates are enhanced by temperature (Grimaud et al., 2017). But above the optimal temperature – the temperature for which the growth is maximized- the structure of some proteins changes, especially in the electron transport chain, leading to a rapid drop in the net growth rate concomitant with a mortality increase (Serra-Maia et al., 2016). Above a maximum temperature, the algae can no longer grow (Grimaud et al., 2017).

Climate control in greenhouses is an option to adapt the thermal conditions to the physiological optimum of the species. Shadowing for reducing photoinhibition can also be implemented in greenhouses, using nets or even photovoltaic panels (Martínez et al., 2017, Morales et al., 2019). Artificial light is also a possibility for mitigating the natural daily and seasonal fluctuations for outdoor production. The low winter temperatures diminish productivity, while in summer extreme light intensities might inhibit growth (Serra-Maia et al., 2016) and elevated temperatures may even lead to culture collapse. Greenhouses thus offer an interesting way to alleviate growth

reduction due to weather, either passively or with additional energy to cool down or warm up depending on the season. The benefit of a greenhouse must then be assessed and compared with the additional costs in terms of energy and infrastructure.

Growth models for microalgae have been successfully used to predict the productivity in response to different conditions and factors, such as nutrient concentration, temperature, and light (Béchet et al., 2013, Bernard, 2011a). Most models are only validated in indoor conditions within a steady environment, which simplifies experimental setting, but extrapolating the use of these models to an outdoor dynamic environment is uncertain (Darvehei et al., 2018). Examples of validated models for raceway cultivation are scarce, even more for long-term production. An artificial neural network model was validated for cultivation in open raceway ponds, while confirming that temperature and light intensity are very important factors influencing productivity (Supriyanto et al., 2019). Some models taking into account hydrodynamics were validated using mass balance of various components, but they do not take into account temperature changes (Fernández et al., 2016, Ranganathan et al., 2017). Furthermore, long-term prediction models that have been developed to account for seasonal and diurnal changes in temperature and light have rarely been validated (Banerjee and Ramaswamy, 2017, Casagli et al., 2021b).

In this work, we explore the effects of seasonal and daily temperature changes in microalgae production using raceways in a greenhouse. The model is used to predict annual production and the influence of greenhouses in the cultivation of microalgae accounting for temperature and light in a dynamic outdoor environment. The model is calibrated and validated using original experiments with *Tetraselmis suecica*, a microalga widely used in aquaculture as a food source, and which can tolerate a variety of temperatures and salinity (Molina et al., 1991, Fabregas et al., 1984). The experiments were carried out in 3 raceways under a greenhouse with three different temperature regimes.

The model is then further consolidated for four other species (*Dunaliella salina*, *Spirulina platensis*, *Phaeodactylum tricornutum* and *Chlorella vulgaris*) based on parameters available in the literature. Productivity simulations for different locations in France are carried out. These simulations exemplify which characteristics of a species are important when cultivated at different seasons and locations.

3.2 Materials and methods

3.2.1 Experiments in raceway ponds

The green microalga *Tetraselmis suecica* AC 254 (Algobank Caen, Université de Caen Normandie, France), was inoculated in three 2.61 m^2 raceway ponds operated in a transparent greenhouse, located in *Villefranche-sur-Mer*, France. The water depth was 16 cm on average, for a total culture volume of 417 L. Fresh water was added daily to compensate evaporation. The temperature of the first raceway was not controlled, while that of the others was controlled to be respectively 5 °C above and 5 °C below the first one. Two experiments are shown in this work, the first one took place between the 12th and 29th of January, which was used to validate the model, and the second one between 06th and 26th of February, which was used to calibrate the model.

Table 3.1: Composition for the culture medium.

Compound	Conc.(M)	Compound	Conc.(M)
$NaNO_3$	8.82 10^{-4}	$FeDTPA$	6.66 10^{-6}
$NaH_2PO_4.H_2O$	3.62 10^{-5}	Mn	2.33 10^{-5}
B	4.12 10^{-5}	Mo	5.62 10^{-7}
Cu	9.44 10^{-7}	Zn	6.48 10^{-6}

3.2.2 Culture medium

The medium consisted in $0.2\mu m$ filtered sea water, enriched with an industrial solution (Kanieltra 15 Fe DTPA) used for agricultural purposes providing trace elements and metals. No vitamins were added. pH was kept at 7.5 by CO_2 injection when pH value was above this setpoint. The concentrations for every element in the medium are listed at Table 5.3.

Temperature was regulated with a Lauda cryostat connected to a thermal exchanger made of inox tubes place in the raceway. A ramp program with a set of 12 temperatures was used for each raceway where the temperature was changed. Temperature and light intensity of the three raceways were recorded every 30 seconds. At the beginning of each batch, culture medium was renewed and raceways were inoculated with the same inoculum of *Tetraselmis suecica* at the same concentration.

3.2.3 Dry weight and cell counting

Biomass dry weight was measured on samples of 50 mL collected in duplicates, centrifuged for 5 minutes at 3000 g and rinsed 3 times with freshwater to reduce the salt concentration. It was verified that *Tetraselmis suecica* cells do not break in fresh water or due to centrifugation. Once the centrifugation supernatant was removed, the cell pellet was resuspended in about 1 mL and filtered through a preweighted 1.2 µm GF/C glass fiber filter. Filters were stored inside an oven at 75°C for 24 hours and then weighted with a precision balance.

The concentration of cells was measured with an optical cell counter (HIAC, model 9703) after dilution of the culture with 0.22 μm filtered seawater. Concentration was calculated counting all particles whose diameters range from 2.65 µm to 20 µm.

3.2.4 Laboratory calibration experiments

Additional experiments were conducted in the laboratory to measure the response of the growth rate to different light intensities and temperatures. For this purpose, we used the MC 1000-OD multicultivator which is equipped with 8 culture flasks placed in a thermostatically controlled bath and individually illuminated by a controlled light source. This device also allows us to measure automatically every 5 minutes the optical density of the cultures at 680 and 720nm. Simultaneous measurements of cell density have shown that, over a given concentration range, the relationship between cell concentration and OD is linear. For each of the 2 experiments, the samples taken from the raceway ponds were diluted so that the initial cell concentrations in 8 culture flasks were identical. For the light and temperature conditions tested (see paragraph

2.6), the maximum growth rate was measured during the exponential growth phase, i.e., when the slope of the data on a logarithmic scale is linear.

3.2.5 Model hypotheses

To represent algal cultivation under a greenhouse submitted to actual environmental fluctuations, the model needs to consider variations in light intensity, light attenuation along depth due to algae, and temperature. The nutriments are introduced in excess so that they are not limiting growth. We consider a modified version of the simple model validated by Bernard and Rémond (2012). The growth rate is represented by the product of two functions of light (I) and temperature (T), considering a multiplicative effect:

$$\mu\ (T, I)\ =\ \phi(T)\mu_{opt}(I) \tag{3.1}$$

$$\phi(T) = \frac{(T-T_{max})(T-T_{min})^2}{(T_{opt}-T_{min})((T_{opt}-T_{min})(T-T_{opt})-(T_{opt}-T_{max})(T_{opt}+T_{min}-2T))} \tag{3.2}$$

$$\mu_{opt}\ (I) = \mu_{max}\frac{I}{I + \frac{\mu_{max}}{\alpha}\left(\frac{I}{I_{opt}} - 1\right)^2} \tag{3.3}$$

This equation represents the growth rate at the depth where the light intensity equals I. The average growth rate over the water column must account for the light attenuation with depth, due to turbidity. We assume that I decreases exponentially with depth, according to the Beer-Lambert law (Benson and Rusch, 2006):

$$I(L)\ =\ I_0 exp(-\sigma X L) \tag{3.4}$$

With I_0 being the light intensity at the surface, σ the extinction coefficient, X the algae concentration and L the depth of water. The analytical expression of the integrated form of Equation (3.3) along depth gives the average growth, $\overline{\mu_{opt}}$. Its mathematical expression was given by Martínez et al. (2018). The extinction coefficient σ is dependent on the biomass. As suggested by Morel (1988) it can be written in the form of a power function, i.e.

$$\sigma = AX^B \tag{3.5}$$

The other modification consists in the addition of a mortality rate function as in Béchet et al. (2017):

$$\lambda(T) = \lambda_0(t) exp(\beta T) \tag{3.6}$$

$\lambda_0(t)$ is considered to take two possible values. One for daytime when photosynthesis is active, and one for nighttime when only respiration takes place. The net growth rate is then given by:

$$\mu_{net}(T, I) = \phi(T)\overline{\mu_{opt}}(I) - \lambda(T) \tag{3.7}$$

The ordinary differential equation describing biomass dynamics is thus:

$$\frac{dX}{dt} = \mu_{net} X \tag{3.8}$$

3.2.6 Model calibration

In order to calibrate the response curves to temperature and light for *Tetraselmis suecica* growth rates obtained at different light intensities and temperatures from the Multi-Cultivator were divided into four data sets. The $\mu_{opt}(I)$ function is calibrated using data from two different temperatures, 26 °C and 30 °C, at 8 different light intensities: 50, 100, 150, 250, 400, 600, 900, 1023 $\mu mol.m^{-2}s^{-1}$. The $\phi(T)$ function is calibrated using data from two different light intensities, 250 and 400 $\mu mol.m^{-2}s^{-1}$, at 5 different temperatures: 15, 21, 26, 30, 33 °C. The biomass concentration in the multicultivator ranged from 0.05 to 0.6 g/L.

Model was calibrated in two steps: A pre-calibration using the Multi-Cultivator data and a final calibration using data from a raceway batch. All numerical calculations were made using *Python* programming language. In the first step, the Trust Region Reflective algorithm - present in the *curve_fit* function from the *Scipy* library - was used to fit the growth function to the data from the Multi-Cultivator. The algorithm to find the parameters functions in a loop, an initial guess for the 6 parameters is used then μ_{max}, I_{opt} and α are calculated applying the *curve_fit* function to $\mu(T, I)$ considering the $\phi(T)$ value from the initial guess, then T_{min}, T_{opt} and T_{max} are calculated applying the *curve_fit* function to $\mu(T, I)$ considering the value of $\mu_{opt}(I)$ from the previously calculated parameters. These steps are repeated until convergence of the parameters. The respiration rate was taken from Béchet et al. (2017).

In the second step T_{min}, α and μ_{max} are reparametrized using data from one of the raceway experiments (calibration experiment). T_{min} is recalculated because temperatures reached in the cold raceway were much lower than what was possible to reach in the Multi-Cultivator, μ_{max} must be recalculated because it is dependent on light availability and the raceways are submitted to day and night cycles. α is also recalculated, after a posteriori verification of its high sensibility - the parameter α was recalculated to fit the experimental data. It was verified that the new recalibrated value was still correctly predicting the data from the Multi-Cultivator. For this second calibration, we use the L-BFGS-B algorithm present in the *minimize* function from the *Scipy* library, starting from many initial guesses. The final set of parameters with the least squared error is chosen.

The function of the extinction parameter σ was calibrated using a similar method from Béchet et al. (2017). A total of 12 values of σ were obtained measuring the light intensity at the surface and at the bottom of the raceway at different biomass concentrations.

3.2.7 Model validation criteria

We validate the model by comparing simulations to measurements acquired in the 3 models. We calculate Theil's inequality coefficient (TIC) (Decostere et al., 2016, Theil, 1961):

$$TIC = \frac{\sqrt{\sum_i (y_i - y_{i,m})^2}}{\sqrt{\sum_i y_i^2} + \sqrt{\sum_i y_{i,m}^2}} \tag{3.9}$$

where y_i is the simulated result and $y_{i,m}$ represents measured data. Values of TIC lower than 0.3 usually indicate a good model performance (Xianmin, 1993).

3.2.8 Sensitivity analysis

We studied the global sensitivity of the model parameters in relation to the simulated results using the input data from the raceways. The sensitivity coefficient was defined as in Bernard et al. (2001):

$$\sigma_y^{\Delta p} = \frac{1}{t_f} \int_0^{t_f} \frac{y(p + \Delta p, x_0, u, \tau) - y(p, x_0, u, \tau)}{y(p, x_0, u, \tau)} d\tau \tag{3.10}$$

where y is the simulated output at time τ with parameter p, initial condition x_0 and input variables u (e.g. light intensity and temperature). Global sensitivity is calculated using Morris's screening method (Morris, 1991), but we replace the standard elementary effect by the sensitivity coefficient defined above. The parameters domain analysed is the region comprised between the base value and $\pm 100\%$ of the standard deviation of the calibrated parameter. The domain of each parameter is divided in 20 points, at each iteration parameters are changed by addition or subtraction of 5/19 of the total length of the domain. Morris algorithm is repeated 100 times, then the mean value of the sensitivity coefficient and its standard deviation are calculated for each parameter.

3.2.9 Water temperature model

To evaluate the performance of raceway production under greenhouse we parametrized an autoregressive model for raceway water temperature inside and outside of the greenhouse. The autoregressive model is based on external air temperature (given by Meteo-France (Gwennaëlle Larvor et al., 2020)) and light intensity. The model for water temperature inside the greenhouse has one additional constant, T_0. Note also that the time delays differ between the two models:

$$T_{w_{in}}(t) = aT_{ext}(t-3) + bT_{ext}(t) + cI(t-1) + dT_{w_{in}}(t-1) + T_0 \tag{3.11}$$

$$T_{w_{out}}(t) = aT_{ext}(t-4) + bT_{ext}(t) + cI(t) + dT_{w_{out}}(t-1) \tag{3.12}$$

where $T_{w_{in}}$ and $T_{w_{out}}$ are, respectively, the water temperature in the raceway inside the greenhouse and the temperature in the raceway outside, T_{ext} is the air temperature outside in degrees Celsius, I is the light intensity reaching the raceway surface in $\mu mol/(m^2 s)$ and t is the time in hours.

Time delays were identified in the parameters calibration phase. First, delays were fixed, and the parameters were determined through minimization of squared errors using the calibration dataset. This was repeated for a variety of time delays combinations. The model with the least error for the validation data was kept. Validation and calibration datasets for greenhouse water temperature are each composed of six weeks, for different months of the year. Only one week of data for calibration and one week for validation were available for the external raceway. Calibrated values can be found in Table 3.2.

Table 3.2: Model parameters to calculate water temperature inside the greenhouse or outside.

Parameter	a	b	c	d	T_0
Greenhouse	0.094	0.209	0.0033	0.660	1.207
Outside	0.058	0.151	0.0018	0.760	-

3.2.10 Solar data

Solar data to estimate annual production was obtained from NASA Langley Research Center (LaRC) POWER Project. Daily ground radiation, H, was calculated with the method of Duffie and Beckman (1991), using daily insolation clearness index and the following equations:

$$H = H_0.K_T \tag{3.13}$$

where K_T is the insolation clearness index obtained from the NASA database, H_0 is the daily extraterrestrial radiation on a horizontal surface in $J.m^{-2}$ given by:

$$H_0 = c(1 + 0.033\frac{360n}{365})(\cos\,\phi\,\cos\,\delta\,\sin\,\omega_s + \frac{\pi\omega_s}{180}\sin\,\phi\,\sin\,\delta) \tag{3.14}$$

where c is a constant equals to 3.76 10^7, n is the n^{th} day of the year, ϕ is the latitude, δ is the sun's declination angle and ω_s is the sunset hour angle. Average hourly irradiance, I, in $W.m^{-2}$ can be calculated by:

$$I = H \times \frac{\pi(a + b\,\cos\,\omega)}{24 \times 3600}\frac{\cos\,\omega - \cos\,\omega_s}{\sin\,\omega_s - \frac{\pi\omega_s}{180}\cos\,\omega_s} \tag{3.15}$$

where a and b are coefficients dependent on sunset solar angle (Duffie and Beckman, 1991) and ω is the hour angle. We assume a conversation factor of 2.02 from $W.m^{-2}$ to $\mu mol.m^{-2}.s^{-1}$ (Hassika et al., 1997). Also, based on our data we consider that the film covering the greenhouse attenuates 20 % of the solar radiation.

3.2.11 Yearly cultivation prediction and greenhouse efficiency

The yearly production was simulated by the model for *Tetraselmis suecica*, *Dunaliella salina* (Béchet et al., 2017), *Spirulina platensis* (Venkataraman, 1997, Qiang et al., 1998), *Phaeodactylum tricornutum* (Bernard and Rémond, 2012, Fernández et al., 1997, Bitaubé Pérez et al., 2008, Wu et al., 2010) and *Chlorella vulgaris* (Béchet et al., 2015). The model parameters for *Chlorella vulgaris* and *Dunaliella salina* were both validated in Béchet et al. (2015) and Béchet et al. (2017), respectively. The parameters for the three other species were extrapolated from data available from several works, as detailed in Appendix Table 2. We simulated yearly production for locations near Nice, Rennes and Paris which represent three different climates in France. The yearly production was simulated considering a continuous system with a dilution rate D. In this case the equation describing the system is:

$$\frac{dX}{dt} = \mu_{net}X - DX \tag{3.16}$$

The resulting average daily productivity P is computed as follows:

$$P = \frac{1}{T}\int_0^T D(t)X(t)dt \tag{3.17}$$

The biomass concentration that maximizes production P for each season, X_{opt}, was computed based on a day representing the average conditions of the season (see Appendix A). X_{opt} is used as the initial condition for each season and D is either equal to μ_{net} or zero if $X(t) < X_{opt}$ in order to maximize productivity P.

The yearly productivity was computed, for each location, considering three scenarios: cultivation outside, inside the greenhouse with free temperature and optimal temperature control inside the greenhouse. For this last case it is assumed that a climate control system could reach any temperature between the two extreme former cases. This means that when the temperature inside the greenhouse is above the optimal temperature, it will be regulated nearby the optimal temperature by increasing the air flux entering (and leaving) the greenhouse. X_{opt} is calculated for each season as well as for each cultivation condition.

We define the greenhouse efficiency as the gain in productivity when growing algae in the greenhouse:

$$e_g = \frac{P_{greenhouse}}{P_{outside}} \tag{3.18}$$

where $P_{greenhouse}$ and $P_{outside}$ are the optimal productivities outside or under greenhouse given the conditions of each season. Greenhouse efficiency is used to analyse the impact of each parameter on the choice of cultivating under greenhouse.

To better understand the factors affecting the greenhouse efficiency we compute its sensitivity with respect to the model parameters:

$$\sigma_{e_g}^{\Delta p} = \frac{e_g(p+\Delta p) - e_g(p)}{e_g(p)} \tag{3.19}$$

and Δp varies between $\pm 10\%$ of the base values for *Tetraselmis suecica*.

3.3 Results

3.3.1 Model calibration

Table 3.3 shows the model parameters for *Tetraselmis suecica*. The optimal temperature for gross growth rate (T_{opt}) was found to be 30.9 °C (when corrected with respiration, the optimal temperature for net growth is 30.4 °C) which is in the literature range (Tredici et al., 2015, Molina et al., 1991, Weiss et al., 1985). The parameter T_{max} was 34.5 °C, in the upper range compared to literature (25-32 °C) (Molina et al., 1991, Weiss et al., 1985). This can be explained by an acclimation of the cells, since the raceways were inoculated at the end of summer when water temperature frequently exceeded 30 °C. Finally, parameter T_{min} was estimated to be -14 °C , but this parameter has been shown to be rather artificial, and this low value means that growth rate is less sensitive to temperatures changes below the optimal rate (*i.e.* that $\frac{d\mu}{dT}$ for temperature below T_{topt}, is closer to zero when T_{min} is very low).

Table 3.3: Growth model parameters for *Tetraselmis suecica*. In parenthesis the associated standard deviation computed from the covariance matrix.

Parameter	Unit	Value (SD)
μ_{max}	(d^{-1})	3.5 (0.1)
I_{opt}	$(\mu mol.m^{-2}s^{-1})$	571 (73)
α	$(m^2.s.\mu mol^{-1}d^{-1})$	0.024 (0.003)
B	-	-0.635 (0.055)
A	$(L^{B+1}.g^{-B-1}.m^{-1})$	35.30 (7.43)
T_{min}	$(°C)$	-14.7 (1.5)
T_{opt}	$(°C)$	30.9 (0.9)
T_{max}	$(°C)$	35.2 (0.4)
β	$(°C^{-1})$	0.0715 (0.0002)
$\lambda_0(day)$	(d^{-1})	0.042 (0.003)
$\lambda_0(night)$	(d^{-1})	0.052 (0.003)

Table 3.4: Global sensitivity of the model parameters for *Tetraselmis suecica*

Parameter	Global Sensibility mean	Variance
μ_{max}	0.179	0.015
α	0.141	0.012
B	0.137	0.011
T_{max}	0.030	0.001
β	-0.003	0.000
I_{opt}	-0.049	0.003
λ	-0.058	0.002
T_{min}	-0.081	0.006
T_{opt}	-0.153	0.007
A	-0.153	0.014

3.3.2 Sensitivity analysis

Results of the global sensitivity analysis are shown in Table 3.4. The global sensitivity mean indicates the level of influence of the parameter on the output, negative values indicate that an increase in the parameter will decrease the output of the model (i.e. final biomass concentration). The parameter with the most influence on the output was μ_{max}, followed by the extinction coefficient parameters, T_{opt} and α. The variance indicates if there are coupling effects between parameters or nonlinear responses. The results show that the variance correlates with the absolute value of the sensibility, indicating little coupling in the analyzed range. The weak coupling between parameters is in accordance with our model since the growth rate function is written as a multiplication of two independent functions depending either on light or temperature. However, nonlinear responses could be expected, for example, when the water temperature is close to one of the cardinal temperatures of the model.

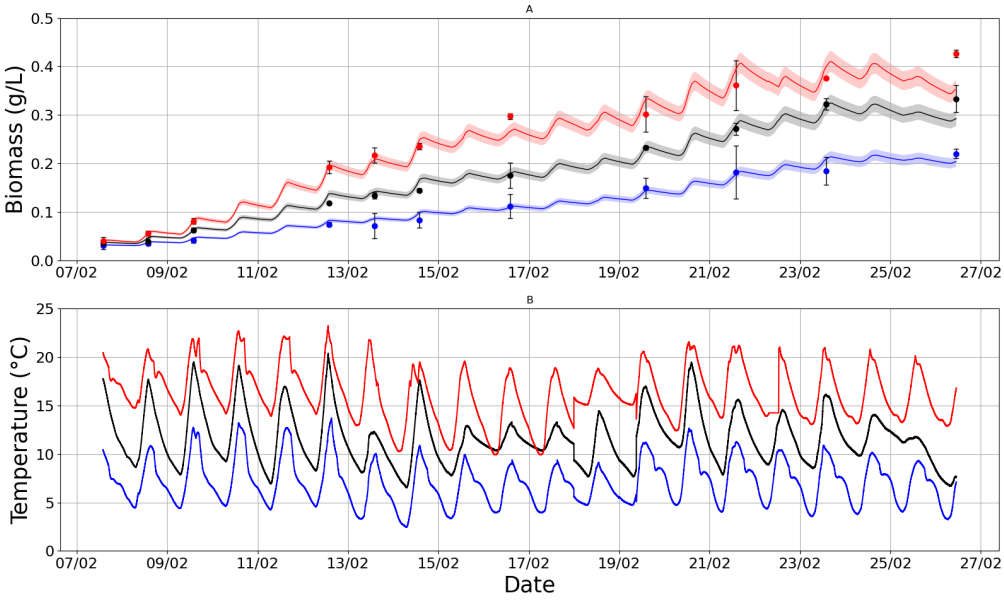

Figure 3.1: Calibration experiment. Points represent the measured dry weight and its confidence interval. A - curves represent the model simulation, the colored areas represent the region inside $\pm$ the standard deviation. B - measured temperature for the raceways. Red, blue and black colors are the data for the $+5°C$, $-5°C$ and at ambient temperature raceways, respectively.

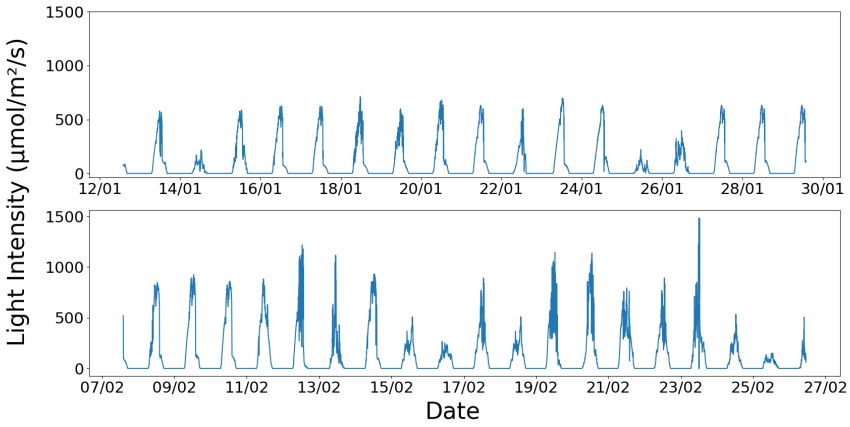

Figure 3.2: Light intensity for the two experimental campaigns

3.3.3 Experiments and simulation

Figures 3.1 and 3.3 show simulated and measured biomass evolution in the 3 ponds during the 2 experiments. The raceway with +5 °C shift was always the most productive, followed by the raceway at ambient temperature. The heated raceway showed a twofold productivity compared to the cold one for the second experiment in February, demonstrating the strong limitation due to temperature in the algal productivity.

Figure 3.2 shows the light intensity received by the raceways during the two set of experiments. For the second experiment, raceway ponds were exposed to a more intense radiation, most of the days light intensity reached at least 750 $\mu mol.m^{-2}s^{-1}$ and the maximum intensity was 1500 $\mu mol.m^{-2}s^{-1}$, while during the first experiment it never exceeded 700 $\mu mol.m^{-2}s^{-1}$. Since there is no marked difference in the temperature, this explains the difference in productivity between the two batches. Productivity was never reduced due to photoinhibition at light intensities higher than I_{opt}, since the average growth rate $\overline{\mu_{opt}}$, even at the lowest biomass concentration, would decrease only for light intensities higher than 1400 $\mu mol.m^{-2}s^{-1}$.

A TIC value of 0.05 was calculated using data from the validation raceway experiment data and a value of 0.08 using data from all raceway batches, which demonstrates the good model predictive capability.

Some situations were less accurately predicted by the model, especially when a lag phase was observed after culture inoculation. Predicting growth in the stress phase immediately after inoculation is often very challenging. The underestimation of biomass concentration in the validation experiment (Figure 3.3) can be due to grazing by ciliates in the heated raceway, as it has been regularly observed at higher temperature.

These results highlight the importance of managing the culture temperature to keep an optimal productivity. Since the experiments were carried out during winter, temperature in all raceways was always below the optimal. Our study also shows that the use of a greenhouse in cold climates is a simple and energy efficient (*i.e.* without requiring heating energy) way of increasing productivity by simply keeping the culture at some degrees above ambient temperature.

3.4 Discussion

3.4.1 Contamination at higher temperature

Bio contamination is one of the most common problems in raceway culture, either with virus, bacteria or other microalgae species. A greenhouse is a barrier around the culturing environment that contributes to reduce the contaminations. This function of a greenhouse, provided that some basic precautions are respected, is the cheapest way to protect non-aseptic raceways from a direct contact with the biodiversity contained in the aerosols or in the rain (Sialve et al., 2015). A raceway under a greenhouse is thus an economical interesting trade-off between outdoor raceway and photobioreactors which can ensure axenic growth conditions, but for a much higher cost. However, the greenhouse results in a temperature shift, which is a double edged sword because it also enhances the growth rate of the biological contaminants present in the culture. We observed most notably contamination by bacteria and ciliates. Ciliates are predators of microalgae, and their presence can rapidly lead to a crash of the culture, wherein all the biomass can be consumed

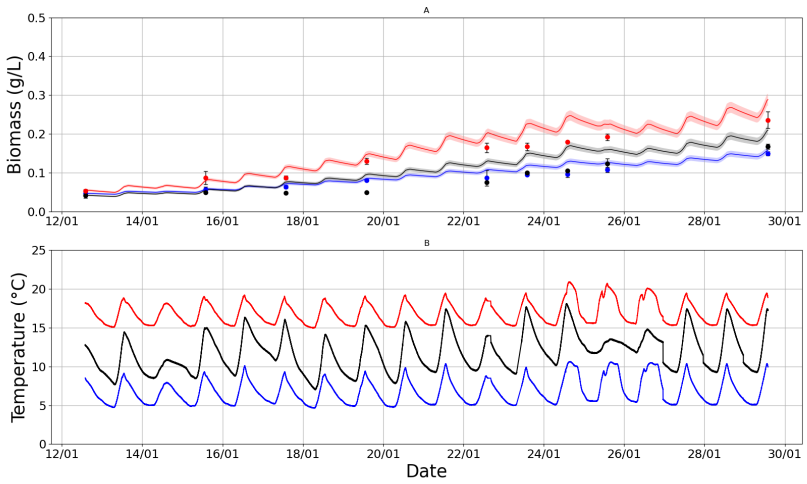

Figure 3.3: Validation experiment. Points represent the measured dry weight and its confidence interval. A - curves represent the simulation of the model, the colored areas represent the region inside $\pm$ the standard deviation. B - measured temperature for the raceways. Red, blue and black colors are the data for the $+5°C$, $-5°C$ and at ambient temperature raceways, respectively.

in a few days. One way to mitigate the growth rate of ciliates despite temperature increase in the greenhouse is to modify salinity (von Alvensleben et al., 2013). *Tetraselmis* is very resistant to changes in salinity and it can grow even in fresh water (Fabregas et al., 1984). We observed under the microscope that dilution with tap water was an effective method to eliminate ciliates while keeping *Tetraselmis* cells mostly intact.

3.4.2 For which species is the greenhouse beneficial?

The model was used to simulate the yearly productivity of five species along the year. The model accounts for the shift in temperature but also for the light attenuation by the greenhouse film. Results for the estimated productivities over the whole year are shown in Table 3.6 (full data can be found at the Appendix A Tables 3, 4 and 5). These predictions are in accordance with the average values reported in the literature for outdoor cultivation (Banerjee and Ramaswamy, 2017). The low productivity for *Phaeodactylum Tricornutum* is probably due to the fact that this species does not grow at high temperatures. Silva Benavides et al. (2013) recorded an average productivity of 11.7 $g.m^{-2}.d^{-1}$ in summer, regulating water temperature close to the optimal value, while the highest productivity we simulated for non optimal temperature conditions was of 8.9 $g.m^{-2}.d^{-1}$. De-Luca et al. (2019) also estimated for *Chlorella Vulgaris* productivities ranging from 37 in summer to 5 $g.m^{-2}.d^{-1}$ in winter using optimal control.

Apart from *Dunaliella*, all microalgae had better productivities during winter when cultivated under greenhouse. On spring and fall, greenhouse cultivation starts to be less effective and, in general, a strong greenhouse aeration by implementing air fluxes with exterior becomes regularly

Table 3.5: Sensibility of the greenhouse efficiency to model parameters (parameter for *Tetraselmis suecica*)

Parameter	$\sigma_{e_g}^{\Delta p}(10^{-2})$	Variance (10^{-2})
T_{opt}	2.176	0.848
T_{min}	0.897	0.086
B	0.324	0.073
α	0.289	0.048
A	-0.003	0.030
λ	-0.021	0.020
I_{opt}	-0.196	0.055
μ_{max}	-0.229	0.050
β	-0.782	0.104
T_{max}	-1.399	0.530

necessary to decrease temperature and thus maintain productivity in the same range as outdoor cultivation. In summer, cooling down the raceway by a high air exchange with the external environment becomes imperative, since water temperature can easily surpass the optimal and even maximum temperature. This explains the close to zero productivity for *Phaeodactylum tricornutum* during summer.

Without temperature control in the greenhouse, productivity for *Tetraselmis suecica* during summer, at a location near Nice, is similar to the average during cold seasons, with the risk of culture collapse on the hottest days (see Table 2 of Appendix A and Table 3.6). On the other hand, during winter the greenhouse can increase the productivity by more than 25%. Because during hot seasons the increase in temperature becomes harmful, the average production over the whole year in and out of the greenhouse are similar. Indeed, the months associated with the highest solar flux are more favourable in terms of light intensity, but the expected higher productivities are not met due to overwarming.

Dunaliella salina is a microalgae tolerant to higher temperatures. Naturally thriving in shallow ponds, it can grow at temperatures higher than 40°C while having an optimal temperature for growth around 34 °C. Despite this, it performs badly under greenhouse. It can be explained by the fact that its optimal growth rate is reached at very high light intensities, being very resistant to photo inhibition. It results that the light loss due to the greenhouse film deeply impacts productivity. The same explanation holds for *Chlorella*, but since it has an even higher optimal temperature, 37°C, productivity stays similar under both cultivation systems.

The cyanobacteria *Spirulina platensis* was the only species to be more productive under greenhouse during all seasons. Its characteristics are perfect for greenhouse production, i.e. high temperature tolerance and optimal growth at low light intensities. Cultivation under greenhouse could increase winter productivity by more than 80 %, and by more than 20 % over the whole year. The predictions for these microalgae point out the necessity to adapt the species to be cultivated to the local climate and cultivation system. As in agriculture, it is pointless to cultivate a single species along the year, and the species must be alternated along the seasons. When considering Table 3.6 (see also Appendix A Tables 3, 4 and 5) it turns out that cultivating

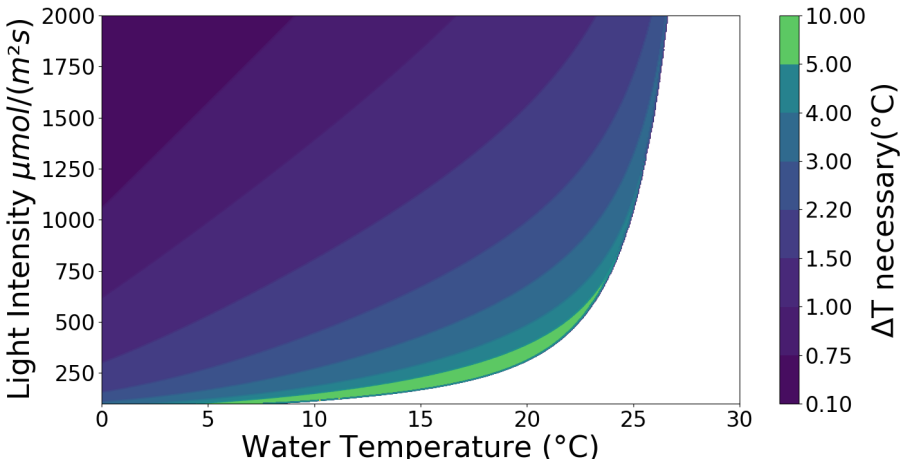

Figure 3.4: Temperature shift (ΔT *necessary*) inside the greenhouse to compensate the growth rate loss due to the 20% attenuation of the solar radiation for *Tetraselmis suecica*. In the white region it is not possible to compensate the loss with an increase in temperature.

Chlorella in winter, *Tetraselmis* in spring and *Spirulina* in summer and fall is the optimal solution for dry mass production for the region of Nice. This crop rotation strategy would improve annual biomass production by at least 7 % and up to 20 % by comparing it to a monoculture of *Spirulina* and *Chlorella*, respectively. This strategy would slightly change for the regions of Paris and Rennes, where it would be more productive to cultivate *Chlorella* during fall. This highlights how the optimal cultivation strategy depends on the local climate. The qualitative value of each different species also has a primary influence on the cultivation strategy. For example the dry biomass of *D. salina* is ten times more expensive than for *S. platensis*.

The greenhouse, when passively used, can extend the culture duration in the cold seasons, but species that are not prone to photoinhibition must be grown in spring and summer without a greenhouse. Furthermore, simulation shows that a system of temperature control inside the greenhouse, simply by air exchange with the external environment, is able to greatly improve productivity compared to a simple greenhouse. Specially during summer, when temperature can surpass the optimal temperature for the microalgae species. The contamination problem can then become an issue and an air filtering system must probably be implemented.

3.4.3 Light attenuation by the greenhouse film: a heavy burden

To better understand the consequence of loosing 20% of incoming light due to the greenhouse film, we computed the temperature shift in the greenhouse necessary (ΔT *necessary*) to compensate this light reduction by providing the same growth rate (see Figure 3.4). The figure presents this necessary shift in temperature for different light intensity scenarios and various water temperature in a raceway without greenhouse. At high light intensities and low temperatures (a situation which is not common), the greenhouse is more effective. A large shift in

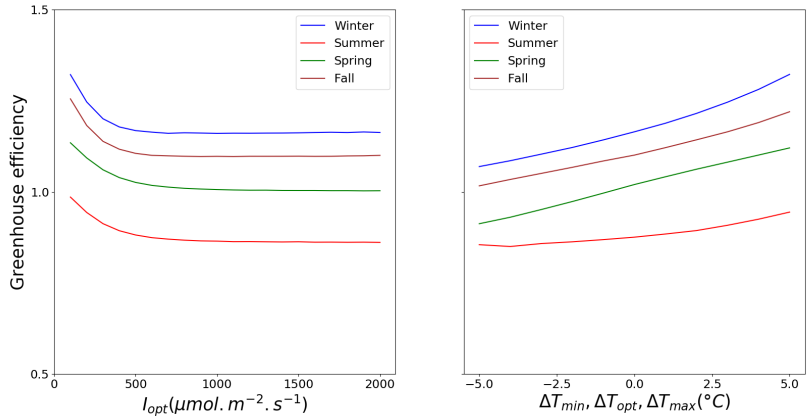

Figure 3.5: Greenhouse efficiency as a function of the parameter I_{opt} (left) and as a function of a shift ΔT in all the cardinal temperatures parameters (right), for *Tetraselmis suecica* along different seasons of the year at a location near Nice.

temperature can result in significant gains of growth rate. Nonetheless, when temperature in outdoor raceways is already high, the greenhouse can no longer compensate the light attenuation.

To sketch which characteristics of a species would maximize greenhouse efficiency, we consider the parameters having the highest impact (see Table 3.5), as deduced from the sensitivity analysis (Equation 19). Furthermore, since variances indicate weak coupling between parameters, we can expect the results to be valid for other species as well. As expected, higher T_{opt} and T_{min} (species adapted to warmer conditions) promote greenhouse usage. Surprisingly, a species with higher T_{max} (but same T_{opt}) would not benefit from the greenhouse. A closer look at equation $\phi(T)$ shows that a higher value of T_{max} reduces $\frac{d\phi(T)}{dT}$ between T_{min} and T_{opt} so that the increase in temperature in the greenhouse results in a less marked gain in growth rate. But as shown in Figure 3.5, assuming a shift in the temperature growth curve (simultaneous increase in T_{min}, T_{opt} and T_{max}), temperature parameters have a positive impact on greenhouse efficiency. The species with higher I_{opt} (subjected to less photoinhibition), are less interesting for greenhouse cultivation. Growth rate decreases when light intensity is higher than I_{opt} due to stronger photoinhibition. For species where I_{opt} is below the average values of light intensity, the greenhouse is more efficient and protects against photoinhibition. Figure 3.5 also demonstrates clearly how greenhouse efficiency increases from summer to winter. These results qualitatively apply to other species, as long as average temperatures stay in the range of T_{min} and T_{max}.

3.5 Conclusions

The first motivation to use a greenhouse is the protection against contamination and potential negative effect of rain. By increasing temperature, greenhouses generally improve productivity during winter and autumn, but this effect is less marked in spring and summer as a consequence

Table 3.6: Simulated productivity ($g.m^{-2}.d^{-1}$) during hot (summer and spring) and cold (winter and fall) seasons for raceway cultivation outside (O.), inside greenhouse without (G.) and with temperature control (G. T c.)

		Cold seasons			Hot seasons			Year		
Species		Nice	Paris	Rennes	Nice	Paris	Rennes	Nice	Paris	Rennes
	G.	5.4	2.7	2.9	13.3	12.2	14.5	9.3	7.5	8.7
T. suecica	O.	5.2	2.5	2.6	**25.9**	17.5	17.9	15.6	10	10.3
	G. T c.	**6.3**	3.1	3.3	25.2	17.9	18.6	**15.8**	10.5	10.9
	G.	7.1	3.5	3.7	24.2	18.6	18.3	15.6	11.1	11
S. platensis	O.	4.5	1.8	1.9	21.5	15.2	14.4	13	8.5	8.1
	G. T c.	**7.3**	3.7	3.9	**24.8**	19.1	18.6	**16**	11.4	11.3
	G.	7.5	5.3	5.6	19.5	15.3	15.7	13.5	10.3	10.6
C. vulgaris	O.	7.7	5.4	5.6	**21.1**	16.6	16.5	**14.4**	11	11.1
	G. T c.	**7.8**	5.6	5.8	20	16	16.1	13.9	10.8	10.9
	G.	3.3	1.9	2	10.9	7.7	8.1	7.1	4.8	5.1
D. salina	O.	**3.8**	2	2.2	**13.9**	9.6	9.8	**8.9**	5.8	6
	G. T c.	3.6	2.1	2.2	11.7	8.3	8.5	7.7	5.2	5.4
	G.	3.3	2.5	2.7	1	2.5	3.2	2.2	2.5	3
P. tricornutum	O.	4.1	2.8	3.1	4.1	5.9	7.3	4.1	4.4	5.2
	G. T c.	**4.5**	3.2	3.5	4	6.1	**7.4**	4.3	4.7	**5.5**

of light attenuation in conjunction with lethal temperatures. Cultivation under greenhouse should be considered when outdoor water temperature is several degrees lower than the optimal growth temperature for species susceptible to photoinhibition. Alternation of the cultivated species is also a good strategy to improve annual productivity.

Chapter 4

Control of zooplankton populations in microalgal cultivation systems

This chapter has been adapted from the published article
Martínez, C., Pessi, B.A., Bernard, O., 2022. Optimal production of microalgae in the presence of grazers. Journal of Process Control 118, 153–164.

Abstract

Zooplankton contamination represents a major constraint in large-scale microalgal cultivation systems. While zooplankton contamination cannot be avoided, their development can be controlled by regulating the dilution rate. However, it is not straightforward to find the best control strategy for the dilution rate. Low dilution rates (or long retention times) favor grazer development and high dilution rates avoid their establishment at the risk of reducing microalgal productivity. Furthermore, the presence of periodic regimes arising from the interaction between predator-prey makes it unclear if a strategy to wash-out the predators must be used. In this work, we study the role of the dilution rate in the control of zooplankton populations and in the optimization of biomass productivity. We show that in the long-term operation (static optimization control problem or SOCP), the optimal constant dilution rate must ensure the eradication of the zooplankton population. In the case of a time-varying dilution rate, we numerically solve an optimal control problem (OCP) over a finite interval of time. We find that the optimal solution approaches the solution for the SOCP most of the time, except when zooplankton actively avoids the pond outflow. Based on these results, we propose a simple sub-optimal feedback control that approximately matches the solution of the OCP when the initial concentration of grazers is low.

4.1 Introduction

Microalgae are a growing natural resource with several commercial applications in the fields of pharmaceuticals, cosmetics up to feedstocks for aquaculture (Rizwan et al., 2018). A challenge when cultivating these organisms, especially in open reactors like raceways, is to limit contamination by other organisms such as viruses, bacteria, fungus, other microalgal species, and grazers (Molina et al., 2019, Molina-Grima et al., 2021). In industrial conditions, it is indeed impossible to operate the process under axenic conditions, and the surrounding environment will permanently bring invaders to the medium. In particular, predators (ciliates, rotifers, daphnia, copepods, etc.) are a poignant issue since they may rapidly develop and lead to a culture crash within a few days (Molina-Grima et al., 2021).

To date, there is no efficient strategy to limit crop loss through zooplanktonic predation, and most of the microalgae grown outdoors in open reactors are extremophiles, which develop in a medium hostile to most of the organisms present in the surrounding ecosystems. Chemical pesticides can limit contamination but they have both high economical and environmental costs. Physical methods are effective, but they are not cost-effective and they can also affect negatively microalgae (Montemezzani et al., 2015). An alternative that has scarcely been explored, is the control of the dilution rate. Since species with a generation rate slower than the dilution rate are unlikely to establish in cultivation systems, high dilution rates reduce grazers abundance (Schlüter and Groeneweg, 1981, Montemezzani et al., 2016).

The dilution rate (the inverse of the hydraulic retention time) is one of the most important operational variables for continuous cultivation systems (Hajinajaf et al., 2021). Different authors have studied how to control the dilution rate to maximize biomass productivity. In the absence of predators, when algal growth is limited only by light, there exists the well-known compensation principle (Mairet et al., 2015, Martínez et al., 2018): the productivity is maximal when at the bottom of the culture, the specific growth rate equals the respiration rate. The existence of an optimal dilution rate has been experimentally shown by many authors in the absence of predators (Tang et al., 2012, Qiang and Richmond, 1996). Other theoretical works have considered the dilution rate varying in time (Grognard et al., 2014, Muñoz-Tamayo et al., 2013). Regarding contaminated systems, there are only a few works concerning the impact of the dilution rate (Flynn et al., 2017, Martínez et al., 2021).

Any control of the dilution rate should consider two important aspects. The first one is that zooplankton often escape the outflow of the pond. For example, cladocerans migrate to near the pond surface at night, resulting in lower densities in the deep water column where the outflow is often located (Montemezzani et al., 2016, Bhuiyan et al., 2011). This has also been observed in chemostat experiments, where cladocerans concentrate near the bottom of the chemostat, remaining below the surface from which the overflow occurs (Sommer, 1992). The critical dilution rate to eradicate some zooplankton populations is therefore higher than their generation rate. Previous theoretical studies assume that microalgae and zooplankton are equally affected by the dilution rate (Flynn et al., 2017, Martínez et al., 2021, Deruyck et al., 2019), and therefore they may underestimate the impact of grazers. The second important aspect is shared by any predator prey model, that is, the existence of limit cycles. Low dilution rates favor the existence of limit cycles (Martínez et al., 2021). This is because of the enrichment paradox: favorable

conditions for the prey may cause the population to destabilize into a limit cycle (Rosenzweig, 1971, Fussmann et al., 2000, Deruyck et al., 2019). Comparing the productivity along limit cycles with that of equilibria is not trivial. In Martínez et al. (2021), we numerically show that in the long-term, productivity is higher in the absence of grazers. Therefore, limit cycles cannot be an optimal regime. However, we did not consider that zooplankton may avoid outflow. As shown in this work, the avoidance of the outflow favors the existence of limit cycles.

In this work, we investigate how to control the dilution rate to maximize the microalgae production in a culture susceptible to predation. We study two cases: (I) the dilution rate is constant and the system is operated in the long-term; (II) the dilution rate is time-varying, and biomass is harvested over a finite interval of time. In both cases, we consider a chemostat model in which microalgae and grazers grow together (predator-prey model), and we assume that only a fraction of predators is diluted. For Case (I), we determine the necessary and sufficient conditions for the coexistence of both populations, and we show that any solution of the model approaches either an equilibrium or a limit cycle. This allows us to state the static optimal control problem: to find the value of the dilution rate that maximizes the biomass productivity in the long-term (*i.e.* when an attracting set is reached). Using the toolbox Matcont for MATLAB (Dhooge et al., 2003), we numerically find the best dilution rate. This is not trivial due to the existence of limit cycles. As we show in this work, for a fixed dilution rate, limit cycles provide higher biomass productivity than equilibria. For case (II), we use the software BOCOP (Bonnans et al., 2011) to numerically solve the optimal control problem of finding the best strategy to maximize the biomass productivity over a finite time interval.

This chapter is organized as follows. In Section 4.2, we describe the microalgae-grazers model. In Section 4.3, we determine the necessary and sufficient conditions for the survival of predators. In Section 4.4, we study the steady-state optimization problem of maximizing biomass productivity. In Section 4.5, we study the optimal control problem of maximizing biomass productivity in a fixed interval of time. In Section 4.6, we discuss our results. Finally, the conclusion is presented in Section 4.7.

4.2 Model description

4.2.1 Mass balance equations

We consider the growth of microalgae (with density x) in a chemostat contaminated by predators of population density y (see Figure 4.1). The dynamics of both populations is given by the following system of ordinary differential equations

$$\begin{aligned} \frac{dx}{dt} &= [\mu(x) - D]x - \frac{1}{\gamma}\nu(x)y, \\ \frac{dy}{dt} &= [\nu(x) - m - \alpha D]y. \end{aligned} \tag{4.1}$$

The terms μ and ν are the specific growth rate of microalgae and predators, respectively. The parameter $\gamma \in [0, 1]$ corresponds to the assimilation efficiency and m is the mortality rate of predators. The term D is the dilution rate that is defined as the inlet flow (F) divided by the

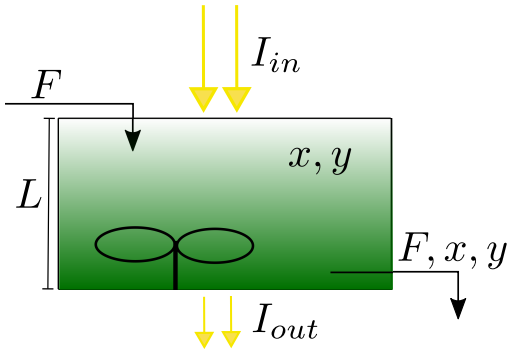

Figure 4.1: Scheme of a continuous microalgae (x) culture contaminated by predators (y). The culture is illuminated from above with an incident light intensity I_{in}. The culture has a depth L. The light intensity at the bottom is I_{out}. The inlet flow F is equal to the output flow.

reactor volume (see Figure 4.1). Finally, the term α is a parameter reflecting the fact that some predators can escape from dilution, for example, by accumulating at a place below the output of the chemostat (Sommer, 1992). This parameter takes values between 0 and 1. If $\alpha = 1$, then algae and predators are equally diluted; however, if $\alpha = 0$, then predators are unaffected by dilution. Along this work, we assume that α is constant in time. While this assumption is reasonable for indoor cultures, where parameters such as light and temperature are kept constant, the main motivation for this assumption is to keep the model simple. Predator-prey models with time-varying parameters may exhibit a chaotic behavior (Gao et al., 2009).

4.2.2 Specific growth rates

The specific growth rate of predators depends on the microalgae concentration as follows:

$$\nu(x) = \nu_{max} \frac{x}{K_x + x}, \tag{4.2}$$

with ν_{max} the maximal growth rate and K_x a half saturation constant.

The growth rate of microalgae follows from the combination of a light response model and a light distribution model. Light intensity decreases as it passes through the microalgae culture due to absorption and scattering by algal cells (Martínez et al., 2018). Let L be the depth of the culture, which is illuminated from above as illustrated in Figure 4.1. In line with standard hypotheses for photolimited photobioreactors (Bernard, 2011b), light is assumed to be attenuated exponentially according to the Lambert-Beer law. Thus, at a distance $z \in [0, L]$ from the illuminated surface, the corresponding light intensity $I(x, z)$ is given by:

$$I(x, z) = I_{in} e^{-kxz},$$

with $k > 0$ the specific light attenuation coefficient of microalgae. Following Huisman et al. (2002), the growth rate of microalgae is obtained integrating the local specific growth rates over all the culture

$$\mu(x) := \frac{1}{L} \int_0^L p(I(z, x))dz - r, \tag{4.3}$$

Table 4.1: Parameters

Parameter	Value	Unit	Reference
p_{max}	1.68	d^{-1}	(Huisman et al., 2002)
K_I	108	μ mol m^{-2} s^{-1}	(Huisman et al., 2002)
r	0.1	d^{-1}	
k	0.2	m^2 g^{-1}	(Deruyck et al., 2019)
ν_{max}	1.4	d^{-1}	(Deruyck et al., 2019)
K_x	219	g m^{-3}	(Deruyck et al., 2019)
m	0.15	d^{-1}	(Deruyck et al., 2019)
γ	0.21		(Deruyck et al., 2019)
L	0.15	m	
I_{in}	1000	μ mol m^{-2} s^{-1}	
D_{max}	2	d^{-1}	

where $p(I)$ corresponds to the light response of microalgae and $r > 0$ is the respiration rate. The function $p(I)$ is described by a Monod model:

$$p(I) = p_{max} \frac{I}{K_I + I}, \tag{4.4}$$

where $K_I > 0$ is a half-saturation constant, and $p_{max} > 0$ is the maximal specific growth rate. Parameters of the model are given in Table 4.1.

The following lemma establishes some basic properties of the specific growth rate of microalgae.

Lemma 1. The function $\mu : [0, \infty) \to \mathbb{R}$ defined by (4.3) is continuous, strictly decreasing, and $\lim_{x\to\infty} \mu(x) = -r$.

Proof. See Appendix C in Martínez et al. (2020). □

From Lemma 1, we have that μ is strictly decreasing (see also Figure 4.2). This property reflects the self shading effect, that is, as microalgae concentration increases, the light availability in the medium decreases, thus reducing the growth rate. This implies that $\mu(0)$ is the maximal hypotetical growth rate of microalgae at which they tend to grow as their concentration decreases and the medium becomes transparent. From (4.3), we have that

$$\mu(0) = p(I_{in}) - r. \tag{4.5}$$

The light intensity I_{in} is assumed to be large enough such that microalgae can grow (gross growth rate larger than respiration), and therefore we assume that

$$\mu(0) > 0. \tag{4.6}$$

When $x > 0$, we can integrate (4.3) to obtain

$$\mu(x) = \frac{\mu_{max}}{kxL} \ln\left(\frac{K_I + I_{in}}{K_I + I_{out}(x)}\right) - r, \quad x > 0. \tag{4.7}$$

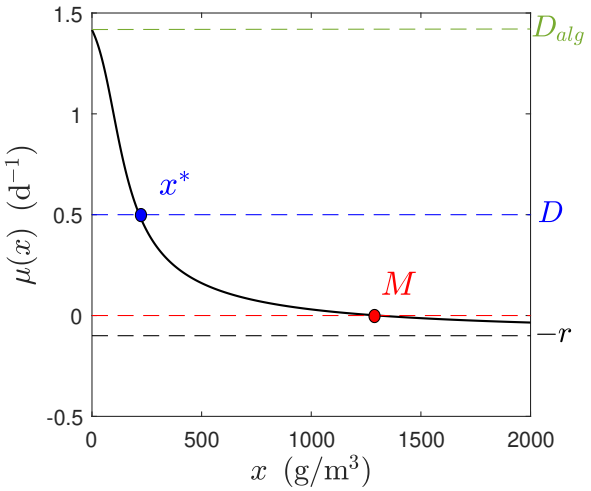

Figure 4.2: Graphical description of M and x^* defined by (4.8) and (4.11), respectively.

with $I_{out}(x) = I(x, L)$ the light intensity at the bottom of the culture.

Using Lemma 1 and (4.6), we have the existence of a unique $M > 0$ such that (see Figure 4.2):

$$\mu(M) = 0. \tag{4.8}$$

If the microalgae concentration is higher than M, then respiration (r) exceeds the average photosynthesis along the column, and the specific growth rate μ becomes negative. Thus, the quantity M represents the maximum population density that can be reached by microalgae at steady state (replace D and y by zero in (4.1)). In this work, we are interested in the case where predators can develop in the reactor, therefore we assume

$$\nu(M) > m. \tag{4.9}$$

If (4.9) does not hold, predators will naturally disappear from the reactor in the long-term.

4.3 Establishment of predators

4.3.1 Dynamics in the absence of grazers

To determine conditions for the establishment of predators in the chemostat, we begin describing the situation in which microalgae grow in the absence of predators, that is, we replace y by zero in (4.1). The dynamics of microalgae is then given by the following one-dimensional differential equation:

$$\frac{dx}{dt} = [\mu(x) - D]x. \tag{4.10}$$

If $\mu(0) > D$, since μ is strictly decreasing (Lemma 1) and $\mu(M) = 0$, there is a unique $x^* > 0$ such that (see Figure 4.2):

$$\mu(x^*) = D. \tag{4.11}$$

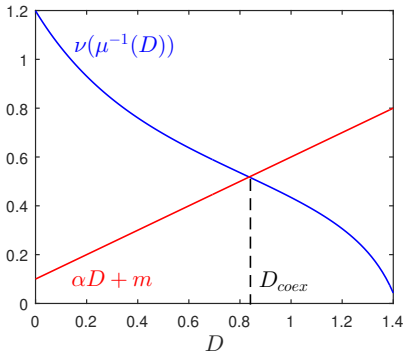

Figure 4.3: Graphical representation of D_{coex}. As α or m decreases, the value of D_{coex} increases.

It is clear that x^* is globally asymptotically stable (GAS) with respect to (4.10) on $(0, \infty)$. Sometimes we will write $x^*(D)$ instead of x^* to emphasize the fact that x^* depends on the dilution rate. On the other hand, if $\mu(0) \leq D$, then any solution to (4.10) converges to 0. Thus, the dilution rate $D_{alg} := \mu(0)$ represents the minimal dilution rate at which microalgae are washed out from the culture (see Figure 4.2). The equilibrium (of (4.1)) characterized by the presence of microalgae and the absence of predators, whenever it exists, will be denoted by

$$E^* = (x^*, 0). \tag{4.12}$$

4.3.2 Coexistence of microalgae and predators

The following proposition answers the question whether microalgae and predators coexist in the long-term.

Proposition .1 (Coexistence)**.** There is a dilution rate $D_{coex} \in (0, D_{alg})$ such that

(a) If $0 < D < D_{coex}$, then there is a unique coexistence equilibrium $E_c = (x_c, y_c)$, and any solution to (4.1) approaches asymptotically either E_c or a positive periodic solution.

(b) If $D_{coex} \leq D < D_{alg}$, then there is no coexistence equilibrium, and any solution to (4.1) approaches E^* asymptotically.

(c) If $D_{alg} < D$, then any solution approaches asymptotically $(0, 0)$.

Proof. See B.1. □ □

Proposition .1 shows the existence of a dilution rate D_{coex} that characterizes the long-term coexistence of microalgae and zooplankton. Note that since zooplankton needs microalgae to grow, the survival of predators is equivalent to the coexistence of both populations.

The value of D_{coex} can be determined from the following system of equations for x^* and D (see the proof of Proposition .1):

$$\begin{aligned} \mu(x^*) &= D, \\ \nu(x^*) &= m + \alpha D. \end{aligned} \tag{4.13}$$

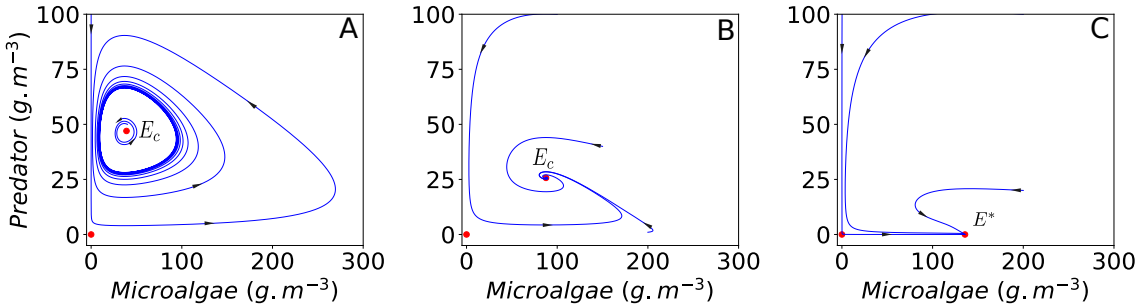

Figure 4.4: Possible assymptotic behaviors of the solutions of (4.1) when $\alpha = 0.5$. **A**. The coexistence equilibrium E_c exists and is unstable ($D = 0.1\, d^{-1}$), and solutions approach a limit cycle. **B**. The coexistence equilibrium E_c exists and is stable ($D = 0.5\, d^{-1}$). **C**. There is no coexistence equilibrium ($D = 0.8\, d^{-1}$)

From Lemma 1, we have that the inverse of μ exists. Thus, from the first equation in (4.13), we can write $x^* = \mu^{-1}(D)$. Then, D_{coex} is obtained as the intersection between the line $m+\alpha D$ and the function $D \longmapsto \nu(\mu^{-1}(D))$ (see Figure 4.3). We note that low values of α (longer retention time of grazers in the reactor) and low values of m (low mortality rate) result in higher values of D_{coex}.

From now on, the coexistence equilibrium, whenever it exists, will be denoted by

$$E_c = (x_c, y_c). \tag{4.14}$$

The following result states some dynamical properties of E_c.

Lemma 2. (Stability of the coexistence equilibrium) Let us define the function $h : (0, \infty) \to \mathbb{R}$ by

$$h(x) := \frac{(\mu(x) - D)x}{\nu(x)}. \tag{4.15}$$

If the coexistence equilibrium E_c exists, then:

(a) if $h'(x_c) < 0$, then E_c is a sink (locally stable),

(b) if $h'(x_c) = 0$, then E_c is globally stable on $(0, \infty) \times (0, \infty)$

(c) if $h'(x_c) > 0$, then E_c is a source (unstable).

Proof. See B.2. □ □

Remark .1 (Existence of limit cycles)**.** Lemma 2 describes the local stability of the coexistence equilibrium. In particular, it states sufficient conditions for the existence of limit cycles, that is, when E_c is unstable. Proving that the instability of E_c is a necessary condition for the existence of limit cycles has been the concern of many authors (Ardito and Ricciardi, 1995, Hesaaraki and Moghadas, 2001, Moghadas and Corbett, 2008). However, most of the results are limited

Table 4.2: Different notations for the dilution rate.

Notation	Description
D_{alg}	Minimum dilution rate at which microalgae go extinct.
D_{coex}	Minimum dilution rate at which predators go extinct.
D_C	Solution of (4.17). Dilution rate at which a pure culture of microalgae reaches its maximal productivity.
D_{SOCP}	Conjectured solution of the SOCP (4.20). Dilution rate at which the contaminated cultures reaches its maximal productivity.
D_{OCP}	Optimal solution of the OCP (4.24).
$\hat{D}$	Suboptimal feedback control proposed in this work (see (4.25)).

to the case when $\mu(x)$ is described by logistic growth. In this work, we do not aim to prove such results for our model, which could be the subject of a completely different work. However, numerical simulations suggest that the instability of E_c is a necessary and sufficient condition for the existence of a limit cycle (see Figure 4.4).

Lemma 3. Let h be the function defined by (4.15). Assume that the coexistence equilibrium E_c exists. If $h'(x_c) > 0$ and the following inequality holds for all $x \in (0, x^*) - \{x_c\}$:

$$\frac{d}{dx}\left(\frac{\nu(x)h'(x)}{\nu(x) - m - \alpha D}\right) \leq 0. \tag{4.16}$$

then (4.1) admits a unique limit cycle, which is globally stable on $(0, \infty) \times (0, \infty)$.

Proof. Direct application of Theorem 2.2 in Hwang (1999). □

Remark .2 (Uniqueness of limit cycles)**.** Following Lemma 3, we can show numerically that when E_c is unstable, (4.1) admits a unique limit cycle that is globally stable (Martínez et al., 2021). Such result is probably not surprising, the multiplicity of limit cycles has only been observed, for example, in the presence of Allee effect on prey (González-Olivares et al., 2011) or non-monotonic functional responses by predators (Xiao and Zhu, 2006).

4.4 Static optimal control problem (SOCP)

4.4.1 Productivity in the absence of grazers

In the absence of predators ($y = 0$), as discussed in the previous section, for any dilution rate $D \in [0, D_{alg}]$, the microalgae concentration converges toward the steady state $x^*(D)$ (defined by (4.11)). We then define the steady state biomass productivity as follows $P(D) := LDx^*(D)$. This term represents the quantity of microalgae that is produced per unit of area and time when a solution of (4.10) reaches its steady state. The units of $P(D)$ are g/m^2/d. The problem of maximizing P can be written as:

$$\begin{array}{ll} \max\limits_{D} & P(D), \\ s.t. & 0 < D < D_{alg}. \end{array} \tag{4.17}$$

It is well known that P reaches the maximum value when the following compensation condition holds (Mairet et al., 2015):

$$p(I_{out}(x^*(D))) = r. \tag{4.18}$$

We will denote by D_C the dilution rate at which the equilibrium x^* verifies the compensation condition (4.18). Thus, D_C is the solution to (4.17) (The different notations for the dilution rate are summarized in Table 4.2.).

4.4.2 Productivity in the presence of grazers

To extend the definition of productivity to a culture contaminated by predators, we must take into account the asymptotic behavior of any solution to (4.1) with positive initial conditions. According to Proposition .1, there is a dilution rate $D_{coex} > 0$ such that for any $D < D_{coex}$ microalgae and predators survive in the long-term. Moreover, they either settle in the coexistence equilibrium E_c or they approach a periodic solution of (4.1) (see Figure 4.4). Following Remark .2, we assume that, when E_c is unstable, there is a unique limit cycle. We will denote the trajectory and period of the limit cycle by (x_p, y_p) and T, respectively.

We define the areal long-term productivity, denoted by $Q(D)$, as

$$\begin{cases} \dfrac{L}{T}\displaystyle\int_0^T Dx_p(t)dt, & if \quad E_c \text{ exists and is unstable,} \\ LDx_c, & if \quad E_c \text{ exists and is stable,} \\ LDx^*, & if \quad D_{coex} \leq D < D_{alg}. \end{cases} \tag{4.19}$$

The definition of Q accounts for three different types of asymptotic behavior in which microalgae are present; limit cycle, coexistence equilibrium, and equilibrium without grazers (see Figure 4.4). Note that when the coexistence equilibrium exists and is unstable, Q corresponds to the daily average biomass productivity during the whole period of the limit cycle. In this way, the units of Q are exactly the same of P, and the productivity along limit cycles can be compared to that provided by equilibria.

We are interested in finding the constant dilution rate that provides the highest value of Q, that is, we want to solve the following optimization problem:

$$\begin{array}{ll} \max\limits_{D} & Q(D), \\ s.t. & 0 < D < D_{alg}. \end{array} \tag{4.20}$$

We will refer to (4.20) as the *static optimal control problem* (SOCP).

4.4.3 Limit cycles are not optimal

The following proposition shows that Q cannot be optimal at a coexistence equilibrium.

Proposition .2. Let D be such that E_c exists and is stable. Then, $Q(D)$ is not the maximum value of Q.

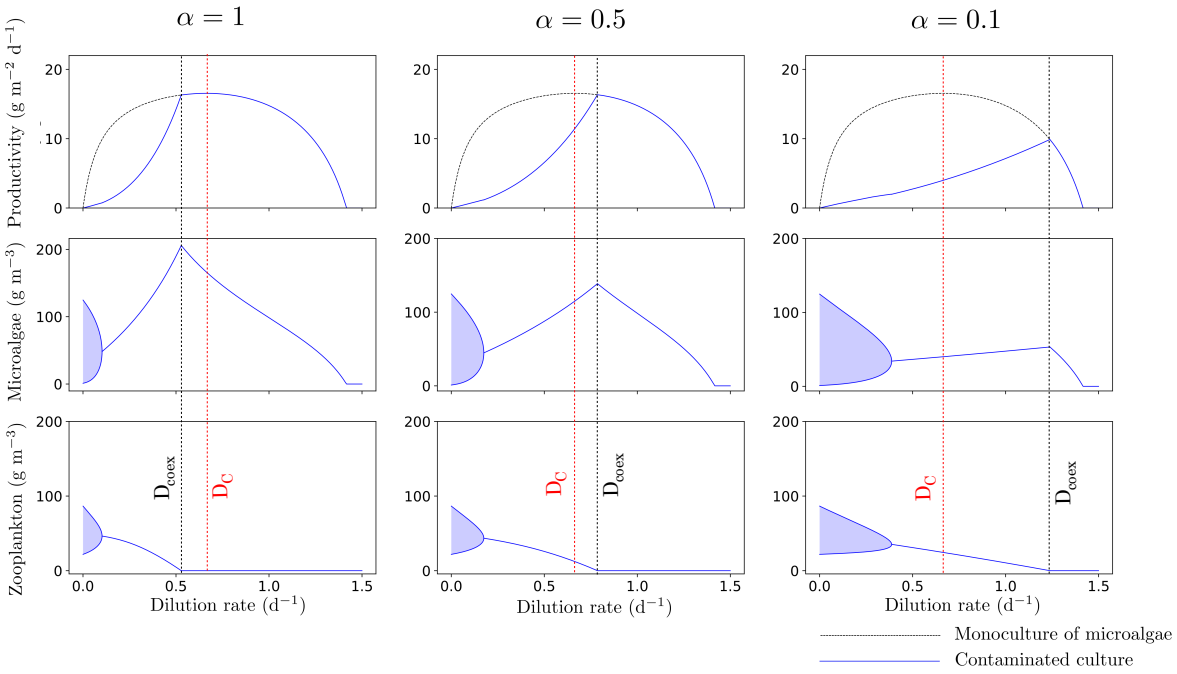

Figure 4.5: Bifurcation diagram and long-term productivity for different values of α.

Proof. Let D be such that $E_c = (x_c, y_c)$ exists and is stable and let us assume that Q reaches the maximum at D. Let x^* be such that $\mu(x^*) = D_{coex}$. From the definition of x_c and from Proposition .1, we have

$$\nu(x_c) = \alpha D + m < \alpha D_{coex} + m.$$

From (4.13), we conclude that

$$\nu(x_c) = \alpha D + m < \nu(x^*).$$

Since ν is strictly increasing, $x_c < x^*$, and consequently $Q(D) < Q(D_{coex})$, which contradicts the hypothesis that Q reaches the maximum at D. □

Proposition .2 states that Q is optimal either at a limit cycle or at an equilibrium without predators (first and third cases in (4.19)). When dealing with limit cycles, it is not clear how Q behaves. The following Proposition shows that for a constant dilution rate, the microalgal biomass along a limit cycle is higher than that of the unstable coexistence equilibrium.

Proposition .3. Let D be such that (4.1) admits an unstable coexistence equilibrium $E_c = (x_c, y_c)$. Let (x_p, y_p) be a limit cycle of (4.1) with period T. We have that

$$x_c < \frac{1}{T}\int_0^T x_p(t)dt. \tag{4.21}$$

Proof. From the second equation in (4.1), after dividing both sides by y, we have that

$$\nu(x_c) = m + \alpha D = \frac{1}{T}\int_0^T \nu(x_p(t))dt.$$

Since ν is strictly concave, applying Jensen inequality (Needham, 1993), we obtain

$$\nu(x_c) < \nu\left(\frac{1}{T}\int_0^T x_p(t)dt\right).$$

Finally, since ν is increasing, we obtain (4.21). □

In terms of productivity, Proposition .3 states that for an unstable equilibrium $E_c = (x_c, y_c)$, we have that

$$DLx_c < Q(D). \tag{4.22}$$

Expression (4.22) is an indicator of the difficulty of arguing that limit cycles cannot be optimal. While we know that coexistence equilibria are not optimal (see Proposition .2), we have no argument to say that the gain in biomass through a limit cycle cannot surpass the biomass production in the absence of predators. An answer to the question of whether limit cycles can be optimal or not can be numerically investigated.

4.4.4 Numerical evaluation of the productivity

We use the toolbox Matcont for MATLAB (Dhooge et al., 2003) to evaluate numerically the productivity Q as a function of the dilution rate. Figure 4.5 shows a bifurcation diagram of (4.1) with respect to the dilution rate and the evaluation of Q. We observe a unique value of D at which a Hopf bifurcation takes place, that is, when the coexistence equilibrium changes its stability and a limit cycle appears (Perko, 2001). We observe that a reduction of α favors the existence of limit cycles and increases the range of dilution rates admitting a coexistence equilibrium. Regarding the productivity, we observe that Q is strictly increasing on $[0, D_{coex}]$, despite the presence of limit cycles. When grazers are equally diluted as microalgae ($\alpha = 1$), D_{coex} is lower than D_C, and D_C is the trivial choice for the optimal dilution rate. This dilution rate not only ensures the washout of grazers, but ensures the highest biomass productivity. When $\alpha = 0.5$ or $\alpha = 0.1$, D_{coex} is higher than D_C. In this case, D_{coex} is the optimal dilution rate, despite the apparent microalgal biomass loss.

Based on our numerical simulations and on the fact that there is no paper citing any advantage of predators in microalgal cultivation, we propose the following conjecture on the solution of the SOCP.

Conjecture .1. Let D_{coex} be the dilution rate given by Proposition .1 and let D_C be the solution to (4.17). Then, the solution to (4.20) is given by

$$D_{SOCP} = \max\{D_{coex}, D_C\}. \tag{4.23}$$

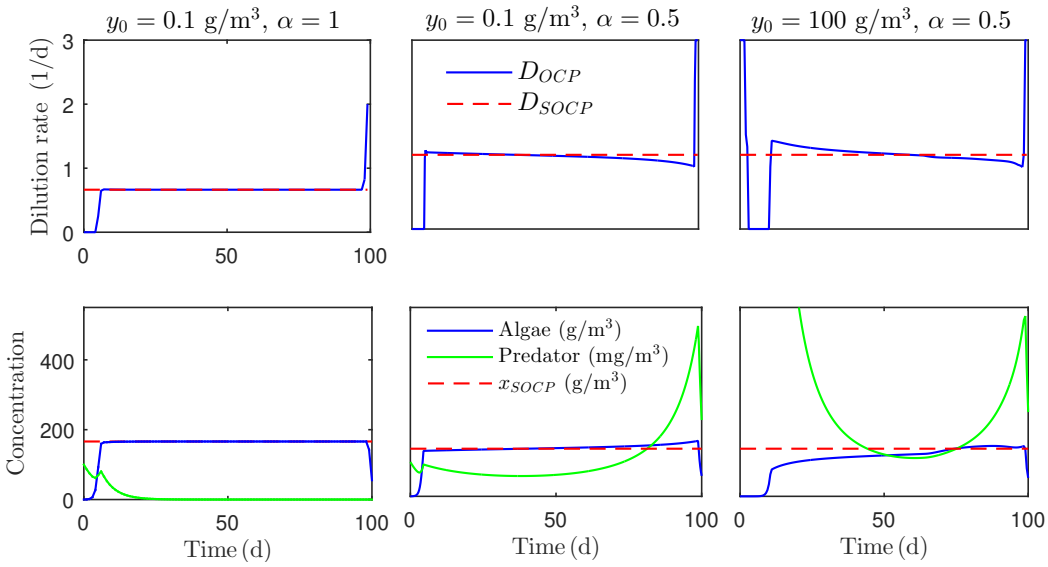

Figure 4.6: Optimal solution for different initial conditions and values of α. Note that predators and microalgae have different concentration units. Parameters are taken from Table 4.1 and $x_0 = 0.1$ g/m^3.

To understand Conjecture .1, let us imagine a chemostat with a pure culture of microalgae that is operated at optimal dilution rate D_C. We allow then the system to reach steady state. Now, let us imagine that a zooplankton population invades the culture. If the growth rate of zooplankton is negative, they will washout and the culture is optimally operated. This corresponds to the case $D_{coex} < D_C$, that is, the optimal dilution rate for the monoculture is too high to allow the development of zooplankton. On the other hand, if the invaders have a positive growth rate, they will develop and remain in the culture in the long-term. In this case, Conjecture .1 states that the dilution rate must be increased at the minimal value ensuring the washout of predators, that is, the dilution rate must be set at D_{coex}.

4.5 Optimal control problem (OCP)

4.5.1 Problem statement

In the previous section, we studied the optimal constant value of the dilution rate in the long-term operation. In this section, the dilution rate is allowed to vary in time, and we want to maximize the quantity of biomass that is harvested on a fixed interval of time $[t_0, t_f]$. We

consider the following optimal control problem (OCP):

$$\begin{aligned} \max_{D} \quad & J := \int_{t_0}^{t_f} LD(t)x(t)dt, \\ s.t. \quad & \frac{dx}{dt} = [\mu(x) - D]x - \frac{1}{\gamma}\nu(x)y, \\ & \frac{dy}{dt} = [\nu(x) - m - \alpha D]y, \\ & 0 \leq D(t) \leq D_{max},\ t \in [t_0, t_f], \end{aligned} \tag{4.24}$$

where D_{max} is the maximal dilution rate allowed. The best policy for $D(t)$ is known as optimal control.

Note that the microalgae productivity is given by J and it represents the quantity of biomass (in grams) that is harvested per meter squared in a given interval of time. This productivity is measured in g/m^2 and not g/m^2/d as the productivity Q defined in Section 4.4. If J is divided by $t_f - t_0$, then we obtain the daily average productivity, which is comparable to Q. However, in this section, we are focus on investigating the structure of the optimal control. For this purpose, it is equivalent to maximizing J or $J/(t_f - t_0)$.

Since D appears linearly in the objective function in (4.24) and on the system (4.1), the optimal control is "bang-bang" type, singular, or a combination of both. This follows from the theory of optimal control and the application of the Pontryagin Maximum Principle (Vinter and Vinter, 2010). When a singular arc takes place, the dilution rate takes intermediate values between 0 and D_{max}. When a bang-bang solution occurs, the optimal control oscillate between 0 and D_{max}.

4.5.2 Numerical solution

We solve numerically the OCP (4.24) with a direct method implemented in the sofware BOCOP (Bonnans et al., 2011) (version 2.21). The problem is discretized by a two-stage Gauss-Legendre method of order 4 with 100-500 time steps. We consider a constant initialization, and the tolerance for IPOPTNLP solver is set at 10^{-12}.

Figure 4.6 shows the optimal solution for different values of α and different initial abundances of grazers. When the initial concentration of grazers is low ($y_0 = 0.1$ g/m^3), the structure of the optimal control is bang-singular-bang. That is, at the beginning, the dilution rate is set to zero (bang), then the dilution rate takes intermediate values between 0 and D_{max} (singular arc), and finally the dilution rate is set to D_{max} (bang). We note that the singular arc is very close to the solution of the SOCP, especially when $\alpha = 1$. When $\alpha = 0.5$ and the initial concentration of predators is high ($y_0 = 100$ g/m^3), the optimal control is of the form bang-bang-singular-bang. In this case, an additional switch time is added and the dilution rate is set at its maximum value at the beginning.

4.5.3 Suboptimal feedback control

Microalgae cultures are generally initiated without zooplankton and with a low concentration of microalgae. In such situations, as suggested by numerical simulations, the solution of the

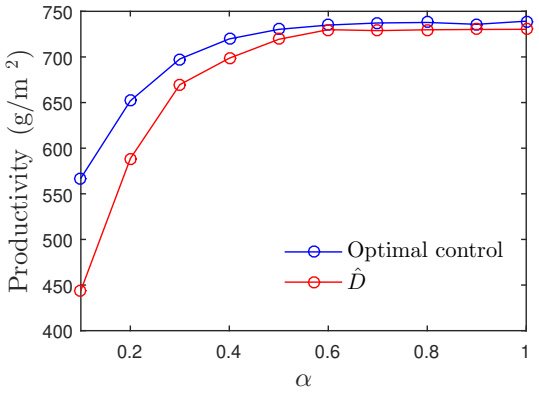

Figure 4.7: Comparison of the productivity obtained with the optimal control and the feedback control defined by (4.25). The initial conditions are $x_0 = 0.1$ g/m^3 and $y_0 = 0.1$ g/m^3 and the interval of time is $[0, 50]$.

OPC (4.8) sets the dilution rate to zero at the beginning (see Figure 4.6). Thus, the microalgae concentration will rapidly increase until reaching a value close to x_{SOCP}. Then, the microalgae concentration stays close to x_{SOCP} while the dilution rate follows a singular arc that is close to D_{SOCP}. Based on this, we propose the following feedback control:

$$\hat{D} = \begin{cases} D_{SOCP} & if \quad x \geq x_{SOCP}, \\ 0 & if \quad x < x_{SOCP}, \end{cases} \tag{4.25}$$

with D_{SOCP} defined by Conjecture .1 and x_{SOCP} defined by the following equation

$$\mu(x_{SOCP}) = D_{SOCP}. \tag{4.26}$$

The feedback control $\hat{D}$ depends on whether zooplankton can develop or not when the compensation condition (4.18) holds. If the zooplankton cannot survive, the feedback control leads the process to rapidly satisfy the compensation condition and keeps the system in that state. However, if zooplankton can survive, the strategy consists of forcing microalgae to rapidly reach the lowest equilibrium concentration (x^*) at which they do not support the development of zooplankton ($\nu(x^*) = m + \alpha D$). Then, the system is kept at equilibrium until the end.

This control is not appropriate for initial conditions with a remarkable dominance of grazers in the culture, the application of $\hat{D}$ can result in an oscillatory behavior in which zooplankton will collapse only after reaching its maximum concentration. As shown in Figure 4.6, an initial phase with high dilution rate is more suitable for cultures that are already highly contaminated.

Figure 4.7 shows the productivity associated with $\hat{D}$ and with the optimal control D_{OCP}. We observe that as α increases, both controls give a similar productivity. However, when α is low, the optimal control clearly outperforms the feedback control $\hat{D}$. When $\alpha = 0.1$, the productivity associated with D_{OCP} is a 27 % higher than the productivity associated with $\hat{D}$. When α is higher than or equal to 0.4, there is an increase of only about $0.5 - 1.5$ %. This is

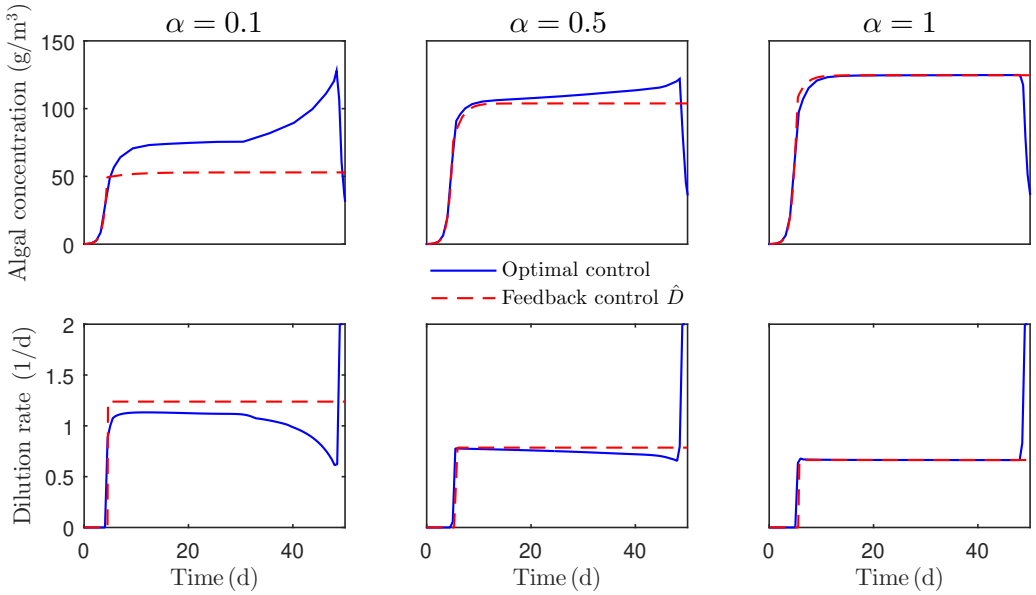

Figure 4.8: Comparison of the optimal control and $\hat{D}$ on the time interval $[0, 50]$ for different values of α. The initial conditions are $x_0 = 0.1$ g/m^3 and $y_0 = 0.1$ g/m^3.

because low values of α are associated with high values of D_{SOCP} [1]. Thus, when the feedback control sets the dilution rate at D_{SOCP}, there is a loss of microalgae due to dilution that cannot be compensated by the eradication of grazers. For this reason, the optimal control D_{OCP} sets a dilution rate lower than $\hat{D}$ (see Figure 4.8). This disagreement between both controls is related to time horizon. The construction of $\hat{D}$ is primarily based on the solution of the steady state optimization problem, that is, when the system is operated in large time. As shown in Figure 4.9, when the time horizon is increased, the optimal control becomes closer to the control $\hat{D}$. This reveals a Turnpike-like property of the optimal control problem (Trélat and Zuazua, 2015): when the optimal control is settled in large time intervals, most of the time the optimal control stays close to the solution of the steady state problem. This property becomes more evident when α approaches 1 as shown in Figure 4.6.

4.6 Discussion

4.6.1 Description of the feedback control $\hat{D}$

We have proposed a feedback control of the dilution rate to increase biomass productivity in microalgal cultivation systems that are susceptible to predation. This strategy is sub-optimal but very close to the optimal control when predators are efficiently diluted (Figure 4.7, $\alpha \approx 1$) and their initial concentration is low. The efficiency of the feedback control reflects a Turnpike-like

[1] Indeed, for low values of α, predators are more likely to survive, that is, D_{coex} is higher (see Figure 4.3). In view of the definition of D_{SOCP} (see (4.23)), it is clear that D_{SOCP} increases with α.

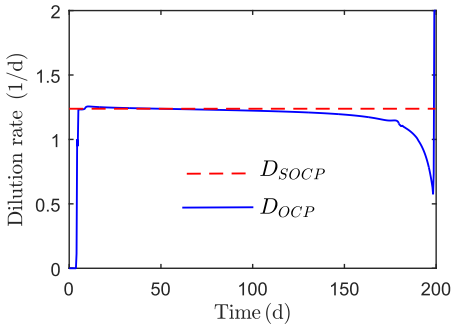

Figure 4.9: Optimal control when $\alpha = 0.1$ on the time interval $[0, 200]$.

property of the optimal control problem (Trélat and Zuazua, 2015): when the optimal control is settled in a large time interval, most of the time the optimal control stays close to the solution of the steady state problem (see Figure 4.9).

The feedback control $\hat{D}$ is not suitable for highly contaminated systems because setting the dilution rate to zero is not always a good way to rapidly increase microalgae concentration. At fixed low dilution rates, solutions of (4.1) oscillate approaching a limit cycle (e.g. see Figure 4.4A). If the initial concentration of grazers is high, the trajectories move counter clock-wise in such a way that the microalgae concentration decreases. Indeed, as shown in Figure 4.6, the optimal control (D_{OCP}) sets the dilution rate to its maximum value to reduce the abundance of grazers and then sets the dilution rate to zero to increase microalgae density to its optimum steady state value. In future work, the minimum-time control problem of reaching the optimal microalgae density of the SOCP should be addressed. This would allow us to propose a feedback control suitable for any state of the cultivation system.

4.6.2 Typical start-up of continuous cultures

In real situations, microalgae continuous cultures generally consist of two phases: an initial batch phase (dilution rate equal to zero) in which microalgae rapidly grow reaching a sufficient cell density, and a second phase in which the reactor is operated at a constant dilution rate (Bougaran et al., 2003, Fernandes et al., 2015). Such control of the dilution rate, henceforth referred to as initial-batch (IB) control (we follow the definition given in Bougaran et al. (2003)), is characterized by two parameters, the microalgae concentration at the end of the batch phase (denoted by x_{batch}) and by a constant dilution rate during the chemostat phase (denoted by D_{fixed}).

The feedback control $\hat{D}$ can be seen as an IB control in which x_{batch} is equal to x_{SOCP} (defined by (4.26)) and D_{fixed} is equal to D_{SOCP} (defined by (4.23)). A natural question is whether $\hat{D}$ is an efficient control among all the possible IB controls. Figure 4.10A shows the productivity that is obtained using the IB control with different values of x_{batch} and D_{fixed}. We observe the existence of an optimal combination of x_{batch} and D_{fixed} that maximizes the productivity. The maximal productivity is 727 g/m^2, and it is reached when $x_{batch} = 109$

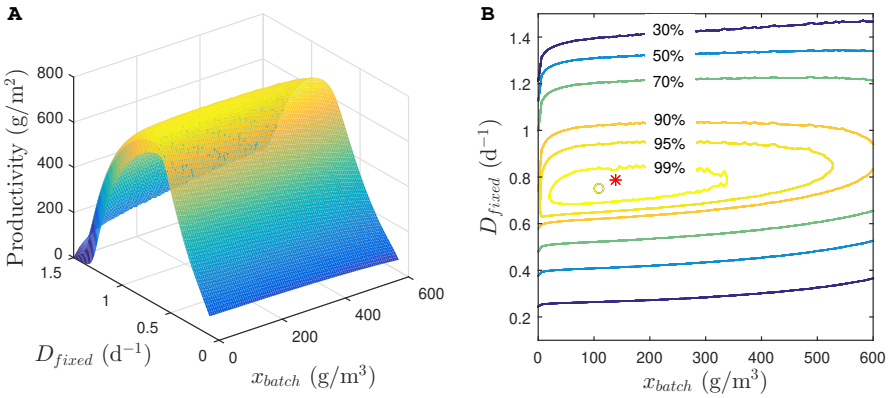

Figure 4.10: Productivity obtained using the IB control (defined in subsection 4.6.2) with different values for x_{batch} and D_{fixed}, with $\alpha = 0.5$. **A**. Productivity surface plot. **B**. Level curves of the productivity. The circle (o) represents the optimum choice of x_{batch} and D_{fixed} for the IB control, which provides productivity of 727 g/m^2. Each level curve represents a percentage of the maximum productivity obtained using the IB control. The star ($\star$) corresponds to the feedback control $\hat{D}$, that is, $x_{batch} = x_{SOCP}$ and D_{SOCP}.

g/m^3 and $D_{fixed} = 0.75$ d^{-1}. Figure 4.10B shows the level curves for productivity obtained using the IB control and the productivity associated with $\hat{D}(x)$. We observe that the feedback control is close to the optimal IB control ($x_{SOCP} = 139$ g/m^3 and $D_{SOCP} = 0.79$ d^{-1}). The feedback control provides productivity of 724 g/m^2, that is, a 0.46 % lower than the maximum productivity with IB control. This shows that the feedback control $\hat{D}$ is a good approximation of the optimal IB control. Simultaneously, this indicates a good agreement between the optimal IB control and the optimal control D_{OCP}.

4.6.3 Real implementation of the feedback control

The implementation of the feedback control $\hat{D}$ requires the estimation of three parameters D_{coex}, D_C, and x_{SOCP}. The value of D_C can be estimated from the specific growth rate curve of the microalgae species, which can be also estimated experimentally (Tang et al., 2012). Estimating the value of D_{coex} is probably the main difficulty. This parameter depends on the growth capacity of the possible predators, that is, the growth rate of predators, the capacity of predators to escape the outflow, the mortality rate of zooplankton, and factors that were not taken into account in this study, such as temperature. Therefore, dedicated experiments may be needed to estimate D_{coex}, or at least an upper bound.

Once the parameters defining $\hat{D}$ are known, the application of this feedback control is rather conventional: the microalgae culture is operated in batch mode until the system reaches the concentration x_{SOCP}, and then the culture is operated at the dilution rate D_{SOCP}. Different techniques exist for online estimation of the microalgae concentration (Havlik et al., 2013). This allows the estimation of the moment at which the microalgae concentration is equal to x_{SOCP}. It is important to note that even if this estimation is not precise, the final productivity will

not be severely affected. The productivity is rather robust with respect to x_{SOCP} (see Figure 4.10B).

One advantage of the feedback control $\hat{D}$ is that it does not depend on the population density of grazers. Early detection of zooplankton is not easy, and by the time zooplankton have been detected, it may be too late to optimally operate the system (Day et al., 2017).

4.6.4 Presence of predators in optimal regime

Under constant operation of the chemostat, the optimal dilution rate D_{SOCP} avoids the development of predators. When the dilution rate is allowed to vary in time and the system is run for a finite time, under optimal operation, the presence of grazers depends on their capacity to avoid outflow. If the retention time of zooplankton is notably higher than that for phytoplankton, then some grazers are allowed to stay in the system (see Figure 4.6 case $\alpha = 0.5$). This is because the benefits of eradicating predators in a short period do not compensate the losses of microalgae due to a high dilution. When grazers are more susceptible to be flushed out from the culture, some grazers may develop but their concentration will remain low for most of the time (see Figure 4.6 case $\alpha = 1$). Thus, while in the long-term operation predators must be completely avoided, in a finite horizon their eradication may not be optimal.

4.6.5 Integrated solutions

Microalgae productivity can be increased by implementing the feedback control $\hat{D}$ with other techniques to reduce zooplankton contamination. Since x_{SOCP} is determined from (4.26), the control $\hat{D}$ is completely determined by the value of D_{SOCP}. As shown in Figure 4.11A, the productivity associated with $\hat{D}$ decreases as the value of D_{SOCP} increases. The reduction of D_{SOCP} is possible by decreasing α or m (see Figure 4.11B). The reduction of α, can be technologically performed by, for example, placing the outflow closer to where grazers have higher concentrations, or just after the paddle wheel, where individuals are uniformly mixed (Montemezzani et al., 2016). Increasing m can be done using methods such as hydrodynamic cavitation (Kim et al., 2017a), increasing CO_2 concentration (Ma et al., 2017), or even biocontrol using predators for the grazers (Montemezzani et al., 2015).

4.7 Conclusion

We have shown the existence of an optimal dilution rate for the steady state operation of continuous microalgae cultures. This optimal dilution rate ensures the eradication of zooplankton, and therefore the productivity cannot be optimal through limit cycles. Another consequence of this is that if predators develop in a microalgae culture, the dilution rate must be increased until avoiding their development. This strategy may not be intuitive because increasing the dilution rate will also negatively affect the microalgae population. However, as shown in this work , the loss in microalgae through dilution compensates for grazing. It is also important to highlight that such a property is mathematically difficult to prove, and therefore, we used numerical methods to corroborate it.

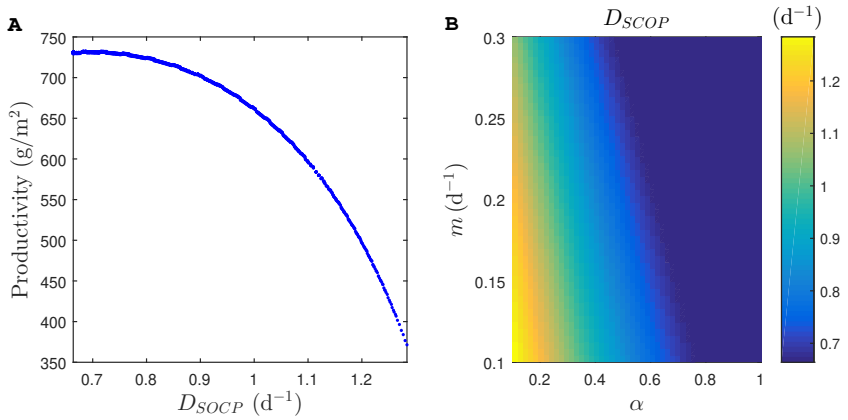

Figure 4.11: **A**. Productivity using $\hat{D}$ with different values of D_{SOCP} on the interval of time $[0, 50]$ d. **B**. Value of D_{SOCP} for different combinations of $a\alpha$ and m.

When the culture is operated during a finite interval of time, we have proposed a simple feedback control for the dilution rate. This feedback control is characterized by an initial batch phase followed by a steady-state operation at optimum dilution rate. This control follows the same structure as typical controls used in real systems, which makes its application possible. Even if this control is not optimal, it is very close to the optimal control when the system is operated for a large period. Finally, this feedback control has the advantage that it depends on the microalgae concentration but not on the grazers concentration. This is important because the early detection of zooplankton is not easy.

Predation during microalgae cultivation has often been neglected while it is one of the most challenging problems at the industrial scale. Its theoretical approach has so far rarely been targeted. It is important to account for a longer residence time of the predators, which makes the problem even more difficult to address. The strategies that we propose are likely to help a practitioner manage this issue and avoid the installation of the grazers in the reactor. They must, however, be associated with a direct treatment, such as biocontrol, to increase the predator mortality rate, favoring the proposed control strategy.

Chapter 5

Reduced metabolic model applied to wastewater treatment

This chapter has been adapted from the published article
Pessi, B.A., Baroukh, C., Bacquet, A., Bernard, O., 2023. A universal dynamical metabolic model representing mixotrophic growth of Chlorella sp. on wastes. Water Research 229, 119388.

Abstract

An emerging idea is to couple wastewater treatment and biofuel production using microalgae to achieve higher productivities and lower costs. This work 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.

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

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

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

During dark fermentation of organic wastes, anaerobic bacteria and archaea convert complex and non-assimilable compounds into Volatile Fatty Acids (VFAs) usable by microalgae. Indeed, VFAs can support heterotrophic growth of microalgae, while they use ammonium and phosphate in the wastewater as source for nitrogen and phosphorus (Baroukh et al., 2017, Turon et al., 2015b,a, Gao et al., 2022). In this perspective, *Chlorella sp.* was selected for its potential in associating biofuel production with effluent treatment (Casagli et al., 2021b, Gao et al., 2022,

Wágner et al., 2016). Indeed, this species can accumulate up to 50% of its dry weight in lipids, essentially in the form of triacylglycerol (TAG). In addition, its capacity to grow in the main dark fermentation effluents has been demonstrated, under heterotrophic or mixotrophic conditions (Turon et al., 2015a, Lacroux et al., 2022). The VFA mixture resulting from dark fermentation is typically made of a blend composed of about 30% acetate and 70% butyrate. Other organic acids can also be found (Rafrafi et al., 2013, Turon et al., 2015a), among which lactate, which is not consumed and does not affect the growth of *Chlorella* (Turon et al., 2015a). Turon et al. (2015a) first proposed a kinetic model of the consumption of butyrate and acetate by *Chlorella*. Later, an extended metabolic model was proposed using the DRUM framework (Baroukh et al., 2017). Here, we extend further this model by including the consumption of glucose and glycerol by *Chlorella* and propose a universal multi-substrate dynamic reduced metabolic model.

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

The model developed here represents, in detail, the growth under different autotrophic, heterotrophic and mixotrophic conditions and for four organic substrates. To our knowledge, this is the first model including such a large range of potential substrates in mixotophic regimes. The model was validated in different cultivation conditions using the abundant literature available on autotrophic, heterotrophic or mixotrophic growth of *Chlorella*. To this end, data from 122 experiments was extracted from 15 publications, amounting to more than 2600 concentration data points (see Table 5.1), thus reaching an unprecedented level of validation. This model is shown to support a strategy to enhance the bioconversion of VFA into microalgal biomass by managing the way the different substrates are supplied.

5.2 Materials and methods

5.2.1 General Principles of the DRUM approach

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

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

Table 5.1: 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)	*C. sorokiniana*	León-Vaz et al. (2019)
		C. sp.	Sen and Martin (2018)
		C. protothecoides	Chen and Walker (2011), O'Grady and Morgan (2011)
		C. vulgaris	Ma et al. (2016)
Glucose	22 (11)	*C. sorokiniana*	León-Vaz et al. (2019), Li et al. (2013, 2014b)
		C. protothecoides	Espinosa-Gonzalez et al. (2014), Chen and Walker (2011)
		C. pyrenoidosa	Ogbonna et al. (1997)
Glucose/Glycerol	2 (2)	*C. protothecoides*	O'Grady and Morgan (2011)
Acetate	40 (25)	*C. sorokiniana*	León-Vaz et al. (2019), Turon et al. (2015a,b), Chen et al. (2017a,b), Xie et al. (2020a)
		C. sacchrarophila	Xie et al. (2020b)
Butyrate	10 (4)	*C. sorokiniana*	Turon et al. (2015b,a)
Acetate/Butyrate	23 (6)	*C. sorokiniana*	Turon et al. (2015b,a)
Autotrophic	13	*C. sorokiniana*	Li et al. (2014b), León-Vaz et al. (2019), Turon et al. (2015b,a)
		C. sp.	Sen and Martin (2018)

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

$$\frac{dM}{dt} = \frac{d}{dt}\begin{pmatrix} S \\ C \\ P \\ B \end{pmatrix} = \begin{pmatrix} K_S \\ K_C \\ K_P \\ K_B \end{pmatrix} .v(M).B - DM + DM_{in} = K.v(M).B - DM + DM_{in} \tag{5.1}$$

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

In metabolic models, intracellular metabolites are generally assumed to be quasi-stationary ($\frac{dC}{dt} = K.v = 0$), i.e. they are assumed to be consumed as soon as they have been synthesised. However, in the case of microalgae, this hypothesis has proven to be false for certain of its metabolites (denoted A) during mixotrophic or autotrophic growth (Baroukh et al., 2014, 2017). The DRUM method (Baroukh et al., 2014), consists in dividing the metabolic network into n quasi-stationary subnetworks ($K = (K_{SN_1}, \ldots, K_{SN_n})$, $K_{SN_i}.v_{SN_i} = 0 \; for \; i \in 1, \ldots, n$). These are linked by the A metabolites that are, in contrast, non-stationary, can accumulate and be later consumed. This division into subnetworks is justified by the presence of metabolic pathways that correspond to metabolic functions, to reaction groups that are regulated simultaneously and to the presence of compartments within the cell. Cellular mechanisms are therefore employed for assessing the subnetwork. Hence, the A metabolites can either be found at the junction of several metabolic pathways, or they can be transported from one compartment to another, or they can be final products that accumulate in the cell. The system of ordinary differential equations (5.1) therefore becomes:

$$\frac{dM'}{dt} = \frac{d}{dt}\begin{pmatrix} S \\ A \\ P \\ B \end{pmatrix} = \begin{pmatrix} K_S \\ K_A \\ K_P \\ K_B \end{pmatrix} .\alpha.B = K'.\alpha.B - DM' + DM'_{in} \tag{5.2}$$

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

5.2.2 Construction of the model

The core of the metabolic network from Baroukh et al. (2017) has been used and modified in order to add the glucose and glycerol consumption pathways (Figure 5.1). This network contains the central autotrophic, mixotrophic and heterotrophic metabolic pathways including photosynthesis, glycolysis, the pentose phosphate pathway, the Krebs cycle, oxidative phosphorylation and the synthesis of chlorophyll, carbohydrates (e.g. starch), amino-acids and nucleotides. The synthesis pathways of macromolecules such as proteins, lipids, starch, DNA, RNA as well as the functional biomass are represented through macroscopic reactions.

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

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

5.2.3 Analysis of the sub-networks

Motivations

In this section, we present the macroscopic reactions which result from the reduction of each subnetwork by the computation of the EFMs. As recommended by Baroukh et al. (2014), the reaction kinetics must be mathematically represented using minimal hypotheses, and when possible applying a mass action law. A list of all sub-networks and the macroscopic reactions can be found at Table 5.2.

Glyoxysome subnetwork

The glyoxysome pathway consists of 26 reactions, from which 8 are exchange reactions. The glyoxysome is the peroxysome compartment where the glyoxylate cycle occurs. Here carbon compounds are converted to succinate, also allowing the production of glucose from lipids. In this compartment, fatty acids can be used as a source of energy and carbon for growth is produced when no photosynthesis takes place. Two EFMs have been achieved for this subnetwork (MR1 and MR2). In the glyoxysome, butyrate and acetate are converted into acetyl-CoA, which is in turn converted, via the glyoxylate cycle, into succinate. The succinate then enters the cytosol

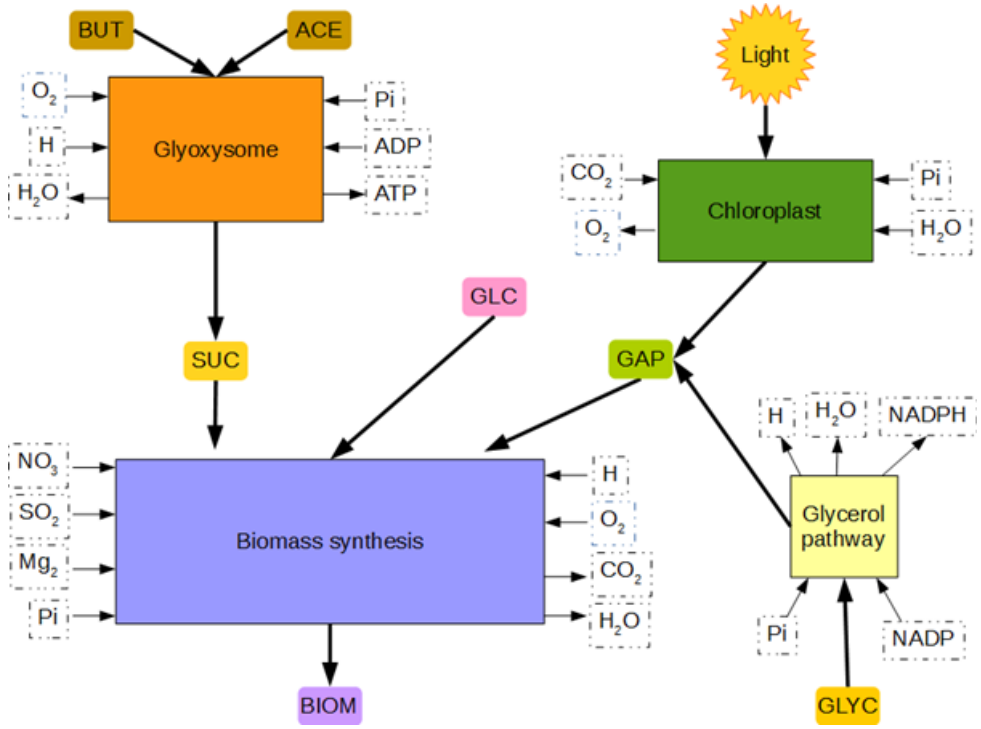

Figure 5.1: Considered metabolic subnetworks to represent growth of *Chlorella* on a mixture of glycerol, glucose, acetate and butyrate

	Subnetwork	Macroscopic reaction
MR_1	Glyoxysome	$2\ \mathrm{ACE} + 3.5\ \mathrm{H} + 0.5\ O_2 \rightarrow SUC + 0.5\ H_2O$
MR_2		$BUT + 7H + 1.5O_2 \rightarrow SUC + 5H_2O$
MR_3	Chloroplast	$Light + 3CO_2 + 2H_2O + Pi \rightarrow GAP + 3O_2$
MR_4	Glycerol pathway	$GLY + Pi \rightarrow GAP + H_2O$
$MR_5(NH_4)$	Biomass synthesis	$4.15\ GAP + 2.54O_2 + 0.99NH_4 + 0.02SO_4 + 0.01Mg_2 \rightarrow B + 0.99H + 2.90H_2O + 3.92CO_2 + 4.02Pi$
$MR_5(NO_3)$		$4.64GAP + 2.04O_2 + 0.99NO_3 + 0.98H + 0.02SO_4 + 0.01Mg_2 \rightarrow B + 5.39CO_2 + 2.90H_2O + 4.51Pi$
$MR_6(NH_4)$	Biomass synthesis	$4.15\ SUC + 7.30H + 4.61O_2 + 0.99NH_4 + 0.12Pi + 0.02SO_4 + 0.01Mg_2 \rightarrow B + 7.04H_2O + 8.06CO_2$
$MR_6(NO_3)$		$4.90SUC 5.28O_2 + 0.99NO_3 + 0.12Pi + 10.78H + 0.02SO_4 + 0.01Mg_2 \rightarrow B + 11.07CO_2 + 8.31H_2O$
$MR_7(NH_4)$	Biomass synthesis	$2.07\ GLC + 2.54O_2 + 0.99NH_4 + 0.12Pi + 0.02SO_4 + 0.01Mg_2 \rightarrow B + 3.91CO_2 + 7.04H_2O + 0.99H$
$MR_7(NO_3)$		$2.34GLC + 2.14O_2 + 0.99NO_3 + 0.12Pi + 0.98H + 0.02SO_4 + 0.01Mg_2 \rightarrow B + 5.49CO_2 + 7.63H_2O$

Table 5.2: List of the macroscopic reactions in each respective subnetwork. For biomass production, the stoichiometric values differ if the nitrogen source is nitrate or ammonium.

and is injected into the Krebs cycle, thus producing the different metabolites necessary for the synthesis of biomass.

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

$$\alpha_{MR_1} = \frac{k_{MR1}.ACE}{KS_{MR1} + ACE} \tag{5.3}$$

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

$$\alpha_{MR2} = \frac{k_{MR2}.BUT}{BUT + \frac{k_{MR2}}{\beta_{MR2}}.(\frac{BUT}{Sopt_{MR2}} - 1)^2} \frac{k_d}{(ACE + k_d)} \tag{5.4}$$

Chloroplast subnetwork

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

For autotrophic growth, light drives the reaction rate. When algae are growing on a turbid medium like wastewater, the average light intensity stays low and the local photoinhibition impact can be neglected (Martínez et al., 2018). Also under mixotrophic conditions with elevated concentration of carbon substrates, large biomass densities can be reached. Dense microalgal cultures strongly attenuate light. In line with Baroukh et al. (2017), photosynthesis rate is assumed to be linearly depending upon the average light intensity I_μ in the culture (see Equation 5.6).

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

$$I\,(z) = I_0.e^{-\alpha.z.B} \tag{5.5}$$

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

$$I_\mu = \frac{I_0}{L} \int_0^L e^{-\alpha.B.z} dz = \frac{I_0(1 - e^{-\beta_{MR3}.B})}{\beta_{MR3}.B} \tag{5.6}$$

The kinetics in the chloroplast subnetwork is finally given by:

$$\alpha_{MR3} = \; k_{MR3}.I_\mu \tag{5.7}$$

Glycerol absorption subnetwork

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

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

$$\alpha_{MR4} = \frac{k_{MR4}.GLY}{GLY + \frac{k_{MR4}}{\beta_{MR4}}.(\frac{GLY}{Sopt_{MR4}} - 1)^2} \tag{5.8}$$

Functional biomass synthesis subnetwork

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

This subnetwork generated 86,167 EFMs, including 3310 that did not produce biomass. Nearly all of the calculated EFMs are part of the biomass synthesis network. They can be sorted by using a similar method to the FBA (Flux Balance Analysis). The standard hypothesis supporting FBA is that evolution has selected metabolisms maximising biomass growth on each substrate (Orth et al., 2010b), or equivalently, minimising the loss of carbon as CO_2. Therefore, for each substrate, the EFM presenting the highest GAP/BIOM, SUC/BIOM and GLC/BIOM yields were selected. In this way, the use of GAP, SUC or GLC for the synthesis of biomass resulted from three macroscopic reactions (MR5, MR6 and MR7). The yield of biomass on the carbon substrate depends on the nitrogen source (NH_4 or NO_3). Table 5.1 shows the resulting macroscopic reactions for both cases.

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

$$\alpha_{ME5} = k_{MR5}.GAP \tag{5.9}$$

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

$$\alpha_{MR6} = k_{MR6}.SUC \tag{5.10}$$

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

$$\alpha_{MR7} = \frac{k_{MR7}.GLC}{GLC + \frac{k_{MR7}}{\beta_{MR7}}.(\frac{GLC}{Sopt_{MR7}} - 1)^2} \tag{5.11}$$

5.2.4 Global dynamics of the network

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

$$\frac{dM^{"}}{dt} = \frac{d\begin{pmatrix} S \\ A \\ B \end{pmatrix}}{dt} = K^{"}.\alpha.B - DM^{"} + DM^{"}_{in} \tag{5.12}$$

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

Moreover, the total biomass comprising the functional biomass and the metabolites A can be described as follows:

$$X_z(t) = \sum_A Z_A.A(t) + Z_B.B(t) \tag{5.13}$$

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

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

$$v = \begin{pmatrix} v_{SN_1} \\ \ldots \\ v_{SN_k} \end{pmatrix} = \begin{pmatrix} E_{SN_1}.\alpha_{SN_1} \\ \ldots \\ E_{SN_k}.\alpha_{SN_k} \end{pmatrix} \tag{5.14}$$

5.2.5 Sensitivity analysis

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

$$\sigma_y^{\Delta p} = \frac{1}{t_f}\int_0^{t_f} \frac{y(p + \Delta p, x_0, u, \tau) - y(p, x_0, u, \tau)}{y(p, x_0, u, \tau)} d\tau \tag{5.15}$$

where y is the simulated output at time τ with parameter set p, initial condition x_0 and input variables u (e.g. light intensity and dilution rate). We calculate the global sensitivity using Morris's sampling method implemented in the SALib Python toolbox (Morris, 1991, Herman and Usher, 2017), replacing the standard elementary effect by the sensitivity coefficient defined

above. We analyse the region between ±20% of the calibrated values of the parameters. The analysis is conducted separately for each of the four carbon substrates in mixotrophic conditions at a light intensity of 500 $\mu mol/(m^2.s)$. $\sigma_y^{\Delta p}$ is calculated as the average of the simulations considering a cultivation of 6 hours and three different initial concentration of the considered carbon substrate (0.1, 1 and 10 g/L). Table 5.3 shows the result of the sensibility analysis for all the parameters.

5.2.6 Reduced model calibration

In order to calibrate, and then validate the model, a large set of experiments from the literature have been used. In total 122 selected experiments (see Table 5.1) gather data on growth i) under autotrophic conditions, without any organic carbon input and submitted to light intensities ranging from 30 to 540 $\mu E.m^{-2}.s^{-1}$ ii) under heterotrophic conditions, without any light, and with varying concentrations in acetate, butyrate, glucose and glycerol, pure or combined iii) under mixotrophic conditions, with light and varying concentrations in acetate, butyrate, glucose and glycerol. Depending on the studies, different combinations of these substrates were tested.

Only the parameters for glucose and glycerol consumption were calibrated. All the other kinetic parameters for the macro reactions are taken from Baroukh et al. (2017). The calibration was done following a two-step process. First a stochastic global optimizer, Differential Evolution algorithm (Storn and Price, 1997), calculates the set of parameters minimizing the relative error between model and experimental data of biomass and substrate concentration over time. This parameter set is, then, used as initial point in a Markov Chain Monte Carlo sampler, which returns the parameters set inside a confidence interval (Foreman-Mackey et al., 2013). Glucose kinetics parameters were calibrated using concentration data from 5 experiments of Li et al. (2013), while glycerol parameters were fitted using data from 6 experiments of Ma et al. (2016). Table 5.3 shows the value of calibrated parameters and Figure 5.2 shows the simulation of the calibrated model together with the experimental data used for calibration.

5.2.7 Optimization of butyrate consumption

We consider the optimal control problem (Harmand et al., 2019), whose objective is to minimize the time t_f, where the chemical oxygen demand 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).

$$\min_{S_0} t_f : S(t_f) \leq \bar{S} \tag{5.16}$$

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

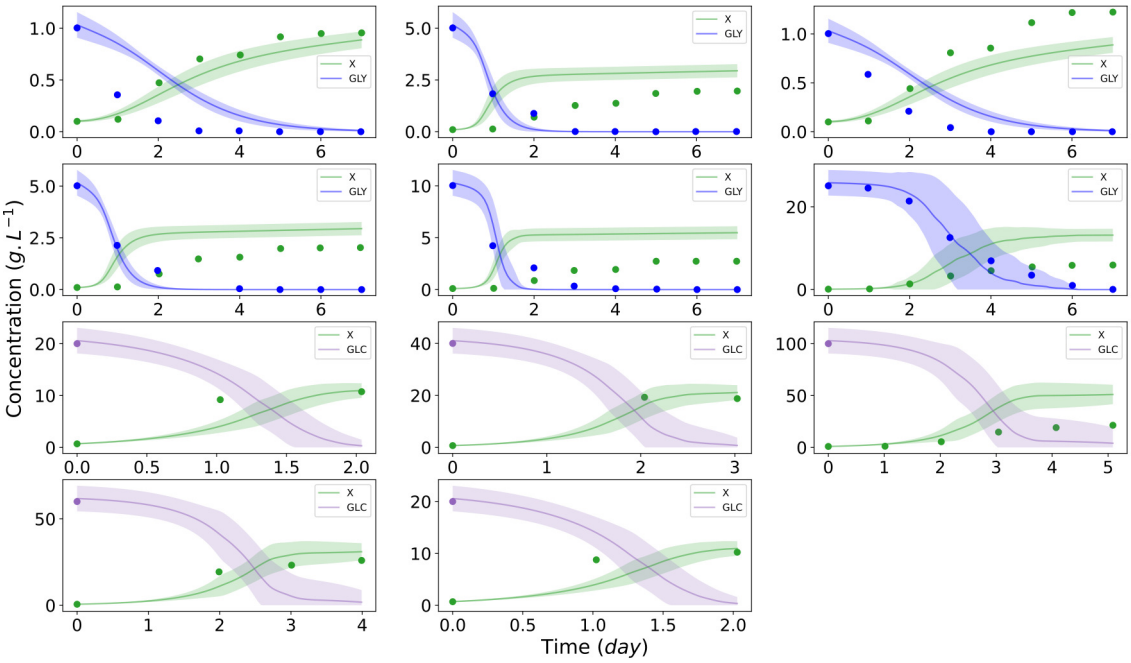

Figure 5.2: Comparison of the model against experimental data (Li et al., 2013, Ma et al., 2016) used for the calibration of glucose and glycerol kinetic parameters. Lines represent the average of 100 simulations and the colored region represents ±1 standard deviation of estimated concentrations using Markov Chain Monte Carlo method.

Table 5.3: 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 10^{-1}	$M.h^{-1}.MB^{-1}$	4.5910^{-03}	2.5210^{-03}	B
KS_{MR1}	5.50 10^{-5}	M	9.2010^{-05}	5.2410^{-03}	B
k_{MR2}	3.60 10^{-2}	$Mh^{-1}.MB^{-1}$	9.9610^{-05}	1.4010^{-03}	B
$Sopt_{MR2}$	1.90 10^{-5}	M	2.4810^{-04}	1.4110^{-03}	B
β_{MR2}	2.58 10^{5}	$h^{-1}.MB^{-1}$	1.8910^{-04}	1.4010^{-03}	B
k_{MR3}	1.90 10^{-3}	$Mh^{-1}.MB^{-1}.(m^2.s.\mu mol^{-1})$	9.2110^{-04}	5.7010^{-04}	B
β_{MR3}	2.48 10^{3}	MB^{-1}	-5.3810^{-04}	8.8210^{-05}	B
k_{MR4}	1.01	$M.h^{-1}.MB^{-1}$	5.5410^{-04}	1.8610^{-03}	A
$Sopt_{MR4}$	7.00 10^{-2}	M	9.2710^{-04}	2.2510^{-03}	A
β_{MR4}	9.05	$h^{-1}.MB^{-1}$	1.1010^{-04}	2.4310^{-03}	A
k_{MR5}	2.82 10^{1}	$h^{-1}.MB^{-1}$	9.7310^{-04}	1.1610^{-03}	B
k_D	5.39 10^{-10}	M	-1.1010^{-04}	1.4310^{-03}	B
k_{MR6}	2.37 10^{5}	$h^{-1}.MB^{-1}$	3.0710^{-06}	3.3310^{-03}	B
k_{MR7}	6.81 10^{-2}	$M.h^{-1}.MB^{-1}$	2.3910^{-03}	2.1610^{-03}	A
$Sopt_{MR7}$	5.07 10^{-2}	M	9.5410^{-05}	3.2210^{-03}	A
β_{MR7}	2.95 10^{+1}	$h^{-1}.MB^{-1}$	4.2410^{-04}	3.2410^{-03}	A

5.3 Results and discussion

5.3.1 Validation of the model

The experimental data not used during the calibration stage were used to validate the model. A coherent set of experiments, representing various experimental conditions, was kept for this validation stage (see Table 5.1). We used a Markov Chain Monte Carlo method to select parameters (Foreman-Mackey et al., 2013). We considered a ±20% uncertainty in the initial concentration of substrates. In Figure 5.3, the results of the model simulations are compared with the experimental data with more than one substrate.

As illustrated in Figure 5.2 for single substrates, the model efficiently predicts the production of biomass and the consumption of substrates, whether in autotrophic, heterotrophic or in mixotrophic conditions. Furthermore, the model is still accurate when there are two substrates (Figure 5.3). More generally, the predictive performance of the model is summarized in the Taylor diagram in Figure 5.4 for the whole data set. This diagram represents at the same time the standard deviation of the biomass prediction error and the Pearson correlation coefficient (Taylor, 2001). It illustrates both the centered and reduced quadratic errors between the experimental data sets and the associated simulations, as well as the correlation between the model and the data. It thus summarizes the degree of resemblance between the data and the simulations for the vast range of considered data. Indeed, the closer a data point is to (0;1), the better the model reproduces the experimental data (Taylor, 2001). Figure 5.5 represents all the data points versus model prediction, also demonstrating the goodness of fit of the model.

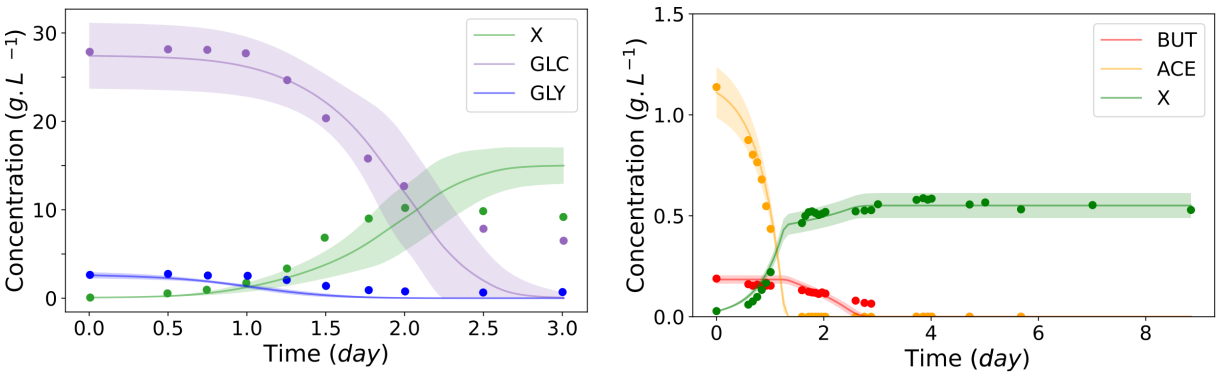

Figure 5.3: Model simulation and experimental data with two substrates. From left to right: glycerol and glucose; butyrate and acetate. The line represents the average of 100 simulations and the colored region represents ±1 standard deviation of estimated concentrations using Markov Chain Monte Carlo method.

The results of the sensibility analysis is also show in Table 5.3. Most parameters have the same order of sensibility (10^{-4}) showing that all have an importance in the model. The high standard deviation of the sensibility coefficient demonstrates the intertwined influence of parameters, for example maximum uptake rate and the optimal concentration of the substrate, also and the dependence of the actual substrate and biomass concentration in the dynamics of the system. This demonstrates the necessity of calibrating the model in a wide range of conditions.

5.3.2 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 5.6 and 5.7). Using only

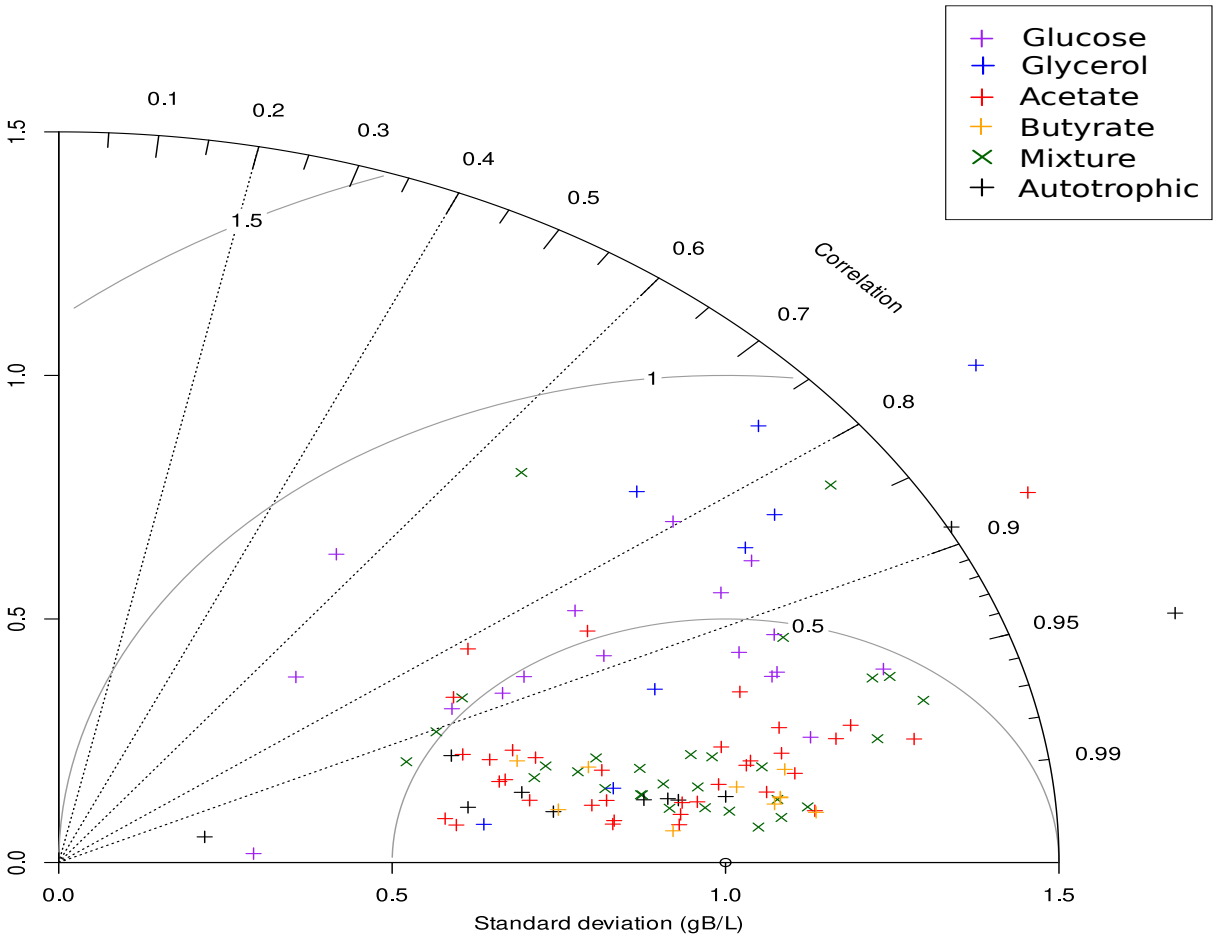

Figure 5.4: Taylor Diagram where each point represents the Pearson correlation coefficient and a normalized standard deviation of one experiment and model simulation. The semi circles centered at standard deviation 1.0 show the root-mean-square error.

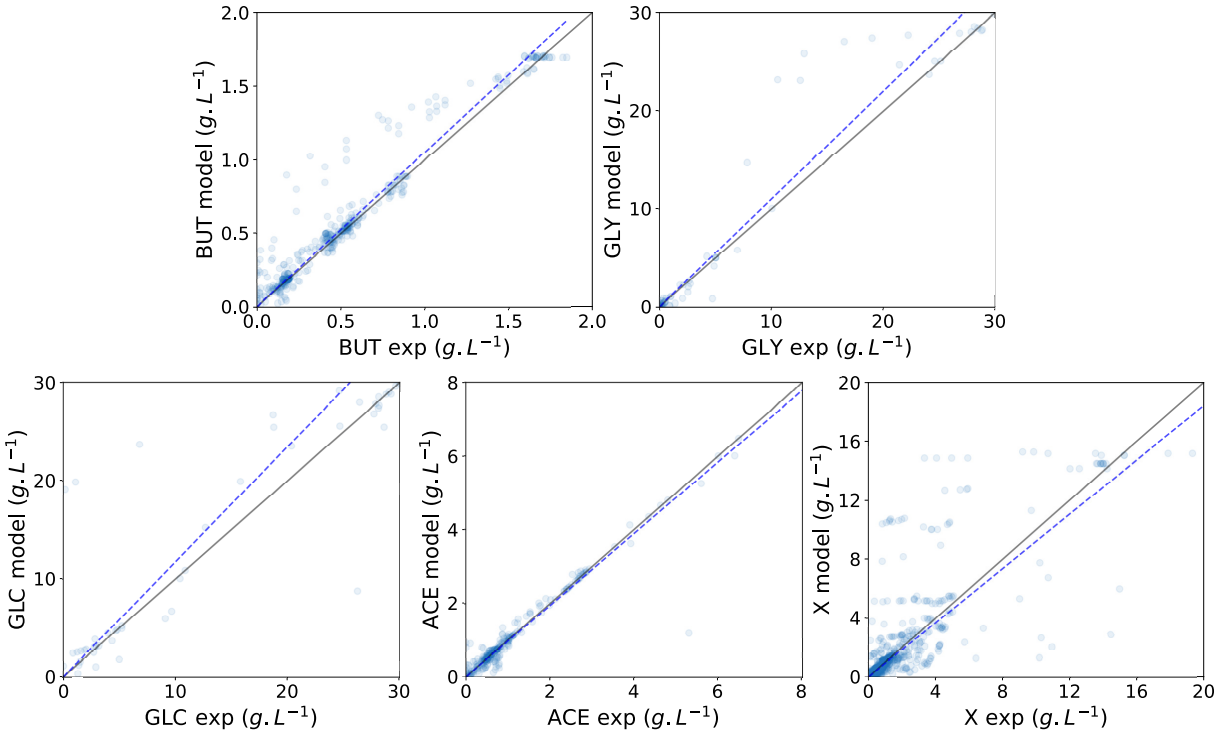

Figure 5.5: Validation of model predictions based on experimental data of butyrate, glycerol, glucose, acetate and biomass. All p-values for the regression are below 10^{-3}. R^2 for the lines are, respectively, 0.97, 0.71, 0.95, 0.97 and 0.64. Darker colors represent a concentration of data points.

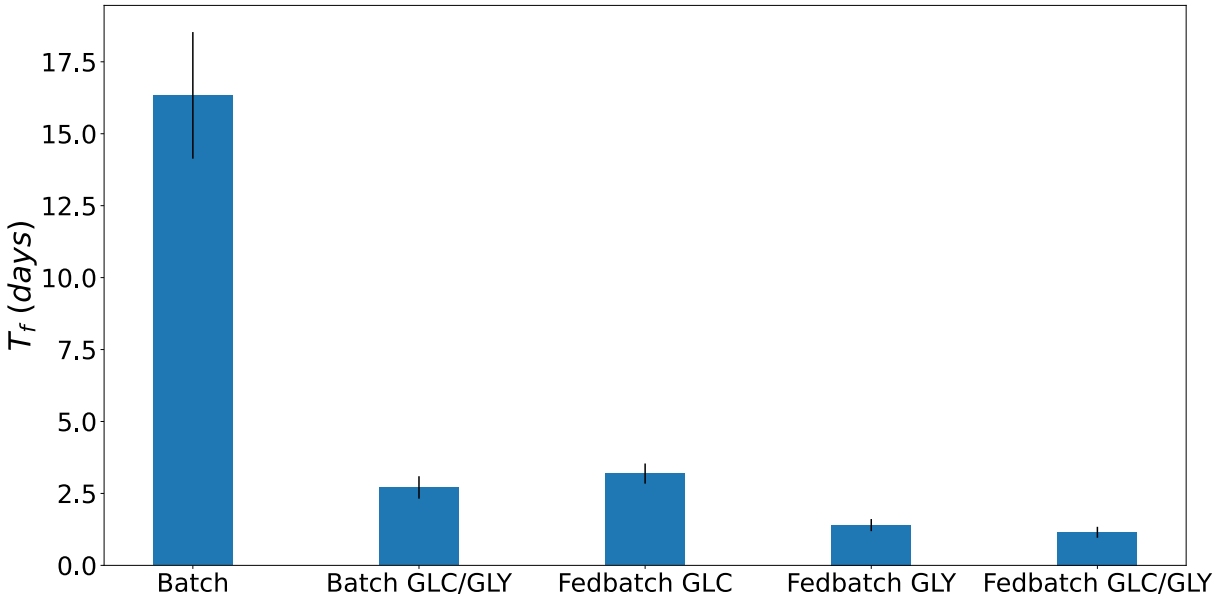

Figure 5.6: Time to reach the regulation threshold of a typical dark fermentation effluent using different conditions of cultivation - Batch (typical waste effluent), Batch GLC/GLY (typical waste with addition of an optimal concentration of glucose and glycerol), Fed-batch GLC, GLY, GLC/GLY (typical waste feeding, respectively, glucose, glycerol and a mixture of both.)

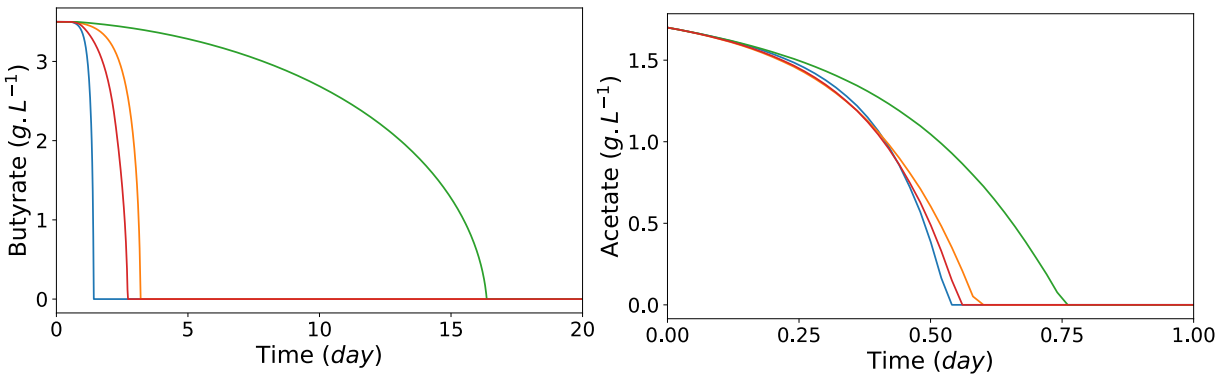

Figure 5.7: Concentrations of butyrate and acetate over time in different cultivation conditions. Optimal fed-batch with glycerol (blue), optimal fed-batch with glucose (red), batch with optimal addition of glycerol and glucose (orange), batch with only butyrate and acetate as substrates (green).

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

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

$$\min_D t_f : S(t_f) \leq \bar{S} \tag{5.17}$$

where D is the dilution rate of the inflow containing only the additional substrates, at a high concentration, so that the volume of the reactor does not change. t_f is the time to reach the regulation threshold. The optimal strategy can be approximated into a sub-optimal strategy, which would maximize the reaction rates for glucose and glycerol consumption. The strategy thus consists in computing the dilution rate such that glucose and glycerol concentrations stay constant close to the values for which the consumption of glucose and glycerol is maximum. The control problem is then reduced to finding the optimal final time of the inlet flux in the cases of a glucose, glycerol, and mixture inlet. In the case of the mixture, a fraction of 0.21 of glucose and 0.79 of glycerol is obtained, keeping glycerol at an optimal concentration. Using this control strategy, the final times for fed-batch cultivation are: 3.2 days for glucose, 1.4 days for glycerol and 1.2 days for a mixture of glucose and glycerol (see Figure 5.6), to be compared to the 16 days whithout inorganic carbon addition.

5.3.3 Analysis of metabolic maps

One of the advantages of metabolic model is that the main metabolic fluxes can be estimated. They are represented in Figure 5.8 considering the exponential growth phase. During autotrophic growth, a strong activity takes place in the chloroplast subnetwork where photosynthesis occurs. The greatest fluxes are associated to the fixation of CO_2 by RUBISCO, the conversion to 3PG (3-phosphoglycerate) and the conversion of 3PG to glycerate-3-phosphate (GAP). GAP is then mainly transported toward the cytoplasm, where it is injected into the glycolysis. From then on, it enables the synthesis of precursor metabolites that are necessary for the synthesis of functional biomass composed of proteins, DNA, RNA, chlorophyll, carbohydrates, and lipids. During pure heterotrophic growth, in the dark, no reaction occurs in the chloroplast subnetwork. Considering a mix of acetate or butyrate, the largest fluxes are concentrated in the glyoxysome subnetwork where carbon substrates are converted into succinate. The synthesized succinate is exported from the glyoxysome and injected into the Krebs cycle. The latter then enables the synthesis of precursor metabolites for the production of functional biomass. To ascend the glycolysis, the anaplerotic pathways are in an upward direction.

No reaction takes place in the glyoxysome, nor in the glycerol utilization pathway for growth on glucose only. Glucose carried from the medium into the cytoplasm is directly injected into the upper glycolysis for the production of precursor metabolites necessary for the synthesis of functional biomass. Glycolysis is therefore entirely in a downward direction, as the anaplerotic pathways that enable the synthesis of oxaloacetate (OXA) fueling the Krebs cycle.

For growth on glycerol only, the greatest fluxes are located in the glycerol utilization pathway subnetwork, where the uptaken glycerol is converted to glycerate-3-phosphate. The glycerate-3-phosphate is then injected into the middle of glycolysis, this time in an upward direction for

the high glycolysis and in a downward direction for the lower glycolysis. The remnants of the fluxes are similar to those observed for glucose.

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

Figure 5.8 shows the metabolic fluxes after 30 hours of batch culture with optimal initial conditions (110 g/l of glucose and 20 g/L of glycerol). At this point, the acetate in the culture is practically depleted, but the concentration of butyrate remains similar to the initial condition. The rate of consumption of butyrate then remains stable for 2 days, and the rate rapidly increases until the concentration threshold is reached. Since the optimal concentration is about 40 times lower than the regulation threshold, the butyrate consumption rate does not reach the maximum value. Glycerol has a much lower concentration, while the flux of glucose concentration is still high. As a result of the higher concentration of biomass and the lower availability of light per cell, the metabolic fluxes in the chloroplast are reduced.

5.3.4 Model limitations and perspectives

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

The effect of mixotrophic growth, here considered as the sum of autotrophic and heterotrophic conditions, can be more subtle in some cases. There is no consensus in the literature for a general model that fits all cases. For example, according to Martínez and Orús (1991), mixotrophic growth is greater than the sum of autotrophic and heterotrophic conditions when the concentration of CO_2 is limiting, since CO_2 produced by respiration can be recycled for the photosynthesis pathway. Recently, circular use of CO_2 and O_2 and preferential consumption of inorganic carbon has also been evidenced for *Chlorella vulgaris*, nonetheless heterotrophic growth was promoted at high 0_2 concentrations (Manhaeghe et al., 2020). Furthermore, under certain conditions an inverse relationship between light intensity and glucose consumption has been recorded (Patel et al., 2019, Wan et al., 2011).

Temperature is one of the most important factors affecting microalgae growth, even if only a minority of the models account for it (Shoener et al., 2019). The experimental data considered here were carried out at various temperatures and accounting for it would increase model accuracy. In the case of metabolic models, including temperature is challenging, and will involve a large number of parameters to characterize the influence of temperature on each individual reaction. The influence of short variations of temperature could be well estimated by the Arrhenius equation, but large variations of temperature typically occurring in outdoor cultivation require

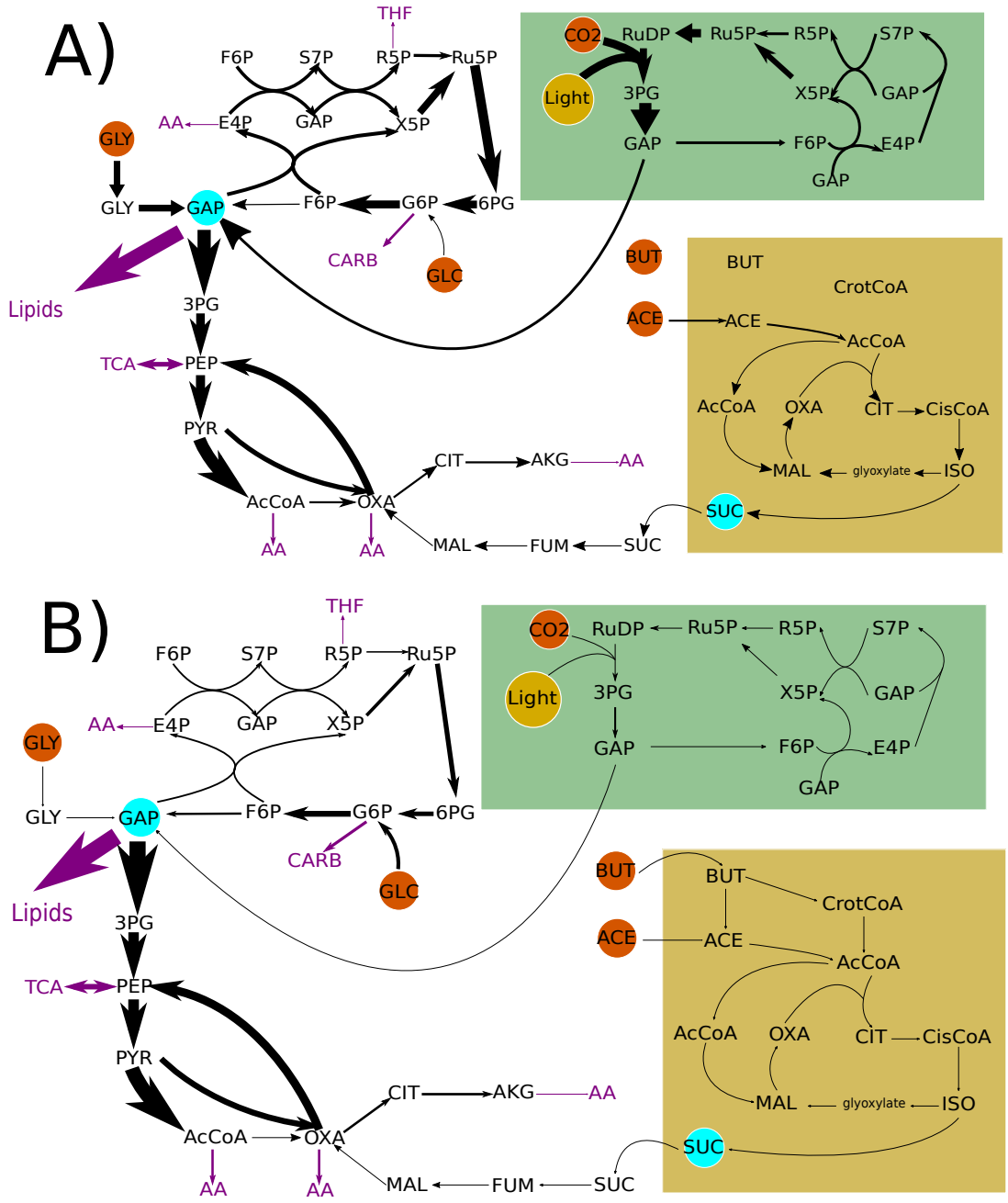

Figure 5.8: Metabolic charts showing the fluxes of an optimized batch. Above, A, at initial concentration of acetate and butyrate with optimal conditions of glucose and glycerol. Below, B, 30 hours after the beginning of the batch. The widths are linearly proportional to the calculated flux.

more advanced models (Casagli and Bernard, 2022, Assis Pessi et al., 2022). For these reasons, calibrating the temperature effect would require a large amount of dedicated experiments.

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

Extending the metabolic model to account for all these mechanisms is beyond the scope of this work. It will require a large number of experiments to further calibrate and validate the model. Our objective was primarily to validate the model across a wide range of conditions, demonstrating a strong foundation for future improvements in the model, knowing that most models of water resource recovery facilities are calibrated, but not validated (Shoener et al., 2019). Overall, despite the simplicity of the model in its present form, it is already very efficient. Especially when accounting for the diversity of strain and experimental conditions considered through the 15 studied papers. The model can then, already in its present form, can be used as a tool for optimizing microalgae growth on a mix of substrates.

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

5.4 Conclusion

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

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

Chapter 6

Modelling and optimization of lactate production via thiamine auxotroph *E. coli*

6.1 Introduction

Lactate is one of the standard and most widespread fermentation products, with a wide range of applications in the industry (Ghaffar et al., 2014). The production of lactate by fermentation processes is generally carried out in two-stages. An aerobic growth stage is followed by an anaerobic production phase (Hartline et al., 2021), where the switch between the two is controlled by the oxygen supply. The dissolved oxygen concentration acts as a natural signal for the cell to switch its metabolic state. Biomass growth is reduced but still continues during the production stage, hampering the overall productivity of the process (Venayak et al., 2018).

The necessity to use renewable sources and to reduce prices pushed for approaches in synthetic biology to improve productivity. For example, with a genetically modified *E. coli* to be a thiamine auxotroph, D-lactate productivity was improved by 10.2% (Tian et al., 2016). With this modified strain, it was possible to control the concentration of thiamine in the medium so as to regulate the flux for biomass production, eventually improving the yield of glucose to D-lactate.

Thiamine is a cofactor that regulates the reaction of pyruvate into acetyl-COA (see Fig. 6.1). In the absence of thiamine, *E. coli* cannot produce biomass. Moreover, under anaerobic conditions, almost all the pyruvate is used for the production of D-lactate. Thiamine auxotroph strains have to scavenge thiamine from the medium. Thiamine is therefore strongly controlling the metabolism, by funneling the carbon flux either to biomass or to product synthesis.

Interdisciplinary strategies combining metabolic engineering and control techniques (Choi et al., 2019) are promising tools that proved spectacular results. Genome-scale metabolic models represent an in-depth set of possible fluxes within the cell, depending on the environmental conditions. The theoretical tools from optimal control theory provide the optimal strategy to maximize the production of a target metabolite, based on the dynamics description embedded in the model.

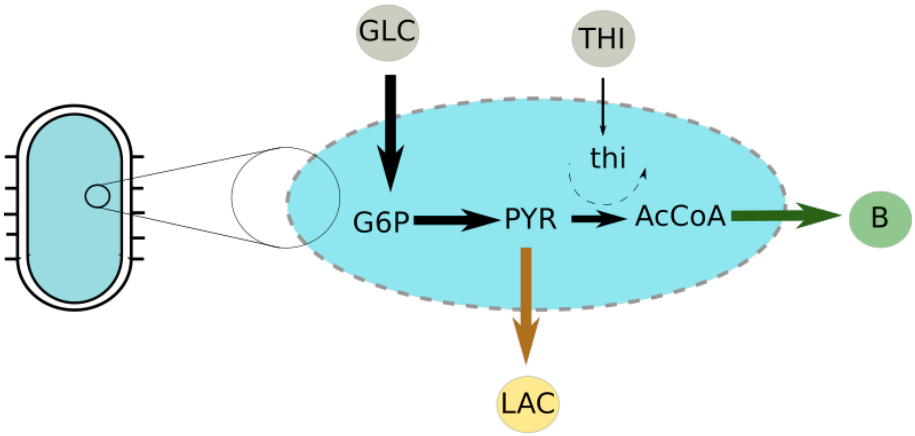

Figure 6.1: Sketch of the metabolic pathway from glucose (GLC) to pyruvate (PYR) which is then used for D-lactate (LAC) or acetyl-COA (AcCoA) production. AcCoA production, regulated by thiamine (THI), is ultimately used for biomass.

In this work, we use a genome-scale metabolic model of *E. coli* (Orth et al., 2011). The prediction power of this metabolic model has recently been demonstrated in the design of an *E. coli* strain using CO_2 as carbon source (Gleizer et al., 2019). Based on this genome-scale metabolic model, we build a reduced dynamical model of a thiamine auxotroph *E. coli* strain overproducing D-lactate. Eventually, the model predicts the reactor productivity over time. After a phase of model reduction using elementary flux modes (EFMs), the reaction rate parameters are calibrated and validated using available data from Tian et al. (2016).

We optimize a two-stage fed-batch process (see Figure 6.2) by controlling the injection of thiamine during the production stage. We also assess productivity in a continuous process and the effect of thiamine extra supply. Under optimal conditions, we predict an additional 55% increase in the production rate of D-lactate with the thiamine auxotroph strain compared to the non auxotroph strain.

6.2 Materials and methods

6.2.1 Metabolic model construction

A metabolic network with n_m metabolites and n_r reactions in a continuous reactor can be modeled by the following system of differential ordinary equations:

$$\frac{dM}{dt} = S.v(M).E - DM + DM_{in} \tag{6.1}$$

In the case of fed-batch cultivation, considering constant volume of the reactor, the following equation is valid:

$$\frac{dM}{dt} = S.v(M).E + DM_{in} \tag{6.2}$$

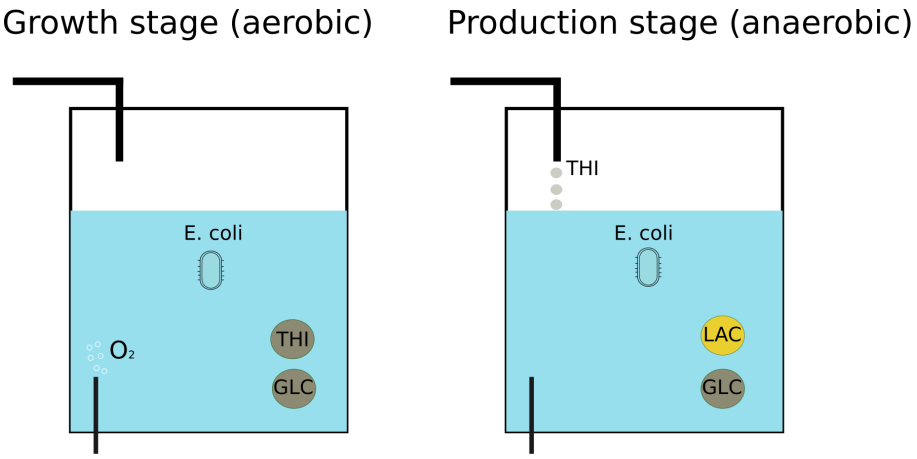

Figure 6.2: Sketch representation of the two-stage production system. Growth stage, with supply of oxygen and initial concentration of glucose (GLC) and thiamine (THI). Production stage in aerobic conditions, allowing the production of D-lactate (LAC) and external supply of thiamine if fed-batch cultivation.

Where $M \in \mathcal{R}^{n_m}$ is the vector of metabolite concentration, E is the biomass of *E coli*. $v \in \mathcal{R}^{n_r}$ is the vector of reaction kinetics, $S \in \mathcal{R}^{n_m \times n_r}$ the stoichiometric matrix, D the dilution rate, $M_{in} \in \mathcal{R}^{n_m}$ the vector of metabolite concentration in the input flux. In this work, we use the metabolic network of Orth et al. (2011) for *E. coli*, encompassing $n_r = 2251$ reactions, and $n_m = 1136$ metabolites.

The metabolic network can be reduced after computing the Elementary flux modes (EFMs). In the current hypothesis that internal metabolites are at steady state, their concentration can be deduced from the external concentration of substrates and products using the EFMs (Schuster et al., 1999). The EFMs can be found by solving the following equation:

$$S.v(M) = 0 \tag{6.3}$$

where M is the concentration of the metabolites inside the cell. The set of EFMs gives the set of chains of reactions with which the cell is at steady-state. Calculating EFMs, leads to macroscopic reactions, i.e. simplified reactions linking the substrates to the final products without the need to represent the intermediate metabolites (Provost and Bastin, 2004, Baroukh et al., 2014).

Here, we focus on glucose (GLC) as substrate and D-lactate (LAC) as product. Biomass (E) and the vitamin thiamine (THI) are also represented. The other substrates are considered to be in excess or not influencing the kinetics of the reactions. We assume that an efficient mechanism of oxygen (O_2) regulation is at play, maintaining the process in the optimal state for aerobic and anaerobic cultivation conditions with a fast switch between these two states. Since this *E. coli* strain is a thiamine auxotroph mutant, no biosynthesis pathway is represented. We consider that the internal concentration of thiamine is at equilibrium with the external concentration (THI).

6.2.2 Model reduction and macro-reactions

In genome-scale metabolic models, the calculation of the EFMs is frequently prohibitive due to their very large number and the lack of computational power (Maton et al., 2022). In order to circumvent this problem, instead of determining the set of all EFMs for the full network, we first reduce the number of reactions allowed to carry flux in the steady state by first applying Flux Balance Analysis (FBA) to the full network.

In order to generate the macro reactions of our system, we proceed with FBA in three different cases. Using as objective function the reaction of biomass production in anaerobic (setting as constraint zero flux for the reaction of uptake of oxygen) and aerobic conditions, then as objective function the production of lactate in anaerobic conditions. The result of the FBA calculation is then used to determine the subset of EFMs that are necessary to describe the growth and lactate production in the context of this work. For each of the three cases, the reactions not carrying any flux are removed from the network, then we proceed with the calculation of EFMs for this subset of the network. After EFM enumeration, when more than one EFM is calculated for the subnetwork, the vector containing the best yield for the objective metabolite is kept and used for the construction of the macro reactions of the system.

In aerobic conditions glucose is consumed with high energetic yield and biomass growth rate. Selecting the EFM with the best yield of GLC to E in aerobic conditions provides the following macro reaction.

$$10.2GLC + 0.2K + 0.3SO_4 + 10.8NH_4 + 17.9O_2 + 1.0Pi \longrightarrow 20.0CO_2 + 46.4H_2O + 9.2H + E \quad (6.4)$$

where the biomass E is expressed in kg of biomass. The yield of biomass on glucose is 0.54 $kg.kg^{-1}$, close to measured values of *E. coli* growing on glucose (Shiloach and Fass, 2005).

When growth takes place in anaerobic conditions, the following macro reaction is obtained from the set of EFMs.

$$52GLC + 11NH_4 + Pi + 0.2K + 0.3SO_4 \longrightarrow 2CO_2 + 29H_2O + 100H + 90LAC + 0.3SUC + E \quad (6.5)$$

In anaerobic conditions, the yield of glucose is considerably reduced, since the redox imbalance must be compensated, funneling a considerable part of the organic carbon to excreted organic acids. According to Orth et al. (2010a), a ratio of 4.1 between aerobic and anaerobic growth is expected when glucose is the substrate. Although, there is production of D-lactate in the macro reaction (6.5), there is a direct pathway to its production from glucose:

$$GLC + 2ADP + 2Pi \longrightarrow 2LAC + 2ATP + 2H^+ \quad (6.6)$$

this macro reaction gives the correct stoichiometric yield to model the production of D-lactate from glucose.

6.2.3 Model kinetics

To model the kinetics for the reaction of biomass production we consider a standard Monod equation for the glucose consumption, but multiplied by a Hill equation:

$$v_{E\ aerobic}(GLC, E, THI) = \mu_{Emax} \frac{GLC}{GLC + K_E} . \frac{1}{1 + (\frac{K_{THI}}{THI/E})^n} \quad (6.7)$$

The Hill equation (Saa and Nielsen, 2017) part represents the constraint of growth due to limitation of thiamine inside the cell. The intracellular concentration of thiamine is supposed to be in equilibrium with the external concentration in the medium, its influence on the kinetics is given by THI/E. In anaerobic conditions, the maximum rate of biomass production is reduced, given by the following equation:

$$v_{E\ anaerobic}(GLC, E, THI) = \mu_{E\ anaerobic}\frac{v_{E\ aerobic}}{\mu_{Emax}} \tag{6.8}$$

The production of D-lactate (LAC) is given by a simple Monod kinetic equation, dependant only on the glucose concentration:

$$v_{LAC}(GLC) = \mu_{LACmax}\frac{GLC}{GLC + K_{lac}} \tag{6.9}$$

We consider that D-lactate production only happens during the production (anaerobic) stage, although in practice there is some production even during growth stage. Therefore, $v_{LAC} = 0$ in aerobic conditions.

Besides the reactions given by the metabolic model, we consider a mortality rate for the biomass, m_E, and a decay rate for thiamine, m_{THI}. The system of differential equations for a continuous reactor is shown below:

$$\frac{dGLC}{dt} = -(\frac{v_E}{\gamma_E} + \frac{v_{LAC}}{\gamma_{LAC}})E + D(GLC_{in} - GLC) \tag{6.10}$$

$$\frac{dLAC}{dt} = v_{LAC}E - D.LAC \tag{6.11}$$

$$\frac{dE}{dt} = v_E E - m_E E + D(E_{in} - E) \tag{6.12}$$

$$\frac{dTHI}{dt} = (THI_{in} - THI)D - m_{THI}.THI \tag{6.13}$$

Here, THI represents the total concentration of thiamine in the system, including internal and external molecules. γ_E and γ_{LAC} are, respectively, the yield of biomass and lactate from glucose.

6.2.4 Model calibration

The model is calibrated using data available from Tian et al. (2016), where experiments were performed under different concentrations of thiamine. We calibrate the model separating the dataset into four sets. First, we select a small part of data where there is no growth, to calculate a mortality rate, m_E for the bacteria. Second, we select data from the growth stage in aerobic conditions to determine the parameters for biomass growth: μ_{Emax}, K_E, K_{THI}, n and m_{THI}. Third, data from a non-auxotroph *E. coli* is used to calibrate the $\mu_{E\ anaerobic}$ in anaerobic conditions. Fourth, data in anaerobic conditions is used to calibrate the parameters for lactate production: μ_{LAC} and K_{LAC}.

The best fitting parameters for each of the four sets of calibration were obtained with the Differential Evolution algorithm (Storn and Price, 1997). This stochastic algorithm identifies the global minimum of a multivariate function. It is available in the *Scipy* package for *Python* programming language. Here, the objective function is to minimize the squared error between model prediction and experimental data.

6.3 Results

6.3.1 Model calibration

Calibrated parameters for the kinetic equations can be found in Table 6.1. The parameters belong to the expected ranges, comparing them to other models found in the literature. In Zhou et al. (2012) maximum specific lactate productivity measured was 11 day^{-1}, while μ_{LACmax} calibrated here was 13.48 day^{-1}. See Varma et al. (1993) for a discussion on *E.coli* growth rates using glucose as substrate and by-production secretion in limited oxygen supply.

Based on the minimum number of molecules of thiamine to maintain growth, 96 molecules per cell, given by Tian et al. (2016), K_{thi} should be below 0.0644 μg of thiamine per g of *E. coli* biomass, the calibrated value of 0.015 is in the range of expected values. The decay rate, m_{THI} for thiamine is much higher than the calibrated value for the mortality rate of biomass, m_E. This is expected because thiamine has a very short half-life (Marrs and Lonsdale, 2021).

The model reproduces accurately the experimental data used for the calibration phase (Figure 6.3). Performance with the validation data set (6.4) is also accurate, although especially at thiamine concentration of $0.31 \mu g/L$ the model underestimates experimental data.

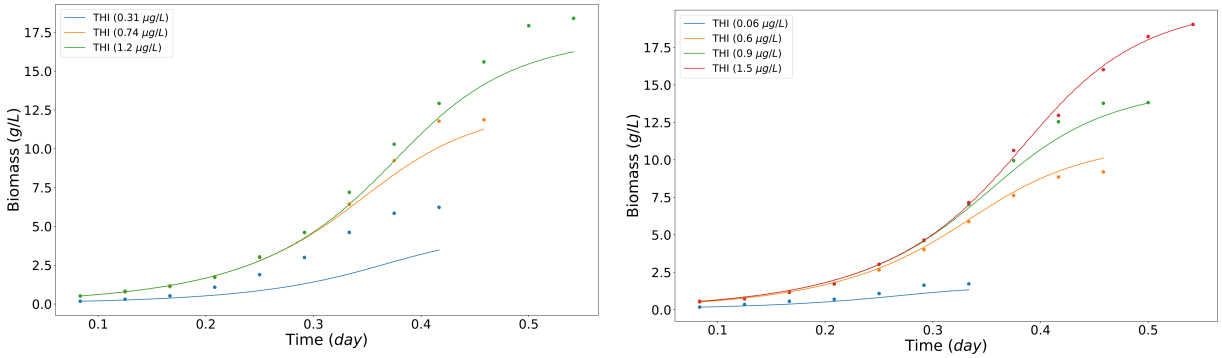

Figure 6.3: Model validation (left) and calibration (right) for different initial concentrations of thiamine

6.4 Discussion

6.4.1 Optimisation of the continuous process

The first objective of this paragraph is to propose a strategy for maximizing, at steady state, the lactate productivity (P) in the reactor (given by $P = D.LAC$). The following objective

-0cm

Parameter	Value	Unit	Description
μ_{Emax}	10.61	d^{-1}	Maximum rate of aerobic growth
$\mu_{anaerobicE}$	0.62	d^{-1}	Maximum rate of anaerobic growth
K_E	0.04	$g.L^{-1}$	Half-saturation constant of growth rate
m_E	0.33	d^{-1}	Mortality rate
K_{thi}	0.015	$\mu g.L^{-1}$	Half-saturation constant of thiamine influence on growth
n	2.14	-	Hill coefficient
m_{THI}	5.5	d^{-1}	Decay rate of thiamine
μ_{LACmax}	13.48	d^{-1}	Maximum production rate of lactate
K_{LAC}	0.97	$g.L^{-1}$	Half-saturation constant of lactate production
γ_E	0.54	$gGLC.gE^{-1}$	Yield of biomass from glucose
γ_{LAC}	1	$gGLC.gLAC^{-1}$	Yield of D-lactate from glucose

Table 6.1: Parameters calibrated for the model.

function must then be maximized:

$$\max_{D,THI_{in}} : P = D.LAC \tag{6.14}$$

Figure 6.5 shows the effect of influent thiamine concentration on the lactate productivity, once the optimal dilution rate has been applied (D such that P is maximized). We consider a two-step process, where the first step is carried out at a high aeration rate to maximize biomass productivity. The flux of *E. coli* biomass, E_{in}, is then injected in the continuous reactor for lactate production.

For this optimization, we consider an influent biomass concentration, E_{in} of $11g/L$ from the first stage and $100g/L$ of influent glucose (GLC_{in}). As we can see in Figure 6.5, the maximum productivity is achieved with a minimal concentration of thiamine of about $5\mu g/L$. Adding more thiamine does not increase productivity anymore. Also, for low influent concentrations of thiamine, the dilution rate increases to high values which are unrealistic in real production conditions. Such a high dilution rate compensates for the lack of thiamine to induce biomass growth.

Now we consider a more accurate criterion where we account for the initial transient at the reactor start-up, for total process duration T. The average productivity on the horizon (0; T) is given by:

$$P = \frac{1}{T}\int_0^T D(t).LAC(t).dt \tag{6.15}$$

Simulations considering both constant and periodic dilution rates were performed. The periodic regime was carried out assuming varying thiamine concentration could induce periods of higher biomass followed by periods with lower thiamine concentration with a better yield of glucose to lactate. Nevertheless, numerical calculations showed that the optimal productivity is obtained at a constant dilution rate. Some simulations with periodic dilution rates showed an increase in productivity, with a minimal gain of about 2%, possibly due to numerical approximation or because the steady-state was reached faster with a varying dilution rate. A

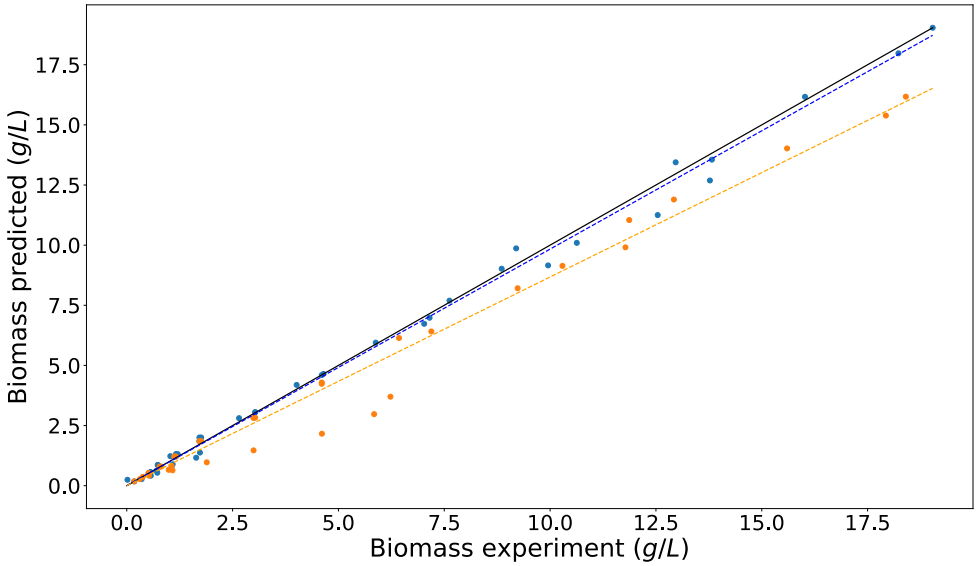

Figure 6.4: Biomass data points showing model prediction and expected value. Blue data points are used for the calibration of the model, and orange data points for validation. Calibration ($R^2 = 0.996$) and validation ($R^2 = 0.951$). Linear regression of calibration dashed blue curve (slope coefficient 0.98), and validation results dashed orange curve (slope of 0.86).

mathematical development using optimal control theory could validate the hypothesis that, after an initial transient, a constant dilution rate is indeed the optimal cultivation condition in continuous operation.

This result contrasts with optimal fed-batch cultivation (see the following Section), where a maximum influent thiamine concentration leads to optimal lactate productivity. This happens because during fed-batch it is necessary to minimize the loss of carbon from glucose to biomass. While during continuous cultivation, productivity is dependent on the instantaneous production rate of lactate, which is not dependent on thiamine concentration.

6.4.2 Optimisation of a two-stage fed-batch system

We proceed with the optimisation of the fed-batch system, considering two cases. First, we focus only on the production stage, then we optimize both stages of the process. We compare our optimised results with experimental data from Tian et al. (2016). When optimising only the second stage, we consider as the initial condition the experimentally measured concentrations at the switching point.

We consider a fed-batch with addition of vitamin. When considering only the production stage, the objective functions is the following:

$$\max_{D_{thi}} : P = \frac{1}{T} LAC(T) \tag{6.16}$$

Here we consider the variable $D_{thi} = D.THI_{in}$. We assume that THI_{in} is concentrated enough

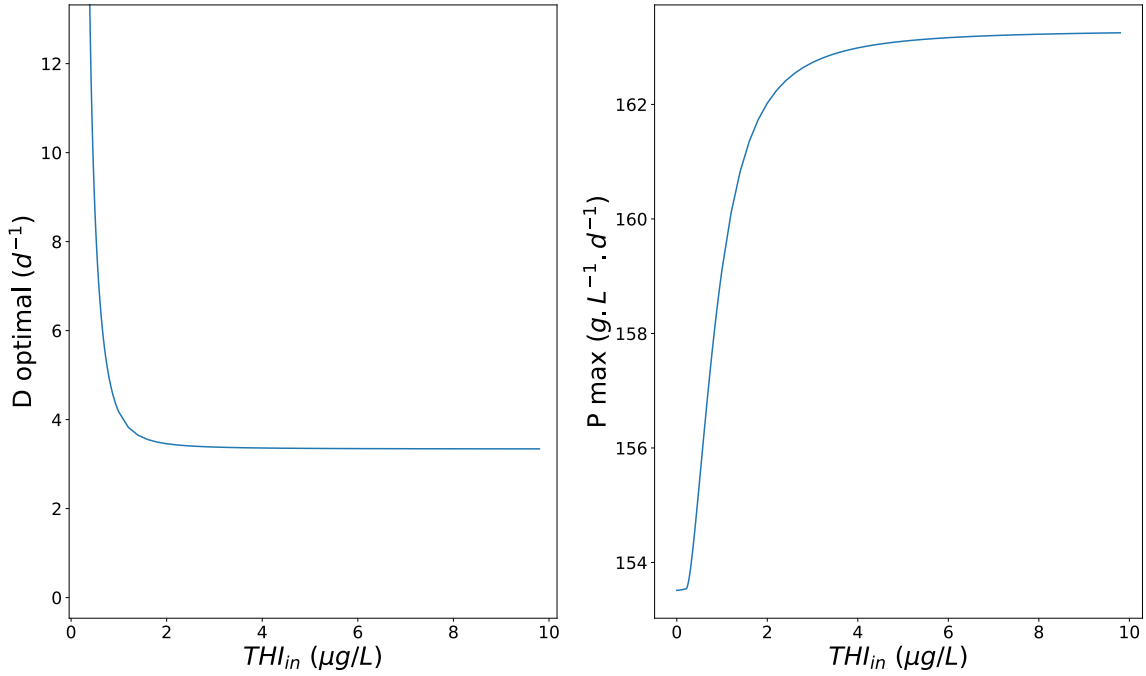

Figure 6.5: Optimal dilution rate (left figure) and maximum productivity (right figure) for different influent thiamine concentrations. Considering E_{in} 11 g/L and GLC_{in} 100 g/L.

so that D does not have a significant impact on the volume of the reactor, therefore D is not present in the dynamics of the other differential equations.

To optimize the two-stage process as a whole, the control variables are the initial concentration of thiamine (THI_0), the time of switch between the growth (aerobic) and production (anaerobic) stage (t_g), and the dilution rate D.

$$\max_{(D,THI_0,t_g)} : P = \frac{1}{T} LAC(T) \tag{6.17}$$

Figure 6.6 shows the concentration of lactate, biomass and thiamine over time for three different cases, together with non-optimized experimental batch. The dotted curve *batch model* represents model simulation of experimental data in non-optimal cultivation conditions. The dashed curve, *fed-batch opt.*, is an optimized fed-batch giving the initial conditions of the non-optimized experimental batch. Finally, the full curve *two-stage opt.* shows the optimized model simulation for the whole duration of the two-stage process.

The optimized fed-batch cultivation, considering only the second stage, slightly increases the concentration of thiamine in the medium, compensating the biomass loss due to the mortality rate. By consequence, at the end of the process there was a slightly higher biomass concentration, enough to increase the instantaneous production of lactate without reducing the yield. The gain with the addition of thiamine was minor. The productivity of the non-optimal batch was 101 $g(L.d)^{-1}$, compared to 106 $g(L.d)^{-1}$ for the optimized fed-batch.

A considerably higher productivity (157 $g(L.d)^{-1}$) was obtained when the two steps of the process were optimized, as seen in Figure 6.6. Here, the strategy was to start the production stage with a higher biomass concentration. For this purpose, the cultivation starts with a higher concentration of thiamine, THI_0, guarantying a faster biomass growth. After the switch to anaerobic production of lactate, the thiamine concentration slowly decreases, but still keeping a concentration higher than in the batch cultivation. The process ends when the glucose in the medium is depleted at 99.9%.

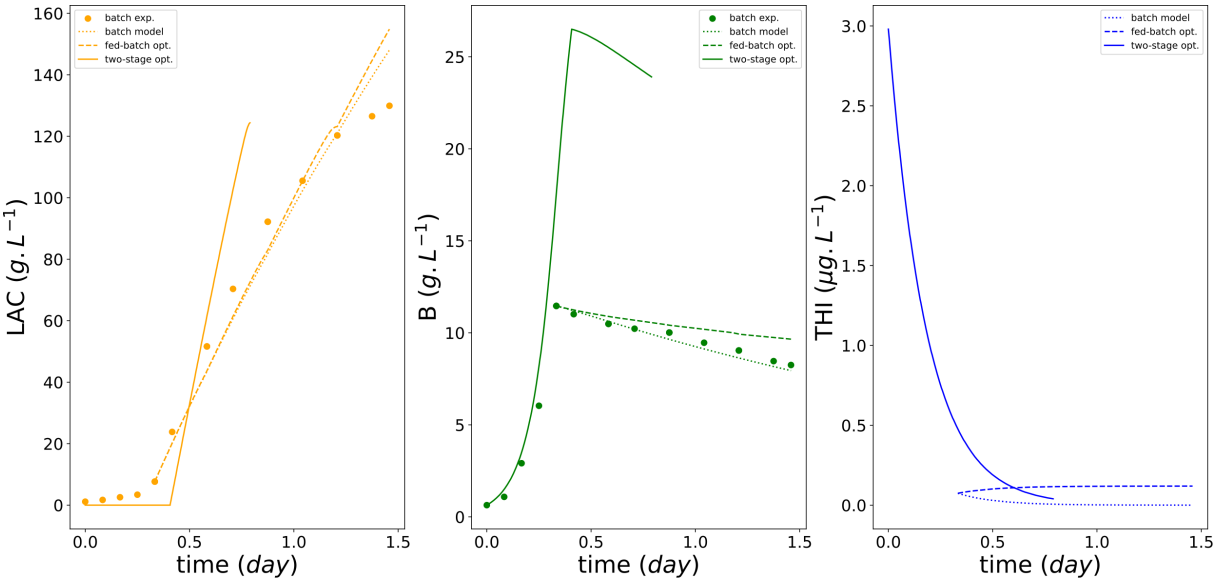

Figure 6.6: Lactate, biomass and thiamine concentration, considering three cultivation conditions. Non optimal batch, fed-batch with optimized vitamin supply for anaerobic stage, optimized two-stage fed-batch.

6.5 Conclusion and model perspectives

The standard strategy for producing lactate with *E. coli* is to use a two-stage process. First an aerobic stage for biomass growth and then a second anaerobic phase for lactate fermentation. This is not the only possible strategy. For instance, Zhou et al. (2012) experimented a process where lactate production occurred also during growth stage, what minimized process time, obtaining a two-fold productivity. As seen in Varma et al. (1993), the availability of oxygen determines the by-products of the fermentation by E. coli; also, when cells grow under conditions of oxygen limitation, the glucose uptake rate (shadow cost) increases to compensate for the lower ATP yield of fermentation. The strategy of coupling oxygen limitation with B1 supply limitation is supposed to increase lactate production, because even less metabolic pathways will be available for *E. coli*, forcing all pyruvate obtained from glycolysis to go into the lactate fermentation pathway to compensate for the redox imbalance (Boecker et al., 2022).

In this work, we show that the dynamical reduction of a genome scale metabolic network can allow to apply the tools of control and optimisation for maximising productivity. These approaches can further guide and improve the results of synthetic biology experiments. In particular, the control of vitamin concentration in the culture medium might guide the metabolic pathway of microorganisms improving the production of a desired product. In Chapter 7, we will expand the results and insights of this work, associating a vitamin producer (a bacteria) and an auxotrophic organism (a microalgae) in the same culture.

Chapter 7

Metabolic modeling of a symbiotic association between microalgae and bacteria: applications to lipid production

7.1 Introduction

Symbiotic interactions between microalgae and bacteria have a great influence on the evolution and development of ecosystems. Recently, the potential of their symbiosis became the aim of scientific research to improve the current design of bio-processes (Ramanan et al., 2016, Zhang et al., 2021, Xiao et al., 2021). In particular, it has been demonstrated that wastewater treatment can benefit from the symbiotic interactions between algae and bacteria to improve the removal of phosphorous and nitrogen from the medium (Casagli et al., 2021b). Advances in the modeling of these complex symbiotic interactions will help to decipher the dynamics of the ecosystem and eventually will pave the way for more efficient and sustainable processes.

The importance of the interactions between microalgae and bacteria is clear when one species necessarily depends on a compound produced by the other. Some vitamins, such as thiamine, are required by all living cells, but not all microalgae are capable of producing them, i.e. they are auxotroph (Croft et al., 2005, Cooper et al., 2019). In some cases, the algae is not completely dependent on a vitamin, but it may take advantage of it when it is supplied from an external source. This is the case of *C. reinhardtii* when it is growing in the presence of an external supply of cobalamin (B12), which enables a B12-dependent pathway for the synthesis of methionine that has a higher reaction rate.

The exogenous supply of vitamins from bacteria to microalgae is an important factor for designing more sustainable cultivation systems (Tandon et al., 2017). The modeling of algae-bacteria interactions will provide guidance for the development of more robust and environmentally friendly processes with reduced need for external nutrients or micronutrients. It will allow a model-based approach for improving the design of the cultivation system and offers the

possibility to set up optimal strategies to improve the overall efficiency of the process. A model can also be a useful tool to guide metabolic and genetic engineering.

Biotin (vitamin B7) is a necessary co-factor in lipid biosynthesis for all organisms, because it acts together with the enzyme acetyl-CoA carboxylase (ACCase) to promote the transformation of acetyl-CoA into malonyl-CoA, which is the primary building block for the synthesis of fatty acids (Huerlimann and Heimann, 2013) (see Fig. 7.1 for a simplified overview of the lipid pathway synthesis).

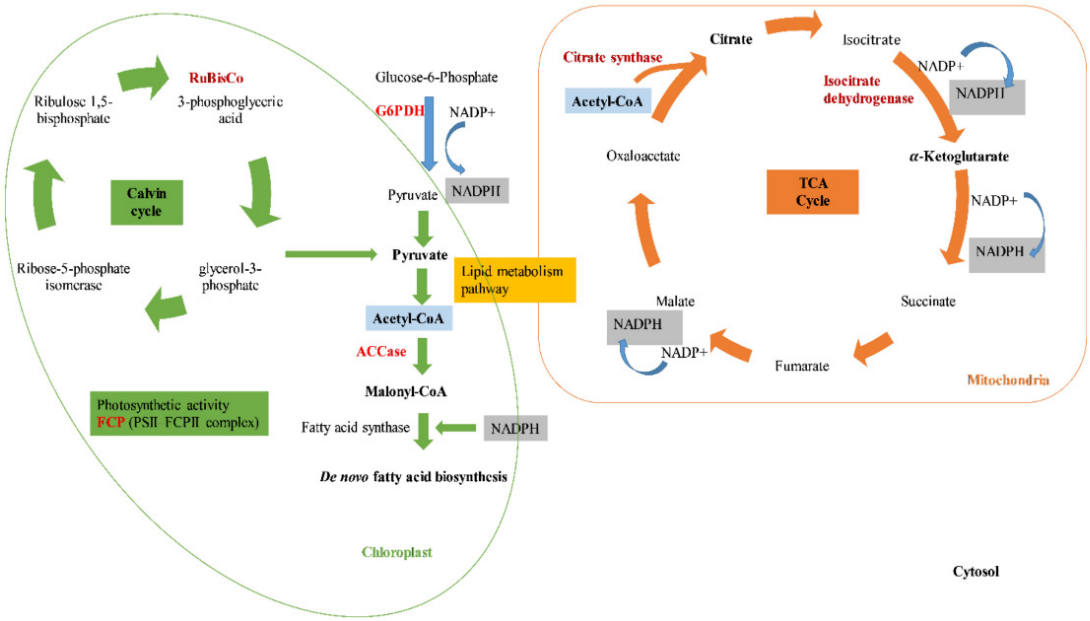

Figure 7.1: Lipid synthesis pathway, adapted from Curcuraci et al. (2022)

Recently, there were several attempts to improve lipid production in microalgae with the help of metabolic and genetic engineering (Sun et al., 2019). Some of those works aimed at the overexpression of genes related to the lipid synthesis pathway, for example, the genes related to the synthesis of ACCase. Tai and Stephanopoulos (2013) by overexpression of ACCase alone increased lipid content by 2-fold in yeast *Yarrowia lipolytica*. Davis et al. (2000) by overproducing the four subunits of ACCase of *E. coli* with the help of a bacteriophage, fatty acid synthesis was increased by 6-fold. Gomma et al. (2015) also via overexpression of ACCase genes in *Scenedesmus quadricauda* improved fatty acid content by 1.6 fold.

Magdouli et al. (2020) tried to improve lipid production with the addition of leucine and biotin for *Y. lipolytica* without genetic modifications. The control medium contained $25\mu g/L$ of biotin, at $100\mu g/L$ there was an improvement of about 15 % in lipid concentration. According to Barth and Gaillardin (1997) *Y. lipolytica* is an auxotroph for thiamine, and at least for some mutants also for biotin. Fazeli Danesh et al. (2018) found for a culture of mixed microalgae species an improvement in lipid accumulation, from 17 to 33 % of dry weight composition, by using a medium supplemented with vitamins biotin, thiamine, cyanocobalamin, and B5.

Sun et al. (2019) do not recommend overexpressing ACCase, mainly based on results from Sheehan et al. (1998) and Li et al. (2014a). In the first, genetic engineering for overexpression of ACCase did not result in lipid accumulation for *Cyclotella cryptica* and *Navicula saprophila*. In the second, a downregulation of ACCase was found for *Nannochloropsis oceanica* under nitrogen depletion conditions. Rismani-Yazdi et al. (2012) found a more nuanced result for *Neochloris oleoabundans*, where genetic expression of cytosolic ACCase was downregulated, while the biotin subunit present only in the plastidic ACCase was upregulated. Nevertheless, Giridhar Babu et al. (2017) working with *Chlorella sorokiniana* found upregulation of ACCase in N limitation conditions in the presence of phytohormones. Demonstrating the necessity of ACCase activity and biotin for lipid accumulation in the genus *Chlorella*.

According to Croft et al. (2006) few microalgae species are biotin auxotrophs, and usually they are also auxotrophs for thiamine or cobalamin. The necessity of biotin for lipid production is well established by our knowledge of metabolic pathways, but most works focused on genetic engineering trying to over synthesize the enzyme ACCase without considering the influence of biotin. The mixed results obtained, regarding the effect of ACCase in lipid accumulation, demonstrate the complexity of the lipid synthesis pathway and its regulation. Also, little attention has been paid to the actual concentration of biotin, and its effect on lipid accumulation (Hakalin et al., 2014, Welter et al., 2013).

The synthesis of biotin has only recently been understood in model organisms (Wei et al., 2021, Lin et al., 2010). The known precursors of biotin are pimelic acid and malonyl-CoA, requiring many enzymatic reactions until the final formation of biotin, and a remarkable slow rate reaction at the last step of the synthesis with biotin synthase (Jeschek et al., 2017). As a consequence of the costly synthesis of biotin and based on the current state of the art regarding lipid synthesis pathway regulation, it is likely that the increase of biotin concentration can impact the accumulation of lipids in *Chlorella*, although it might also require the overexpression of ACCase and other enzymes.

In this work, we consider this possible pathway of regulation of lipid production via biotin, by considering a co-culture of a biotin over-producer *E. coli* (Wei et al., 2021) and of *Chlorella*. We adapt the metabolic model from Chapter 5, by including the lipid content of *Chlorella* and by considering its dependency on the biotin availability. We assess, via simulations, whether the model supports our strategy accounting for the uncertainties in the parameter representing the influence of biotin in the rate of lipid synthesis.

7.2 Methods

7.2.1 Model

The proposed model is based on the metabolic models of *Chlorella* and of *E. coli* from Chapters 5 and 6. Figure 7.2 presents a simplified scheme of the symbiotic model. The model is derived from these metabolic networks after a step of dynamical reduction following the DRUM approach (Baroukh et al., 2014, 2016). Here, we only consider the autotrophic growth of *Chlorella* and detail the lipid synthesis pathway. We consider an *E. coli* biotin over-producer as described in Wei et al. (2021). The accumulation variables for which the dynamics are described are

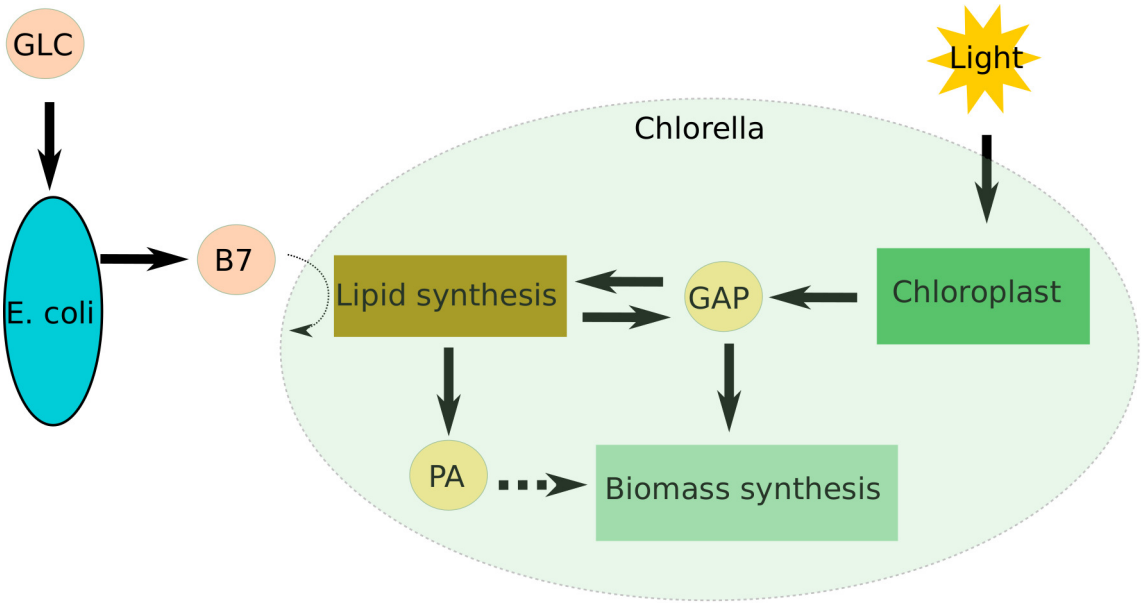

Figure 7.2: Simplified representation of the metabolic network, showing the 6 state variables and reactions.

glyceraldehyde 3-phosphate (GAP), in cellular concentration (grams per grams of biomass), and phosphatidic acid (PA), representing the cell lipids also in cellular concentration. Four additional state variables are considered: E, *E. coli* biomass concentration; GLC, glucose concentration; $B7$, vitamin biotin concentration; B, *Chlorella* functional biomass concentration. The dynamical metabolic model results from the subsystem of ordinary differential equations described below, for a continuous reactor. The value and description of the parameters are shown in Table 7.1.

$$\frac{dE}{dt} = \mu_{Emax}\frac{GLC}{GLC + K_E}E - (D + m_E)E \tag{7.1}$$

E. coli grows heterotrophically, using glucose as a carbon substrate.

$$\frac{dGLC}{dt} = -\gamma_E.\mu_{Emax}\frac{GLC}{GLC + K_E}E + D(GLC_{in} - GLC) \tag{7.2}$$

Glucose is consumed only by *E. coli* and is supplied in the influent at a concentration, GLC_{in}.

$$\frac{dB7}{dt} = \mu_{B7_{max}}.\frac{GLC}{GLC + K_{B7}}E - (m_{B7} + D).B7 \tag{7.3}$$

Biotin is produced by *E. coli* at a maximum rate of $\mu_{B7_{max}}$, following a Monod relation depending on the concentration of glucose.

$$\frac{dGAP}{dt} = \mu_{GAP}(I, X) + \mu_{PA^{-1}_{max}}.PA - \gamma_{GAP}\mu_B - \mu_{PA_{max}}.GAP.\frac{B7/B}{B7/B + K_{PA}} - \mu_B.GAP \tag{7.4}$$

Glyceraldehyde 3-phosphate (GAP) is synthesized through the photosynthetic pathway and is used as a reactant for functional biomass production and for the production of lipids, at a rate regulated by the concentration of biotin.

$$\frac{dPA}{dt} = \gamma_{PA_{GAP}}.\mu_{PA_{max}}.GAP.\frac{B7/B}{B7/B + K_{PA}} - \gamma_{PA_{GAP}}.\mu_{PA^{-1}_{max}}PA - \mu_B.PA \tag{7.5}$$

Phosphatidic acid acts as a carbon reserve, therefore it is produced by a reversible reaction, from GAP, and its consumption back to GAP is regulated by its own concentration.

$$\frac{dB}{dt} = (\mu_B - D)B \tag{7.6}$$

where,

$$\mu_B = \mu_{B_{max}}.GAP \tag{7.7}$$

Chlorella biomass production rate is regulated by the internal concentration of GAP. B is the *functional biomass*, the total microalgal biomass (or dry weight biomass), X, is given by the following equation:

$$X = \frac{B}{1 - PA - GAP} \tag{7.8}$$

The rate of synthesis of GAP from photosynthesis is given by the average growth in the reactor (Béchet et al., 2015):

$$\mu_{GAP(I,X)} = \frac{\mu_{GAP_{max}}}{\sigma.(X+E).L} ln(\frac{K_I + \sigma I_0}{K_I + \sigma I_0 exp(-\sigma.(X+E).L)}) \tag{7.9}$$

where the light attenuation due to biomass absorption and scattering is given by the Beer-Lambert equation:

$$I = I_0 exp(-(X+E)\sigma L) \tag{7.10}$$

where I_0 is the light intensity at the surface of the reactor, L the depth of the reactor, and σ the extinction coefficient is given by the average of the extinction coefficients of E. *coli* and *Chlorella*:

$$\sigma = \frac{a.X^{1-b} + \sigma_E.E}{E + X} \tag{7.11}$$

7.2.2 Parameter calibration

Model parameters related to the dynamics of biotin - μ_{B7}, K_{B7} and m_{B7} - are calculated using data reported by Wei et al. (2021) for the growth of an *E. coli* biotin over-producer. This work presents the growth of *E. coli* together with glucose and biotin concentration over time. The parameters are determined by minimizing the error between the model predicted and experimentally measured biotin concentration in the medium over time, using the Differential Evolution optimization algorithm (Storn and Price, 1997). Figure 7.3 shows the calibration

Parameter	Value	Unit	Description
μ_{Emax}	10.6	d^{-1}	Bacteria maximum growth rate
K_E	0.04	$g.L^{-1}$	Half-saturation constant for glucose consumption
m_E	0.33	d^{-1}	Mortality rate of bacteria
γ_E	1.84	$gGLC/gE$	Yield of glucose to bacteria biomass
μ_{B7}	0.05	d^{-1}	Maximum production rate of biotin (vitamin B7)
K_{B7}	3.8	$g.L^{-1}$	Half-saturation constant of biotin production
m_{B7}	0.38	d^{-1}	Decay rate of vitamin B7
μ_{PAmax}	23.3	d^{-1}	Maximum production rate of lipids
K_{PA}	10^{-5} *	$gB7/gB$	Half-saturation constant of lipid production
γ_{PA_GAP}	3.87	$gPA/gGAP$	Yield of lipid production from GAP
γ_{PA_B}	0.31	gPA/gB	Yield of lipids to B
$\mu_{PA^{-1}}$	2.75	d^{-1}	Maximum rate of GAP production from stored lipids
K_I	182	$\mu mol.g^{-1}s^{-1}$	Light half-saturation constant
σ_E	285	m^2g^{-1}	E. coli extinction coefficient
a	117.2	-	Light extinction coefficient
b	0.2	-	Light extinction power coefficient
$\mu_{GAP_{max}}$	1.15	d^{-1}	Maximum synthesis rate of GAP from photosynthesis
γ_{GAP}	3.79	$gGAP/gB$	Yield of GAP to microalgae biomass
$\mu_{B_{max}}$	3.64	d^{-1}	Maximum growth rate of microalgae biomass

Table 7.1: Parameters calibrated for the model. $^*K_{PA}$: first guess value

results, by comparing the simulations to the experimental data. The calibrated parameters are: $\mu_{B7} = 0.05d^{-1}$, $K_{B7} = 3.8g.L^{-1}$ and $m_{B7} = 0.38d^{-1}$. *E. coli* average extinction coefficient, σ_E, was calculated based on optical density data and average cell composition information (Hu et al., 2017, Mira et al., 2022).

The parameters related to the synthesis of lipids were modified from Baroukh et al. (2014) since here we consider the cellular concentration (grams of metabolite per grams of biomass) and not the total concentration (grams per liter). Especially, $\mu_{PA^{-1}}$ and μ_{PAmax} were modified. This is done by multiplying the old parameter with the average value of biomass of the experiments used to calibrate the previous model - and also by multiplying it with the average concentration of other internal accumulating metabolites not considered in this current model.

The half-saturation constant of biotin regulating the synthesis of lipids, K_{PA} is the only undetermined parameter of the model. We consider as a first guess the value of $10^{-5}gB7/gB$, based on the experimental conditions of Magdouli et al. (2020) where biotin concentrations from 25 to 200 $\mu g/L$ were used for a culture of *Yarrowia lipolytica* reaching biomass concentrations up to 30 g/L. Additional experiments are definitely needed to more accurately assess the value of this parameter.

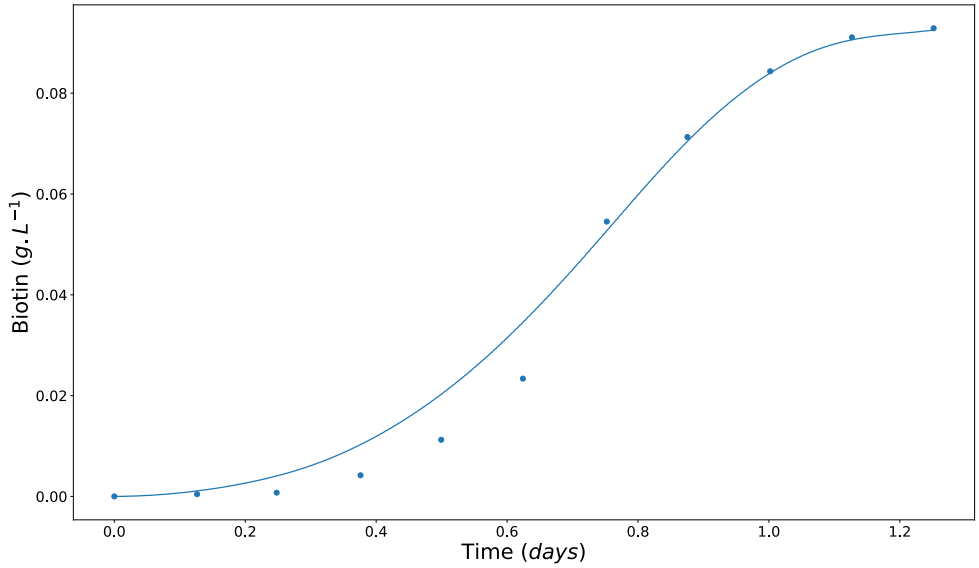

Figure 7.3: Model calibration for biotin production. Experimental data points from Wei et al. (2021)

7.2.3 Optimisation

Here we consider the lipid productivity P in the reactor outlet:

$$P = PA.X.D \tag{7.12}$$

The objective is to determine the optimal conditions for maximizing the lipid production rate at steady state, playing on the dilution rate D and the influent glucose concentration GLC_{in}. The values of GLC, E, $B7$, and GAP at steady state can be directly computed from equations 7.1, 7.2, 7.3 and 7.6, respectively. The value of B at steady state is determined numerically from the root of the equation, using Newton's algorithm. Finally, PA is derived from the value of B with Equation 7.5. The values of D and GLC_{in} maximizing P at steady state are obtained using the Differential Evolution optimization algorithm (Storn and Price, 1997).

7.3 Results and discussion

7.3.1 Model dynamical behaviour

Figure 7.4 shows the dynamical behavior at two different values of GLC_{in}. It shows that the concentrations of *E. coli* and biotin rapidly reach their steady state, due to the faster heterotrophic growth rate. The response of *Chlorella* is different, due to the slower growth in autotrophic conditions. Several days are needed for the microalgal biomass to reach its steady state. The lipid content is dependent on the light availability since lipids constitute a way to store carbon. Under normal day-night conditions, microalgae store carbon during the day in the

form of lipids, and later consume them during the night (Lacour et al., 2012). More generally, higher light intensity is known to stimulate lipid accumulation (Maltsev et al., 2021).

For a low glucose concentration ($GLC_{in} = 0.1g/L$), *Chlorella* biomass increases over time, reducing light availability, and as a consequence the lipid content decreases over time. When GLC_{in} equals $0.5g/L$, a higher concentration of *E. coli* is reached at the beginning of the cultivation and as a consequence, reducing then the light transmitted to the culture. The microalgal biomass concentration thus decreases over time. It is important to note also, that since we consider an equilibrium between external and internal concentration of biotin, as the biomass decreases the internal biotin is higher ($B7/B$), and thus the internal lipids per biomass unit, PA, increases. This factor aggregates with the dynamics of light availability, since internal concentration of biotin will also determine the content of lipids at steady state.

This analysis highlights one of the most important behaviors of the model, the trade-off between light availability and biotin production. A higher glucose influent concentration increases the concentration of bacteria which reduces the transmitted light, and thus the average light intensity. Because of this behavior, as the production of biotin increases, there is eventually a loss of productivity due to reduced light availability. This is an important limitation of the co-culture when considering the autotrophic growth of microalgae. Given the probable range of K_{PA}, a concentration of *E. coli* moderately impacting light availability could be reached, while supplying the minimum required biotin to enhance lipid accumulation.

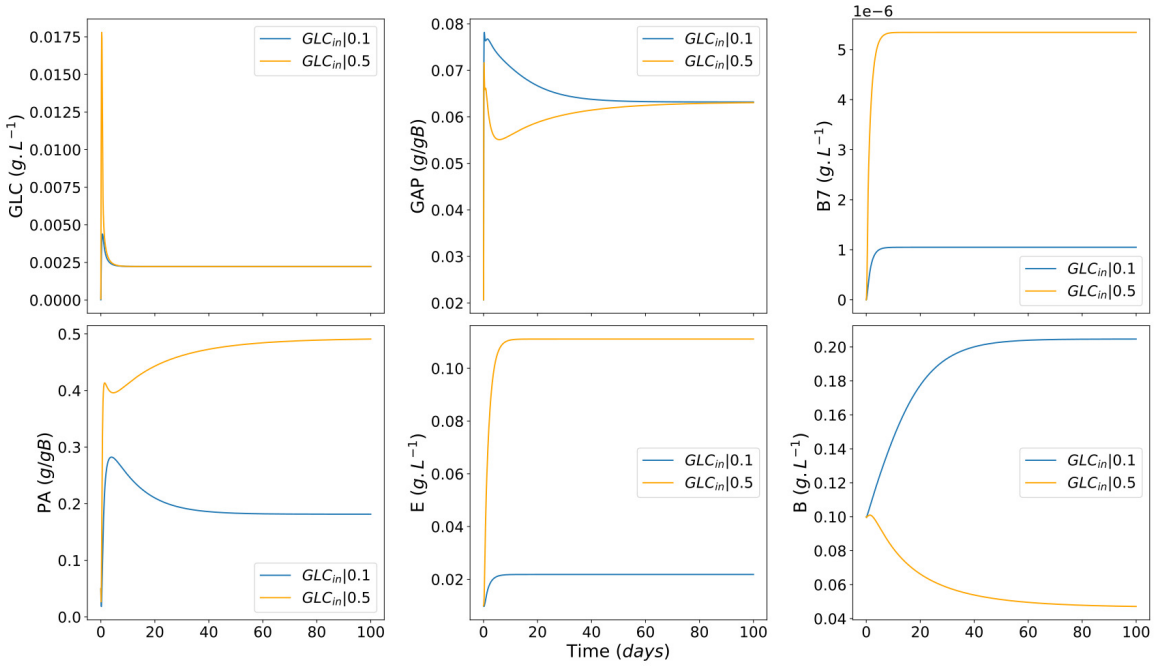

Figure 7.4: Dynamics of the system at constant dilution rate ($D = 0.23d^{-1}$), for two different influent concentrations of glucose. GLC_{in} : $0.1g/L$ (blue) and $0.5g/L$ (orange).

Figures 7.5 and 7.6 show the influence of parameter K_{PA} in the production of lipids. Figure 7.6 presents the optimal productivity of lipids, optimal dilution rate, and the optimal input of glucose with respect to the value of K_{PA}. As the value of K_{PA} increases, more biotin is necessary to keep the same rate of lipid synthesis. In order to increase the production of biotin by *E. coli* the supply of glucose needs to be increased. This can be achieved by increasing the dilution rate and the influent concentration of glucose. GLC_{in} has a direct impact in the concentration of *E. coli* at steady state, therefore increasing the production of biotin. As shown in Figure 7.6, the optimal dilution rate increases with K_{PA} until it reaches a stable value, when an increase of the dilution rate does not result in a productivity gain, as it goes close to the wash-out point of *Chlorella* biomass concentration. In Figure 7.5 lipid productivity contour lines are shown for four different values of K_{PA}, it demonstrates that as the value of K_{PA} increases, the optimal operational conditions get closer to the wash-out region, strongly affecting the productivity.

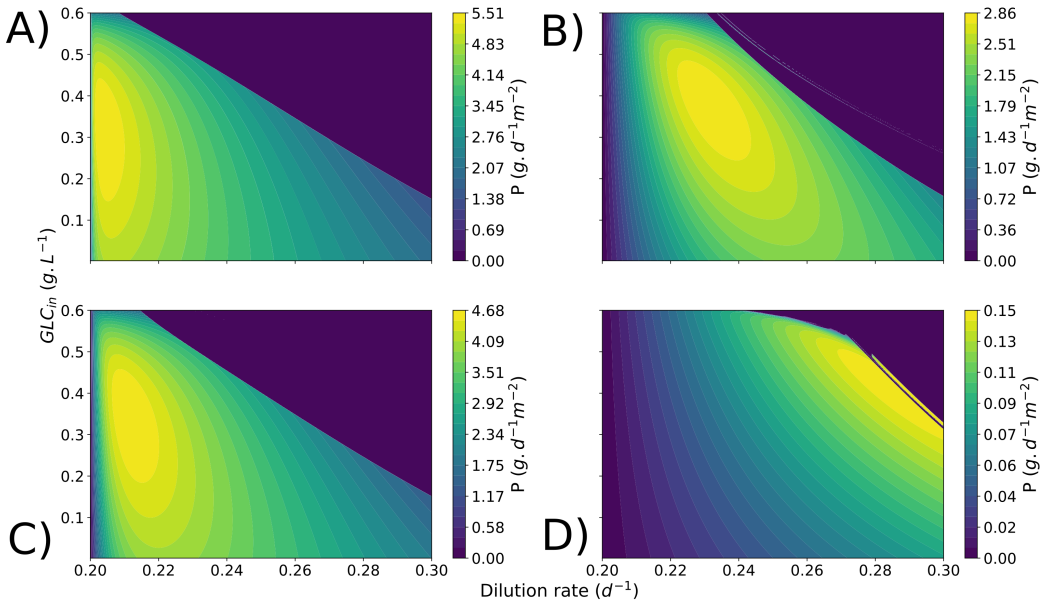

Figure 7.5: Contour lines showing how lipid productivity (P) changes as a function of the dilution rate (D) and the concentration of influent glucose (GLC_{in}) for four different values of K_{PA} ($gB7/gB$). A) $K_{PA} = 10^{-7}$. B) $K_{PA}= 10^{-5}$. C) $K_{PA}= 10^{-6}$. D) $K_{PA}= 10^{-3}$.

Figure 7.7 shows the behavior of the model at steady state considering three cases of glucose influent concentration. Optimal influent glucose concentration (GLC_{in} *opt*) maximizing lipid production for a fixed dilution rate, and two different inputs, 0.8 and 1.2 times this optimal value. Figure 7.7-A shows the productivity of lipids at different dilution rates, highlighting the existence of the optimal dilution rate giving the maximum productivity. Note also that an excess or a lack of glucose in the influent produces similar productivity losses. Figure 7.7-C shows

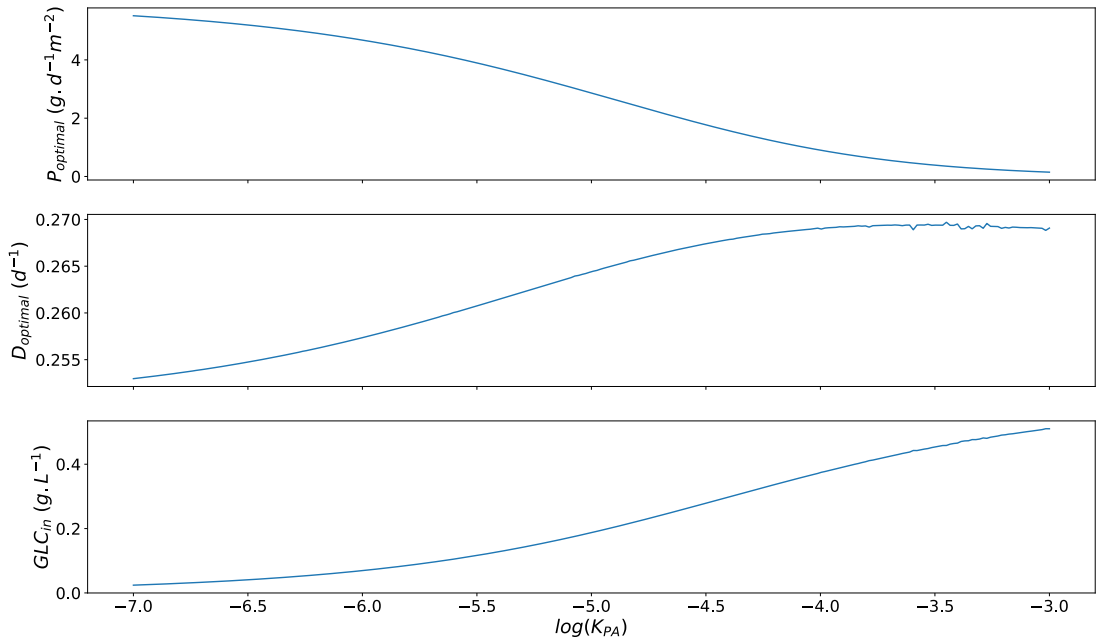

Figure 7.6: Optimal productivity of lipids ($P_{optimal}$) at the steady state and the respective dilution rate and the concentration of influent glucose (GLC_{in}) to obtain it, considering different values of K_{PA}.

how the ratio of *Chlorella* per *E. coli* biomass changes with the dilution rate, demonstrating the existence of the optimal ratio where the trade-off between biotin production and light availability is optimal. Plots 7.7-B and 7.7-D show how lipid content and biotin availability ($B7/B$) varies with the concentration of *E. coli* at steady state. Most importantly, it shows how an excess of glucose leads to higher biotin availability and as a consequence higher lipid content, but with lower algal biomass.

7.3.2 Model limitations and perspectives

The main uncertainty of the model is the relationship between the lipid content and the internal concentration of biotin. According to our model, the content of lipids at steady state is regulated by the internal concentration of biotin. The consideration of a Monod equation multiplying the kinetics of lipid synthesis is a first approximation, reducing the number of unknown parameters to only one unknown. It may be necessary to consider, instead of multiplying the effects, a minimal relation which would better model the case of excess of biotin, or even using a Hill equation, as it was used for thiamine in Chapter 6.

Here, we did not represent the internal production of biotin by *Chlorella*. It is reasonable to assume that the production of biotin by *Chlorella* is negligible compared to the quantity produced by the *E. coli* mutant. Regarding the model, the major change would be that as the concentration of external biotin goes to zero, there would still be a concentration of biotin for lipid production. Also, it would be important to determine experimentally if the external biotin

will down regulate the production of internal biotin by *Chlorella*. As seen in Kazamia et al. (2012), a vitamin B12 optional *C. reinhardtii* gene expression was regulated in the presence of a B12-producing bacteria. Since biotin synthesis is costly (Jeschek et al., 2017), we could hypothesize a secondary effect where the growth of *Chlorella* would improve, which is not considered in the current state of the model.

Lipid accumulation is linked with limitations in the supply of nitrogen, which is not described in the kinetics of the model. Since there could be upregulation of ACCase in *Chlorella* in nitrogen deplete conditions, it is likely that supplementation of biotin will have a greater effect in the accumulation of lipids (Giridhar Babu et al., 2017). Adding vitamins enhanced lipid accumulation, but to different extents depending if the culture is nitrogen depleted or replete (Fazeli Danesh et al., 2018). The inclusion of the dynamics and kinetic influence of nitrogen is thus one of the most important upgrades expected for the model in the future.

In metabolic models, the biomass reaction is generally built considering an average composition based on experimental measurements. In this case, when modeling the internal composition of the cell, besides modeling the energy expenditure for the growth, the reaction should consider the varying cell stoichiometry. Therefore, including the dynamics of other macromolecules, such as carbohydrates and proteins may be necessary to more accurately represent the dynamics of lipid content.

7.4 Conclusions

This metabolic model is the first, to our knowledge, to include the dynamical influence of biotin in the accumulation of lipids. Additional experiments are now necessary to further validate the model and adapt it to the case of nitrogen limitation. It demonstrates the usefulness of the DRUM framework in modeling the dynamics of complex metabolic interactions, even in the case of a multi-species culture. Following the consolidation of the biological knowledge of how metabolic pathways are regulated, more robust and accurate mathematical models are expected. Here we considered the interaction between two organisms through a particular vitamin, but in nature, this interaction is due to a large palette of chemical compounds. Representing such interactions, which are for most of them still unknown, will be a difficult challenge in the future. Being able to correctly model the metabolic interactions between bacteria and microalgae will make possible not only the optimization of current processes but also open new possibilities and new designs of bio-processes.

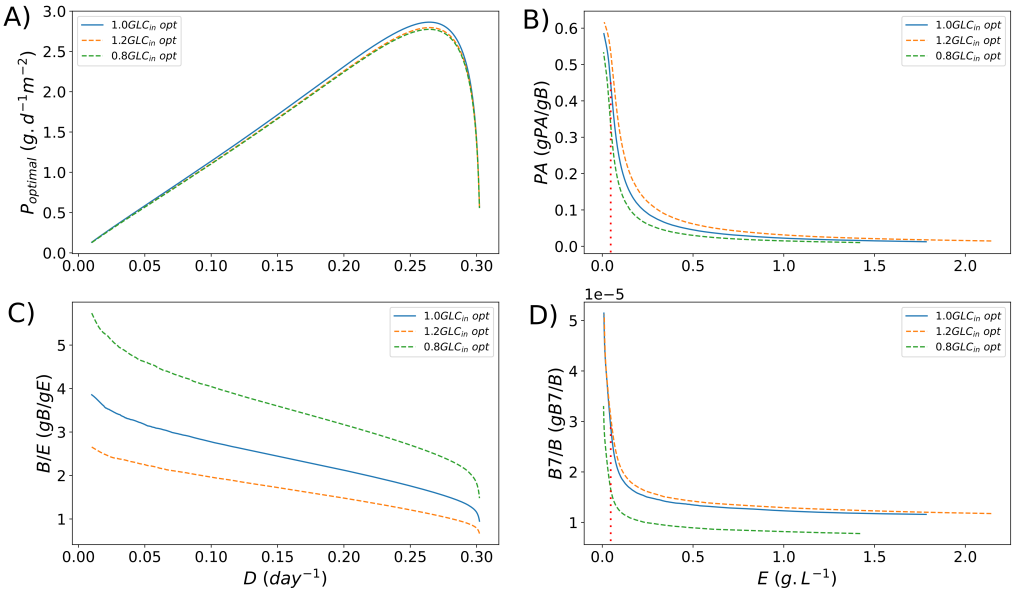

Figure 7.7: Characteristics of the system at steady state at three different influent glucose concentrations, using as a basis the optimal concentration (GLC_{in} *opt*) for a given dilution rate, blue solid curve, and for 80% and 120% of GLC_{in} *opt*, green dashed curve and orange dashed curve, respectively. A: Optimal lipid productivity for a fixed D. B: Lipid content. C: Ratio of *Chlorella* biomass per *E. coli* concentration. D: Biotin concentration per biomass. Red dotted vertical line indicates E at maximal $P_{optimal}$.

Chapter 8

Conclusion and Perspectives

Optimizing the production of microalgae is challenging and requires considering a large variety of processes at different scales. In this PhD, several key aspects impacting productivity were considered, but for which there were so far not so many modeling developments. First, we considered macroscopic effects, namely the overall impact of temperature on microalgae cultivation and then the loss of productivity due to predation. Later, at metabolic scale, the mixotrophic use of organic substrates was studied, and finally the possibility to enhance productivity by associating a bacterial population to the microalgae culture.

In Chapter 3, we analyze how temperature and light intensity affect microalgae production in the context of cultivation under greenhouse. The result of the analysis demonstrates the necessity of adapting the species to climate conditions, especially its thermal niche. Microalgae characteristics must fit the season, but also the process design for large-scale cultivation systems. Using a greenhouse has a positive effect to protect against contamination and predation, but it also significantly shifts the medium temperature, and this trade-off must be carefully considered, especially for most of the species thriving in ecosystems whose temperatures are much lower.

In Chapter 4 we focus on the issue of predation on microalgal cultivations. We proposed a model based on a classical predator-prey formulation to mathematically describe the cultivation of microalgae in the presence of grazers. We determine that optimal productivity takes place when there is extinction of the predators of microalgae, and we proceed to formulate a control strategy to optimize biomass production.

The last Chapters of this thesis consider the metabolic scale. The metabolism was represented using the DRUM approach to account for the dynamics related to storage and reuse in the cell. The goal was to use this dynamical metabolic description to optimize microalgae productivity. First, by constructing a metabolic model for microalgae *Chlorella* growing mixotrophically in the presence of various substrates and especially volatile fatty acids (VFAs). The model is then used to propose a cultivation strategy to treat in minimum time a waste containing butyrate and acetate.

The last two Chapters included the effects of vitamins thiamine and biotin on the metabolism. A metabolic model was proposed including the pathway related to these vitamins. In Chapter 6, by constructing a model of bacteria *E. coli* which depends on an external supply of thiamine to support its growth, we proposed a strategy to maximize the production of lactate via fermen-

tation. The validation of this model using experimental data available in the literature, allows us to use it as a basis for the model in Chapter 7, where vitamin biotin produced by a bacteria influences the lipid composition of microalgae.

The model proposed in Chapter 7 will require experimental validation. First, it will be necessary to set up experiments where the growth and lipid composition of microalgae *Chlorella* is monitored at different external concentrations of biotin. Second, it will be required to verify the influence of cultivation conditions, such as nitrogen supply, since it is expected that the demand for biotin will be higher when nitrogen supply is limited. After confirmation of the influence of the external supply of biotin for *Chlorella*, then the next step will require the co-cultivation of microalgae and the bacteria overproducing biotin. The goal of this experiment, of course, will not be to demonstrate a cultivation system that will alone improve lipid productivity to the levels of economical viability, but this strategy should be incorporated into existing processes.

One of the major desirable improvements for these metabolic models will be the consideration of temperature in the metabolic reactions. As we demonstrated in Chapter 3 temperature deeply impacts the cultivation dynamics. A link between the model at macroscopic scale and metabolic models is therefore expected to more finely represent the effect of temperature on the metabolism dynamics. But in order to limit the number of new parameters to be taken into account when considering temperature, it will be required to develop a modeling strategy, or framework, that will be able to represent mathematically the global effect of temperature on growth, instead of representing the effect for every reaction. The development of this temperature model for metabolic networks would facilitate the calibration of the model and greatly improve accuracy.

Furthermore, as we gather more knowledge about enzymes and their reaction rates, the decision of compartmentalization and which accumulating metabolites should be considered in the DRUM framework will be facilitated. For example, in a recently developed framework, the kinetic properties of hundreds of enzymes of the central metabolic network of a plant were characterized (Küken et al., 2020). This knowledge would make it possible to apply analytical techniques of network reduction based on the reaction rates (López Zazueta et al., 2019). Noticing that the goal of the DRUM framework is to limit the number of variables required to describe the system, in order to facilitate the development and calculation of control strategies. It will be possible to use the knowledge gathered from enzyme reaction rates to construct a reduced metabolic model with the desired accuracy - enough to control and optimize the cultivation system.

Also, the incorporation of new techniques of metabolic network reduction, such as those developed in Küken et al. (2021) could be applied together with the DRUM framework. It would be possible to obtain a simpler model; by applying it before the determination of the subnetworks, facilitating the determination of which reactions belong to the given subnetwork, or after the determination of the subnetworks, facilitating even more the calculation of EFMs.

Metabolic models will allow for an unprecedented level of control and design of bioprocesses and biorefineries. As it has been demonstrated that metabolic models can predict the specific volatile fatty acids (VFAs) produced during fermentation by acidogenic bacteria, as a function of available organic substrates and medium characteristics (Regueira et al., 2020). Incorporating both steps, VFA production by bacteria and consumption by microalgae for biomass would be a clear route to expanding the work done in this thesis. This would allow for the creation

and design strategies of a biorefinery, where production of volatile fatty acids, hydrogen, and also biomass production would be possible. Also, the prediction of the system response must cope with the permanent fluctuations in substrate composition and concentration. This is very important in wastewater treatment, since the influent quality and flux are always varying, on a daily and seasonal scale.

For the next generation of models, it will be necessary to consider together macroscopic effects and metabolic modeling. This global approach, once validated, will support strategies for maximizing productivity. These approaches will be necessary for the expansion of microalgae as an energy source. For example, the analysis done in Chapter 3 using the knowledge about growth rate dependence on light intensity and temperature allows us to anticipate the response to weather and climate, and eventually propose an efficient system design. It enables to design the system and choosing the best locations for the cultivation of a given species, if more than one species should be cultivated during different seasons of the year, and the necessity of cultivating under greenhouse or not.

Afterward, meteorological predictions, together with models for the raceway temperature, can be used as inputs for the metabolic model. Since cellular metabolic fluxes are dependent on the availability of substrates, the prediction of substrates concentration and temperature of the medium might provide information for the way the metabolic model is reduced. This means that, depending on medium conditions, a different reduced metabolic model might be used, choosing the best trade-off between computational time for online process control and model accuracy.

Appendix A

Supplementary Information of Chapter 3

A.1 Extinction coefficient calibration for *Tetraselmis suecica*

In order to calibrate the equation of the extinction coefficient we measured light intensity at the surface and at the bottom of the raceway pond. This procedure was repeated for a variety of biomass concentrations, then we calculate the extinction coefficient by the following equation, obtained from the Beer-Lambert equation:

$$\sigma = -\frac{1}{LX} ln(\frac{I_L}{I_0}) \tag{A.1}$$

where I_L and I_0 are the light intensities measured at the bottom and at the surface of the raceway respectively, and L the depth of water. Then, the extinction coefficient can be calibrated in relation to biomass concentration (see Figure A.1), by the power function:

$$\sigma = AX^B \tag{A.2}$$

A.2 Water temperature model calibration

Figures A.2 and A.3 show the validation of the water temperature model, respectively, for a raceway pond under greenhouse and without greenhouse, as described in the paper.

A.3 Light attenuation of greenhouse film

The value of 20% of light attenuation due to the greenhouse transparent film was determined by plotting 5171 points of measured light intensity close to the raceway under the greenhouse and from another probe outside the greenhouse at the same time. The value of the slope in Figure A.4 is about 0.8.

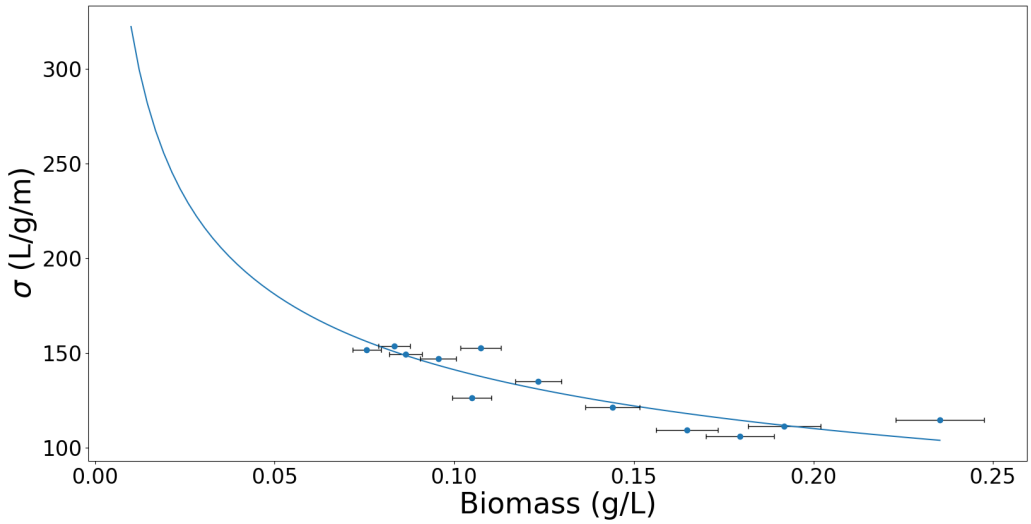

Figure A.1: Data points represent measured extinction coefficients at the raceway pond. Errors bars represent the standard deviation of biomass concentration based on an average coefficient of variation. The blue curve is the fitted power function to the data.

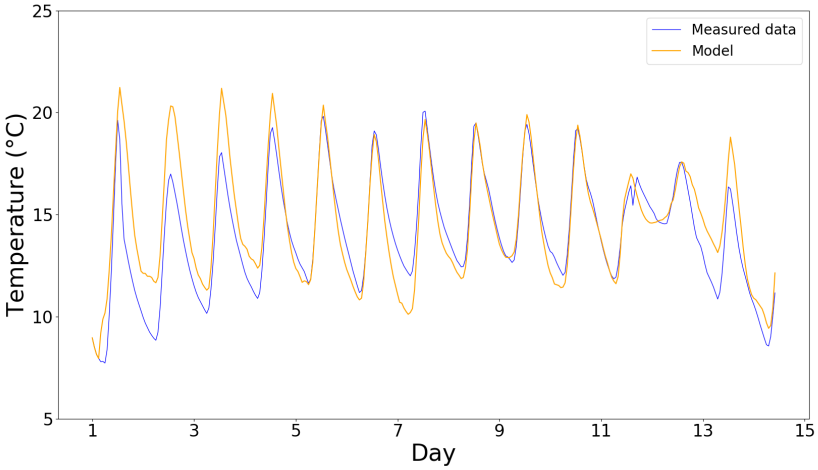

Figure A.2: Raceway water temperature under greenhouse. The blue curve represents the measured water temperature in the raceway and the orange curve represents model predictions.

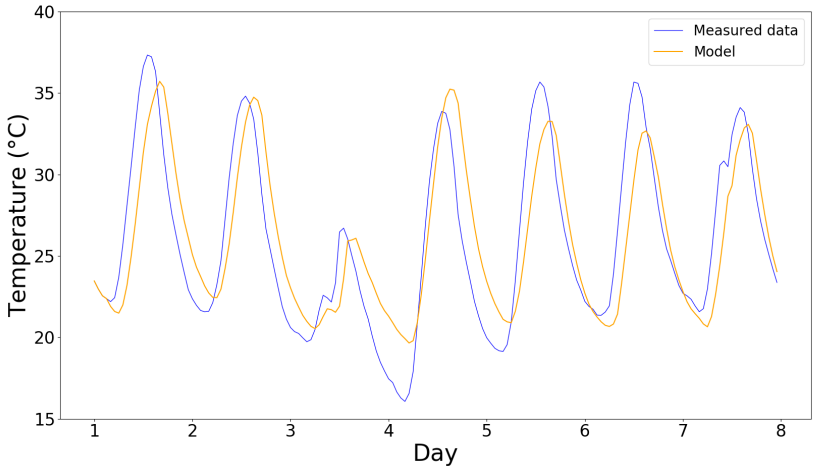

Figure A.3: Raceway water temperature without greenhouse. The blue curve represents the measured water temperature in the raceway and the orange curve represents model predictions.

A.4 Productivity optimization based on average seasonal conditions

The continuous productivity is optimized based on a day representing the average hourly temperature and light intensity for each season and location. To do so, we calculate the biomass concentration, X_{opt}, that optimizes productivity for each of these average days. X_{opt} is calculated given the following control strategy where D is either equal to μ_{net} or zero if $X(t) < X_{opt}$. Average light intensity ($I_{average}$), minimal and maximal water temperature inside or outside greenhouse, for days representing the average conditions for each season and location, can be found at Table A.1.

A.5 Microalgae parameters

Parameters for *Spirulina platensis*, *Phaeodactylum tricornutum*, *Dunaliella salina* and *Chlorella vulgaris* can be found at Table A.2. For *Spirulina* parameters were adapted from Venkataraman (1997), Qiang et al. (1998) and for *Phaeodactylum* from Bernard and Rémond (2012), Fernández et al. (1997), Bitaubé Pérez et al. (2008), Wu et al. (2010). Full model and parameters for *Dunaliella salina* and *Chlorella vulgaris* can be found at Béchet et al. (2017) and Béchet et al. (2015) respectively, where photoinhibition is not considered ($I_{opt} > 1000$).

A.6 Annual simulation

Complete information about yearly productivity for locations near Nice, Paris and Rennes can be found, respectively, at Tables A.3, A.4 and A.5.

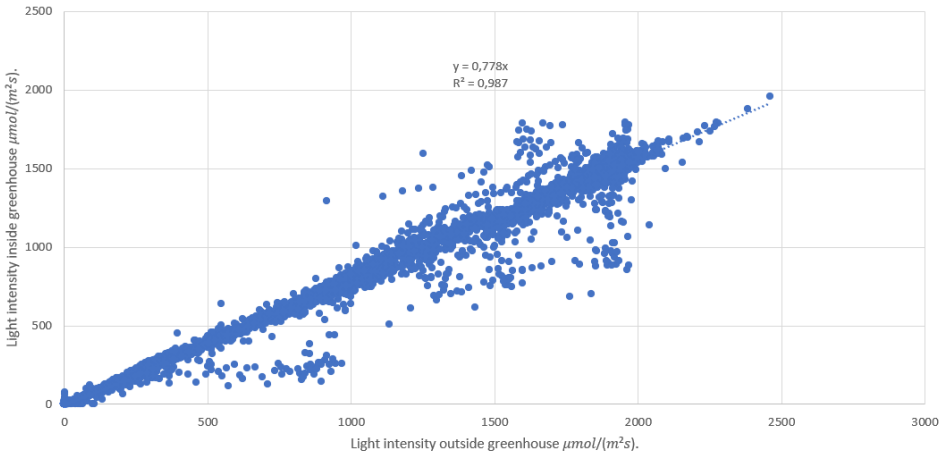

Figure A.4: 5171 points of light intensity measured outside and inside the greenhouse at the same instant. The slope represents the average fraction of light that passes through the transparent film.

Table A.1: Average conditions for each location and season. Temperature in $°C$ and I in $\mu mol/(m^2 s)$.

Location	Season	$I_{average}$	$T_{w_{out}}min$	$T_{w_{out}}max$	$T_{w_{in}}min$	$T_{w_{in}}max$
	Winter	126	3.6	8.4	6.9	13.3
	Spring	283	8.2	18.4	11.2	23.9
Paris	Summer	261	14	27.5	16.9	33.6
	Fall	127	5.8	12.6	9.1	17.9
	Winter	185	7	13.1	10.4	18.2
	Spring	393	12.1	23.3	15.3	29
Nice	Summer	356	18.5	31.2	21.9	37.2
	Fall	184	10.4	17.2	13.9	22.5
	Winter	134	4.7	9.3	8	14.2
	Spring	293	8.7	18.3	11.7	23.8
Rennes	Summer	266	14	25.8	17	31.8
	Fall	134	7	12.7	10.4	17.8

Table A.2: Growth model parameters for *Spirulina platensis*, *Phaeodactylum tricornutum*, *Dunaliella salina* and *Chlorella vulgaris*.

Parameter	Unit	*Spirulina*	*Ph. trico.*	*D. salina*	*Ch. vulga.*
μ_{max}	(d^{-1})	1.15	1.8	13.5	5.6
I_{opt}	$(\mu mol.m^{-2}s^{-1})$	142	357	-	-
α	$(m^2.s.\mu mol^{-1}d^{-1})$	0.06027	0.0933	-	-
B	-	-0.130	-0.149	-0.37	-0.2
A	$(L^{1-B}.g^{B-1}.m^{-1})$	59.694	166	79.1	117.4
T_{min}	(°C)	0	-27.7	-7.8	-12.2
T_{opt}	(°C)	36	22.5	34	37.2
T_{max}	(°C)	50	25.2	43	42
β	(C^{-1})	0.0636	0.0715	0.0715	0.0357
$\lambda_0(day)$	(d^{-1})	0.0093	0.0418	0.0418	0.044
$\lambda_0(night)$	(d^{-1})	0.0118	0.0522	0.0522	0.0544

Table A.3: Nice simulated productivity $(g.m^{-2}.d^{-1})$ at different seasons of the year.

Species		Winter	Spring	Summer	Fall	Year
	Greenhouse	4.5	20.8	5.7	6.2	9.3 ±0.5
Tetraselmis suecica	Outside	4.1	22.9	28.9	6.4	15.6 ±0.8
	Greenhouse T c.	5.3	24.0	26.4	7.3	15.8 ±0.8
	Greenhouse	5.4	21.2	27.2	8.7	15.6 ±0.8
Spirulina platensis	Outside	3.0	16.7	26.3	6.0	13.0 ±0.7
	Greenhouse T c.	5.6	21.5	28.1	8.9	16.0 ±0.8
	Greenhouse	7.2	18.4	20.6	7.9	13.5 ±0.7
Chlorella vulgaris	Outside	7.2	18.9	23.2	8.3	14.4 ±0.8
	Greenhouse T c.	7.4	18.6	21.4	8.2	13.9 ±0.7
	Greenhouse	3.1	10.7	11.1	3.5	7.1 ±0.4
Dunaliella salina	Outside	3.5	12.6	15.2	4.2	8.9 ±0.5
	Greenhouse T c.	3.4	11.0	12.5	3.7	7.7 ±0.4
	Greenhouse	4.2	2.0	0.0	2.4	2.2 ±0.1
Phaeodactylum tricornutum	Outside	4.1	7.9	0.4	4.1	4.1 ±0.2
	Greenhouse T c.	4.6	7.7	0.4	4.4	4.3 ±0.2

Table A.4: Paris simulated productivity ($g.m^{-2}.d^{-1}$) at different seasons of the year.

Species		Winter	Spring	Summer	Fall	Year
	Greenhouse	1.9	13.9	10.5	3.5	7.5 ±0.4
Tetraselmis suecica	Outside	1.6	14.2	20.8	3.4	10.0 ±0.5
	Greenhouse T c.	2.2	15.6	20.3	4.0	10.5 ±0.6
	Greenhouse	2.2	14.7	22.6	4.9	11.1 ±0.6
Spirulina platensis	Outside	0.7	10.3	20.1	2.8	8.5 ±0.5
	Greenhouse T c.	2.3	14.8	23.4	5.0	11.4 ±0.6
	Greenhouse	4.9	14.1	16.5	5.8	10.3 ±0.5
Chlorella vulgaris	Outside	4.8	14.4	18.8	5.9	11.0 ±0.6
	Greenhouse T c.	5.1	14.3	17.7	6.1	10.8 ±0.6
	Greenhouse	1.6	7.1	8.3	2.2	4.8 ±0.3
Dunaliella salina	Outside	1.6	8.1	11.2	2.5	5.8 ±0.3
	Greenhouse T c.	1.8	7.3	9.3	2.4	5.2 ±0.3
	Greenhouse	2.6	4.8	0.3	2.4	2.5 ±0.1
Phaeodactylum tricornutum	Outside	2.4	7.8	4.1	3.3	4.4 ±0.2
	Greenhouse T c.	3.0	8.1	4.1	3.5	4.7 ±0.2

Table A.5: Rennes simulated productivity ($g.m^{-2}.d^{-1}$) at different seasons of the year.

Species		Winter	Spring	Summer	Fall	Year
	Greenhouse	2.1	15.1	13.9	3.6	8.7 ±0.5
Tetraselmis suecica	Outside	1.8	14.7	21.1	3.4	10.3 ±0.5
	Greenhouse T c.	2.5	16.3	20.9	4.1	10.9 ±0.6
	Greenhouse	2.5	14.8	21.9	4.9	11.0 ±0.6
Spirulina platensis	Outside	1.0	10.2	18.5	2.9	8.1 ±0.4
	Greenhouse T c.	2.7	14.9	22.3	5.1	11.3 ±0.6
	Greenhouse	5.1	14.5	16.8	6.0	10.6 ±0.6
Chlorella vulgaris	Outside	5.1	14.7	18.3	6.2	11.1 ±0.6
	Greenhouse T c.	5.4	14.7	17.5	6.2	10.9 ±0.6
	Greenhouse	1.8	7.5	8.7	2.3	5.1 ±0.3
Dunaliella salina	Outside	1.8	8.5	11.0	2.6	6.0 ±0.3
	Greenhouse T c.	2.0	7.8	9.3	2.5	5.4 ±0.3
	Greenhouse	2.8	5.6	0.8	2.7	3.0 ±0.2
Phaeodactylum tricornutum	Outside	2.7	8.5	6.2	3.5	5.2 ±0.3
	Greenhouse T c.	3.1	8.9	6.0	3.8	5.5 ±0.3

Appendix B

Supplementary Information of Chapter 4

B.1 Proof of Proposition .1

To prove Proposition .1, we need the following lemma on the boundedness of solutions of (4.1).

Lemma 4 (Boundedness)**.** Solutions to (4.1) are bounded.

Proof. Let (x, y) be a solution of (4.1) with $x(0), y(0) > 0$ and let $\bar{x}$ be such that

$$\frac{d\bar{x}}{dt} = (\mu(\bar{x}) - D)\bar{x}, \quad \bar{x}(0) = x(0).$$

It is clear that $\bar{x}(t) \leq b := \max\{x(0), x^*\}$ with x^* defined by (4.11). From a comparison argument, it follows that $x(t) \leq \bar{x}(t)$ for all $t \geq 0$, then $x(t)$ is bounded from above by b. Now, let us define the variable $z = \gamma x + y$. Then we have

$$\frac{dz}{dt} = \gamma\mu(x)x - my - Dz + (1 - \alpha)Dy.$$

Since, $y(t) \leq z(t)$ for all $t \geq 0$, we obtain that

$$\frac{dz}{dt} \leq \gamma\mu(0)b - Dz.$$

Then $z(t) \leq b' := \max\{z(0), \gamma\mu(0)b/D\}$ for all $t \geq 0$. It is clear that b' is an upper bound for y which completes the proof. □ □

Proof. (of Proposition .1) For any $D \geq 0$, let us define $x^*(D)$ by means of (4.11). Now define $\varphi(D) = \nu(x^*(D)) - m - \alpha D$. Note that φ is strictly decreasing, and that $\varphi(0) = \nu(M) - m > 0$ and $\varphi(D_{alg}) = -m - \alpha D_{alg} < 0$. Then, there is a unique $D_{coex} \in (0, D_{alg})$ such that $\varphi(D_{coex}) = 0$. For the part (a), assume that $D < D_{coex}$, then there is $x_c \in (0, x^*(D_{coex}))$ such that

$\nu(x_c) = D$. Since μ is strictly decreasing, we have that $\mu(x_c) > \mu(x^*(D_{coex})) = D_{coex} > D$. Consequently,

$$y_c := \gamma \frac{(\mu(x_c) - D)x_c}{\nu(x_c)} > 0.$$

Then the coexistence equilibrium is given by $E_c = (x_c, y_c)$. The uniqueness of E_c follows directly from the monotonicity of ν and μ.

The Jacobian matrix associated with (4.1) is given by

$$\begin{bmatrix} \mu(x) - D + \mu'(x)x - \frac{1}{\gamma}\nu'(x)y & -\frac{1}{\gamma}\nu(x) \\ \nu'(x)y & \nu(x) - m - \alpha D \end{bmatrix}. \tag{B.1}$$

It is straightforward to verify that $E_0 := (0, 0)$ and E^* are saddle points. Using a stable manifold theorem argument Perko (2001), E_0 and E^* can only be reached by solutions starting on $(\{0\} \times \mathbb{R}_+) \cup (\mathbb{R}_+ \times \{0\})$. From Lemma 4, any solution to (4.1) is bounded. Thus, using the Poincaré-Bendixon Theorem, we conclude that any solution starting on the interior of $\mathbb{R}^2_+$ approaches asymptotically either E_c or a periodic cycle. For part (b), by contradiction, let $E_c = (x_c, y_c)$ be a coexistence equilibrium. Then

$$\mu(x_c) - \mu(x^*) = \frac{\nu(x_c)}{\gamma x_c} y_c > 0,$$

from where $x_c < x^*$. Now, since $D \geq D_{coex}$ we have

$$0 = \nu(x_c) - m - \alpha D < \nu(x^*) - m - \alpha D = \varphi(D) \leq 0,$$

which is a contradiction. Then, there is no coexistence equilibrium. Hence, there is no limit cycle. Consequently, any solution with positive initial conditions approaches either E_0 or E^*. Again, since E_0 is a saddle point, using a stable manifold theorem argument, we conclude that E_0 can only be reached by solutions starting on $\{0\} \times \mathbb{R}_+$. Which completes the proof of (b). Finally, if $D > D_{alg}$, E_0 is the unique equilibrium that (4.1) admits. Using again the Poincaré-Bendixon Theorem, we conclude that any solution approaches asymptotically E_0 and the part (c) is proved. □

B.2 Proof of Lemma 2

Proof. The Jacobian matrix associated to (4.1) evaluated at E_c is given by (see (B.1))

$$J(E_c) = \begin{bmatrix} \mu(x_c) - D + \mu'(x_c)x_c - \frac{1}{\gamma}\nu'(x_c)y & -\frac{1}{\gamma}\nu(x_c) \\ \nu'(x_c)y_c & 0 \end{bmatrix}.$$

Then the trace of $J(E_c)$, denoted τ, and the determinant of $J(E_c)$, denoted δ, are given by

$$\begin{aligned} &\tau = \mu(x_c) - D + \mu'(x_c)x_c - \tfrac{1}{\gamma}\nu'(x_c)y \quad \text{and} \\ &\delta = \tfrac{1}{\gamma}x_c y_c f(x_c)\nu'(x_c) > 0. \end{aligned}$$

Thus, if $\tau < 0$, E_c is a sink, if $\tau > 0$, then E_c is a source. If $\tau = 0$, then E_c is a center for the linear system $\frac{d}{dt}(x, y)^T = J(E_c)(x, y)^T$. Then, according to Theorem 5 in Chapter 2.10 in the book of Perko (2001), E_c is either a focus, a center, or a center-focus for (4.1). Choosing appropriately b in Theorem 2.1 in Hwang (1999), we conclude that there are no limit cycles when $\tau = 0$. Hence, E_c is focus, and consequently stable. From Proposition .1, we conclude the global stability of E_c. Finally, it is straightforward to prove that τ and $h'(x_c)$ have the same sign. This completes the proof. □

Appendix C

Supplementary Information of Chapter 5

C.1 List of reactions

Photosynthesis
R1 8 Light + 3 ADP + 3 Pi + H + 2 cNADP → O2 + H2O + 2 cNADPH + 3 ATP
R2 CO2 + H2O + cRu15DP → 2 cG3P
R3 ATP + cG3P → ADP + H + c13DPG
R4 H + cNADPH + c13DPG ↔ cNADP + Pi + cGAP
R5 cGAP ↔ cDHAP
R6 cDHAP + cGAP ↔ cF16P
R7 H2O + cF16P ↔ Pi + cF6P
R8 cF6P + cGAP ↔ cE4P + cX5P
R9 H2O + cE4P + cGAP ↔ Pi + cS7P
R10 cGAP + cS7P ↔ cR5P + cX5P
R11 cX5P ↔ cRu5P
R12 cR5P ↔ cRu5P
R13 ATP + cRu5P → ADP + H + cRu15DP
R14 cGAP ↔ GAP
Glycerol pathway
R15 $GLYC_{ext}$ + NADPH + H ↔ GLYC + NADP
R16 DHAP + H2O ↔ DHA + Pi
R17 DHA + H + NADPH ↔ GLYC + NADP
R18 ATP + GLYC → ADP + GLYC3P + H
R19 DHAP ↔ GAP
Glyoxyzome
R20 BUT + 4 ATP ↔ BUT_g + 4 ADP + 4 Pi
R21 BUTg + *AcCoA*g ↔ BUTylCoA + ACE_g
R22 $BUTylCoA_g$ + *O2*→ CrotonylCoAg + H2O2g
R23 CrotonylCoAg + *H2O* ↔ *3-HydroxyBUTylCoA*g

R24 3-HydroxyBUTylCoAg + *NAD* ↔ *AceAcCoA*g + NADH
R25 AceAcCoA$_g$ + CoA$_g$ ↔ 2 AcCoA$_g$
R26 2 H2O2$_g$→ O2 + 2 H2O
R27 ACE + 0.25 ATP ↔ ACE$_g$ + 0.25ADP + 0.25Pi
R28 ACEg + *ATP* ↔ *ACEP*g + ADP
R29 ACEPg + *CoA*g ↔ AcCoA$_g$ + Pi
R30 AcCoAg+ *H2O* + *OXA*g ↔ CITg + *CoA*g
R31 AcCoAg + *H2O* + *glyoxylate*g ↔ MALg + *CoA*g
R32 ISOg ↔ *SUC*g + glyoxylate$_g$
R33 CITg ↔ *cisAconitate*g + H2O
R34 cisAconitate$g$ + *H2O* ↔ *ISO*g
R35 MALg + *NAD* ↔ *OXA*g + NADH
R36 SUC ↔ SUC$_g$
R37 3.5 H + 2.5 ADP + 2.5 Pi + NADH + 0.5 O2 → NAD + 2.5 ATP + 3.5 H2O
Glycolysis
R38 G6P ↔ G1P
R39 F6P ↔ G6P
R40 ATP + F6P → ADP + F16P + H
R41 F16P + H2O → F6P + Pi
R42 DHAP + GAP ↔ F16P
R43 DHAP ↔ GAP
R44 GAP + NAD + Pi ↔ 13DPG + H + NADH
R45 13DPG + ADP ↔ 3PG + ATP
R46 3PG ↔ 2PG
R47 2PG ↔ H2O + PEP”
R48 ADP + H + PEP ↔ ATP + PYR
Tricarboxylic acid cycle
R49 CoA + NAD + PYR → AcCoA + CO2 + NADH
R50 AcCoA + H2O + OXA ↔ CIT + CoA + H
R51 CIT + NAD ↔ AKG + CO2 + NADH
R52 AKG + CoA + NAD → CO2 + NADH + SUCCoA
R53 ADP + Pi + SUCCoA ↔ ATP + CoA + SUC
R54 FAD + SUC ↔ FADH2 + FUM
R55 FUM + H2O ↔ MAL
R56 FAD + MAL ↔ FADH2 + OXA
R57 ATP + CO2 + H2O + PYR → ADP + OXA + Pi + 2 H
R58 ATP + OXA → ADP + CO2 + PEP
R59 CO2 + H2O + PEP ↔ H + OXA + Pi
Pentose phosphate pathway
R60 G6P + H2O + NADP ↔ 6PG + NADPH + 2 H
R61 6PG + NADP ↔ CO2 + NADPH + RU5P
R62 RU5P ↔ R5P
R63 RU5P ↔ X5P

R64 R5P + X5P ↔ GAP + S7P
R65 GAP + S7P ↔ E4P + F6P
R66 F6P + GAP ↔ E4P + X5P
Glycerol synthesis
R67 GLYC3P + NAD ↔ DHAP + H + NADH
N fixation
R68 H + NADH + NO3 ↔ H2O + NAD + NO2
R69 5 H + 3 NADPH + NO2 ↔ NH4 + 2 H2O + 3 NADP
S fixation
R70 ATP + SO4 → APS + PPi
R71 APS + NADH → AMP + NAD + SO3
R72 5 H + 3 NADPH + SO3 ↔ H2S + 3 H2O + 3 NADP
Oxidative phosphorylation
R73 1.5 ADP + 1.5 H + 1.5 Pi + FADH2 + 0.5 O2 → FAD + 1.5 ATP + 2.5 H2O
R74 3.5 H + 2.5 ADP + 2.5 Pi + NADH + 0.5 O2 → NAD + 2.5 ATP + 3.5 H2O
R75 H2O + PPi → H + 2 Pi
R76 AMP + ATP → 2 ADP
R77 ATP + H2O → ADP + H + Pi + MAINT
Amino acids and protein synthesis
R78 AKG + H + NADPH + NH4 → GLT + H2O + NADP
R79 ATP + GLT + NH4 → ADP + GLN + H + Pi
R80 AKG + GLN + H + NADPH ↔ NADP + 2 GLT
R81 3PG + GLT + H2O + NAD ↔ AKG + H + NADH + Pi + SER
R82 SER → NH4 + PYR
R83 AcCoA + H2S + SER ↔ Ace + CYS + CoA + H
R84 ATP + Ace + CoA → ADP + AcCoA + Pi
R85 GLT + PYR → AKG + ALA
R86 H + THR ↔ 2-oxobutan + NH4
R87 2-oxobutan + GLT + H + NADPH + PYR ↔ AKG + CO2 + H2O + ILE + NADP
R88 2 H + ALA + NADPH + PYR ↔ CO2 + H2O + NADP + VAL
R89 2 PYR + AcCoA + GLT + H + NAD + NADPH ↔ AKG + CoA + LEU + NADH + NADP + 2 CO2
R90 2 PEP + ATP + E4P + NADPH → ADP + CHO + NADP + 4 Pi
R91 CHO ↔ PRE
R92 GLT + H + PRE ↔ AKG + CO2 + H2O + PHE
R93 GLT + NAD + PRE ↔ AKG + CO2 + NADH + TYR
R94 CHO + GLN ↔ ANTH + GLT + H + PYR
R95 ANTH + H + PRPP + SER ↔ CO2 + GAP + PPi + TRYP + 2 H2O
R96 3 H2O + 2 NAD + ATP + GLN + PRPP → AICAR + AKG + HIS + Pi + 2 NADH + 2 PPi + 5 H
R97 GLT + OXA ↔ AKG + ASP
R98 ASP + ATP + GLN + H2O → ADP + ASN + GLT + H + Pi
R99 2 ATP + 2 H2O + CO2 + GLN → CaP + GLT + Pi + 2 ADP + 3 H

R100 2 GLT + ASP + ATP + CaP + NADH → AKG + AMP + ARG + FUM + H2O + NAD + PPi + Pi
R101 3 H + 2 NADH + GLT ↔ PRO + 2 H2O + 2 NAD
R102 AKG + O2 + PRO ↔ CO2 + HydPro + SUC
R103 ASP + ATP + H + NADPH → ADP + ASA + NADP + Pi
R104 2 H + ASA + GLT + NADH + PYR ↔ AKG + DAP + H2O + NAD
R105 DAP ↔ CO2 + H + LYS
R106 ASA + H + NADPH ↔ HSER + NADP
R107 ATP + H2O + HSER → ADP + H + Pi + THR
R108 AcCoA + CYS + H2O + HSER ↔ Ace + CoA + HCYS + H + NH4 + PYR
R109 HCYS + MTHF ↔ H + MET + THF
R110 4.306 ATP + 3.306 H2O + 0.111 ALA + 0.092 GLY + 0.09 LEU + 0.061 VAL + 0.06 LYS + 0.056 PRO + 0.056 THR + 0.054 SER + 0.052 ARG + 0.052 GLN + 0.052 GLT + 0.047 ASN + 0.047 ASP + 0.041 PHE + 0.037 ILE + 0.03 TYR + 0.024 MET + 0.017 HIS + 0.012 CYS + 0.009 HydPro + 0.001 TRYP → PROTEIN + 4.306 ADP + 4.306 Pi + 4.319 H
R111 GLY + H + PYR ↔ ALA + glyoxylate
R112 SER + glyoxylate ↔ GLY + HydPyr
R113 GLY + H2O + METHF ↔ SER + THF
R114 GLY + NAD + THF ↔ CO2 + METHF + NADH + NH4
R115 H + HydPyr + NADH ↔ Glycerate + NAD
R116 ATP + Glycerate → ADP + 2 H + 3PG
THF metabolism
R117 ATP + R5P → AMP + H + PRPP
R118 5FTHF + H ↔ H2O + MYLTHF
R119 H2O + MYLTHF ↔ H + N10FTHF
R120 ATP + FORM + THF → ADP + N10FTHF + Pi
R121 MYLTHF + NADPH ↔ METHF + NADP
R122 H + METHF + NADPH ↔ MTHF + NADP
R123 5FTHF + ATP + H2O → ADP + H + N10FTHF + Pi
R124 FORM + H + THF ↔ H2O + N10FTHF
R125 DHF + H + NADPH ↔ NADP + THF
Lipids synthesis
R126 ACP + AcCoA + H ↔ AcACP + CoA
R127 ATP + AcCoA + CO2 + H2O ↔ ADP + H + MalCoA + Pi
R128 ACP + MalCoA ↔ CoA + MalACP
R129 10 H + 10 NADPH + 5 MalACP + AcACP ↔ C12:0ACP + 5 ACP + 5 CO2 + 5 H2O + 10 NADP
R130 12 H + 12 NADPH + 6 MalACP + AcACP ↔ C14:0ACP + 6 ACP + 6 CO2 + 6 H2O + 12 NADP
R131 14 H + 14 NADPH + 7 MalACP + AcACP ↔ C16:0ACP + 7 ACP + 7 CO2 + 7 H2O + 14 NADP
R132 C16:0ACP + H + NADH + O2 ↔ C16:1ACP + NAD + 2 H2O
R133 C16:1ACP + H + NADH + O2 ↔ C16:2ACP + NAD + 2 H2O

R134 C16:2ACP + H + NADH + O2 ↔ C16:3ACP + NAD + 2 H2O
R135 16 H + 16 NADPH + 8 MalACP + AcACP ↔ C18:0ACP + 8 ACP + 8 CO2 + 8 H2O + 16 NADP
R136 C18:0ACP + H + NADH + O2 ↔ C18:1ACP + NAD + 2 H2O
R137 C18:1ACP + H + NADH + O2 ↔ C18:2ACP + NAD + 2 H2O
R138 C18:2ACP + H + NADH + O2 ↔ C18:3ACP + NAD + 2 H2O
R139 GLYC3P + 0.474 C16:0ACP + 0.446 C18:3ACP + 0.276 C18:2ACP + 0.253 C16:3ACP + 0.16 C18:1ACP + 0.148 C16:2ACP + 0.104 C12:0ACP + 0.051 C14:0ACP + 0.048 C18:0ACP + 0.04 C16:1ACP ↔ PA + 2 ACP + 2 H

Nucleic acids synthesis

R140 4 ATP + 2 GLN + 2 H2O + ASP + CO2 + GLY + N10FTHF + PRPP → AICAR + FUM + PPi + THF + 2 GLT + 4 ADP + 4 Pi + 7 H
R141 ASP + CaP + H + O2 + PRPP ↔ CO2 + H2O + H2O2 + PPi + Pi + UMP
R142 2 H2O2 ↔ O2 + 2 H2O
R143 ATP + UMP → ADP + UDP
R144 ATP + UDP ↔ ADP + UTP
R145 ATP + GLN + H2O + UTP → ADP + CTP + GLT + Pi + 2 H
R146 ATP + CDP ↔ ADP + CTP
R147 AICAR + N10FTHF ↔ H2O + IMP + THF
R148 ATP + H2O + IMP + NAD + NH4 → AMP + GMP + NADH + PPi + 3 H
R149 ATP + GMP → ADP + GDP
R150 ATP + GDP ↔ ADP + GTP
R151 ASP + GTP + IMP ↔ AMP + FUM + GDP + Pi + 2 H
R152 ATP + H + METHF + NADPH + UDP → ADP + DHF + H2O + NADP + dTTP
R153 ATP + CDP + H + NADPH → ADP + H2O + NADP + dCTP
R154 ATP + GDP + H + NADPH → ADP + H2O + NADP + dGTP
R155 ATP + H + NADPH ↔ H2O + NADP + dATP
R156 2.372 H2O + 1.372 ATP + 0.18 dATP + 0.18 dTTP + 0.32 dCTP + 0.32 dGTP → DNA + PPi + 1.372 ADP + 1.372 Pi + 2.372 H
R157 1.4 H2O + 0.56 ATP + 0.34 GTP + 0.16 UTP + 0.34 CTP → 0.4 ADP + 0.4 H + 0.4 Pi + PPi + RNA

Chlorophyll synthesis

R158 12 H + 8 ATP + 8 GLT + 8 NADPH + 2.5 O2 → PPorphyrin + 4 NH4 + 6 CO2 + 8 AMP + 8 NADP + 8 PPi + 13 H2O
R159 18 H + 15 NADPH + 8 ATP + 4 GAP + 4 PYR → Phytyl-PP + 4 ADP + 4 AMP + 4 CO2 + 7 PPi + 8 H2O + 15 NADP
R160 ATP + H2O + MET → AdMET + H + PPi + Pi
R161 AdHCYS + H2O ↔ Ad + HCYS
R162 ATP + Ad → ADP + AMP + H
R163 4 NADPH + 2.5 O2 + 2 ATP + AdMET + Mg2 + PPorphyrin + Phytyl-PP → AdHCYS + Chlorophyll + PPi + 2 ADP + 2 H2O + 2 Pi + 3 H + 4 NADP

Carbohydrate synthesis

R164 G1P ↔ CARB + Pi

Biomass synthesis
R165 5.5595 ATP + 5.5595 H2O + 0.6025 PROTEIN + 0.2641 CARB + 0.0876 PA + 0.0011 DNA + 0.0101 Chlorophyll + 0.0329 RNA → Biomass + 5.5595 H + 5.5595 ADP + 5.5595 Pi
Glucose utilisation
R 166 GLC + ATP → G6P + ADP + H
Transport reactions
R167 Biomass →
R168 ↔ CO2
R169 ↔ O2
R170 ↔ H2O
R171 ↔ Pi
R172 ↔ SO4
R173 ↔ NH4
R174 ↔ Mg2
R175 → Light
R176 ↔ H
R178 → BUT
R179 → ACE
R180 MAINT →
R181 → GLC
R182 → $GLYC_{ext}$
R183 ↔ ATP
R184 ↔ ADP
R185 → GAP
R186 → SUC
R187 ↔ NADPH
R188 ↔ NADP

C.2 List of metabolites

M1	13DPG	1,3-diPhosphoglycerate
M2	2-oxobutan	2-Oxobutanoate
M3	2PG	2-Phosphoglycerate
M4	3-HydroxyBUTylCoA$_g$	3-Hydroxybutylryl-CoEnzyme A in the glyoxysome
M5	3PG	3-Phosphoglycerate
M6	5FTHF	5-Formyl-THF
M7	6PG	6-Phosphogluconate
M8	ACE	Acetate
M9	ACEP$_g$	Acetyl Phosphate in the glyoxysome
M10	ACE$_g$	ACE in the glyoxysome
M11	ACP	Acetyl-Carrier Protein
M12	ADP	Adénosine diphosphate
M13	AICAR	5-Aminoimidazole-4-carboxamide ribonucleine

M14	AKG	2-Oxoglutarate (alpha-ketoglutarate)
M15	ALA	Alanine
M16	AMP	Adenosine monophosphate
M17	ANTH	Anthranilate
M18	APS	Adenylyl sulfate
M19	ARG	Arginine
M20	ASA	L-Aspartic semialdehyde
M21	ASN	Asparagine
M22	ASP	Aspartate
M23	ATP	Adenosine triphosphate
M24	AcACP	Acetyl-ACP
M25	AcCoA	Acetyl-CoA
M26	$AcCoA_g$	Acetyl-CoA of the glyoxysome
M27	Ace	Acetate - intracellular
M28	$AceAcCoA_g$	AcetoAcetyl-CoEnzyme A in the glyoxysome
M29	Ad	Adenosine
M30	AdHCYS	S-Adenosyl-L-homocysteine
M31	AdMET	S-Adenosyl-L-methionine
M32	BUT	Butyrate
M33	BUT_g	Butyrate in the glyoxysome
M34	Biomass	Funtional biomass
M35	$BUTylCoA_g$	Butyryl-CoEnzyme A in the glyoxysome
M36	C12:0ACP	Dodecanoyl-ACP (Lauric acid)
M37	C14:0ACP	Tetradecanoyl-ACP (Myristic acid)
M38	C16:0ACP	Hexadecanoyl-ACP (Palmitic acid)
M39	C16:1ACP	Trans-Hexadec-2-enoyl-ACP (Palmitoleic acid)
M40	C16:2ACP	Hexadecadienoic acid -ACP
M41	C16:3ACP	Hexadecatrienoic acid -ACP
M42	C18:0ACP	Octadecanoyl-ACP (Stearic acid)
M43	C18:1ACP	Cis-11-ocadecanoate-ACP (Oleic acid)
M44	C18:2ACP	Linoleic acid -ACP
M45	C18:3ACP	Alpha-linoleic acid -ACP
M46	CARB	Carbohydrate
M47	CDP	Cytidine diphosphate
M48	CHO	Chorismate
M49	CIT	Citrate
M50	CIT_g	Citrate in the glyoxysome
M51	CO2	Carbon dioxide
M52	CTP	Cytidine triphosphate
M53	CYS	Cysteine
M54	CaP	Carbamoyl phosphate
M55	Chlorophyll	Chlorophyll
M56	CoA	Coenzyme A

M57	CoA_g	Coenzyme A in the glyoxysome
M58	$CrotonylCoA_g$	Crotonul-CoEnzyme A in the glyoxysome
M59	DAP	Diaminopimelate
M60	DHAP	Dihydroxyacetone-P
M61	DHF	Dihydrofolate
M62	DNA	Deoxyribonucleic acid
M63	E4P	Erythrose 4-phosphate
M64	F16P	Fructose 1,6-bisphosphate
M65	F6P	Fructose 6-phosphate
M66	FAD	Flavin adenine dinucleotide oxidized
M67	FADH2	Flavin adenine dinucleotide reduced
M68	FORM	Formic acid
M69	FUM	Fumarate
M70	G1P	Glucose 1-phosphate
M71	G6P	Glucose 6-phosphate
M72	GAP	Glyceraldehyde 3-phosphate
M73	GDP	Guanosine diphosphate
M74	GLN	Glutamine
M75	GLT	Glutamate
M76	GLY	Glycine
M77	GLYC3P	Glycerol 3-phosphate
M78	GMP	Guanosine monophosphate
M79	GTP	Guanosine triphosphate
M80	Glycerate	Glycerate
M81	H	Proton
M82	H2O	Water
M83	H2O2	Hydrogen peroxyde
M84	$H2O2_g$	Hydrogen peroxyde in the glyoxysome
M85	H2S	Hydrogen sulfur
M86	HCYS	Homocysteine
M87	HIS	Histidine
M88	HSER	Homoserine
M89	HydPro	Hydroxyproline
M90	HydPyr	3-Hydroxyproline
M91	ILE	Isoleucine
M92	IMP	Inosine monophosphate
M93	ISO_g	Isocitrate in the glyoxysome
M94	LEU	Leucine
M95	LYS	Lysine
M96	Light	Photons
M97	MAINT	Maintenance term
M98	MAL	Malate
M99	MAL_g	Malate in the glyoxysome

M100	MET	Methionine
M101	METHF	5,10-Methylene-THF
M102	MTHF	Methyl-THL
M103	MYLTHF	5,10-Methenyl-THF
M104	MalACP	Malonyl-ACP
M105	MalCoA	Malonyl-CoA
M106	Mg2	Magnesium
M107	N10FTHF	10-Formyl-THF
M108	NAD	Nicotinamide oxidized
M109	NADH	Nicotinamide reduced
M110	NADP	Nicotinamidephosphate oxidized
M111	NADPH	Nicotinamidephosphate reduced
M112	NH4	Ammonium
M113	NO2	Nitrite
M114	NO3	Nitrate
M115	O2	Oxygen
M116	OXA	Oxaloacetate
M117	OXA_g	Oxaloacetate in the glyoxysome
M118	PA	Phosphatidic acid
M119	PEP	Phosphoenolpyruvate
M120	PHE	Phenylalanine
M121	PPi	Pyrophosphate
M122	PPorphyrin	Protoporphyrine
M123	PRE	Prephanate
M124	PRO	Proline
M125	PROTEIN	Protein
M126	PRPP	Phosphorybosylpyrophosphate
M127	PYR	Pyruvate
M128	Phytyl-PP	Phytyl-diphosphate
M129	Pi	Orthophosphate
M130	R5P	Ribose 5-phosphate
M131	RNA	Ribonucleic acid
M132	RU5P	Ribulose 5-phosphate
M133	S7P	Sedoheptulose 7-phosphate
M134	SER	Serine
M135	SO3	Sulphite
M136	SO4	Sulphate
M137	SUC	Succinate
M138	SUCCoA	Succinyl Coenzyme A
M139	SUC_g	Succinate in the glyoxysome
M140	THF	Tetrahydrofolate
M141	THR	Threonine
M142	TRYP	Tryptophane

M143	TYR	Tyrosine
M144	UDP	Uridine diphosphate
M145	UMP	Uridine monophosphate
M146	UTP	Uridine triphosphate
M147	VAL	Valine
M148	X5P	Xylulose 5-phosphate
M149	c13DPG	chloroplast 13DPG
M150	cDHAP	chloroplast DHAP
M151	cE4P	chloroplast E4P
M152	cF16P	chloroplast F16P
M153	cF6P	chloroplast F6P
M154	cG3P	chloroplast G3P
M155	cGAP	chloroplast GAP
M156	cNADP	chloroplast NADP
M157	cNADPH	chloroplast NADPH
M158	cR5P	chloroplast R5P
M159	cRu15DP	chloroplast Ru15DP
M160	cRu5P	chloroplast Ru5P
M161	cS7P	chloroplast S7P
M162	cX5P	chloroplast X5P
M163	$\text{cisAconitate}_{\text{g}}$	glyoxysome cisAconitate
M164	dATP	Deoxy ATP
M165	dCTP	Deoxy CTP
M166	dGTP	Deoxy GTP
M167	dTTP	Deoxy TTP
M168	glyoxylate	glyoxylate
M169	$\text{glyoxylate}_{\text{g}}$	glyoxysome glyoxylate
M170	GLU	alpha-d-glucose
M171	GLYC	glycerol
M172	GLYC_{ext}	extracellular glycerol
M173	GLYC3P	GLYC3P in glycerol pathway

Bibliography

Arashiro, L.T., Josa, I., Ferrer, I., Van Hulle, S.W.H., Rousseau, D.P.L., Garfí, M., 2022. Life cycle assessment of microalgae systems for wastewater treatment and bioproducts recovery: Natural pigments, biofertilizer and biogas. Science of The Total Environment 847, 157615. doi:10.1016/j.scitotenv.2022.157615.

Ardito, A., Ricciardi, P., 1995. Lyapunov functions for a generalized gause-type model. Journal of Mathematical Biology 33, 816–828. doi:https://doi.org/10.1007/BF00187283.

Assis Pessi, B., Pruvost, E., Talec, A., Sciandra, A., Bernard, O., 2022. Does temperature shift justify microalgae production under greenhouse? Algal Research 61, 102579. doi:10.1016/j.algal.2021.102579.

Azma, M., Mohamed, M.S., Mohamad, R., Rahim, R.A., Ariff, A.B., 2011. Improvement of medium composition for heterotrophic cultivation of green microalgae, Tetraselmis suecica, using response surface methodology. Biochemical Engineering Journal 53, 187–195. doi:10.1016/j.bej.2010.10.010.

Banerjee, S., Ramaswamy, S., 2017. Dynamic process model and economic analysis of microalgae cultivation in open raceway ponds. Algal Res. 26, 330–340. doi:10.1016/j.algal.2017.08.011.

Baroukh, C., Muñoz-Tamayo, R., Bernard, O., Steyer, J.P., 2016. Reply to the Comment on "Mathematical modeling of unicellular microalgae and cyanobacteria metabolism for biofuel production" by Baroukh et al. [Curr. Opin. Biotechnol. 2015, 33:198–205]. Current Opinion in Biotechnology 38, 200–202. doi:10.1016/j.copbio.2016.02.018.

Baroukh, C., Muñoz-Tamayo, R., Steyer, J.P., Bernard, O., 2014. DRUM: A new framework for metabolic modeling under non-balanced growth. Application to the carbon metabolism of unicellular microalgae. PLOS ONE 9, 1–15. doi:10.1371/journal.pone.0104499.

Baroukh, C., Turon, V., Bernard, O., 2017. Dynamic metabolic modeling of heterotrophic and mixotrophic microalgal growth on fermentative wastes. PLOS Computational Biology 13, 1–18. doi:10.1371/journal.pcbi.1005590.

Barsanti, L., Gualtieri, P., 2018. Is exploitation of microalgae economically and energetically sustainable? Algal Research 31, 107–115. doi:10.1016/j.algal.2018.02.001.

Barth, G., Gaillardin, C., 1997. Physiology and genetics of the dimorphic fungus Yarrowia lipolytica. FEMS microbiology reviews 19, 219–237. doi:10.1111/j.1574-6976.1997.tb00299.x.

Batstone, D., Hülsen, T., Oehmen, A., 2019. Metabolic modelling of mixed culture anaerobic microbial processes. Current Opinion in Biotechnology 57, 137–144. doi:10.1016/j.copbio.2019.03.014.

Baudry, G., Delrue, F., Legrand, J., Pruvost, J., Vallée, T., 2017. The challenge of measuring biofuel sustainability: A stakeholder-driven approach applied to the french case. Renew. Sustain. Energy Rev. 69, 933–947. doi:10.1016/j.rser.2016.11.022.

Béchet, Q., Chambonnière, P., Shilton, A., Guizard, G., Guieysse, B., 2015. Algal productivity modeling: A step toward accurate assessments of full-scale algal cultivation. Biotechnol. Bioeng. 112, 987–996. doi:10.1002/bit.25517.

Béchet, Q., Moussion, P., Bernard, O., 2017. Calibration of a productivity model for the microalgae dunaliella salina accounting for light and temperature. Algal Res. 21, 156–160. doi:10.1016/j.algal.2016.11.001.

Béchet, Q., Shilton, A., Guieysse, B., 2013. Modeling the effects of light and temperature on algae growth: State of the art and critical assessment for productivity prediction during outdoor cultivation. Biotechnol. Adv. 31, 1648–1663. doi:10.1016/j.biotechadv.2013.08.014.

Benson, B.C., Rusch, K.A., 2006. Investigation of the light dynamics and their impact on algal growth rate in a hydraulically integrated serial turbidostat algal reactor ({HISTAR}). Aquac. Eng. 35, 122–134. doi:10.1016/j.aquaeng.2005.09.005.

Bernard, O., 2011a. Hurdles and challenges for modelling and control of microalgae for CO2 mitigation and biofuel production. Journal of Process Control 21, 1378–1389. doi:10.1016/j.jprocont.2011.07.012.

Bernard, O., 2011b. Hurdles and challenges for modelling and control of microalgae for CO2 mitigation and biofuel production. Journal of Process Control 21, 1378–1389. doi:https://doi.org/10.1016/j.jprocont.2011.07.012.

Bernard, O., Hadj-Sadok, Z., Dochain, D., Genovesi, A., Steyer, J.P., 2001. Dynamical model development and parameter identification for an anaerobic wastewater treatment process. Biotechnol. Bioeng. 75, 424–438. doi:10.1002/bit.10036.

Bernard, O., Rémond, B., 2012. Validation of a simple model accounting for light and temperature effect on microalgal growth. Bioresource Technology 123, 520–527. doi:10.1016/j.biortech.2012.07.022.

Bhagea, R., Bhoyroo, V., Puchooa, D., 2019. Microalgae: The next best alternative to fossil fuels after biomass. A review. Microbiology Research 10. doi:10.4081/mr.2019.7936.

Bhuiyan, A., Akhter, S., Quddus, M., 2011. Diurnal vertical migration of some cladocerans in relation to the physico-chemical factors in a fish pond. Dhaka University Journal of Biological Sciences 20, 147–154. doi:https://doi.org/10.3329/dujbs.v20i2.8975.

Bitaubé Pérez, E., Caro Pina, I., Pérez Rodríguez, L., 2008. Kinetic model for growth of {Phaeodactylum} tricornutum in intensive culture photobioreactor. Biochem. Eng. J. 40, 520–525. doi:10.1016/j.bej.2008.02.007.

Boecker, S., Espinel-Ríos, S., Bettenbrock, K., Klamt, S., 2022. Enabling anaerobic growth of Escherichia coli on glycerol in defined minimal medium using acetate as redox sink. Metabolic Engineering 73, 50–57. doi:10.1016/j.ymben.2022.05.006.

Bonnans, J., Grelard, V., Martinon, P., 2011. Bocop, the optimal control solver, open source toolbox for optimal control problems. URL http://bocop. org .

Bougaran, G., Le Déan, L., Lukomska, E., Kaas, R., Baron, R., 2003. Transient initial phase in continuous culture of isochrysis galbana affinis tahiti. Aquatic Living Resources 16, 389–394. doi:https://doi.org/10.1016/S0990-7440(03)00053-6.

Burgard, A.P., Vaidyaraman, S., Maranas, C.D., 2001. Minimal Reaction Sets for Escherichia coli Metabolism under Different Growth Requirements and Uptake Environments. Biotechnology Progress 17, 791–797. doi:10.1021/bp0100880.

Burgess, M.G., Ritchie, J., Shapland, J., Pielke, R., 2020. IPCC baseline scenarios have over-projected CO2 emissions and economic growth. Environmental Research Letters 16, 014016. doi:10.1088/1748-9326/abcdd2.

Casagli, F., Bernard, O., 2022. Simulating biotechnological processes affected by meteorology: Application to algae–bacteria systems. Journal of Cleaner Production 377, 134190. doi:10.1016/j.jclepro.2022.134190.

Casagli, F., Rossi, S., Steyer, J.P., Bernard, O., Ficara, E., 2021a. Balancing Microalgae and Nitrifiers for Wastewater Treatment: Can Inorganic Carbon Limitation Cause an Environmental Threat? Environmental Science & Technology 55, 3940–3955. doi:10.1021/acs.est.0c05264.

Casagli, F., Zuccaro, G., Bernard, O., Steyer, J.P., Ficara, E., 2021b. ALBA: A comprehensive growth model to optimize algae-bacteria wastewater treatment in raceway ponds. Water Res. 190, 116734. doi:10.1016/J.WATRES.2020.116734.

Castillo, T., Ramos, D., García-Beltrán, T., Brito-Bazan, M., Galindo, E., 2021. Mixotrophic cultivation of microalgae: An alternative to produce high-value metabolites. Biochemical Engineering Journal 176, 108183. doi:10.1016/j.bej.2021.108183.

Chen, C.Y., Ho, S.H., Liu, C.C., Chang, J.S., 2017a. Enhancing lutein production with Chlorella sorokiniana Mb-1 by optimizing acetate and nitrate concentrations under mixotrophic growth. J. Taiwan Inst. Chem. Eng. 79, 88–96. doi:10.1016/j.jtice.2017.04.020.

Chen, J.H., Chen, C.Y., Chang, J.S., 2017b. Lutein production with wild-type and mutant strains of Chlorella sorokiniana MB-1 under mixotrophic growth. J. Taiwan Inst. Chem. Eng. 79, 66–73. doi:10.1016/j.jtice.2017.04.022.

Chen, Y.H., Walker, T.H., 2011. Biomass and lipid production of heterotrophic microalgae Chlorella protothecoides by using biodiesel-derived crude glycerol. Biotechnol. Lett. 2011 3310 33, 1973–1983. doi:10.1007/S10529-011-0672-Y.

Choi, K.R., Jang, W.D., Yang, D., Cho, J.S., Park, D., Lee, S.Y., 2019. Systems Metabolic Engineering Strategies: Integrating Systems and Synthetic Biology with Metabolic Engineering. Trends in Biotechnology 37, 817–837. doi:10.1016/j.tibtech.2019.01.003.

Cooper, M.B., Kazamia, E., Helliwell, K.E., Kudahl, U.J., Sayer, A., Wheeler, G.L., Smith, A.G., 2019. Cross-exchange of B-vitamins underpins a mutualistic interaction between Ostreococcus tauri and Dinoroseobacter shibae. The ISME Journal 13, 334–345. doi:10.1038/s41396-018-0274-y.

Croft, M.T., Lawrence, A.D., Raux-Deery, E., Warren, M.J., Smith, A.G., 2005. Algae acquire vitamin B12 through a symbiotic relationship with bacteria. Nature 438, 90–93. doi:10.1038/nature04056.

Croft, M.T., Warren, M.J., Smith, A.G., 2006. Algae need their vitamins. Eukaryot. Cell 5, 1175–1183.

Cuevas, D.A., Edirisinghe, J., Henry, C.S., Overbeek, R., O'Connell, T.G., Edwards, R.A., 2016. From DNA to FBA: How to Build Your Own Genome-Scale Metabolic Model. Front. Microbiol. 7, 907. doi:10.3389/FMICB.2016.00907.

Curcuraci, E., Manuguerra, S., Messina, C., Arena, R., Renda, G., Ioannou, T., Amato, V., Hellio, C., Barba, F., Santulli, A., 2022. Culture Conditions Affect Antioxidant Production, Metabolism and Related Biomarkers of the Microalgae Phaeodactylum tricornutum. Antioxidants doi:10.3390/antiox11020411.

Darvehei, P., Bahri, P.A., Moheimani, N.R., 2018. Model development for the growth of microalgae: {A} review. Renew. Sustain. Energy Rev. 97, 233–258. doi:10.1016/j.rser.2018.08.027.

David, L., Nelson, D.L., Cox, M.M., Stiedemann, L., McGlynn Jr, M.E., Fay, M.R., 2000. Lehninger principles of biochemistry .

Davis, M.S., Solbiati, J., Cronan, J.E., 2000. Overproduction of Acetyl-CoA Carboxylase Activity Increases the Rate of Fatty Acid Biosynthesis in Escherichia coli *. Journal of Biological Chemistry 275, 28593–28598. doi:10.1074/jbc.M004756200.

Day, J.G., Gong, Y., Hu, Q., 2017. Microzooplanktonic grazers–a potentially devastating threat to the commercial success of microalgal mass culture. Algal Research 27, 356–365. doi:https://doi.org/10.1016/j.algal.2017.08.024.

De-Luca, R., Bezzo, F., Béchet, Q., Bernard, O., 2019. Meteorological data-based optimal control strategy for microalgae cultivation in open pond systems. Complexity 2019, 4363895. doi:10.1155/2019/4363895.

Decostere, B., Craene, J.D., Hoey, S.V., Vervaeren, H., Nopens, I., Hulle, S.W.H.V., 2016. Validation of a microalgal growth model accounting with inorganic carbon and nutrient kinetics for wastewater treatment. Chem. Eng. J. 285, 189–197. doi:10.1016/j.cej.2015.09.111.

Deruyck, B., Nguyen, K.H.T., Decaestecker, E., Muylaert, K., 2019. Modeling the impact of rotifer contamination on microalgal production in open pond, photobioreactor and thin layer cultivation systems. Algal Research 38, 101398. doi:https://doi.org/10.1016/j.algal.2018.101398.

Dhooge, A., Govaerts, W., Kuznetsov, Y.A., 2003. Matcont: a matlab package for numerical bifurcation analysis of odes. ACM Transactions on Mathematical Software (TOMS) 29, 141–164. doi:https://doi.org/10.1145/779359.779362.

Duffie, J.A., Beckman, W.A., 1991. Solar engineering of thermal processes. A {Wiley}-{Interscience} Publication, Wiley.

Erdrich, P., Steuer, R., Klamt, S., 2015. An algorithm for the reduction of genome-scale metabolic network models to meaningful core models. BMC Systems Biology 9, 48. doi:10.1186/s12918-015-0191-x.

Espinosa-Gonzalez, I., Parashar, A., Bressler, D.C., 2014. Heterotrophic growth and lipid accumulation of Chlorella protothecoides in whey permeate, a dairy by-product stream, for biofuel production. Bioresour. Technol. 155, 170–176. doi:10.1016/J.BIORTECH.2013.12.028.

Fabregas, J., Abalde, J., Herrero, C., Cabezas, B., Veiga, M., 1984. Growth of the marine microalga Tetraselmis suecica in batch cultures with different salinities and nutrient concentrations. Aquaculture 42, 207–215. doi:10.1016/0044-8486(84)90101-7.

Fachet, M., Witte, C., Flassig, R.J., Rihko-Struckmann, L.K., McKie-Krisberg, Z., Polle, J.E.W., Sundmacher, K., 2020. Reconstruction and analysis of a carbon-core metabolic network for Dunaliella salina. BMC Bioinformatics 21, 1. doi:10.1186/s12859-019-3325-0.

Fazeli Danesh, A., Mooij, P., Ebrahimi, S., Kleerebezem, R., van Loosdrecht, M., 2018. Effective role of medium supplementation in microalgal lipid accumulation. Biotechnology and Bioengineering 115, 1152–1160. doi:10.1002/bit.26548.

Fernandes, B.D., Mota, A., Teixeira, J.A., Vicente, A.A., 2015. Continuous cultivation of photosynthetic microorganisms: approaches, applications and future trends. Biotechnology Advances 33, 1228–1245. doi:https://doi.org/10.1016/j.biotechadv.2015.03.004.

Fernández, F.G.A., Camacho, F.G., Pérez, J.A.S., Sevilla, J.M.F., Grima, E.M., 1997. A model for light distribution and average solar irradiance inside outdoor tubular photobioreactors for the microalgal mass culture. Biotechnol. Bioeng. 55, 701–714. doi:10.1002/(SICI)1097-0290(19970905)55:5<701::AID-BIT1>3.0.CO;2-F.

Fernández, I., Acién, F.G., Guzmán, J.L., Berenguel, M., Mendoza, J.L., 2016. Dynamic model of an industrial raceway reactor for microalgae production. Algal Res. 17, 67–78. doi:10.1016/j.algal.2016.04.021.

Flynn, K.J., Kenny, P., Mitra, A., 2017. Minimising losses to predation during microalgae cultivation. Journal of applied phycology 29, 1829–1840. doi:https://doi.org/10.1007/s10811-017-1112-8.

Flynn, K.J., Raven, J.A., 2017. What is the limit for photoautotrophic plankton growth rates? Journal of Plankton Research 39, 13–22. doi:10.1093/plankt/fbw067.

Foreman-Mackey, D., Hogg, D.W., Lang, D., Goodman, J., 2013. Emcee : The MCMC Hammer. Publications of the Astronomical Society of the Pacific 125, 306–312. doi:10.1086/670067.

Fussmann, G.F., Ellner, S.P., Shertzer, K.W., Hairston Jr, N.G., 2000. Crossing the hopf bifurcation in a live predator-prey system. Science 290, 1358–1360. doi:10.1126/science.290.5495.1358.

Gao, F., Han, L., 2012. Implementing the Nelder-Mead simplex algorithm with adaptive parameters. Computational Optimization and Applications 51, 259–277. doi:10.1007/s10589-010-9329-3.

Gao, M., Shi, H., Li, Z., 2009. Chaos in a seasonally and periodically forced phytoplankton–zooplankton system. Nonlinear Analysis: Real World Applications 10, 1643–1650. doi:https://doi.org/10.1016/j.nonrwa.2008.02.005.

Gao, Y., Guo, L., Jin, C., Zhao, Y., Gao, M., She, Z., Wang, G., 2022. Metagenomics and network analysis elucidating the coordination between fermentative bacteria and microalgae in a novel bacterial-algal coupling reactor (BACR) for mariculture wastewater treatment. Water Research 215, 118256. doi:10.1016/j.watres.2022.118256.

García-Camacho, F., Sánchez-Mirón, A., Molina-Grima, E., Camacho-Rubio, F., Merchuck, J.C., 2012. A mechanistic model of photosynthesis in microalgae including photoacclimation dynamics. J. Theor. Biol. 304, 1–15. doi:10.1016/j.jtbi.2012.03.021.

Ghaffar, T., Irshad, M., Anwar, Z., Aqil, T., Zulifqar, Z., Tariq, A., Kamran, M., Ehsan, N., Mehmood, S., 2014. Recent trends in lactic acid biotechnology: A brief review on production to purification. Journal of Radiation Research and Applied Sciences 7, 222–229. doi:10.1016/j.jrras.2014.03.002.

Ghimire, A., Frunzo, L., Pirozzi, F., Trably, E., Escudie, R., Lens, P.N., Esposito, G., 2015. A review on dark fermentative biohydrogen production from organic biomass: Process parameters and use of by-products. Appl. Energy 144, 73–95. doi:10.1016/j.apenergy.2015.01.045.

Giridhar Babu, A., Wu, X., Kabra, A.N., Kim, D.P., 2017. Cultivation of an indigenous Chlorella sorokiniana with phytohormones for biomass and lipid production under N-limitation. Algal Research 23, 178–185. doi:10.1016/j.algal.2017.02.004.

Gleizer, S., Ben-Nissan, R., Bar-On, Y.M., Antonovsky, N., Noor, E., Zohar, Y., Jona, G., Krieger, E., Shamshoum, M., Bar-Even, A., Milo, R., 2019. Conversion of escherichia coli to generate all biomass carbon from CO2. Cell 179, 1255–1263.e12. doi:10.1016/j.cell.2019.11.009.

Gomma, A.E., Lee, S.K., Sun, S.M., Yang, S.H., Chung, G., 2015. Improvement in Oil Production by Increasing Malonyl-CoA and Glycerol-3-Phosphate Pools in Scenedesmus quadricauda. Indian Journal of Microbiology 55, 447–455. doi:10.1007/s12088-015-0546-4.

González-Olivares, E., Meneses-Alcay, H., Gonzalez-Yanez, B., Mena-Lorca, J., Rojas-Palma, A., Ramos-Jiliberto, R., 2011. Multiple stability and uniqueness of the limit cycle in a gause-type predator–prey model considering the Allee effect on prey. Nonlinear Analysis: Real World Applications 12, 2931–2942. doi:https://doi.org/10.1016/j.nonrwa.2011.04.003.

Grimaud, G.M., Mairet, F., Sciandra, A., Bernard, O., 2017. Modeling the temperature effect on the specific growth rate of phytoplankton: A review. Rev Env. Sci Biotechnol 16, 625–645. doi:10.1007/s11157-017-9443-0.

Grognard, F., Akhmetzhanov, A.R., Bernard, O., 2014. Optimal strategies for biomass productivity maximization in a photobioreactor using natural light. Automatica 50, 359–368. doi:https://doi.org/10.1016/j.automatica.2013.11.014.

Gwennaëlle Larvor, Berthomier, L., Chabot, V., Le Pape, B., Pradel, B., Perez, L., 2020. Meteonet, an open reference weather dataset by meteo-france.

Hajinajaf, N., Mehrabadi, A., Tavakoli, O., 2021. Practical strategies to improve harvestable biomass energy yield in microalgal culture: A review. Biomass and Bioenergy 145, 105941. doi:https://doi.org/10.1016/j.biombioe.2020.105941.

Hakalin, N., Paz, A.P., Aranda, D., Moraes, L.M., 2014. Enhancement of cell growth and lipid content of a freshwater microalga Scenedesmus sp. by optimizing nitrogen, phosphorus and vitamin concentrations for biodiesel production. doi:10.4236/NS.2014.612095.

Hameri, T., Fengos, G., Hatzimanikatis, V., 2021. The effects of model complexity and size on metabolic flux distribution and control: Case study in Escherichia coli. BMC Bioinformatics 22, 1–25. doi:10.1186/s12859-021-04066-y.

Harmand, J., Lobry, C., Rapaport, A., Sari, T., 2019. Optimal Control in Bioprocesses: Pontryagin's Maximum Principle in Practice. volume 3 of *Chemical Engineering Series - Chemostat and Bioprocesses*. Wiley.

Hartline, C.J., Schmitz, A.C., Han, Y., Zhang, F., 2021. Dynamic control in metabolic engineering: Theories, tools, and applications. Metabolic Engineering 63, 126–140. doi:10.1016/j.ymben.2020.08.015.

Hassika, P., Berbigier, P., Bonnefond, J.M., 1997. Measurement and modelling of the photosynthetically active radiation transmitted in a canopy of maritime pine. Ann. For. Sci. 54, 715–730. doi:10.1051/forest:19970803.

Havlik, I., Lindner, P., Scheper, T., Reardon, K.F., 2013. On-line monitoring of large cultivations of microalgae and cyanobacteria. Trends in biotechnology 31, 406–414.

Herman, J., Usher, W., 2017. SALib: An open-source Python library for Sensitivity Analysis. The Journal of Open Source Software 2. doi:10.21105/joss.00097.

Hesaaraki, M., Moghadas, S., 2001. Existence of limit cycles for predator–prey systems with a class of functional responses. Ecological Modelling 142, 1–9. doi:https://doi.org/10.1016/S0304-3800(00)00442-7.

Höök, M., Tang, X., 2013. Depletion of fossil fuels and anthropogenic climate change—A review. Energy Policy 52, 797–809. doi:10.1016/j.enpol.2012.10.046.

Hu, B., Min, M., Zhou, W., Du, Z., Mohr, M., Chen, P., Zhu, J., Cheng, Y., Liu, Y., Ruan, R., 2012. Enhanced mixotrophic growth of microalga Chlorella sp. on pretreated swine manure for simultaneous biofuel feedstock production and nutrient removal. Bioresour. Technol. 126, 71–79. doi:10.1016/j.biortech.2012.09.031.

Hu, Y., Zhao, N., Gan, T., Duan, J., Yu, H.J., Meng, D., Liu, J., Liu, W., 2017. Analytic Method on Characteristic Parameters of Bacteria in Water by Multiwavelength Transmission Spectroscopy. Journal of Spectroscopy 2017, e4039048. doi:10.1155/2017/4039048.

Huerlimann, R., Heimann, K., 2013. Comprehensive guide to acetyl-carboxylases in algae. Critical Reviews in Biotechnology 33, 49–65. doi:10.3109/07388551.2012.668671.

Huisman, J., Matthijs, H.C., Visser, P.M., Balke, H., Sigon, C.A., Passarge, J., Weissing, F.J., Mur, L.R., 2002. Principles of the light-limited chemostat: theory and ecological applications. Antonie van Leeuwenhoek 81, 117–133. doi:https://doi.org/10.1023/A:1020537928216.

Hwang, T.W., 1999. Uniqueness of the limit cycle for Gause-type predator–prey systems. Journal of Mathematical Analysis and Applications 238, 179–195. doi:https://doi.org/10.1006/jmaa.1999.6520.

Jeschek, M., Bahls, M.O., Schneider, V., Marlière, P., Ward, T.R., Panke, S., 2017. Biotin-independent strains of Escherichia coli for enhanced streptavidin production. Metabolic Engineering 40, 33–40. doi:10.1016/j.ymben.2016.12.013.

Jevremović, D., Boley, D., 2013. Finding minimal generating set for metabolic network with reversible pathways. Bio Systems 112, 31–36. doi:10.1016/j.biosystems.2013.02.003.

Kaleta, C., de Figueiredo, L.F., Schuster, S., 2009. Can the whole be less than the sum of its parts? Pathway analysis in genome-scale metabolic networks using elementary flux patterns. Genome Research 19, 1872–1883. doi:10.1101/gr.090639.108.

Kao, P.H., Ng, I.S., 2017. CRISPRi mediated phosphoenolpyruvate carboxylase regulation to enhance the production of lipid in Chlamydomonas reinhardtii. Bioresource Technology 245, 1527–1537. doi:10.1016/j.biortech.2017.04.111.

Kazamia, E., Czesnick, H., Nguyen, T.T.V., Croft, M.T., Sherwood, E., Sasso, S., Hodson, S.J., Warren, M.J., Smith, A.G., 2012. Mutualistic interactions between vitamin B12-dependent algae and heterotrophic bacteria exhibit regulation. Environ. Microbiol. 14, 1466–1476. doi:10.1111/J.1462-2920.2012.02733.X.

Ketheesan, B., Nirmalakhandan, N., 2013. Modeling microalgal growth in an airlift-driven raceway reactor. Bioresour. Technol. 136, 689–696. doi:10.1016/j.biortech.2013.02.028.

Kim, D., Kim, E.K., Koh, H.G., Kim, K., Han, J.I., Chang, Y.K., 2017a. Selective removal of rotifers in microalgae cultivation using hydrodynamic cavitation. Algal research 28, 24–29. doi:https://doi.org/10.1016/j.algal.2017.09.026.

Kim, W.J., Kim, H.U., Lee, S.Y., 2017b. Current state and applications of microbial genome-scale metabolic models. Current Opinion in Systems Biology 2, 10–18. doi:10.1016/j.coisb.2017.03.001.

Klamt, S., Stelling, J., 2003. Two approaches for metabolic pathway analysis? Trends Biotechnol. 21, 64–69. doi:10.1016/S0167-7799(02)00034-3.

Knoop, H., Zilliges, Y., Lockau, W., Steuer, R., 2010. The Metabolic Network of Synechocystis sp. PCC 6803: Systemic Properties of Autotrophic Growth. Plant Physiology 154, 410–422. doi:10.1104/pp.110.157198.

Küken, A., Gennermann, K., Nikoloski, Z., 2020. Characterization of maximal enzyme catalytic rates in central metabolism of Arabidopsis thaliana. The Plant Journal: For Cell and Molecular Biology 103, 2168–2177. doi:10.1111/tpj.14890.

Küken, A., Wendering, P., Langary, D., Nikoloski, Z., 2021. A structural property for reduction of biochemical networks. Scientific Reports 11, 17415. doi:10.1038/s41598-021-96835-1.

Lacour, T., Sciandra, A., Talec, A., Mayzaud, P., Bernard, O., 2012. Diel variations of carbohydrates and neutral lipids in nitrogen-sufficient and nitrogen-starved cyclostat cultures of isochrysis sp. 1. Journal of Phycology 48, 966–975.

Lacroux, J., Jouannais, P., Atteia, A., Bonnafous, A., Trably, E., Steyer, J.P., van Lis, R., 2022. Microalgae screening for heterotrophic and mixotrophic growth on butyrate. Algal Research 67, 102843. doi:10.1016/j.algal.2022.102843.

Lacroux, J., Trably, E., Bernet, N., Steyer, J.P., van Lis, R., 2020. Mixotrophic growth of microalgae on volatile fatty acids is determined by their undissociated form. Algal Res. 47, 101870. doi:10.1016/j.algal.2020.101870.

Lee, T.M., Lin, J.Y., Tsai, T.H., Yang, R.Y., Ng, I.S., 2023. Clustered regularly interspaced short palindromic repeats (CRISPR) technology and genetic engineering strategies for microalgae towards carbon neutrality: A critical review. Bioresource Technology 368, 128350. doi:10.1016/j.biortech.2022.128350.

León-Vaz, A., Léon, R., Díaz-Santos, E., Vigara, J., Raposo, S., 2019. Using agro-industrial wastes for mixotrophic growth and lipids production by the green microalga Chlorella sorokiniana. New biotechnology 51, 31–38. doi:10.1016/j.nbt.2019.02.001.

Li, J., Han, D., Wang, D., Ning, K., Jia, J., Wei, L., Jing, X., Huang, S., Chen, J., Li, Y., Hu, Q., Xu, J., 2014a. Choreography of Transcriptomes and Lipidomes of Nannochloropsis Reveals the Mechanisms of Oil Synthesis in Microalgae. The Plant Cell 26, 1645–1665. doi:10.1105/tpc.113.121418.

Li, T., Zheng, Y., Yu, L., Chen, S., 2013. High productivity cultivation of a heat-resistant microalga Chlorella sorokiniana for biofuel production. Bioresour. Technol. 131, 60–67. doi:10.1016/j.biortech.2012.11.121.

Li, T., Zheng, Y., Yu, L., Chen, S., 2014b. Mixotrophic cultivation of a Chlorella sorokiniana strain for enhanced biomass and lipid production. Biomass and Bioenergy 66, 204–213. doi:10.1016/j.biombioe.2014.04.010.

Liang, Y., Sarkany, N., Cui, Y., 2009. Biomass and lipid productivities of Chlorella vulgaris under autotrophic, heterotrophic and mixotrophic growth conditions. Biotechnol. Lett. 2009 317 31, 1043–1049. doi:10.1007/S10529-009-9975-7.

Lin, S., Hanson, R.E., Cronan, J.E., 2010. Biotin synthesis begins by hijacking the fatty acid synthetic pathway. Nature Chemical Biology 6, 682–688. doi:10.1038/nchembio.420.

López Zazueta, C., Bernard, O., Gouzé, J.L., 2019. Dynamical reduction of linearized metabolic networks through quasi steady state approximation. AIChE Journal 65, 18–31. doi:10.1002/aic.16406.

Lotz, K., Hartmann, A., Grafahrend-Belau, E., Schreiber, F., Junker, B.H., 2014. Elementary flux modes, flux balance analysis, and their application to plant metabolism. Methods in Molecular Biology (Clifton, N.J.) 1083, 231–252. doi:10.1007/978-1-62703-661-0_14.

Ma, M., Yuan, D., He, Y., Park, M., Gong, Y., Hu, Q., 2017. Effective control of *Poterioochromonas malhamensis* in pilot-scale culture of *Chlorella sorokiniana* gt-1 by maintaining co2-mediated low culture pH. Algal Research 26, 436–444. doi:https://doi.org/10.1016/j.algal.2017.06.023.

Ma, X., Zheng, H., Addy, M., Anderson, E., Liu, Y., Chen, P., Ruan, R., 2016. Cultivation of Chlorella vulgaris in wastewater with waste glycerol: Strategies for improving nutrients removal and enhancing lipid production. Bioresour. Technol. 207, 252–261. doi:10.1016/j.biortech.2016.02.013.

Magdouli, S., Guedri, T., Rouissi, T., Brar, S.K., Blais, J.F., 2020. Sync between leucine, biotin and citric acid to improve lipid production by Yarrowia lipolytica on crude glycerol-based media. Biomass and Bioenergy 142, 105764. doi:10.1016/j.biombioe.2020.105764.

Mairet, F., Muñoz-Tamayo, R., Bernard, O., 2015. Adaptive control of light attenuation for optimizing microalgae production. Journal of Process Control 30, 117–124. doi:https://doi.org/10.1016/j.jprocont.2015.03.010.

Maltsev, Y., Maltseva, K., Kulikovskiy, M., Maltseva, S., 2021. Influence of light conditions on microalgae growth and content of lipids, carotenoids, and fatty acid composition. Biology 10, 1060.

Manhaeghe, D., Blomme, T., Van Hulle, S.W., Rousseau, D.P., 2020. Experimental assessment and mathematical modelling of the growth of Chlorella vulgaris under photoautotrophic, heterotrophic and mixotrophic conditions. Water Res. 184, 116152. doi:10.1016/J.WATRES.2020.116152.

Marrs, C., Lonsdale, D., 2021. Hiding in Plain Sight: Modern Thiamine Deficiency. Cells 10, 2595. doi:10.3390/cells10102595.

Martínez, C., Assis Pessi, B., Bernard, O., 2021. Dynamics and productivity of microalgae in presence of predators. IFAC-PapersOnLine 54, 673–678. doi:10.1016/j.ifacol.2021.08.319.

Martínez, C., Bernard, O., Mairet, F., 2017. Maximizing microalgae productivity by shading outdoor cultures. IFAC-PapersOnLine 50, 8734–8739. doi:10.1016/j.ifacol.2017.08.1725.

Martínez, C., Bernard, O., Mairet, F., 2018. Maximizing microalgae productivity in a light-limited chemostat. IFAC-PapersOnLine 51, 735–740. doi:https://doi.org/10.1016/j.ifacol.2018.04.001.

Martínez, C., Mairet, F., Bernard, O., 2018. Theory of turbid microalgae cultures. J. Theor. Biol. 456, 190–200. doi:10.1016/j.jtbi.2018.07.016.

Martínez, C., Mairet, F., Bernard, O., 2020. Dynamics of the periodically forced light-limited droop model. Journal of Differential Equations doi:ttps://doi.org/10.1016/j.jde.2020.03.020.

Martínez, F., Orús, M.I., 1991. Interactions between Glucose and Inorganic Carbon Metabolism in Chlorella vulgaris Strain UAM 101. Plant Physiol. 95, 1150–1155. doi:10.1104/PP.95.4.1150.

Mata, T.M., Martins, A.A., Caetano, N.S., 2010. Microalgae for biodiesel production and other applications: {A} review. Renew. Sustain. Energy Rev. 14, 217–232. doi:10.1016/j.rser.2009.07.020.

Maton, M., Bogaerts, P., Vande Wouwer, A., 2022. A systematic elementary flux mode selection procedure for deriving macroscopic bioreaction models from metabolic networks. Journal of Process Control 118, 170–184. doi:10.1016/j.jprocont.2022.09.002.

Mira, P., Yeh, P., Hall, B.G., 2022. Estimating microbial population data from optical density. PloS One 17, e0276040. doi:10.1371/journal.pone.0276040.

Mobin, S., Alam, F., 2017. Some Promising Microalgal Species for Commercial Applications: A review. Energy Procedia 110, 510–517. doi:10.1016/j.egypro.2017.03.177.

Moghadas, S., Corbett, B., 2008. Limit cycles in a generalized gause-type predator–prey model. Chaos, Solitons & Fractals 37, 1343–1355. doi:https://doi.org/10.1016/j.chaos.2006.10.017.

Molina, D., de Carvalho, J.C., Júnior, A.I.M., Faulds, C., Bertrand, E., Soccol, C.R., 2019. Biological contamination and its chemical control in microalgal mass cultures. Applied Microbiology and Biotechnology 103, 9345–9358. doi:https://doi.org/10.1007/s00253-019-10193-7.

Molina, E., Martínez, M.E., Sánchez, S., García, F., Contreras, A., 1991. Growth and biochemical composition with emphasis on the fatty acids of {Tetraselmis} sp. Appl Microbiol Biotechnol 36, 21–25. doi:10.1007/BF00164692.

Molina-Grima, E., García-Camacho, F., Acién-Fernández, F.G., Sánchez-Mirón, A., Plouviez, M., Shene, C., Chisti, Y., 2021. Pathogens and predators impacting commercial production of microalgae and cyanobacteria. Biotechnology Advances , 107884doi:https://doi.org/10.1016/j.biotechadv.2021.107884.

Montemezzani, V., Duggan, I.C., Hogg, I.D., Craggs, R.J., 2015. A review of potential methods for zooplankton control in wastewater treatment High Rate Algal Ponds and algal production raceways. Algal Research 11, 211–226. doi:https://doi.org/10.1016/j.algal.2015.06.024.

Montemezzani, V., Duggan, I.C., Hogg, I.D., Craggs, R.J., 2016. Zooplankton community influence on seasonal performance and microalgal dominance in wastewater treatment High Rate Algal Ponds. Algal Research 17, 168–184. doi:https://doi.org/10.1016/j.algal.2016.04.014.

Morales, M., Collet, P., Lardon, L., Hélias, A., Steyer, J.P., Bernard, O., 2019. Chapter 20 - Life-cycle assessment of microalgal-based biofuel, in: Pandey, A., Chang, J.S., Soccol, C.R., Lee, D.J., Chisti, Y. (Eds.), Biofuels from Algae (Second Edition). Elsevier. Biomass, Biofuels, Biochemicals, pp. 507–550. doi:10.1016/B978-0-444-64192-2.00020-2.

Morel, A., 1988. Optical modeling of the upper ocean in relation to its biogenous matter content (case {I} waters). J. Geophys. Res. Ocean. 93, 10749–10768. doi:10.1029/JC093iC09p10749.

Morris, M.D., 1991. Factorial Sampling Plans for Preliminary Computational Experiments. Technometrics 33, 161–174. doi:10.2307/1269043.

Muñoz-Tamayo, R., Mairet, F., Bernard, O., 2013. Optimizing microalgal production in raceway systems. Biotechnology Progress 29, 543–552. doi:https://doi.org/10.1002/btpr.1699.

Needham, T., 1993. A visual explanation of Jensen's inequality. The American Mathematical Monthly 100, 768–771. doi:https://doi.org/10.1080/00029890.1993.11990484.

Oddsdóttir, H.Æ., Hagrot, E., Chotteau, V., Forsgren, A., 2015. On dynamically generating relevant elementary flux modes in a metabolic network using optimization. Journal of Mathematical Biology 71, 903–920. doi:10.1007/s00285-014-0844-1.

Ogbonna, J.C., Masui, H., Tanaka, H., 1997. Sequential heterotrophic/autotrophic cultivation–an efficient method of producing Chlorella biomass for health food and animal feed. J. Appl. Phycol. 9, 359–366. doi:10.1023/A:1007981930676.

O'Grady, J., Morgan, J.A., 2011. Heterotrophic growth and lipid production of Chlorella protothecoides on glycerol. Bioprocess and Biosystems Engineering 34, 121–125. doi:10.1007/s00449-010-0474-y.

Orth, J.D., Conrad, T.M., Na, J., Lerman, J.A., Nam, H., Feist, A.M., Palsson, B.Ø., 2011. A comprehensive genome-scale reconstruction of Escherichia coli metabolism—2011. Molecular Systems Biology 7, 535. doi:10.1038/msb.2011.65.

Orth, J.D., Fleming, R.M.T., Palsson, B.Ø., 2010a. Reconstruction and Use of Microbial Metabolic Networks: The Core Escherichia coli Metabolic Model as an Educational Guide. EcoSal Plus 4. doi:10.1128/ECOSALPLUS.10.2.1.

Orth, J.D., Thiele, I., Palsson, B.Ø., 2010b. What is flux balance analysis? Nat. Biotechnol. 28, 245–248. doi:10.1038/nbt.1614.

Palmer, M.D., Harris, G.R., Gregory, J.M., 2018. Extending CMIP5 projections of global mean temperature change and sea level rise due to thermal expansion using a physically-based emulator. Environmental Research Letters 13, 084003. doi:10.1088/1748-9326/aad2e4.

Patel, A.K., Joun, J.M., Hong, M.E., Sim, S.J., 2019. Effect of light conditions on mixotrophic cultivation of green microalgae. Bioresour. Technol. 282, 245–253. doi:10.1016/J.BIORTECH.2019.03.024.

Perko, L., 2001. Differential equations and dynamical systems. volume 7. Springer Science & Business Media. doi:https://doi.org/10.1007/978-1-4613-0003-8.

Pfeiffer, T., Sánchez-Valdenebro, I., Nuño, J.C., Montero, F., Schuster, S., 1999. METATOOL: For studying metabolic networks. Bioinformatics (Oxford, England) 15, 251–257. doi:10.1093/bioinformatics/15.3.251.

Pielke, R., Ritchie, J., 2021. Distorting the view of our climate future: The misuse and abuse of climate pathways and scenarios. Energy Research & Social Science 72, 101890. doi:10.1016/j.erss.2020.101890.

Provost, A., Bastin, G., 2004. Dynamic metabolic modelling under the balanced growth condition. J. Process Control 14, 717–728. doi:10.1016/j.jprocont.2003.12.004.

Qiang, H., Richmond, A., 1996. Productivity and photosynthetic efficiency of *Spirulina platensis* as affected by light intensity, algal density and rate of mixing in a flat plate photobioreactor. Journal of Applied Phycology 8, 139–145. doi:https://doi.org/10.1007/BF02186317.

Qiang, H., Zarmi, Y., Richmond, A., 1998. Combined effects of light intensity, light-path and culture density on output rate of {Spirulina} platensis ({Cyanobacteria}). Eur. J. Phycol. 33, 165–171. doi:10.1080/09670269810001736663.

Rafrafi, Y., Trably, E., Hamelin, J., Latrille, E., Meynial-Salles, I., Benomar, S., Giudici-Orticoni, M.T., Steyer, J.P., 2013. Sub-dominant bacteria as keystone species in microbial communities producing bio-hydrogen. Int. J. Hydrogen Energy 38, 4975–4985. doi:10.1016/j.ijhydene.2013.02.008.

Ramanan, R., Kim, B.H., Cho, D.H., Oh, H.M., Kim, H.S., 2016. Algae–bacteria interactions: Evolution, ecology and emerging applications. Biotechnology Advances 34, 14–29. doi:10.1016/j.biotechadv.2015.12.003.

Ranganathan, P., Amal, J.C., Savithri, S., Haridas, A., 2017. Experimental and modelling of {Arthrospira} platensis cultivation in open raceway ponds. Bioresour. Technol. 242, 197–205. doi:10.1016/j.biortech.2017.03.150.

Regueira, A., Bevilacqua, R., Lema, J.M., Carballa, M., Mauricio-Iglesias, M., 2020. A metabolic model for targeted volatile fatty acids production by cofermentation of carbohydrates and proteins. Bioresource Technology 298, 122535. doi:10.1016/j.biortech.2019.122535.

Rismani-Yazdi, H., Haznedaroglu, B.Z., Hsin, C., Peccia, J., 2012. Transcriptomic analysis of the oleaginous microalga Neochloris oleoabundans reveals metabolic insights into triacylglyceride accumulation. Biotechnology for Biofuels 5, 74. doi:10.1186/1754-6834-5-74.

Ritchie, J., Dowlatabadi, H., 2017. Why do climate change scenarios return to coal? Energy 140, 1276–1291. doi:10.1016/j.energy.2017.08.083.

Rizwan, M., Mujtaba, G., Memon, S.A., Lee, K., Rashid, N., 2018. Exploring the potential of microalgae for new biotechnology applications and beyond: A review. Renewable and Sustainable Energy Reviews 92, 394–404. doi:https://doi.org/10.1016/j.rser.2018.04.034.

Röhl, A., Bockmayr, A., 2017. A mixed-integer linear programming approach to the reduction of genome-scale metabolic networks. BMC Bioinformatics 18, 2. doi:10.1186/s12859-016-1412-z.

Röhl, A., Bockmayr, A., 2019. Finding MEMo: Minimum sets of elementary flux modes. Journal of Mathematical Biology 79, 1749–1777. doi:10.1007/s00285-019-01409-5.

Rosenzweig, M.L., 1971. Paradox of enrichment: destabilization of exploitation ecosystems in ecological time. Science 171, 385–387. doi:10.1126/science.171.3969.385.

Saa, P.A., Nielsen, L.K., 2017. Formulation, construction and analysis of kinetic models of metabolism: A review of modelling frameworks. Biotechnology Advances 35, 981–1003. doi:10.1016/j.biotechadv.2017.09.005.

Sajjadi, B., Chen, W.Y., Raman, A.A.A., Ibrahim, S., 2018. Microalgae lipid and biomass for biofuel production: A comprehensive review on lipid enhancement strategies and their effects on fatty acid composition. Renewable and Sustainable Energy Reviews 97, 200–232. doi:10.1016/j.rser.2018.07.050.

Schade, S., Meier, T., 2019. A comparative analysis of the environmental impacts of cultivating microalgae in different production systems and climatic zones: {A} systematic review and meta-analysis. Algal Res. 40, 101485. doi:10.1016/j.algal.2019.101485.

Schlüter, M., Groeneweg, J., 1981. Mass production of freshwater rotifers on liquid wastes: I. The influence of some environmental factors on population growth of *Brachionus rubens* Ehrenberg 1838. Aquaculture 25, 17–24. doi:https://doi.org/10.1016/0044-8486(81)90095-8.

Schuster, S., Dandekar, T., Fell, D.A., 1999. Detection of elementary flux modes in biochemical networks: A promising tool for pathway analysis and metabolic engineering. Trends Biotechnol. 17, 53–60. doi:10.1016/S0167-7799(98)01290-6.

Sen, R., Martin, G.J.O., 2018. Glycerol and nitrate utilisation by marine microalgae Nannochloropsis salina and Chlorella sp. and associated bacteria during mixotrophic and heterotrophic growth. Algal Res. 33, 298–309. doi:10.1016/j.algal.2018.06.002.

Serra-Maia, R., Bernard, O., Gonçalves, A., Bensalem, S., Lopes, F., 2016. Influence of temperature on {Chlorella} vulgaris growth and mortality rates in a photobioreactor. Algal Res. 18, 352–359. doi:10.1016/j.algal.2016.06.016.

Sheehan, J., Dunahay, T., Benemann, J., Roessler, P., 1998. Look Back at the U.S. Department of Energy's Aquatic Species Program: Biodiesel from Algae; Close-Out Report. Technical Report NREL/TP-580-24190. National Renewable Energy Lab., Golden, CO. (US). doi:10.2172/15003040.

Shiloach, J., Fass, R., 2005. Growing E. coli to high cell density—A historical perspective on method development. Biotechnology Advances 23, 345–357. doi:10.1016/j.biotechadv.2005.04.004.

Shoener, B.D., Schramm, S.M., Béline, F., Bernard, O., Martínez, C., Plósz, B.G., Snowling, S., Steyer, J.P., Valverde-Pérez, B., Wágner, D., Guest, J.S., 2019. Microalgae and cyanobacteria modeling in water resource recovery facilities: A critical review. Water Research X 2, 100024. doi:10.1016/j.wroa.2018.100024.

Shuba, E.S., Kifle, D., 2018. Microalgae to biofuels: 'Promising' alternative and renewable energy, review. Renewable and Sustainable Energy Reviews 81, 743–755. doi:10.1016/j.rser.2017.08.042.

Sialve, B., Gales, A., Hamelin, J., Wery, N., Steyer, J.P., 2015. Bioaerosol emissions from open microalgal processes and their potential environmental impacts: What can be learned from natural and anthropogenic aquatic environments? Current Opinion in Biotechnology 33, 279–286. doi:10.1016/j.copbio.2015.03.011.

Silva Benavides, A.M., Torzillo, G., Kopecký, J., Masojídek, J., 2013. Productivity and biochemical composition of {Phaeodactylum} tricornutum ({Bacillariophyceae}) cultures grown outdoors in tubular photobioreactors and open ponds. Biomass and Bioenergy 54, 115–122. doi:10.1016/j.biombioe.2013.03.016.

Singh, D., Lercher, M.J., 2020. Network reduction methods for genome-scale metabolic models. Cellular and Molecular Life Sciences 77, 481–488. doi:10.1007/s00018-019-03383-z.

Singh, D., Sharma, D., Soni, S.L., Sharma, S., Kumar Sharma, P., Jhalani, A., 2020. A review on feedstocks, production processes, and yield for different generations of biodiesel. Fuel 262, 116553. doi:10.1016/j.fuel.2019.116553.

Sommer, U., 1992. Phosphorus-limited daphnia: intraspecific facilitation instead of competition. Limnology and Oceanography 37, 966–973. doi:https://doi.org/10.4319/lo.1992.37.5.0966.

Storn, R., Price, K., 1997. Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. Journal of global optimization 11, 341–359. doi:10.1023/A:1008202821328.

Sun, X.M., Ren, L.J., Zhao, Q.Y., Ji, X.J., Huang, H., 2019. Enhancement of lipid accumulation in microalgae by metabolic engineering. Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids 1864, 552–566. doi:10.1016/j.bbalip.2018.10.004.

Supriyanto, Noguchi, R., Ahamed, T., Rani, D.S., Sakurai, K., Nasution, M.A., Wibawa, D.S., Demura, M., Watanabe, M.M., 2019. Artificial neural networks model for estimating growth of polyculture microalgae in an open raceway pond. Biosyst. Eng. 177, 122–129. doi:10.1016/j.biosystemseng.2018.10.002.

Tabe-Bordbar, S., Marashi, S.A., 2013. Finding elementary flux modes in metabolic networks based on flux balance analysis and flux coupling analysis: Application to the analysis of Escherichia coli metabolism. Biotechnology Letters 35, 2039–2044. doi:10.1007/s10529-013-1328-x.

Tai, M., Stephanopoulos, G., 2013. Engineering the push and pull of lipid biosynthesis in oleaginous yeast Yarrowia lipolytica for biofuel production. Metabolic Engineering 15, 1–9. doi:10.1016/j.ymben.2012.08.007.

Tan, X.B., Lam, M.K., Uemura, Y., Lim, J.W., Wong, C.Y., Lee, K.T., 2018. Cultivation of microalgae for biodiesel production: A review on upstream and downstream processing. Chinese J. Chem. Eng. 26, 17–30. doi:10.1016/j.cjche.2017.08.010.

Tandon, P., Jin, Q., Huang, L., 2017. A promising approach to enhance microalgae productivity by exogenous supply of vitamins. Microbial Cell Factories 16, 219. doi:10.1186/s12934-017-0834-2.

Tang, H., Chen, M., Simon Ng, K., Salley, S.O., 2012. Continuous microalgae cultivation in a photobioreactor. Biotechnology and bioengineering 109, 2468–2474. doi:https://doi.org/10.1002/bit.24516.

Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. J. Geophys. Res. Atmos. 106, 7183–7192. doi:10.1029/2000JD900719.

Terzer, M., Stelling, J., 2008. Large-scale computation of elementary flux modes with bit pattern trees. Bioinformatics 24, 2229–2235. doi:10.1093/bioinformatics/btn401.

Theil, H., 1961. Economic forecasts and policy .

Tian, K., Niu, D., Liu, X., Prior, B.A., Zhou, L., Lu, F., Singh, S., Wang, Z., 2016. Limitation of thiamine pyrophosphate supply to growing Escherichia coli switches metabolism to efficient d-lactate formation. Biotechnol. Bioeng. 113, 182–188. doi:10.1002/BIT.25699.

Tibocha-Bonilla, J.D., Zuñiga, C., Godoy-Silva, R.D., Zengler, K., 2018. Advances in metabolic modeling of oleaginous microalgae. Biotechnology for Biofuels 11, 241. doi:10.1186/s13068-018-1244-3.

Tredici, M.R., Bassi, N., Prussi, M., Biondi, N., Rodolfi, L., Chini Zittelli, G., Sampietro, G., 2015. Energy balance of algal biomass production in a 1-ha "green wall panel" plant: How to produce algal biomass in a closed reactor achieving a high net energy ratio. Appl. Energy 154, 1103–1111. doi:10.1016/j.apenergy.2015.01.086.

Trélat, E., Zuazua, E., 2015. The turnpike property in finite-dimensional nonlinear optimal control. Journal of Differential Equations 258, 81–114. doi:https://doi.org/10.1016/j.jde.2014.09.005.

Turon, V., Baroukh, C., Trably, E., Latrille, E., Fouilland, E., Steyer, J.P., 2015a. Use of fermentative metabolites for heterotrophic microalgae growth: Yields and kinetics. Bioresour. Technol. 175, 342–349. doi:10.1016/j.biortech.2014.10.114.

Turon, V., Trably, E., Fouilland, E., Steyer, J.P., 2015b. Growth of Chlorella sorokiniana on a mixture of volatile fatty acids: The effects of light and temperature. Bioresour. Technol. 198, 852–860. doi:10.1016/j.biortech.2015.10.001.

Varma, A., Boesch, B.W., Palsson, B.O., 1993. Stoichiometric interpretation of Escherichia coli glucose catabolism under various oxygenation rates. Applied and Environmental Microbiology 59, 2465–2473.

Venayak, N., Raj, K., Jaydeep, R., Mahadevan, R., 2018. An Optimized Bistable Metabolic Switch To Decouple Phenotypic States during Anaerobic Fermentation. ACS Synthetic Biology doi:10.1021/acssynbio.8b00284.

Venkataraman, L.V., 1997. Spirulina platensis (arthrospira): Physiology, cell biology and biotechnologym, edited by avigad vonshak. J. Appl. Phycol. 9, 295–296. doi:10.1023/A:1007911009912.

Vinter, R.B., Vinter, R., 2010. Optimal control. Springer. doi:https://doi.org/10.1007/978.0.8176.8086.2.

von Alvensleben, N., Stookey, K., Magnusson, M., Heimann, K., 2013. Salinity tolerance of picochlorum atomus and the use of salinity for contamination control by the freshwater cyanobacterium pseudanabaena limnetica. PLoS One 8, 1–12. doi:10.1371/journal.pone.0063569.

Wágner, D.S., Valverde-Pérez, B., Sæbø, M., Bregua de la Sotilla, M., Van Wagenen, J., Smets, B.F., Plósz, B.G., 2016. Towards a consensus-based biokinetic model for green microalgae – The ASM-A. Water Research 103, 485–499. doi:10.1016/j.watres.2016.07.026.

Wan, M., Liu, P., Xia, J., Rosenberg, J.N., Oyler, G.A., Betenbaugh, M.J., Nie, Z., Qiu, G., 2011. The effect of mixotrophy on microalgal growth, lipid content, and expression levels of three pathway genes in Chlorella sorokiniana. Appl. Microbiol. Biotechnol. 91, 835–844. doi:10.1007/s00253-011-3399-8.

Wei, P.P., Zhu, F.C., Chen, C.W., Li, G.S., 2021. Engineering a heterologous synthetic pathway in Escherichia coli for efficient production of biotin. Biotechnology Letters 43, 1221–1228. doi:10.1007/s10529-021-03108-y.

Weiss, V., Gromet-Elhanan, Z., Halmann, M., 1985. Batch and continuous culture experiments on nutrient limitations and temperature effects in the marine alga {Tetraselmis} suecica. Water Res. 19, 185–190. doi:10.1016/0043-1354(85)90197-6.

Welter, C., Schwenk, J., Kanani, B., Van Blargan, J., Belovich, J.M., 2013. Minimal medium for optimal growth and lipid production of the microalgae Scenedesmus dimorphus. Environmental Progress & Sustainable Energy 32, 937–945. doi:10.1002/ep.11835.

Williams, P.J.l.B., Laurens, L.M.L., 2010. Microalgae as biodiesel & biomass feedstocks: Review & analysis of the biochemistry, energetics & economics. Energy Environ. Sci. 3, 554–590. doi:10.1039/B924978H.

Wu, Y., Gao, K., Riebesell, U., 2010. CO 2-induced seawater acidification affects physiological performance of the marine diatom Phaeodactylum tricornutum. Biogeosciences 7, 2915–2923. doi:10.5194/bg-7-2915-2010.

Wu, Z., Shi, X., 2007. Optimization for high-density cultivation of heterotrophic Chlorella based on a hybrid neural network model. Lett. Appl. Microbiol. 44, 13–18. doi:10.1111/j.1472-765X.2006.02038.x.

Xianmin, Z., 1993. A new method with high confidence for validation of computer simulation models of flight systems. J. Syst. Eng. Electron. 4, 43–52.

Xiao, D., Zhu, H., 2006. Multiple focus and Hopf bifurcations in a predator-prey system with nonmonotonic functional response. SIAM Journal on Applied Mathematics 66, 802–819. doi:https://doi.org/10.1137/050623449.

Xiao, Z., Zheng, Y., Gudi, C.R., Liu, Y., Liao, W., Tang, Y.J., 2021. Development of a kinetic model to describe six types of symbiotic interactions in a formate utilizing microalgae-bacteria cultivation system. Algal Research 58, 102372. doi:10.1016/j.algal.2021.102372.

Xie, Y., Li, J., Ho, S.H., Ma, R., Shi, X., Liu, L., Chen, J., 2020a. Pilot-scale cultivation of Chlorella sorokiniana FZU60 with a mixotrophy/photoautotrophy two-stage strategy for efficient lutein production. Bioresour. Technol. 314, 123767. doi:10.1016/j.biortech.2020.123767.

Xie, Z., Lin, W., Liu, J., Luo, J., 2020b. Mixotrophic cultivation of Chlorella for biomass production by using pH-stat culture medium: Glucose-Acetate-Phosphorus (GAP). Bioresour. Technol. 313, 123506. doi:10.1016/j.biortech.2020.123506.

Zhang, C., Li, S., Ho, S.H., 2021. Converting nitrogen and phosphorus wastewater into bioenergy using microalgae-bacteria consortia: A critical review. Bioresource Technology 342, 126056. doi:10.1016/j.biortech.2021.126056.

Zhou, L., Tian, K.M., Niu, D.D., Shen, W., Shi, G.Y., Singh, S., Wang, Z.X., 2012. Improvement of D-lactate productivity in recombinant Escherichia coli by coupling production with growth. Biotechnology Letters 34, 1123–1130. doi:10.1007/s10529-012-0883-x.