MASTERARBEIT | MASTER'S THESIS

Titel | Title

A Dynamic Flux Balance Analysis (dFBA) model for a Leguminous Plant–Rhizobia Genome-Scale Reconstruction Based on the ViNE Model

verfasst von | submitted by

Lino Böhler

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien | Vienna, 2025

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 066 875
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Bioinformatik
Betreut von | Supervisor:	Ass.-Prof. DI Dr. -Ing. Steffen Waldherr

Abstract - Deutsch

In der vorliegenden Arbeit wurde ein dynamic Flux Balance Analysis (dFBA) Modell implementiert, basierend auf dem von diCenzo et al. (2020) publizierten ViNE-Modell. Dieses Model ist eine genomweite metabolische Rekonstruktion von *Medicago truncatula* (Klee) und *Sinorhizobium meliloti* (Bakterium). Das Modell umfasst drei verschiedene Gewebe: Wurzel, Spross und Knöllchen. Die Knöllchen vereinen das metabolische Netzwerk, von sowohl der Pflanze als auch dem Bakterium. Die Knöllchen sind in fünf Entwicklungszonen unterteilt. Das Model besteht aus insgesamt 7297 Reaktionen (diCenzo et al., 2020).
Das dFBA-Modell beinhaltet vier gewöhnliche Differenzialgleichungen (ODEs), die drei dynamisch ausbalancierten Metaboliten – N_2, NH_4 und Stärke – sowie die Biomasseproduktion des Modells beschreiben. Zudem wurde ein das Modell mit einem Tag-Nacht-Zyklus implementiert, welcher über eine anpassbarer Tageslänge verfügt. Die Parameter für die ODEs und die Beschränkungen der Aufnahme Reaktionen wurden aus der verfügbaren Literatur abgeschätzt. Zur Lösung des dFBA-Problems wurde ein direkter Ansatz verwendet (DA), der eine hierarchische Optimierungsstruktur sowie ein LP-Feasibility-Problem nutzt. Daraufhin folgend wurde eine finale Simulation über 13.000 simulierten Stunden durchgeführt. Mit Ende dieser Simulation hatte der aus der Systemumwelt verfügbare Stickstoff N_2 eine Gleichgewichtszustand erreichte. Die mittlere Wachstumsrate über die gesamte Simulation betrug 0,028 Tag^{-1}, während im Gleichgewichtszustand von N_2 eine Rate von etwa 0,012 Tag^{-1} beobachtet werden konnte. Über den gesamten Simulationszeitraum zeigte das Model einen Stärkeanteil an der Biomasse von etwa Durchschnittlichen 2,5 % auf. Mit Ende der Simulation erreichte das Model einen Stärkeanteil von 40 % in der Gesamtbiomasse. Die in der Simulation ermittelte CO_2-Freisetzung pro fixiertem N_2 in den Stickstoff fixirenden Knöllchen des Modells lag im Bereich von 3,5 – 5 g C g^{-1} N, mit einem von etwa 5 g C g^{-1} N am Ende der Simulation.

Abstract – English

In this work, a dynamic Flux Balance Analysis (dFBA) model was implemented based on the ViNE model published by (diCenzo et al., 2020), a genome-scale metabolic reconstruction of *Medicago truncatula* (clover) and *Sinorhizobium meliloti* (bacterium). The model comprises of three distinct tissues: root, shoot and nodules. The nodules combine both plant and bacteria metabolism and are subdivided in 5 differential zones. In total the model consists of 7297 reactions (diCenzo et al., 2020).

The dFBA model includes four ODEs, modelling three dynamically balanced metabolites N_2, NH_4, and starch as well as the biomass produced by the model. The model implements a day-night cycle with an adjustable day length. The parameters for the ODEs and uptake constraints were estimated from data available in the literature. Moreover, for solving the dFBA problem, a direct approach was used, utilising a hierarchical optimisation structure and an LP feasibility problem. Thereafter, a final simulation was run spanning 13,000 simulation hours, during which environmentally available N_2 reached equilibrium by the end of the simulation. The median growth rate over the simulation was 0.028 day^{-1}, with a rate of approximately 0.012 day^{-1} at the equilibrium level of N_2. The model exhibited a starch biomass fraction averaging about 2.5% over the simulation span, accumulating up to 40% at the end of the simulation. The yield of CO_2 evoked per N_2 fixed in mature nodules of the model, obtained within the simulation, was within the range of 3.5 - 5 g C g^{-1} N, showing a yield of about 5 g C g^{-1} N at the end of the simulation.

Für meine Eltern die mir mein Studium ermöglicht haben.

Acknowledgements

I would like to thank my supervisor Ass.-Prof. DI Dr. -Ing. Steffen Waldherr for being so patient and supportive. I also want to thank everyone else at MoSys and the 3rd floor, I had a wonderful time being here.

Inhalt

1. Introduction

The aim of the master's thesis is the implementation of a dynamic Flux Balance Analysis (dFBA) model. The starting point is an already published Genome-Scale Metabolic Model (GEM), a model of this type is a mathematical description of a metabolic reaction network. In particular, 'genome scale' means that it is a representation of the entire metabolic network of an organism. Such models can describe the metabolism of individual species, but also of entire communities of different species (Mayali et al., 2016). The model on which this work is based is called 'Virtual Nodule Environment' (ViNE) (diCenzo et al. 2020). It describes the metabolism of *Medicago truncatula* (clover) and *Sinorhizobium meliloti*, a bacterium that is in symbiosis with this plant.

A Flux Balance Analysis (FBA) is a method for predicting the reaction flux distribution in a cell. In other words, an algorithm to quantitatively determine the reaction rates in a metabolic network. An advantage of FBA simulations is that it can be calculated efficiently, with low hardware demands. Moreover, it does not require any usually hard to measure kinetic parameters. However, simulations performed using FBA only represent the system's current state under the given constraints. It does not allow to draw any conclusions about time-dependent dynamics, or the concentration of metabolites involved (Orth et al., 2010). An extension of this application is the dFBA, where, under the basic assumptions of an FBA problem inside the system, reactions that are in exchange with the environment are modelled by means of differential equations. The general motivation for using dynamic modelling is that it allows to study time dependent changes, changes in the state of the model apart from steady state. More specifically, dFBA allows the prediction of changes in metabolite concentrations (Mahadevan et al., 2002).

1.1 Aim of the Thesis

The focus of the project is on setting up the necessary equations, researching parameters in the literature and databases and carrying out simulations with ViNE model. As a result, the model can react to changes in the system environment that occur over the simulated time. Therefore, it is possible to predict the changes in concentration of the selected metabolites, as well as the change in biomass (i.e. growth rate) over a defined time span. The uptake of nitrogen is of particular interest, as the symbiont is a nitrogen-fixing bacterium. The aim of this work is also to compare the results of the simulations with published data. However, it is not the goal of this thesis to accurately estimate the parameters of the model. The model's predictions are expected to fall within an appropriate range relative to the compared metabolite.

2. Background

2.1 The modelled Organisms

Most of Earth's nitrogen, in the form of inert N_2 is in accessible to plants. To break these strong molecular bonds and make it bioavailable, enzymes like nitrogenase are required. However, Nitrogenase can only be found in prokaryotes. Legumes (*fabacea*) can excess this type of N source, by forming a symbiotic relationship with such bacteria. The legume *M. Truncatula* has been established as model organism for legumes.
Rhizobia are a group of bacteria, freely living in the soil. When a root nodule of the legume is infected by a rhizobia both will undergo a differentiation. The bacterium will grow and replicate and eventually differentiate into N_2-fixing Bacteroide and the nodule will develop into a mature Nodule where N is fixed (Udvardi & Poole, 2013).
There are plenty different rhizobia, however not all part of the genus rhizobium. They all interact specifically with different hosts, mostly legumes. Consequently, the number of host symbiont relationships is limited, the rhizobia *Sinorhizobium meliloti*, is a well-studied symbiont of *M. Truncatula* (Pfau et al., 2018).

2.2 The Nitrogen cycle and relation to Legumes

Nitrogen is one of the most important (the most in terms of total amount used) fertilizers, used in agriculture. So much so that about 40 % of the food globally, could not be produced without the use of artificial N-fertilisation (Rizwan et al., 2024). It is thus crucial for ensuring food security. The N-cycle, flowing through the biosphere, moving N form the air through all ecosystems and back into the atmosphere, can be subdivided into 5 steps:

<u>N-fixation</u>
The process of making inert N_2 gas in the air bioavailable N-species. In the case of biological Nitrogen fixation, this can be only done by prokaryotes with the enzyme nitrogenase, some of the are free some symbiotically living. A small part of Nitrogen can also be fixed abiotically by lightning; however, most is fixed biologically. The equation below describes the conversion of N_2 to NH_3. The reaction consumes 8 e^-, thus an energy demanding reaction.

$$\mathrm{N_2 + 8\,H^+ + 8\,e^- \rightarrow 2\,NH_3 + H_2}$$

<u>Mineralisation or Ammonification:</u>
Mineralization is the process of decomposition, which occurs when organisms die or produce waste. These biological remains are then metabolized by prokaryotes and fungi. The ammonia released as a metabolic byproduct of these decomposers is a crucial element in the nitrogen cycle, where it can be taken up by plants and microbiota (Bernhard, 2010).
However, not all of the nitrogen made available this way can be taken up by plants, as some is immobilized by the soil microbiota. This phenomenon is attributed to the utilization of nitrogen by microbes for their own sustenance, resulting in its immobilization and subsequent unavailability for plant uptake. Furthermore, the nitrogen content in undisturbed

soils is mostly (95%) bound in organic material, therefore inaccessible to plant life (Walworth, 2013).

Nitrification

The Process of microbial conversion of NH_3 to NO_3 .

Denitrification

where Nitrate is converted back to gaseous N_2 as an ultimate product, this happens over intermediate products like N_2O, a strong greenhouse gas. It is thus a reaction that removes bioavailable N from soil or more general the ecosystem and returns it to the atmosphere. Denitrification happens under aerobic conditions, promoted by prokaryotes. (Bernhard, 2010)

All the processes of nitrification and denitrification are also regulated by factors which are secreted by plants and microbes, Rhizosphere exudates. These exudates act as biological nitrification and denitrification inhibitors. Thus, regulating the amount of N in the ecosystem, by controlling the amount of N runoff into water bodies and emission of N_2O into the atmosphere(Lu et al., 2024).

Nitrogen is crucial for ensuring food security, yet the excessive usages of fertilizers at this scale, impacts the environment and ecosystems enormously, both locally as well as on a global scale. Furthermore, too much N is accepted as one of the main drivers of biodiversity changes across the globe (Sala et al., 2000)

The surplus of reactive Nitrogen species, introduced to the biosphere by extensive usage of fertilizer and other anthropogenic activities, pollutes the air and water. Further reduces biodiversity, harms Human health and accelerates climate change, with NO_2 having a around 300 times higher global warming potential then CO_2 over a retention time of 100 years. (Lammel & Graßl, 1995; Rizwan et al., 2024)

The overabundance of bioavailable N or reactive N species in soil, effects diversity by favouring “N-loving species” or species that would otherwise be more restricted by the availability of N. Also, the over usage of fertilizer shifts microbial community composition in soils, thus shifting the balance towards nitrification and denitrification. Ultimately, more emission of NO_2 into the atmosphere and leaching of NO_3 from the soil.

Most of the fertilizer applied is not utilized by the plants but lost to the environment. Resulting in the volatilization of NH_3 to NH_4 into the Air or the leaching of NO_3 in to ground water harming health of animals and local population (methemoglobinemia) and causing eutrophication in rivers and bodies of water.
Overfertilization effects the environment and climate directly by evocation of reactive Nitrogen species into air and water further even causing acid rain leading to over acidification of soil and water. It effects the climate directly by NO_x species acting as greenhouse gas but also indirectly since 2% of the global commercially used energy is being

used by the Haber-Bosch process to produce NH3 from N_2 and H_2. (Fuller et al., 2022; Rizwan et al., 2024)

Given that legumes, do not relay that much on N fertilization, it is of most relevance to understand the Metabolic interaction between legumes and rhizobia. Since, this could e.g. pave the way to genetical engineering approaches. For instance, improving the efficiency of N-fixation, or maybe even engineer other crops like cereals to form symbiosis with rhizobia. (Udvardi & Poole, 2013)

It has been hypothesized that legumes will benefit from rising [CO_2] levels, as their ability to exchange carbon for nitrogen with their symbionts suggests they can enhance growth through increased photosynthetic activity—provided they do not experience any nutrient scarcities or drought conditions. Experiments indicate that legumes increase their nodule biomass (either by producing more or larger nodules) at elevated [CO_2] levels. However, there is no evidence that they enhance nitrogenase activity (Libault, 2014; Rogers et al., 2009).

2.3 Flux Balance Analysis (FBA)

2.3.1 What is FBA

FBA is an algorithm used to calculate the flux distribution, of the reactions through a metabolic network. These metabolic networks can be, partial network e.g. central carbon metabolism of a plant, whole genome reconstructions of organisms (GEMs) or even the metabolic network of entire communities of different species.
The stoichiometries of the reaction network can be represented as matrix S.

$$S \cdot v = 0 \tag{1}$$

with $S \in R^{m \times n}$
and $v \in R^{n \times 1}$

In the above equation S is the stochiometric matrix (size *m*n*) and v is the vector (with length *n*) of all reaction fluxes. Where ***m*** is the number of metabolites and ***n*** the number of reactions in the metabolic network. The n reactions span a n dimensional flux space; by requiring mass balance, we can further constrain it to a flux subspace, which is the null space of the matrix S. Most important a central assumption of FBA is the steady state assumption. This is implied by eq.(1) , with every v in the null space of S, satisfying eq.(1). In other words, every of the m metabolites participating in the reaction of S is mass balanced. No metabolite can deplete or accumulate with in the network, this means that the amounts of Carbon or Nitrogen for example that go in into the system also must leave the system.

S is a sparse matrix with every metabolite not involved in a reaction represented by a 0 in the corresponding rection column n of the matrix. Further metabolites which are produced are represented with a positive and integer and metabolites consumed by a reaction denoted as negative integer in the matrix S.

Since there are usually more reactions than metabolites ($n > m$), hence more variables than equations, the system is underdetermined. There is no unique solution for the system of equations given by eq.(1) .

Since some reactions are irreversible, we can further constrain this subspace into a flux cone**.** By limiting the capacity of reaction further, we would obtain (geometrically explained) a solution resembling a convex flux polytope.

(Orth et al., 2010)

Ultimately, we want to determine the "optimal" flux distribution, under a given objective function. The underlying assumption is the optimality hypothesis, that is have evolved to optimize their metabolism for an objective function e.g. Biomass (Berkhout et al., 2012). The maximisation of the biomass reaction would be a common choice for such an objective function. An alternative objective function, that has been suggested, where the optimisation for energetic efficiency, for example ATP yield per substrate or minimization of fluxes (Harcombe et al., 2013; Schuetz et al., 2007). The Biomass reaction is a pseudo reaction estimating the growth rate, or biomass production by the FBA model. A biomass reaction can be constructed by defining the macromolecular composition of the cell (proteins, RNA, DNA, carbohydrates) to estimate the proportional requirements necessary for productions of building block precursors. To further detail these equations for the formation of biomass, biosynthetic energy requirements are factored in. Finally, a growth rate μ can be calculated by using a biomass objective function, which accounts for NGM costs (non-growth associated maintenance costs) and the growth associated maintenances costs of the cell, for a given substrate uptake rate from the environment into the system (Feist & Palsson, 2010).

This objective function can be given by:

$J = c^T \cdot v$

Where c^T is the vector of weights for each of the reactions fluxes in v, that is to say how much the max- or minimisation of a certain flux participates to the optimisation of J.
As for example when only the biomass reaction is being optimized, the vector c would have a 0 at every position of the vector, except for the position of the biomass flux, where it would equal 1.

(Orth et al., 2010)

So, in other words, with FBA, we aim to determine the flux values of all reactions in the network, represented by **v**, given the steady-state constraints $\mathbf{S \cdot v = 0}$ imposed by the stoichiometry of S and additional constraints limiting the rate and direction of reactions.

These constraints are based on knowledge derived from thermodynamic, kinetic, and biochemical data. For example, factors such as the limited solvent capacity of cells or the fact that many cellular processes operate near the diffusion limit (Bordbar et al., 2014) serve as important biological constraints. Considering all these constraints further restricts the solution space, which is necessary to exclude optimal solutions that are inconsistent with actual biological reality (Schuetz et al., 2007).

The problem, therefore, involves solving a system of linear equations $S \cdot v = 0$ optimizing a given objective J, and satisfying an additional set of constraints, resulting in an optimization problem that can be solved using linear programming (LP) (Orth et al., 2010).

An FBA problem can be formalized as:

$$\max\ J(v)$$

$$\text{s.t. } Sv = 0$$

$$Lb\ v_i \leq v_i \leq Ub\ v_i$$

$$S \in R^{m\,x\,n}$$
$$v \in R^{n\,x\,1}$$

An FBA problem is infeasible if no solution can be achieved under the given constrains. Conventionally, uptake reactions are represented by a negative flux value, so they are consumed by the cell from the environment. Consequently, secretion reactions have positive flux. Therefore, a directionality for reaction is given, forward reactions are positive while reverse reactions are described by negative flux values. However, in the ViNE model uptake reactions are implemented with positive fluxes and secretion reactions with negative reaction fluxes.

2.3.2 Limitations of FBA

Genome scale reconstructions of metabolic networks often incomplete, there are mostly gaps in the knowledge about a modelled organism, thus usually there are reactions missing. Therefore, FBA can be used to predict missing reactions and identify gene essentiality. (Bordbar et al., 2014). FBA mostly relies on constrains, thus an advantage is that only the Stoichiometries of the Reactions and bounds for these reactions are needed. Where other modelling approaches require much more in-depth knowledge. For instance, many kinetic parameters of reactions or concentrations of participating metabolites, which are difficult to measure (Orth et al., 2010).These advantages due to the conceptual simplicity of FBA, on the other side imply limitations of the predictive capabilities of this method. Thus, FBA is not able to make any predictions about concentrations metabolites or any other time dependent dynamics (Mahadevan et al., 2002). Another limitation of FBA is that it does not account for changes in gene expression or any explicit gene regulation. (Orth et al., 2010).

2.3.3 Dynamic Fux Balance Analysis (dFBA):

The dynamic Flux Balance Analysis (dFBA) is an extension of the FBA method, where under the assumptions of a FBA problem inside the system, reactions in exchange with the environment are modelled using differential equations. Where these reactions are constrained dynamically using kinetic parameters, usually using Micahelis-Menten kinetics. Let u_0 be the vector of the initial concentrations of the dynamically balanced metabolites.

$$\frac{\mathrm{d}\,u}{\mathrm{d}\,t} = \mathrm{f}\big(t, \mathrm{h}(u)\big)$$

$$u_0 = t_0$$

Where u is the concentration of the metabolites balanced dynamically, in the system environment, h being function mapping the concentrations of the metabolites in the environment u to the optimal fluxes of the exchange reactions determined by linear programming. Lastly t being the time of the simulation. Furthermore h(u) can be obtained by solving the linear programming problem or FBA problem below.

$$\max\ J(v)$$

$$\text{s.t. } Sv = 0$$

$$Lb\ v_i\ (\mathrm{u(t)}) \leq v_i \leq Ub\ v_i\ (\mathrm{u(t)})$$

Here $Lb\ v_i\ (\mathrm{u(t)})$ and $Ub\ v_i\ (\mathrm{u(t)})$ are the lower and upper bounds of the exchange reactions with the environment, as a function of metabolite concentrations(Gomez et al., 2014). The under-laying assumption is that reactions inside a cell or organism happen at a much faster rate than in the surrounding environment, especially since they are catalysed by enzymes. Consequently, concentrations of metabolites inside the metabolism are assumed to be in steady state. While the concentrations of metabolites outside the metabolism, react much slower and much more time is needed to reach an equilibrium. It is thus possible to predict the metabolite concentrations in the environment, within the simulated time frame. Two types of constraints are used in dFBA modelling. Steady-state constraints apply to fluxes that are not dynamically balanced, such as intracellular metabolites. These constraints are the same as those used in the underlying FBA model. In contrast, dynamic constraints are applied to dynamically balanced metabolites and are often modelled using Michaelis-Menten kinetics. Thus, they often depend on concentration of environmental metabolites

dFBA has been used to study the switch between alternative metabolic pathways, like diauxic growth of *E. coli*. (Mahadevan et al., 2002). What it more is, of striking relevance is the capability of dFBA to predict community compositions in community models with two or more different species. Thus, ecological interactions like predation, competition cooperation

can be modelled. However, for the scope of this Thesis interesting would be modelling of multi tissue models (Mayali et al., 2016).

2.3.4 Solving the dFBA problem - direct approach (DA)

This section will discuss three approaches, that have been formalized to solve the dFBA problem. Further, the approach used for the implementation in this thesis, is the direct approach (DA).

Static optimization approach (SOA):
In this approach, the total simulation time span is divided into smaller time steps. At the start of each time step, the optimization problems are solved, and a flux distribution is determined. The ODEs are then integrated over this time interval, with the flux distribution obtained at the beginning of the time step remaining fixed. The changes in the system's state, calculated during one iteration, are subsequently used to constrain the optimization problems at the start of the next time interval. An approach like this can easily, be implemented using a Euler forward method. However, since most dFBA problem are stiff it is necessary to use an implicit algorithm, thus a Euler backward scheme. Thus, the explicit solvers used in this approach will most likely fail or require a to small step size in order to maintain stability (Gomez et al., 2014).

Direct approach (DA):
For the direct approach the LP solver is directly include in the right-hand side of the ODE evaluator function. Every time the right-hand side evaluator function is called, which is the function describing the model, one (or more) LP problems need to be solved. In this thesis, this will be defined as an iteration, with each iteration involving the solution of three FBA problems. This approach has the advantage that existing integrator functions can be used with adaptive step size and error control, integrators implementing implicit methods such as, backward Euler schemes. Thus, increasing speed and stability of the computations (The Economic Cell Collective (2025) et al., 2025, p. 166).

Dynamic optimisation approach (DOA):
Objective function dependent on the dynamic states over complete time horizon is formalized and then transformed to an NLP problem which is solved once, optimizing over the entire time horizon of the simulation. Thereby also dynamic objective functions can be formulated. However, the NLP problem can become intractable, since it doesn’t scale well with size of network and the fines of time discretization. Thus, computational too expensive for large networks like GEMs (The Economic Cell Collective (2025) et al., 2025, p. 166)

3. Methods and the Model

3.1 The Model

A central part of this Thesis was the implementation of a model for dynamic Flux Balance Analysis, based on the genome scale metabolic model ViNE, published by di Cenzo and colleges (diCenzo et al., 2020). The general structure of the dFBA Modell can be subdivided as follows:

1) Uptake kinetics and Uptake constrains
2) Intracellular Metabolism
3) ODEs for extracellular and dynamically balanced metabolites
   (The Economic Cell Collective (2025) et al., 2025)

3.1.1 Uptake kinetics and constrains:

The constraints on the uptake reactions which are implemented dynamically, are modelled using Michaelis-Menten and Hill kinetics.

Nitrogen uptake

The Nitrogen uptake by the model, is limited to two reactions, that is N_2 and NH_4 uptake in the root area. The constrains on these can be formulated by the following 3 equations:

$$Ub\ NH_4 = \frac{V_{maxNH_4} \cdot c(NH_4)^2}{K_{mNH_4} + c(NH_4)^2} \tag{2}$$

$$\text{with } c(NH_4) > 0$$

$$Ub\ N_2 = \frac{\overline{V}_{\max N_2} \cdot c(N_2)^2}{K_{m\,N_2} + c(N_2)^2} \tag{3}$$

with

$$\overline{V}_{\max N_2} = \begin{cases} v_{N-switch} + \dfrac{V_{\max N_2} - v_{N-switch}}{\kappa_{slope} \cdot\ Ub\ NH_4 + 1} & \text{if } c(NH_4) \leq K_{m\,NH_4} \\ V_{\max N_2} & \text{if } c(NH_4) \leq 0 \\ 0 & \text{if } c(NH_4) > K_{m\,NH_4} \textbf{ or } c(N_2) \leq 0 \end{cases} \tag{4}$$

Where Ub N_2 and Ub NH_4 are the upper boundaries for the N_2 and NH_4 exchange reactions. Furthermore, $c(N_2)$ and $c(NH_4)$ being the extracellular concentrations, that is the availability of these Metabolites in the Soil. Lastly K_m being the Michaelis-Menten constant and V_{max}

being the maximal reaction speed (of an enzyme, catalysed reaction). In this case, it's the maximal uptake rate for N_2 and NH_4 respectively. Lastly $v_{N-switch}$ is the initial N_2 up take flux of $\bar{V}_{\max N_2}$, and κ_{slope} being a parameter controlling the slope of $\bar{V}_{\max N_2}$ function. A summary of all parameters used can be found in chapter 4.1, p.26.

Thus eq. (2) describes the uptake kinetics of NH_4, following a Hill kinetic. Analogical eq. (3) describes the uptake kinetics of N_2, following a Hill kinetic.

The eq. 2) and 3) together, describe an "N-switch", which ensures a smooth transition between the two different N-sources, at low NH_4 concentration levels. Here $\bar{V}_{\max N_2}$ of eq. (3) is described by eq. (4) and not fixed as V_{maxNH_4} in eq. (2). Thus $\bar{V}_{\max N_2}$ is dynamically changing with the current upper bound on NH_4 uptake, thus dependent on the extracellular concentration of NH_4. Hence, the intend was that at low NH_4 concentration the uptake rate of N_2 increases slowly and continuously with no abrupt jumps in the uptake rate, this will be discussed in more detail in section 4.2.3.
The 2nd case of eq. (4) describes the situation where there is no NH_4 available anymore and N_2 uptake follows Hill Kinetics. For any case where the extracellular concentration for NH_4 or N_2 reach zero, the upper bound for that respective reaction while be set to zero. While the concentration of extracellular NH_4 is at a higher level, the uptake of bioavailable Nitrogen (NH_4) is prioritised over the uptake of N_2 from the soil, this is also implied by the first and third case of eq. (4).

Day night cycle

The model follows a day-night cycle. Thus, the light uptake is modelled as depicted in eq.(3.7)

Let $i(t)$ be the light intensity at hour t, which is the fraction of the maximum light uptake possible at simulation hour t. Then is $i(t) \in [0,1]$, with $i(t) = 0$ total darkness and $i(t) = 1$ being the maximum possible solar irradiance. The light intensity is modelled by a sinoid function:

$$i(t) = \max\left[(1-a)\cdot sin\left(\frac{\pi}{12h}\cdot t\ + c\right) + a, 0\right] \quad (5)$$

With $\boldsymbol{t}$ being the simulation time in hours and $\boldsymbol{c}$ being the horizontal shift of the function along the t-axis. The parameter a is used to determine the length of the day. A formula for the calculation of the length of the day, considering latitude and day of the year has been described in (Jenkins, 2013).

The relationship between length of the day and parameter **a** can be described by:

$$a = \frac{I_0 + cos\left(\pi - \frac{\Delta t \pi}{24}\right)}{1 + cos\left(\pi - \frac{\Delta t \pi}{24}\right)} \tag{6}$$

$$with\ I_0 \in [0,1]$$

$$\Delta t \in [0,24]$$

Where I_0 , is the light intensity at down or dusk respectively. That is the chosen intensity level defined to begin and end the day light phase.

Thus, the Photon uptake rate of the model in [nmol/ gDW h] is given by:

$$Lb\ Light\ \leq\ v_{Light}\ \leq\ Ub\ Light \tag{3.7}$$

$$Ub\ Light = \phi^{-1} \cdot \mathrm{v_{CO_2 fix}} \cdot i(\mathrm{t})$$

$$Lb\ Light\ \geq 0$$

With ϕ^{-1} being the Photon / CO_2 turnover ratio (Zakhartsev et al., 2016).

Starch Metabolism:

The degradation and synthesis of immobilized Starch is constraint by eq. (8) below. It is important to note that, Starch is a metabolite, occurring solely within in the organism. The reaction was implemented as exchange reaction with the environment. However, Starch is not being exported into the environment but rather stored in an abstract starch compartment of the model. For more details see sub-chapter 3.1.3. or 3.1.2 . Further its should be noted that starch uptake was modelled with Michaelis Menten Kinetic, not with a Hill kinetic.

$$Lb\ Starch = \begin{cases} -1 \cdot \dfrac{\mathrm{V_{max\ Amylase}} \cdot \mathrm{c(Starch)}}{\mathrm{K_{m\ Amylase}} + \mathrm{c(Starch)}} & if c(Starch) > 0 \\ 0 & otherwise \end{cases} \tag{8}$$

Where $Lb\ Starch$ is the lower boundary of the Strach sink reaction of the model. $V_{max\ Amylase}$ being the maximal rate of Amylase catalysed Starch degradation and $K_{m\ Amylase}$ being the Michaelis Menten constant of Amylase. Lastly c(Strach) is the concentration or

amount of immobilized Starch available to the Model.

While Ub Starch is limited by Ub CO_2 = 180000 $\frac{nmol}{gDWh}$ which is the upper limit for the CO_2 uptake rate of the model.

3.1.2 Intracellular Metabolism – The ViNE Modell

The Virtual Nodule Environment (ViNE), is a genome scale metabolic reconstruction, consisting of 7297 reactions and 6702 Metabolites.
The Intracellular metabolism of this dFBA implementation can be fully described by the ViNE model. The ViNE model is a lumped community model, that means this model embodies the whole intertwined metabolism of two different species *Medicago truncatula* (Clover) and *Sinorhizobium meliloti* (Bacterium), repented as one entity. This means one (compartmentalized) metabolic network with one combined biomass reaction for all tissues and both organisms. The structural composition of the model is comprised of three distinct tissues: root, shoot and nodules. The nodule tissue integrates both plant and bacterial metabolism, the nodule are further subdivided into five distinct developmental zones (I, IId, IIp, IZ and III). Where Zone I is the compartment of uninfected Nodules, the Zones IId, IIp, IZ the differentiating Nodules. Lastly, Zone III represents mature N_2 fixing Nodules (diCenzo et al., 2020).

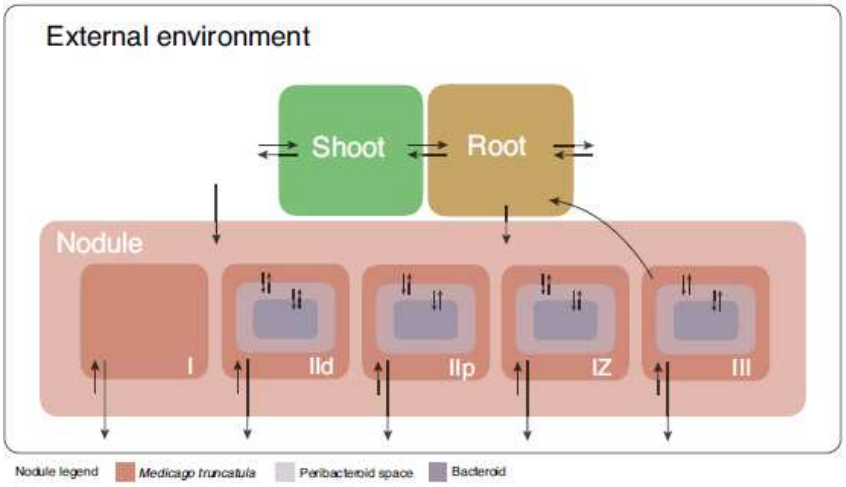

Fig. 3.1 Visual depiction of the ViNE model taken from (diCenzo et al., 2020) *)*

On the Starch metabolism:

To the existing ViNE model one sink reaction has been added, which consumes Starch metabolite from the Root compartment, this reaction however is reversible. This means consumption of starch by the reaction, thus removal of metabolite from the system, is equivalent to the synthesis of storage starch as a carbohydrate reservoir.
Consequently, the uptake of Starch by the system, from the pool (which is the reverse reaction) is the equivalent of amylase reaction or the mobilisation of glucose from

immobilized starch. Moreover, starch is only mobilized to inform of Glucose metabolite and never degraded to maltose, since it is the only degradation reaction present in the model.

Further notes regarding the ViNE model:

Against conventions, uptake reactions by the ViNE Model are defined as positive fluxes. Further NH4 is the only form of bioavailable nitrogen that can be taken up by the model from the Environment. Thus, the only N-sources for the Model are N_2-Gas and NH_4 from the Soil and no uptake of e.g. NO_3.

3.1.3 ODEs for Extracellular and dynamically modelled Metabolites

The eq. (6)- (12) below, describe the ODEs coupled to the FBA model. The concentrations of N_2 and NH_4 in the system environment are given by c(N_2) and c(NH_4). Further, the amount of immobilised Starch in the Plant is given by c(Starch), used as an energy deposit.
These three concentrations are all given in $\frac{mg}{L}$, with m_{N_2}, m_{NH_4} and m_{Gluc} being the molecular weights of N_2, NH_4 and Glucose. Lastly X is the amount of biomass in the system environment expressed as $\frac{g}{L}$.
Since the fluxes in are given with respect to biomass, the eq. (10) - (12) need to be multiplied by X to get the amount of metabolite with respect to system volume.

$$\frac{dX}{dt} = \mu \cdot X \quad \left[\frac{mg}{gDW\ h}\frac{g}{L}\right] \tag{9}$$

$$\text{with X and } \mu \geq 0$$

The first equation *eq.* (6) describes changes of the biomass, where X is the concentration or the amount of biomass given in $\frac{g}{L}$, in the system. The parameter μ in $\frac{mg}{gDW\ h}$ is the growth rate of the model, which is flux of the biomass reaction. Since this growth rate is not given as $\frac{1}{\mathrm{h}}$, we need to dived *eq.* (6) by 10^3, such that the derivative $\frac{dX}{dt}$ is given in $\frac{g}{L\ h}$. which is necessary since the Biomass, X is given in $\frac{g}{L}$.

$$\frac{d\ c(N_2)}{dt} = -\ v_{N_2} \cdot m_{N_2} \cdot X + K_L a \cdot \left(\ c^*(N_2) - c(N_2)\right) \quad \left[\frac{nmol}{gDW\ h}\ \frac{mg}{nmol}\ \frac{g}{L} + \frac{1}{h}\ \cdot \frac{mg}{L}\right] \tag{10}$$

$$with\ c(N_2), v_{N_2}\ and\ X\ \geq 0$$

$$m_{N_2},\ K_L a\ \ and\ \ c^*(N_2)\ constants$$

The equation in (10), describes the c(N_2) level in the environment of the system. Here we have an additional diffusion term, which defines the rate at which N_2 replenished in the

external environment. The diffusion of gases is modelled by the Henry law, the term is thus given by:

$$K_L a \cdot \left(c^*(N_2) - c(N_2) \right)$$

Where $c^*(N_2)$ is the saturation of N_2 in water, given the partial pressure of N_2 in air under standard conditions. Lastly, $K_L a$ is the volumetric gas transfer coefficient (The Economic Cell Collective (2025) et al., 2025, p. 169).

$$\frac{d\, c(NH_4)}{dt} = -\, v_{NH_4} \cdot m_{NH_4} \cdot X \quad \left[\frac{nmol}{gDW\ h} \ \frac{mg}{nmol} \ \frac{g}{L}\right] \tag{11}$$

$$\text{with } c(NH_4),\ v_{NH_4}, X \ \geq 0$$

$$m_{NH_4}\ constant$$

The availability of NH_4 for the Model, by the environment is given by eq.(11) above. It thus describes changes in the c(NH_4) in the system environment.

$$\frac{d\, c(\text{Starch})}{dt} = v_{\text{Starch}} \cdot m_{\text{Gluc}} \cdot X \quad \left[\frac{nmol}{gDW\ h} \ \frac{mg}{nmol} \ \frac{g}{L}\right] \tag{12}$$

$$\text{with } c(\text{Starch})\ , X \ \geq 0$$

$$Lb\ amylase \leq\ v_{\text{Starch}} \ \leq\ Ub\ CO_2$$

The eq.(12 formulates Starch synthesis and degradation. Starch is an intracellular metabolite, however in this model, immobilised starch is implemented as an extracellular metabolite. Thus, the uptake flux of the Strach reaction v_{Starch} into the system would be the equivalent of starch mobilization or amylase activity, which is the degradation of immobilized Starch. Whereas the, the revers reaction a secretion by this sink reaction is the equivalent of starch synthesis from Glucose.

3.2 Methods used for Constructing the Model and executing Simulations.

The Model was implemented using Python (v.3.9.20) and more specific the python package cobrapy (v.0.29.0), a toolbox for constrained based modelling. The FBA optimisation problem, that is the Linear programming problem - has been solved using the Gurobi solver (v.11.0.3,(Gurobi Optimization LLC, 2024)).

Moreover, for the implementation of the dFBA a direct approach (DA)(has been chosen. Consequently, a solver for ordinary differential equations had to be chosen. For most of the simulations SciPy's solver for initial value problems solve_ivp (v.1.13.1) was used. To be more precise LSODA algorithm (Hindmarsh, 1982), a maximal step size of 0.5 hours, relative tolerance *(rtol)* of 1e-3 and absolute tolerance of 1e-6 (atol). LSODA is a numerical method

with adaptive step size and stiffness detection with auto switching, with BDF method for stiff problems and Adams for non-stiff problems.

3.2.1 Hierarchical optimisation:

A big problem of dFBA is that the linear programming problems don't have unique solutions. Thus, for every optimal solution of the objective, there is usually more than one flux distribution achieving that exact solution. This is problematic since non unique flux distributions, implies non unique fluxes for exchange reactions including those appearing in the right-hand side of the ODEs. Thus, for every iteration we could have several possible fluxes for the dynamically balanced metabolites, satisfying the optimal solution.

To overcome the problem of nonunique optimal flux distributions a Hierarchical optimisation structure was implemented. The fluxes appearing in the right-hand site of the ODEs are ordered in a priority list, then the model is optimized for these metabolites in the order of that list. After each solved optimisation problem, the objective value is used to constrain the next exchange reaction in the list. Not every flux distribution for a given optimal objective is unique, but every resulting optimal objective function value is unique. Therefore, every flux in the priority list will be unique and changes in flux will happen continuously with respect to time(Gomez et al., 2014).
Moreover, in this thesis the order in the priority list will be described as levels in the range [0,2] with Level 0 being highest priority and level 2 being the last optimisation in the list. For the implementation of the hierarchical optimisation procedure, the cobrapy function fix_objective_as_constraint need to be slightly modified. Witch is a function implementing the hierarchical optimization method. The function calls the method slim_optimize which unlike the usually use optimize function, raises an error if optimization failed, thereby terminating the whole simulation. More precisely optimisation fails, when the LP problem is infeasible this happens when no solution can be found under the given constraints. Therefore, the error_value parameter of the function needs to be set to 0 instead of Na. Thus, the function while return 0 instead of NA when an error occurs, which should not be occurring with the added LP-Feasbility Problem.

3.2.2 LP-Feasbility Problem:

Another problem is that the optimisation may become infeasible while the simulation moves on. There are two cases where this can happen, on the one hand side the problem is truly not feasible, that can be due to the current boundary condition of the model. However, on the other hand the solution can be very close to becoming infeasible, so the LP problem becomes infeasible in consequence of a lack of numerical precision.

In order to deal with this problem, the slack variables s_+ and s_- are introduced such that a LP-feasibility problem can be constructed (Gomez et al., 2014). Therefore, the hierarchical optimisation approach used with added LP-feasibility problem can be described in the following way.

Level 0:

$$S^* = \min \sum_{i=1}^{m} (s_{+i} + s_{-i})$$

$$\text{s.t. } Sv + s_+ - s_- = 0$$

$$LB\ v_i \le v_i \le Ub\ v_i$$

$$s_{+i}, s_{-i} \ge 0$$

For each metabolite in the Stoichiometric matrix S the two slack variables s_{+i} and s_{-i} are added and the sum over all slack variables minimized. Such that $S^* = 0$ if and only if there is a v satisfying $Sv = 0$ So in essence it finds a solution and returns a distance from a feasible solution.

The slack variables resulting from this optimization are being added as a constraint, to the model for the next two optimisation steps.

Level 1:

$$\mu^* = \max \mu$$

$$S^* \ge \sum_{i=1}^{m} (s_{+i} + s_{-i})$$

$$\text{s.t. } Sv + s_+ - s_- = 0$$

$$Lb\ v_i \le v_i \le Ub\ v_i$$

$$\mu \ge 0;$$

$$v \in R^n$$

Optimization of the growth rate μ. With the given slack variables introduced in level 0, each reaction is mass balanced and always feasible. That is under the given constraints for the lower and upper bounds of the exchange reactions. The resulting growth rate μ^* is added as a lower bound constraint to the model.

Level 2:

$$\max\ v_{\text{starch}}$$

$$S^*\ \geq \sum_{i=1}^{m}(s_{+i} + s_{-i})$$

$$\text{s.t. } Sv\ + s_+\ -\ s_- = 0$$

$$Lb\ v_i \leq v_i \leq Ub\ v_i$$

$$\mu\ =\ \mu^*\ \cdot 0.99$$

$$S \in R^{m \times n}$$

$$s_+\ , s_-\ \in\ R^m$$

$$v \in\ R^n$$

In the last step – Level 2 of the hierarchical optimization procedure, the model is optimized for starch production, while being constraint to achieve at least the growth rate of the optimisation step in level 1. More precisely, this constrained was slightly softened with the model being restricted to obtain only 0.99 of the growth rates. The consideration behind this was that in some case the constraint was too strict. This is for instance when N is not limiting, thus the model attaining very high growth rates starch would only be consumed never produced.

4. Results

4.1 Parameters of the Model

The tables below, lists all parameters chosen for the model. Furthermore, these are the parameters used for the simulation discussed in chapter 4.4 Results of the simulation.

Model parameters relating to **Hill & Michaelis Menten Kinetics**				
Symbol	Explanation	Value	Unit	Source
$V_{max\ NH_4}$	Maximal uptake rate for NH_4 from the environment	30000	$\frac{\text{nmol}}{\text{gDWh}}$	arbitrarily[1]
$\text{K}_{\text{m NH}_4}$	Michaelis-Menten constant for NH_4 uptake reaction	1	$\frac{\text{mg}}{L}$	arbitrarily[1]
$V_{\text{max } N_2}$	Maximal uptake rate for N_2 from the environment	7000	$\frac{\text{nmol}}{\text{gDWh}}$	Experimentally det. (Zamani et al., 2017)
$\text{K}_{\text{m N}_2}$	Michaelis-Menten constant for N_2 uptake reaction	1	$\frac{\text{mg}}{L}$	arbitrarily[1]
$V_{\text{max Amylase}}$	Maximal starch degradation rate	23400	$\frac{\text{nmol}}{\text{gDWh}}$	Calculated
$\text{K}_{\text{m Amylase}}$	Michaelis-Menten constant for amylase	5800	$\frac{\text{mg}}{L}$	Calculated (Doehlert et al., 1982)
Ub CO_2	Upper bound for CO_2 uptake by the model	180000	$\frac{\text{nmol}}{\text{gDWh}}$	Calculated
$v_{N-switch}$	the initial N_2 uptake flux of $\bar{V}_{\text{max } N_2}$	0	$\frac{\text{nmol}}{\text{gDWh}}$	Estimate taken from simulation

Model parameters regarding the Day - Night Cycle				
Symbol	Explanation	Value	Unit	Source
a	Parameter controlling day length	12	-	Calculated
c	Shift of the sin function along the time axis		-	
$v_{CO_2\ fix}$	Rate of CO_2 fixation in leaves ie. Flux / activity of Rubisco	230 000	$\frac{\text{nmol}}{\text{gDWh}}$	Calculated
I_0	Light intensity at down or dusk. Thus, start and end of daytime	0.125	-	Minimal light uptake necessary for optimization to be feasible estimated from ViNE model
Δt	Number of daylight hours	12	hours	-

ϕ^{-1}	Photon / CO_2 turnover ratio	9	mol mol^{-1}	(Zakhartsev et al., 2016)

Model parameters relating to the ODEs				
Symbol	Explanation	Value	Unit	Source
m_{N_2}	molecular weight N_2	$28 \cdot 10^{-6}$	$\frac{\text{mg}}{\text{nmol}}$	PubChem: 28.014 g/mol
m_{NH_4}	molecular weight NH_4	$18 \cdot 10^{-6}$	$\frac{\text{mg}}{\text{nmol}}$	PubChem: ammonium 18.039g/mol
m_{Gluc}	molecular weight glucose	$180 \cdot 10^{-6}$	$\frac{\text{mg}}{\text{nmol}}$	PubChem: D-Glucose 180.16 g/mol
$c^*(N_2)$	N_2 saturation concentration in water, under consideration of partial pressure in air	14	$\frac{\text{mg}}{\text{L}}$	(Battino et al., 1984)
$c^*(NH_4)$	NH_4 concentration found in soil	12.6	$\frac{\text{mg}}{\text{L}}$	(Pinton et al., 2016)
$K_L a$	Volumetric Mass Transfer Coefficient	12.5	$\frac{1}{\text{h}}$	arbitrarily[1] / (Promma et al., 2024)

[1]arbitrarily chosen parameter value

4.2 Calculations involved in construction of the model and for estimating the parameters

4.2.1 Determining Ub CO2

To limit the Amount of starch produced by the model an upper bound for the CO_2 uptake of the model was chosen. For estimating the parameter Ub CO_2 based on experimental data available in the literature 3 different approaches haven been considered. Finally, the arithmetic mean of the 3 parameter estimates, was chosen for the upper bound Ub CO_2.

a) Calculating the Biomass to leaf area ratio and then determine the flux based on infrared measurements of leave gas exchange.

Biomass to leave ratio B_{LA}, was calculated based on measurements by (Moreau et al., 2012) Using a tool to extract data from plots WebPlotDigitizer (Rohatgi, 2021). data for the low and high N-treated groups of *M. Truncatula* was extracted (n= 175). The data extracted from the graphs where, total biomass per plant B_{Plant} $[g/plant]$ and the leave area per plant LA [cm^2 / plant]. Which was obtained for both treatment groups.

The calculated average leaf area based on the extracted data (based on n=11, data points each) was:

Avg. LA cm² /plant	
high N	low N
823.5414	502.9916

The average total Biomass per Plant was:

Total Biomass g / Plant	
high N	low N
5.731039	3.722327

The Biomass to leave area ratio was thus calculated by:

$$B_{LA} = \frac{avg.\,LA}{B_{Plant}}$$

B_{LA} [cm² / gDW]:	
high N	low N
143.6984	135.1283

The photosynthetic CO2 assimilation rate, i.e. The CO_2 uptake of *M. truncatula* is given by $A_{ambient}$ = 33,32[$\mu mol/s\ m^2$], measured at PFD = 1800 [$\mu mol\ photons\ /\ s\ m^2$] and 5 weeks after inoculation of the plants. As described by Adolfsson *et al.* (Adolfsson et al., 2015).

Ub CO_2 was thus calculated by:

$$Ub\ CO_2 = B_{LA} \cdot A_{ambient}$$

$$Ub\ CO_2 = 172\ 000 \frac{nmol}{gDW\ h}$$

b) Based on growth and maintenance respiration measurements, conducted by ^{13}C labelling.

The specific net assimilation rate of CO_2, as determined by Lötscher and colleagues for *Medicago sativa,* is denoted as a_n. Moreover, a_n describes the net assimilation rate of new C taken up by the plant per unit C already assimilated into the Shoot of the plant. (Lötscher et al., 2004)

$$a_n = 0.16 \left[\frac{g\ C}{g\ C\ \cdot\ day}\right]$$

According to (Adolfsson et al., 2015) the carbon content of *M. truncatula* plants is 0.24 [g C / plant], where the average weight of the shoot is 0.61 [gDW / plant]. Thus, the Carbon share of the biomass dry weight B_C is given by:

$$B_c = \frac{0.24}{0.61} = 0.393 \frac{g\ C}{gDW}$$

A CO_2 uptake flux can thus be calculated by:

$$v_{g\ CO_2} = a_n \cdot B_C$$

$$v_{g\ CO_2} = \ 0.0629\ \frac{g\ C}{gDW\ \cdot\ day}$$

$$v_{mol\ CO_2} = 0.0525 \frac{mol}{gDW\ day}$$

$$v_{\ CO_2} = 5.25 \cdot 10^6\ \frac{nmol}{gDW\ day}$$

$$v_{\ CO_2} = 2.2 \cdot 10^5\ \frac{nmol}{gDW\ h}$$

$$Ub\ CO_2\ =\ 2.2 \cdot 10^5\ \frac{nmol}{gDW\ h}$$

The data on *M. truncatula* taken from to (Adolfsson et al., 2015) refers to the control group cohort, 5 weeks post inoculation of the plants. The Plants were grown in a 16h light /dark cycle with PDF = 350 -400 [$\mu mol\ photons/s\ m^2$] .

c) C-uptake calculated from the Growth rate of the ViNE model

The growth rate μ of the ViNE model, was reported as 0.044 [g $day^{-1}$$gDW^{-1}$] by diCenzo *et al.* (diCenzo et al., 2020).

Which is equivalent to:

$$\mu = 1.83 \cdot 10^{-3} \frac{g\ CO_2}{gDWh} = 1.53 \cdot 10^5 \frac{nmol\ CO_2}{gDW\ h}$$

The final solution for $UB\ CO_2$, is given by mean of the three solutions in a) , b) and c) .

$$Ub\ CO_2 = \frac{2.2 \cdot 10^5 +\ 1.72\ 10^5\ +\ 1.53 \cdot 10^5}{3}$$

$$Ub\ CO_2\ =\ 1.81 \cdot 10^5\ \frac{nmol}{gDW\ h}$$

4.2.2 Calculation of the CO2 fixation rate $v_{CO_2 fix}$

Based on a theoretical model for Photosynthesis, Poolman *et. al.* (Poolman et al., 2000), described the flux of the rubisco enzyme in [µmol / mg Chl h] as a function of the light activity [µmol / mg Chl h]. The fluxes of rubisco and the amount of irradiation (PFD) in (Poolman et al., 2000) where expressed with respect to the amount of Chlorophyl in mg. Let thus PFD_{chl} be the photon flux density and $v_{rubisco\,(Chl)}$ be the rubisco flux with respect to the amount of chlorophyll, both in [µmol / mg Chl h]. The respective data was extracted using WebPlotDigitizer (Rohatgi, 2021). Based on work of (Athanasiou et al., 2010) on the acclimation of *A. thaliana* to different light levels, the relationship of mg Chlorophyl per m² leave surface is given by:

$$Chl_{LA} = 134 \frac{\text{mg Chl}}{\text{m}^2}$$

For 4-week-old plants grown at light level of PFD =200 µmol / h m²

Let B_{Chl} be the chlorophyl share of the biomass in leaves and with the estimate of biomass to leave ratio, B_{LA} described at page 27 this is given by:

$$B_{Chl} = B_{LA} \cdot Chl_{LA}$$

$$= 1.43 \cdot 10^{-2} \cdot 134 \frac{\text{m}^2}{\text{gDW}} \frac{\text{mg Chl}}{\text{m}^2}$$

$$B_{Chl} = 1.95 \frac{\text{mg Chl}}{\text{gDW}}$$

The light intensity for a “usual” controlled environment of a *M. truncatula* experiment, is within the range of PFD = 200 – 600 [µmol photons/s m^2] according to (Barker et al., 2006). The PFD under which the plants have been grown in the experiments of Adolfson *et al.* was 350 – 400 [µmol photons/s m^2]. Which are the experimental conditions, for which the other parameters have been derived in previous calculations. The Rubisco activity was thus estimated, by firstly finding the PFD value with respect to leave surface area at about 400 µmol /h m² given the extracted data:

$$\text{PFD} = PFD_{chl} \cdot Chl_{LA}$$

$$\text{PFD} = 1089 \cdot 134 \frac{\text{µmol}}{\text{mg Chl h}} \frac{\text{mg Chl}}{\text{m}^2} \approx 1.46 \cdot 10^5 \frac{\text{µmol}}{\text{h m}^2} = 405 \frac{\text{µmol}}{\text{s m}^2}$$

The respective $v_{rubisco\,(Chl)}$ flux at $\text{PFD}_{\text{chl}} = 1089$ was:

$$v_{rubisco\,(Chl)} = 117 \frac{\text{µmol}}{\text{mg Chl h}}$$

The rubisco flux with respect to biomass at this light activity is thereby:

$$v_{CO_2 fix} = \ v_{rubisco\ (Chl)} \cdot \mathrm{B_{Chl}}$$

$$= 117 \cdot 1.95 \ \frac{\mu\mathrm{mol}}{\mathrm{mg\ Chl\ h}} \frac{\mathrm{mg\ Chl}}{\mathrm{gDW}}$$

$$v_{CO_2 fix} = 2.28 \cdot 10^5 \frac{\mathrm{nmol}}{\mathrm{gDW\ h}}$$

Other calculation supporting this estimate are given by:

For a PDF of $350 - 400$ [μmol photons / s $\mathrm{m^2}$], as described in the experiments of Adolfson *et al.*
this is equivalent to a photon uptake rate of $1.8 \cdot 10^6$ to $2.07 \cdot 10^6 [\mathrm{nmol/gDW\ h}]$.
With a maximal photon up take rate of $2.07 \cdot 10^6 [\mathrm{nmol/gDW\ h}]$, we would expect according to our light uptake model, a CO_2 fixation rate of:

With $\phi^{-1} = 9 \ \frac{\mathrm{nmol\ Photons}}{\mathrm{nmol\ CO_2}}$

$$\boldsymbol{v_{CO_2\,fix}} = \frac{2.07 \cdot 10^6}{9} \frac{\mathrm{nmol\ Photons}}{\mathrm{nmol\ CO_2}} \frac{\mathrm{CO_2\ nmol}}{\mathrm{gDW\ h}} = 2.3 \cdot 10^5 \frac{\mathrm{nmol}}{\mathrm{gDW\ h}}$$

These calculations can be further supported by findings of (Ma et al., 2014), indicating that the carboxylation activity of rubisco is about 1.28 times the net CO_2 assimilation rate observed in *A. thaliana.*

$$v_{CO_2} \cdot 1.28 = 1.8 \cdot 1.28 \ \frac{\mathrm{nmol}}{\mathrm{gDW\ h}} = 2.3 \cdot 10^5 \ \frac{\mathrm{nmol}}{\mathrm{gDW\ h}}$$

4.2.3 N-switch Parameter determination

Determination of the $v_{N-switch}$ and κ_{slope} parameters of the "N-switch".

The N-switch are the constrains on N_2 uptake, effective during the transition between NH_4 and N_2 as nitrogen source, when NH_4 approaches depletion. Let nsw_a be a tuple, defined as the cartesian product, formed by the combination of the parameters $v_{N\text{-switch}}$ and κ_{slope} of the N-switch described in the model at (eq. (3, p. 17).

$$nsw_a = \left\{ \left(v_{N\text{-switch}}, \kappa_{\text{slope}}\right) \mid v_{N\text{-switch}} \in \left\{\frac{v_{max}}{2}, \frac{v_{max}}{4}, 1\right\},\ \kappa_{\text{slope}} \in \{1, 0.1, 0.01, 0.005, 0.001\} \right\}$$

Simulations of all 15 combinations of nsw_a have been performed, runtime and results of have been compared. All test simulations have been performed with starting condition $u_0 = (1.0, 14.0, 12.6, 50.0)$. Moreover, the simulations were terminated when the concentration of NH_4 dropped below 0.01 mg/L. The average runtime was mean=79.5 and median=81.1minutes. The minimum runtime of all tested combinations was 72.9 min, which differs 2 σ or 2.5 σ from mean and median respectively (n=15, σ =3.26 min). Therefore, the runtime of $nsw_a = \{(1,\ 1),\ (1,\ 0.1)\}$ differed significantly from the other tested parameters. However, for most of the tested parameters there was no visible difference. The only combinations with $v_{\text{N-switch}}$ = 1 and κ_{slope} values in the range of [1,0.005], showed a difference in simulation results.

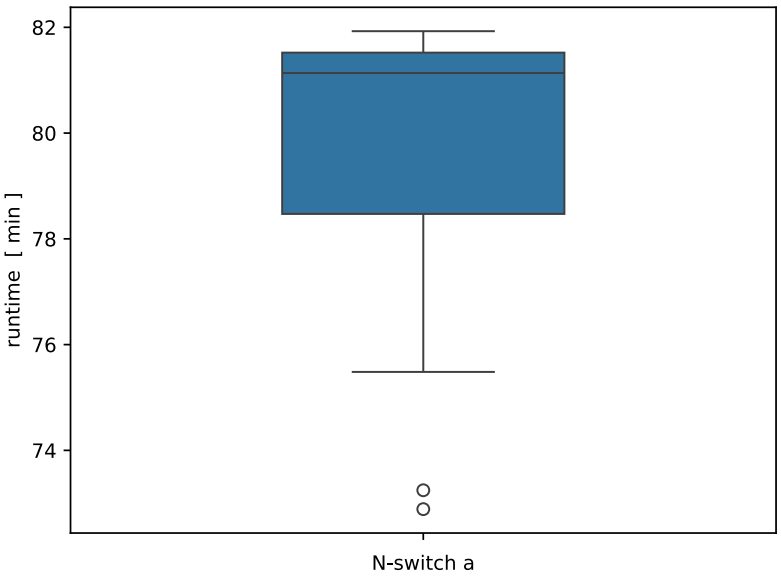

Fig. 4.1 Boxplot of execution time of N-switch a. The y-axis represents runtime in minutes until simulation termination, while the x-axis indicates the N-switch function type. The bar in the blue boxplot represents the median

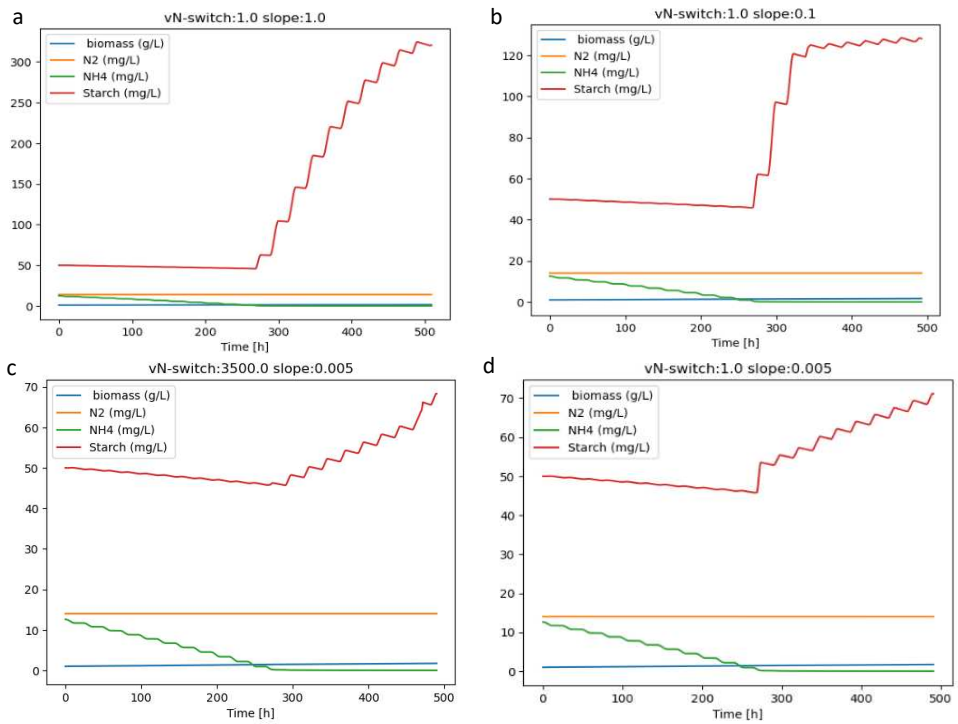

Fig. 4.3 Simulation result with different parameters tested. The figure a - d show the concentrations of dynamically balanced metabolites (x-axis), until simulation was terminated after 500 simulated hours (y-axis). The used parameters for nsw_a are in fig. titles. **a)** simulation result of the parameters nsw_a= (1, 1). **b)** simulation result of the parameters nsw_a = (1, 0.1). **c)** Simulation result is identical to all other nsw_a combinations not represented in this figure. d) simulation result of the parameters nsw_a = (1, 0.005)

For higher κ_{slope} values an increase of starch production is visible.

More precisely, with an increase of κ_{slope}, we can observe more and more step increase in starch production. Since, a higher slope parameter, delays the increase of the $\bar{V}_{max\,N_2}$ function value. The rise of starch production for these parameters' pairs, indicates an unwanted and artificial limitation of N-uptake during the time span where uptake is constrained by the N-switch. Thus, limiting the biomass production resulting in more starch production. As can be seen in the figures Fig. 4.2.b and Fig. 4.2.c below, for the parameters, $nsw_a = \{(1,\ 1),\ (1,\ 0.1)\}$ the biomass production is visible impaired at around 300 h simulated.

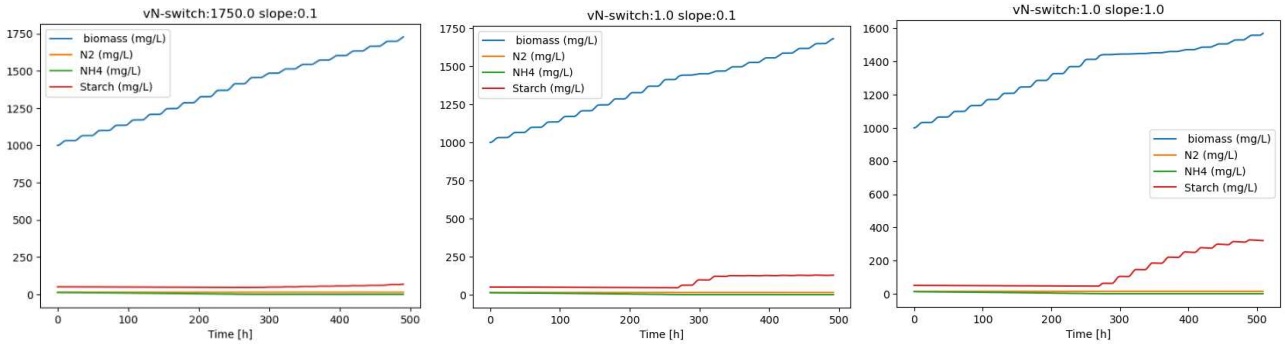

Fig. 4.2 Simulation result, depicting biomass production with different parameters tested. The figure a - c show the concentrations of dynamically balanced metabolites (x-axis), until simulation was terminated after 500 simulated hours (y-axis). The Biomass was expressed in [mg/L], to compare dynamics to starch. The used parameters for nsw_a are in fig. titles **a)** Simulation result is identical to all other nsw_a combinations not represented in this figure. **b)** simulation result of the parameters nsw_a (1, 0.1). **c)** simulation result of the parameters nsw_a (1, 1).

The significant difference in runtime can be explained by the limitation on $V_{max\ N_2}$, the uptake of N_2 is thus far delayed as there is no change in the N_2 concentration. Consequently, with reducing the number of ODEs the solver takes less time to solve the problem. From the reaming eligible parameter combination $\kappa_{slope} = 0.001$ with $v_{\text{N-switch}} = 1$. Seems to be the best choice, since there is no notable difference in comparison to the other parameters. Finally setting $v_{\text{N-switch}} = 0$ is the more parsimonious modelling approach, dropping an unnecessary parameter.

Alternative N-switch function:

After a more thorough consideration, it became apparent that the N-switch function as describe in the model (3), is not continuous over the entire interval in question as it was intended. Hence, a new function was designed. Let the function described in the model be function a, thus the newly described N-switch function will be called function b.

$$b = \frac{Ub_{NH_4}}{0.5 \cdot \mathrm{V}_{\max \mathrm{NH}_4}} \tag{4.1}$$

$$Ub_{NH_4} \geq 0$$

$$\overline{V}_{\max N_2} = \frac{\mathrm{V}_{\max \mathrm{N}_2}}{b^n + 1} - \frac{\mathrm{V}_{\max \mathrm{N}_2}}{b^n + 1} \cdot b^n \tag{4.2}$$

$$\overline{V}_{\max N_2} = \frac{\mathrm{V}_{\max \mathrm{N}_2}}{b^n + 1} \cdot (1 - b^n) \tag{4.3}$$

$$n > 0$$

The given b-function above fulfils all requirements set for function defining $\overline{V}_{\max N_2}$. First, it is continuous everywhere on the interval [0, $\mathrm{V}_{\max\ \mathrm{N}_2}$]. Second, the function increases faster and earlier then all nsw_a tuples would allow. This, of course given an appropriate value for the parameter n of the b function has been chosen.
Thus, a second experiment, has been conducted to compare parameters of the two N-switch functions.

The initial concentration u_0 and the parameters nsw_a for function-a were chosen as described in the first experiment. However, this time the simulations were terminated once $c(NH_4) \leq 0.001$ mg L^{-1}, therefore the simulation runtime was longer for all tested parameters. The b-function has been tested with the n parameters n = (0.5,1, 1.5, 2, 3, 4). Lastly, a simulation with no implemented N-switch function has been conducted as a control.

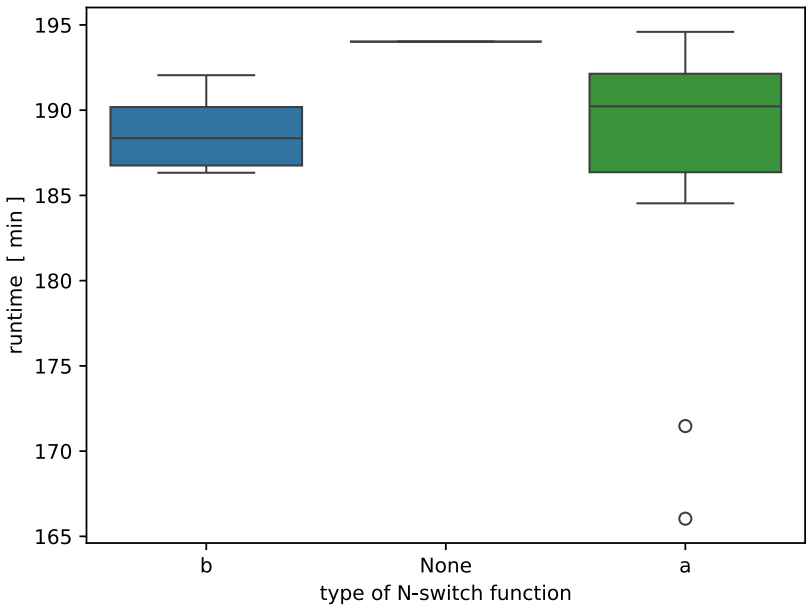

Fig. 4.4 Boxplot comparing the execution time of N-switch functions. The y-axis represents runtime in minutes until simulation termination, while the x-axis indicates the N-switch function type. The blue boxplot represents the b-function runtimes, the green boxplots correspond to type a-functions, and the bar between two boxplots depicts the runtime of the control with no switch. Whiskers generated by IQR·1.5.

From the results depicted in Fig. 4.4 we can generally conclude, with respect to runtime the simulations with N-Switch function outperformed, the model without such a function. Furthermore, the function b performed on average better than the functions in group a. That is if one compares the blue and green boxes, with group a showing more spread, and a skewed distribution as the median bar is not in the middle of the box, indicating that most of the values are located closer to the 75% quartile than to the 25 % quartile of the box plot. Lastly, in group a-we can see two outliers of the parameter combinations nsw_a = (1,1) and (1,0.1). This is consistent with the findings in the previous experiment. With respect to simulation results thus now particular differences between the simulations where visible. With exception for nsw_a = (1,1) and (1,0.1) where we can see same behaviour as in the previous experiment, depicted in Fig. 4.3.a.

A follow up experiment was performed where same set of functions and under identical conditions were tested, with the exception that the simulations were terminated when $c(NH_4) \leq 0$. However, because of a technical failure test series was aborted prematurely and the data stored lost due to a disk failure. Nevertheless, what collude be observed was that the execution time for each simulation series was in the rage of 15 -16 hours runtime. The simulation for all b-functions, the control group and function a with nsw_a = (1,1), was performed before apportioning of the test series. Most strikingly was that nsw_a = (1,1) which was consistently an outlier with respect to runtime in previous experiments. Did not differ

visible but was somewhere with range runtimes of the other functions tested. Lastly, the control with no N-switch function, was the only function exhibiting erroneous results thus starch concentrations spiking and dropping notably below 0. Most likely caused by numerical issues encountered by the solver. This absolutely fits the expectation and confirms the purpose of these N-switch functions. Nevertheless, with all function of group a and b are expected to by continues in the interval $[0.5 \cdot V_{\max N_2}, V_{\max N_2}]$, thus continuous at the end of the tested simulation interval. It is unexpected that function a outperforms function b at the beginning of the simulation. With only b function having a $\overline{V}_{\max N_2}$ defined continuous at the interval at the beginning of the simulation. it can be speculated that this higher performance of this function $nsw_a = (1,1)$, at beginning is due to the virtual absence of N uptake. There for the most computation intensive steps are just delayed. Or it is the behaviour of these function very close to 0 or the machine precision.

However, all these findings suggest that the purpose of these functions is not performance but stability, thus correctness of the simulation results. Unfortunately, the results of the a-function, could not be validated or compared anymore.

4.2.4 The parameter a of the function modelling day night cycle.

Given the function describing the light intensity during the day:

$$i(t) = \max\left[(1-a)\cdot sin\left(\frac{\pi}{12}\cdot t\ + c\right) + a, 0\right]$$

With c = 0

$$i(t) = (1-a)\cdot sin\left(\frac{\pi}{12}\cdot t\ \right) + a$$

$$sin\left(\frac{\pi}{12}\cdot t\ \right) = \frac{I_0 - a}{1-a}$$

$$\frac{\pi}{12}\cdot t\ = arcsin\left(\frac{I_0 - a}{1-a}\right)$$

$$\frac{\pi}{12}\cdot t\ = arcsin\,(k)$$

With t = 12

$$\frac{\pi}{12}\cdot 12\ = arcsin\,(k)$$

$$\pi = \ arcsin\ (k)$$

With t_1 being the start of the day and t_2 being the end of the day.

$$\frac{\pi}{12} \cdot t_2 = \pi \ - \ arcsin(k)$$

$$\Delta t = t_2 - t_1$$

$$\frac{\pi}{12} \cdot \Delta t \ = \ \pi \ - \ arcsin(k) \ - \ arcsin(k)$$

With arcsine(k) = z

$$\Delta t = \frac{12(\pi - 2\ z)}{\pi}$$

$$\Delta t\pi \ - \ 12\pi \ = -24z$$

$$z \ = \frac{12\pi \ - \Delta t\pi}{24}$$

$$z = \frac{\pi}{2} - \frac{\Delta t\pi}{24}$$

$$sin(z) \ = \ k$$

$$sin(z) \ = \frac{I_0 - a}{1 - a}$$

$$a \ = \frac{I_0 - sin(z)}{1 - sin(z)}$$

$$sin\left(\frac{\pi}{2} - \frac{\Delta t\pi}{24}\right) = -\,cos\left(\pi - \frac{\Delta t\pi}{24}\right)$$

$$a = \frac{I_0 + cos(z)}{1 + cos(z)}$$

$$a = \frac{I_0 + cos\left(\pi - \frac{\Delta t\pi}{24}\right)}{1 + cos\left(\pi - \frac{\Delta t\pi}{24}\right)}$$

Thus, describing the relationship between length of the day and parameter **a**, as shown in eq.(5) on page 18 .

4.2.5 Vmax Amylase

In their work on rhythmic oscillations of starch concentrations in *M. Sativa* nodules, Henson and colleagues (Henson et al., 1986) measured the ß-amylase activity. The measurements were conducted on Plants exposed to a dark light cycle of 12h each. Thus
The MESOR of the analysed amylase activity was given by:

$$V_{amylase} = \frac{2.3\ \mu\text{mol maltose}}{h\ gFW}$$

Which is (stoichiometrically) equivalent to:

$$V_{amylase} = \frac{4.6\ \mu\text{mol glucose}}{h\ gFW}$$

The MESOR, (Midline Estimating Statistic of Rhythm) is a measure of central tendency adjusted to a function of circadian rhythm, thus to a 24-hour period(Cornelissen, 2014).
To Calculate the fluxes with respect to dry weight (gDW), $B_{FW/DW}$ the fraction of FW that is DW, was calculated. This has been done based on the measurements of Hidra *et al.* (Hdira et al., 2021). The reported average Biomass [g] for the *M. truncatula* control group was given by

avg. [g]	Fresh weight (FW)	Dry weight (DW)
Arial	4.21	1.05
Root	2.88	0.41
Total Plant	7.09	1.46

Where AFW and RFW are the avg. Arial and Root fresh weight of a plant [g FW / plant]. Consequently, ADW and RDW are the Arial and Root dry weight [gDW / plant].

$$\mathrm{FW_{Total}} = \mathrm{AFW} + \mathrm{RFW}$$

$$\mathrm{DW_{Total}} = \mathrm{ADW} + \mathrm{RDW}$$

$$B_{FW/DW} = \frac{\mathrm{FW_{Total}}}{\mathrm{DW_{Total}}} = 0.2\ \frac{gFW}{gDW}$$

Moreover, the Arial $B_{FW/DW}$ was 0.25 [gFW/gDW] and the Root $B_{FW/DW} = 0.14$ [gFW/gDW] . Additionally, other experimental values indicate that the relative water content [%] of *M. truncatula* leaves is about 80% (Yousfi et al., 2010).

Thus, assuming 1 gFW = 0.2 gDW then a V_{max} for the starch mobilization rate is estimated to be:

$$V_{max\ Amylase} = \frac{23\ 400\ \mathrm{n}mol\ glucose}{h\ gDW}$$

Correction:

$$B_{FW/DW} = \frac{\mathrm{FW_{Total}}}{\mathrm{DW_{Total}}} = 4.86\ \frac{\mathrm{gFW}}{\mathrm{gDW}}$$

$$V_{max\ Amylase} = \frac{5.75\ \cdot\ 10^5\ \mathrm{nmol\ glucose}}{\mathrm{h\ gDW}}$$

However, this is not the value used in the model!

4.3 Construction of the model

The first simulations performed with the model included 3 metabolites Biomass, N_2 and O_2. Further Major changes the in development of the model where:

- Changed numerical method for solving ODEs
- Removed dynamically balanced O_2 uptake from Model
- Rescaling the whole model

In the development process of the model the numerical method to solve the ODEs was changed from BDF to LSODA. Moreover, a benchmarking of numerical integration methods for ODE models of biological systems (Städter et al., 2021), indicated that although for most models BDF outperformed LSODA, the later one had almost 3 times lower failure rate then BDF. This is consistent with the experiences in the development of this model. Although initially a switch from BDF to LSODA was taken because of better performance, at later model development steps BDF, failed to solve the problem and was not able to correctly replicate the simulation results obtained with LSODA. Further, the authors highlighted absolute and relative tolerances as important hyperparameters. Recommending tolerances between 10^{-8} and 10^{-16} as appropriate and suggesting that tolerance should not be relaxed more then to a value of 10^{-6}. However, when more strict tolerances where applied most simulations tested failed or showed almost intractable executions times.

Rescaling of the model.

Since fluxes in the ViNE model where assumed to be given in mmol gdW^{-1} h^{-1} , thus dFBA model was scaled accordingly. However, the fluxes are given in nmol gdW^{-1} h^{-1} and growth rate μ is given in mg gdW^{-1} h^{-1} instead of h^{-1}.

Dynamically balanced O_2 uptake

Within the process of constructing the model attempts of modelling other metabolites like O_2 have been taken. The dynamically balanced O_2 uptake reaction was removed from the model, since its servery destabilised the simulation. A Flux variability analysis revealed that, at the time steps which where most problematic for the solver, there where multiple O_2 uptake fluxes for fulfilling the objective, even though the range of possible fluxes was very tight. The absolute difference of lower and upper bound of that range of possible fluxes was 0.002 nmol gDW^{-1} h^{-1}. Also worth noticing is that initially CPLEX was used to solve the LP problem but later switched to the use of Gurobi, dependent on software installation of the used computer.

4.4 Results of the simulation

The simulation was performed with the parameters as described in chapter 4.1. Let u_0 be the vector of the initial concentrations of the dynamically modelled metabolites

$$u_0 = \big(x, c(N_2), c(NH_4), c(starch)\big)$$

Then, the initial concentration is given by u_0= (0.1, 14.0, 12.6, 5.0) $[\frac{g}{L}, \frac{mg}{L}, \frac{mg}{L}, \frac{mg}{L}]$.
The simulated time span was, $t_{sim} = (0;\ 1.3 \cdot 10^4)\ [h]$. The overall runtime of the simulation was 4d 11h 18 min. The concentration at the end of the simulation was given by $u_{t_{end}} = (2.1 \cdot 10^4, 0.2, 0, 1.2 \cdot 10^7)$.

	X $\frac{g}{L}$	$c(N_2)\frac{mg}{L}$	$c(NH_4)\frac{mg}{L}$	$c(starch)\frac{mg}{L}$
u_0	0.1	14	12.6	5
$u_{t_{end}}$	$2.1 \cdot 10^4$	0.2	0	$1.2 \cdot 10^7$

Further the solver needed to go trough $1.6 \cdot 10^5$ $i$terations, to run the simulation under conditions described above. Moreover, an Iteration is here defined as every time the solver calls the function defining the dFBA model, to define the state of the model when the solver takes a step. More precisely one iteration involves the solution of 3 FBA problems.

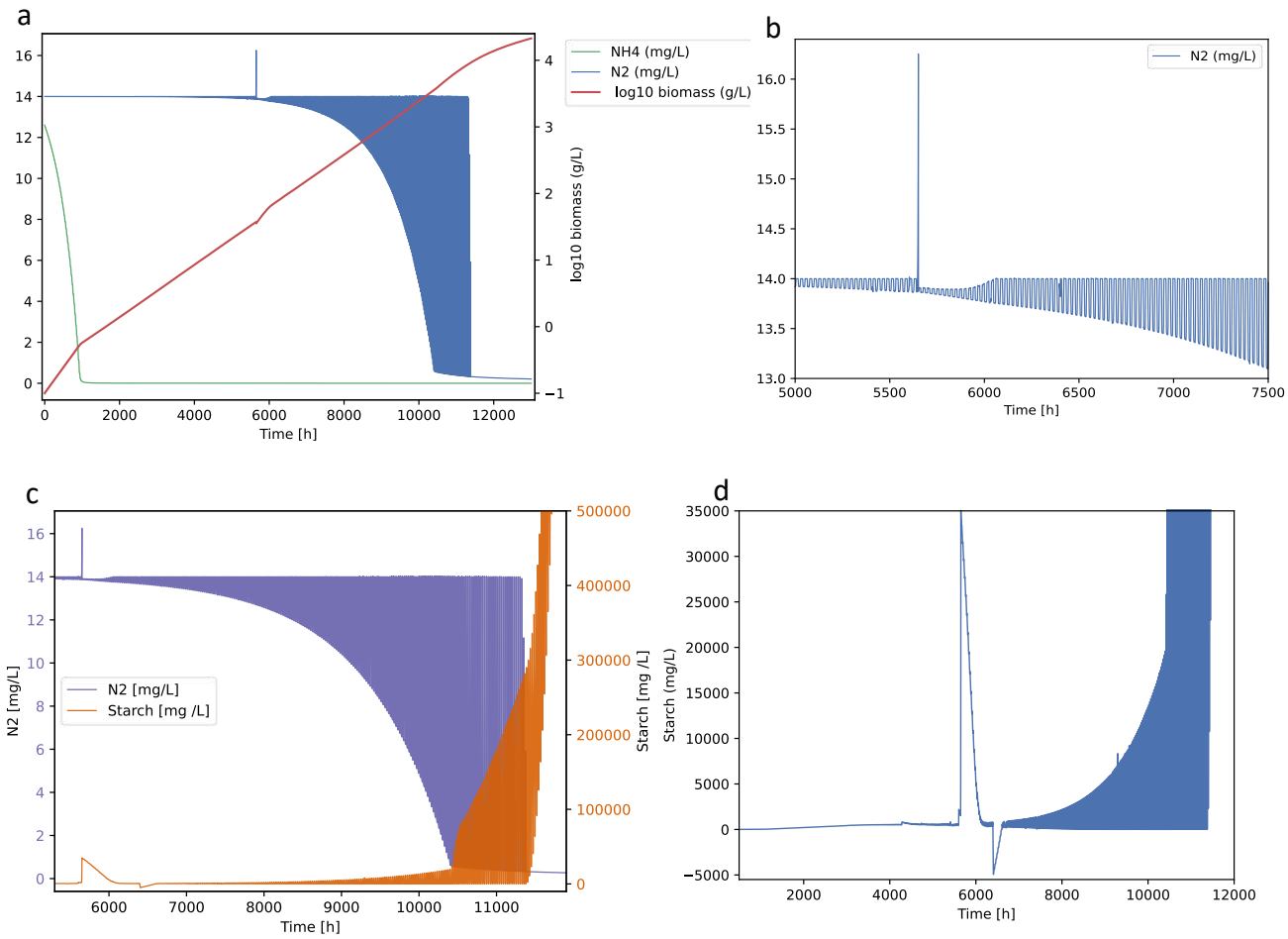

Fig. 4.5 The results of the simulation. The plots show the concentrations of the dynamically balanced metabolites in the system environment (y-axis), over the respective simulation time in simulated hours (x-axis). **a)** N-source and Biomass, the scale on the left y-axis depicts the concentration of N_2 and NH_4 in [mg / L], the right scale shows the log transformed Biomass concentration in [g/L]. The plot illustrates the whole simulation time of 13,000 hours. **b)** A magnified view of the N_2 peak observed in (a) at 6000 h, with the y-axis representing N_2 concentration in [mg/L]. **c)** Starch and N_2, scale on the left shows the N_2 (purple) right the starch concentration (orange) both in [mg/L]. The plot shows section of the results, scale depicting starch is limited to $5{\cdot}10^5$ [mg/L] instead of the full $1.2{\cdot}10^7$ [mg/L], at the end of the simulation. plot shows only segment of simulation time **d)** A magnified view of the starch peak observed in c) around 6000 h, with the y-axis representing starch concentration in [mg/L].

In Fig. 4.5a, the concentration of the two N-Sources NH_4 and N_2 and the log transformed biomass are shown. The NH_4 concentration depletes after ca. 1000 h (green) of simulation, this is also reflected by the biomass, showing a slight decrease in its slope at $t_{sim} \approx 1000$ h. The blue area, which is the N_2 concentration shows the fluctuating pattern of the N_2 level in the day-night cycle. The oscillations of the curve exhibit increasing amplitude. The replenishing of N_2 in the system environment is solely diffusion controlled; It is thus recovering back to $c(N_2)^*$ at nighttime, when there is no or only very little growth. Furthermore, as shown in Fig. 4.5c the daytime N_2 levels decrease further as biomass increases, with the minima of the daily oscillating N_2 curve following a sigmoidal trend. Which is to be expected since reactions described hill kinetics, which usually also describes cooperative binding substrates, follow a sigmoidal trend. While reactions modelled by MM-kinetics follow a hyperbolic trend.

The plot in Fig. 4.5b shows a spike in $c(N_2)$ far exceeding $c(N_2)^*$, which is only possible in the

case of a solver error. Additionally, the daily maxima of the oscillating curve—representing $c(N_2)$ at night—drop below $c(N_2)^*$. Over several day-night cycles (~500 hours simulated), the N_2 level does not recover over night to its former level $c(N_2)^*$.

As depicted in Fig. 4.5c, the relationship between starch production and N_2 consumption can thus be characterized as follows: starch production increases when nitrogen availability decreases. The production of starch, i.e., the level of the starch pool, rises significantly when the minima of the fluctuating $c(N_2)$ curve approach the equilibrium $[N_2]$.

Furthermore, when starch production exceeds nighttime consumption, this leads to a starch overproduction, thus its concentration never reaches zero. This overproduction occurs precisely when the N_2 concentration ceases to oscillate in a diurnal cycle ($t_{sim} \approx 11{,}500$ h). A stop of oscillation in N_2 concentration at exactly that moment, is most likely to be explained as there is sufficiently starch available at nighttime, to support growth in the dark phase. Consequently, N is taken up by the system constantly and never replenished back to its saturation state $c(N_2)^*$, as shown in Fig. 4.5a and Fig. 4.5c

In Fig. 4.5c a sudden peak in starch level appears, at approx. same time as in c(N). It is however not clear if it is reflecting the current situation of the system in particular the level of N_2 or, if it is due to a failure of the solver. However, the consecutive drop of starch concentration below zero Fig. 4.5d can only be caused by a solver failure, probably caused by numerical instability of the method.

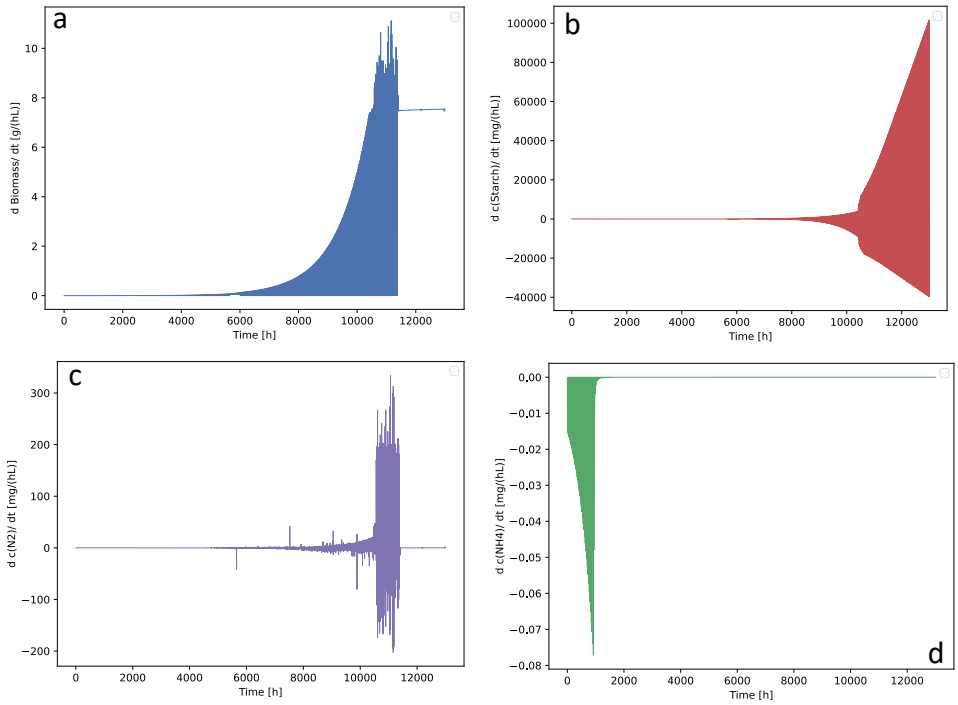

Fig. 4.6 Derivatives of dynamically modelled metabolite with respect to time.
a) derivative of Biomass concentration in system environment with respect to time (y-axis). **b)** derivative of Starch concentration in system environment with respect to time (y-axis). **c)** derivative of N_2 concentration in system environment with respect to time (y-axis). **d)** derivative of NH_4 concentration in system environment with respect to time (y-axis). The derivatives are the current hourly change rates of state of the system (concentrations), at respective hour simulated (x-axis).

Fig. *4.6* , shows Derivatives of dynamically modelled metabolites, thus the hourly change rates of metabolite concentrations in system environment. A important difference of the derivatives as shown above in Fig. 4.6 and fluxes as shown e.g. in Fig. *4.7*, is that the derivatives are scaled with respect to the systems volume. In Fig. *4.6*.a we see the change rate of the biomass, reflecting the exponential growth of the biomass. It expresses the production of biomass per hour in absolute numbers. Since there is not enough starch available at nighttime yet, we see strong fluctuations in that time span of t_{sim}=[6000,11500]. The change rate drops with the N2 concentration approaching its equilibrium; Then, the rate stops oscillating since the system has reached a state where there is enough starch available at nighttime, to support growth. The Fig. 4.6.b shows the change rate of the starch concentration. At about t_{sim} = 11000 the starch consumption and production accelerates strongly however the speed of production increases faster, then the consumption, thus initiating over production of Starch. Moreover, the change rate for the concentration of N_2 is described by Fig. *4.6*.c. The larger amplitude observed here indicates a faster N-uptake at the daytime, when it is utilized leading to greater depletion of the N pool at daytime as can be seen in Fig. 4.5a and Fig. 4.5c. Consequently, more N can be replenished by diffusion at nighttime, which explains an increase in N change rate. In other words, during the period of high amplitude between t_{sim} = [10,000, 11,500], an increase of the N_2 change rate does not indicate that additional nitrogen is being produced; rather, more nitrogen is required or able to diffuse into the system environment at night to compensate for the increased daytime depletion. Most importantly, at the end phase of the simulation starting at about t_{sim}=11500 h, the change rate remaining at zero indicates that the N concentration int environment has reached its equilibrium concentration [N_2]. Thus, the rate of N-uptake equals the rate of N diffusion replenishing the Nitrogen in the environment. Finally, Fig. 4.6.d illustrates the model's NH_4 consumption. The data show that NH_4 uptake increases as biomass in the system environment grows, until NH_4 is fully depleted at approximately t_{sim} = 1,000.

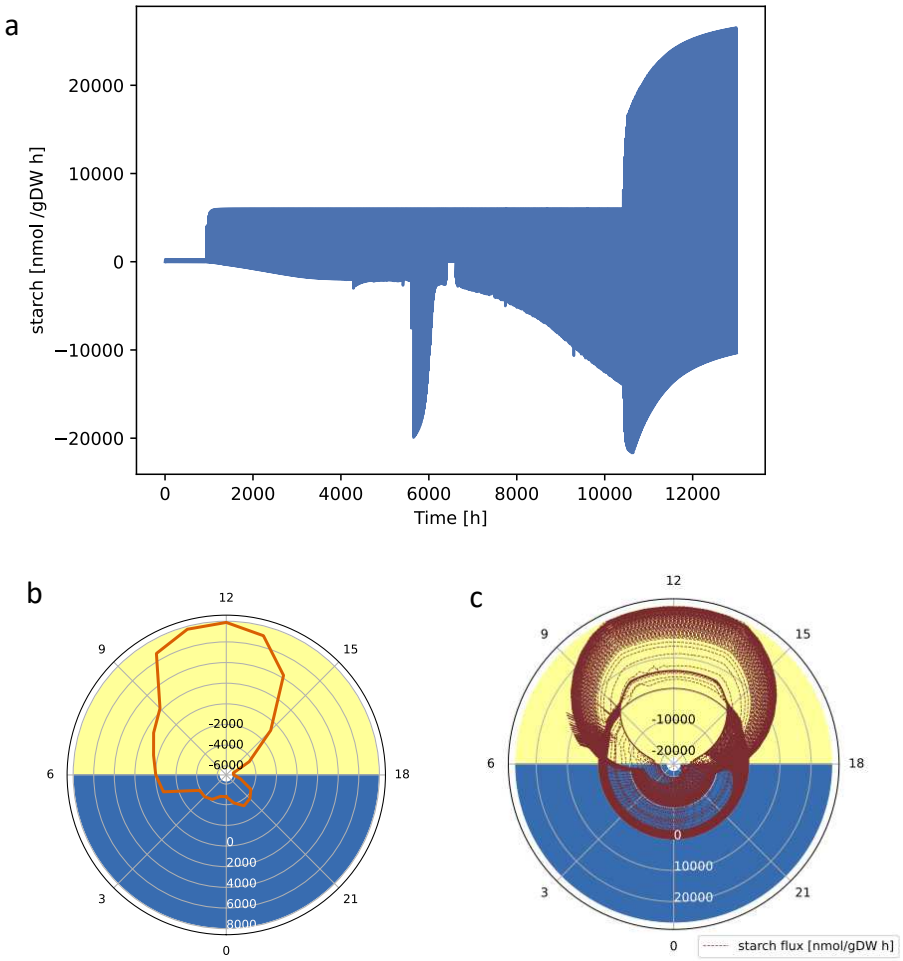

Fig. 4.7 Strach fluxes over span of the whole simulation. The figures show the Starch fluxes of the ViNE model, positive fluxes are starch production, negative fluxes are starch degradation. **a)**. starch fluxes in nmol gDW^{-1} h^{-1} (y-axis), at given simulated hour (x-axis). **b)** Mean starch flux grouped by daytime. The line in orange depicts the avg. flux in in nmol gDW^{-1} h^{-1}. **c)** The line in brown shows starch flux in span of whole simulation at respective day time. For b) and c) the daytime within simulation is given by the numbers around edge of the radial plots. The blue shaded area represents nighttime, the yellow highlighted area the daytime.

In Fig. 4.7.a, the increase in starch production around t = 1,000 is caused by NH_4 depletion in the system environment. Additionally, the drop in starch flux at t = 6,000 hours, which corresponds to a sudden spike in starch uptake, appears to be a reaction to an error—specifically, the abrupt increase in starch concentration visible in Fig. 4.5d at the same time point. Furthermore, the sudden drop in starch flux around t = 11,000 hours in Fig. *4.7*.a suggests that more starch may be utilized at night in response to a corresponding increase in starch production during the day. The availability of carbon at night leads to increased growth rates and nitrogen consumption, preventing nitrogen replenishment. Consequently, starch consumption follows a hyperbolic decrease at night, likely due to reduced nitrogen availability, which limits growth and, in turn, lowers starch consumption. At the same time, starch production exhibits a hyperbolic increase, likely due to nitrogen limitation on growth—causing excess carbon to be converted into starch instead of biomass.
The radial plots in Fig. 4.7 illustrates the average starch flux at different times of the day. The

flux values from the simulation (160,000 data points over 541 days) were grouped into 1-hour bins, with the mean flux calculated for each bin and plotted Fig. 4.7b-c.
In Fig. 4.7.b, a sharp decrease in starch flux is observed at 15:00, indicating that consumption begins before nighttime, while production appears to start at dawn. Moreover, starch consumption peaks around dusk, with the uptake flux being lower on average for the rest of the night. This might suggest that starch is not consumed at a constant rate throughout the night. But might just be caused by variations in consumption at nighttime occurring across the simulation. The plot in Fig. 4.7.b depicts all starch fluxes over the simulated time, the biggest variation is visible around dusk, particularly between 15:00 and 19:00.

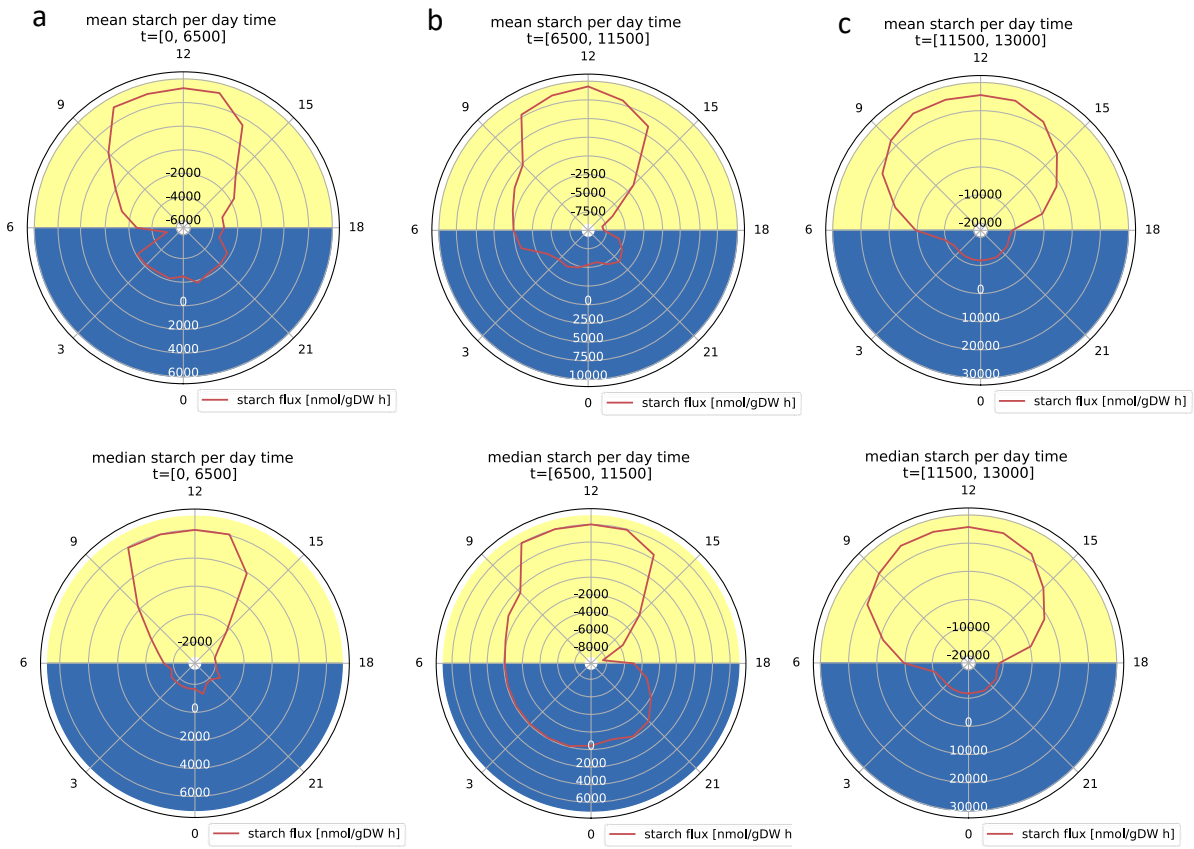

Fig. 4.8 Mean and median starch flux grouped by daytime in different phases of the simulation. The line in red depicts the avg. flux in in nmol gDW^{-1} h^{-1}. The daytime within the simulation is given by the numbers around edge of the radial plots. The blue shaded area represents nighttime, the yellow highlighted area the daytime. **a)** Mean and median starch flux grouped by daytime over simulation interval of t_0=0 to the 6500th hour simulated. **b)** Mean and median starch flux grouped by daytime beginning with 6500th hour simulated until hour 11500. **c)** Mean and median starch flux grouped by daytime of the interval starting with t_{sim} > 11500 until end of simulation at 13000 hours simulated.

In the early phase of the simulation (Fig. 4.8.a) one can see that starch is being utilized over the whole span of the nighttime, however the production of starch starts relatively late in day time, that is starch is being consumed till about 2h after dawn. Except for the outlier visible in Fig. *4.8*.a at t_{sim}= 6000 hours, there is no visible difference for the two radial plots of the early simulation phase. In the middle phase of the simulation (Fig. *4.8*.b), a difference between median and mean radial plot at can be noticed at nighttime. In the plot where

starch was grouped by median (bottom), we can see that starch uptake already stops at midnight. Indicating that there is not enough starch being produced at daytime, to support biomass production throughout the nighttime.

Nighttime is where light intensity is that low such that no or not enough CO_2 can be fixed thus no growth is possible. Consequently, no feasible solution can be achieved and there is no metabolic activity of the model possible without uptake of starch as carbon source
The minimal required light uptake by the model, is the amount of light (nmol photons / gDW h) or any other C-source necessary for the optimization to be feasible. The minimal uptake required (without any starch available), by the model was determined to be 12.5% of maximal light intensity. Consequently, dawn and dusk are the two points on the sine function where it equals i(t)=0.125 within each period. Thus, the sine function was calculated such that the distance between these two points exactly equals the day length Δt. However as e.g. shown by Fig. *4.7*.b, the avg. starch flux by daytime indicates, that in fact not the whole timespan of $\Delta t = 12\ \mathrm{h}$ is being utilized by the model to produce starch. Rather the time span of 6:00 to 15:00 (9 h), shows a positive starch flux that is a net production, as opposed to consumption of starch represented by a negative flux value.
That behaviour, of starch degradation starting at late afternoon can observed throughout most of the simulation. This is most likely caused by the way hierarchical optimisation is structured with biomass production being prioritized over starch production. Therefore, to achieve the highest growth rate possible starch is being consumed if available even though not it is not necessary, since photosynthetic activity provides enough carbon to sustain metabolism. Consequently, starch is not only being consumed when strictly necessary and depletes too early in the night. This is of course not as intended, but it is however questionable if this effects the simulation result anyway. To be more specific it could be argued that this is just relocation of biomass production potential. More growth in twilight phase in return for less or no productivity in the second half of the night. In this dFBA implementation the growth rate is only optimized instantaneously, that is for the current time step and not optimized over the whole time horizon of the simulation. Thus, using dynamic objective function as possible with dynamic optimization approach, we might see a different behaviour. With the optimization of the model not being blind for future.

In the Late phase (Fig. 4.8.c) we can see that starch consumption starts later in the day, a possible reason for that might be with N being more limiting to growth. Consequently, less carbon uptake is necessary to support biomass production, achievable under given N limitation. Moreover, we can see that the production of starch precisely starts at dawn at 6:00. Considering the starch over production at the end phase of the simulation this unlikely caused by an exhaustion of starch reserve, but most likely due to the fact that the low growth rate at end of the simulation can already be sustained with I_0 = 0.125 light intensity available at dawn. Consequently, we can also see that the uptake phase of starch starts later in the day then in earlier phase of the simulation.

4.4.1 Solved Problems per hour simulated

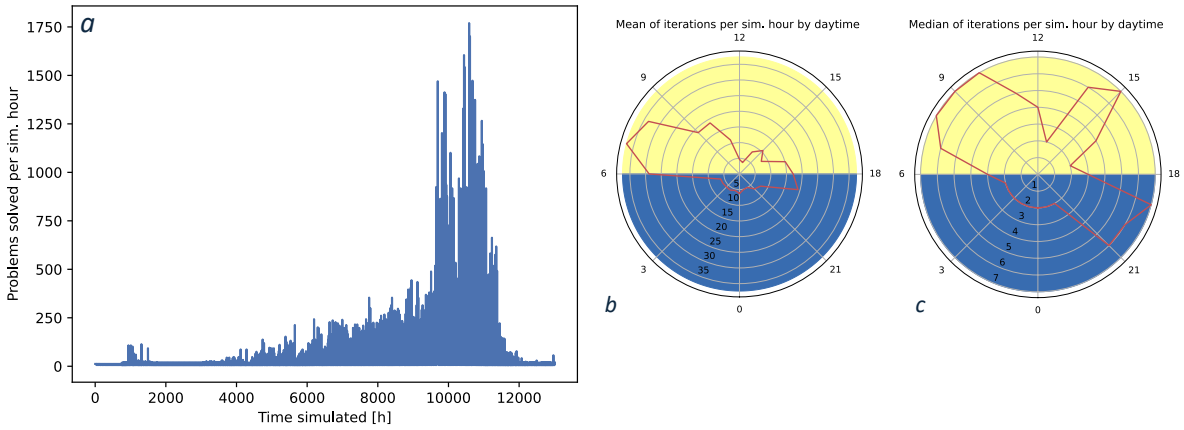

Fig. 4.9 Number of iterations per simulated hour.
a) the number of iterations per bin of size 1h (x-axis), plotted against respective simulation time, the simulated hour (y-axis). The binned data from a) was used to generate plots in b) and c). The line in red depicts the mean or median number of iterations per daytime over the whole span of time simulated. The daytime within the simulation is given by the numbers around edge of the radial plots. The blue shaded area represents nighttime, the yellow highlighted area the daytime. **b)** mean number of iterations grouped by daytime. **c)** median number of iterations grouped by daytime.

The runtime behaviour of the simulation was analysed, that is with respect of increases in iterations necessary to calculate the current state of the system. Furthermore, to locate "regions" with in the span of the simulation that posed problematic for the solver.
In order to pursue this analysis, the documented iterations were grouped with respect to their simulation time in bins of size 1h. Worth noting is that the solver can take step back in time to find the right trajectory for the solution. The plot in Fig. 4.9.a shows the total number of iterations per simulated hour. To the left of the plot at about 1000 - 2000 h a strong increase in computations is visible. This coincides with the depletion of NH_4 and the switch to N_2 as Nitrogen source. This is consistent with the fluctuations in growth rate in Fig. 4.10.b. The number of iterations necessary for the solver per simulated hour, increase strongly within the time span t_{sim}= [4000, 10000], at t_{sim} = 4000 h is where the daily growth rate shows an sudden increase as seen in Fig. 4.10.b, at around t_{sim}= 5500 the amplitude of the N_2 fluctuations start to increase notably, as shown in . Moreover, a drop an increase in the trend is also visible at t_{sim} = 6000 h this is where the solution of the simulation shows unreasonable spikes in concentration of environmental metabolites, indicating possibly numerical problems of the solver. However, the plot in Fig. 4.9.a doesn't indicate extreme changes in computations necessary at this step in the simulation. More general speaking a local increase is visible, but it is not an event differentiating from the global trend.
The biggest increase can be seen at t_{sim} = [10000, 12000]. This is where we can see an increase in oscillation in starch and N_2 levels. A rapid decline in iterations taken can be seen at the end of the simulation where N_2 is close to equilibrium.
In the radial plots of Fig. *4.9*.b and Fig. *4.9*.c we can see the mean and median number of iterations by daytime over the whole simulation time. In comparison the plot to the left Fig. *4.9*.b, shows a much higher average of iterations in the first quarter of the daytime, then the plot to the right. Therefore, we can say that not all calculations in the span of the simulation happen uniformly during same time of day. A subdivision of the simulation time into shorter

intervals, depicting different phase of the simulation with a dedicated radial plot would be beneficial for further inspecting the differences in Fig. 4.9.b and Fig. 4.9.c . #However, due to a technical failure almost all data necessary for further analysing the results of the simulation is lost.
Of course, the mean as a measure of central tendency is more sensitive to extreme values then the median. Therefore, the outlier Fig. 4.9.b between 6:00 and 9:00 might indicate that much of the activity at t_{sim} = [10000, 12000] shown in Fig. *4.9*.a, happens early in the day. Most likely, this is caused by rather abrupt decline in N_2 concentration in the environment, coupled with starch production and increased biomass production.

Overall, most of the activity happens during the daytime, also visible in both plots b and c is a dent during the noon. A behaviour that that can also be seen in the amount of CO_2 evoked in N-fixing nodules (see Fig. 4.14). Worth noticing is that the max step size of the solver is set to 0.5, this means that the lower bound iterations per simulated hour is exactly 2. As can be seen at nighttime in Fig. *4.9*.c.

4.4.2 The growth rate of the model

The fluctuations of the growth rate in the day night-cycle can be seen in Fig. 4.10

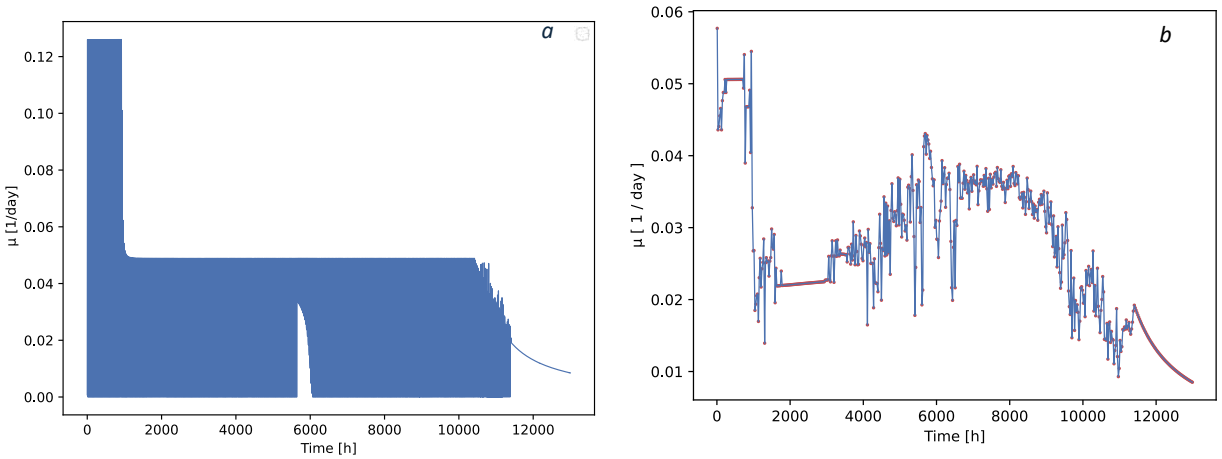

Fig. 4.10 Growth rate of the model. a) The growth rate of the model at each iteration [day^{-1}], calculated from the growth rate in mg gDW^{-1} h^{-1} achieved by the ViNE model, thus showing fluctuations in day night-cycle (y-axis) plotted against the simulated time in hours (x-axis). **b)** avg. growth rate from a) grouped by day, expressing daily growth rate (y-axis) at respective day in simulation (x-axis)

The model exhibits a daily growth rate of about 0.05 day^{-1}, in the time span where NH_4 is available as N-source and is primarily used as such by the model. As can be seen in Fig. 4.10.a the model achieves a peak growth rate of 0.125 day^{-1} which is slightly higher than the experimental determined growth rate for *M. Truncatula* of 0.1 day^{-1} (Lötscher et al., 2004). However, the daily average is about half of this rate indicating, that the starch available at nighttime is insufficient to support this growth rate.
At around 1000 hours simulated a significant drop in the daily growth rate can be observed in Fig. 4.10.b. Followed by increased fluctuations in growth rate probably within the time span simulation where both N_2 and NH_4 is available, with NH_4 close to 0.
A sudden increase on growth rate is visible at around 3500 h, further the daily growth rat

oscillates strongly however the overall trend shows an increase in growth until about 6000h. Moreover, after around 7000h simulated the oscillations stabilize, fluctuating slightly below the growth rate value of 0.04 day^{-1}.

In Fig. 4.10.a, one can see in the Interval of t_{sim} = [1000,10000], that the upper limit for the growth rate is about 0.044 [1/d] at day time, which is the reported growth rate of the ViNE model in (diCenzo et al., 2020). After about 9000 h simulated, the trend of the growth rate curve is declining. It is not clear what is causing this decline, since the growth rate is yet no limited by N uptake. At around 11500 h we can see that growth drops and raises very quickly, the sudden raise might indicate that growth is no longer limited by carbon availability, with the onset of starch overproduction. In the end phase of the simulation in the interval t_{sim}= [11500,13000], a gradual decline of growth in accordance with the concentration of N_2 in the system environment. Thus, it marks the transition of growth being limited by carbon to being limited by Nitrogen.

In Fig. *4.11*.we can see that growth is only possible during daytime since not enough starch is available in the pool at night. In Fig. *4.11*.b we can see some growth at nighttime but not enough to support biomass production over the whole night period. It also shows a slight dent in the avg. growth around noon. This complies with number of iterations in Fig. 4.9.b and c and observation of the Nodule III CO_2 secretion at noon in Fig. *4.14*.

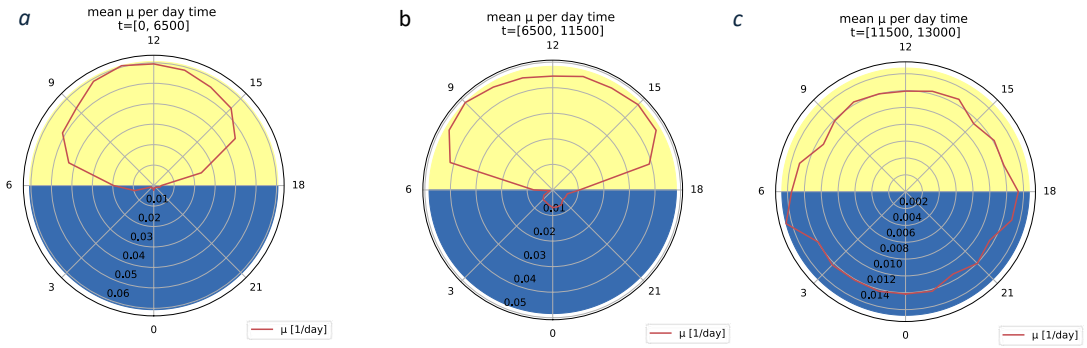

Fig. 4.11 Mean growth rate grouped by daytime. The line in red depicts the avg. growth rate [day^{-1}] calculated from the rate mg gDW^{-1} h^{-1} given by the ViNE model. The daytime within the simulation is given by the numbers around edge of the radial plots. The blue shaded area represents nighttime, the yellow highlighted area the daytime. **a)** Mean growth rate grouped by daytime over simulation interval of t_0=0 to the 6500th hour simulated. **b)** Mean growth rate grouped by daytime beginning with 6500th hour simulated until hour 11500. **c)** Mean growth rate grouped by daytime of the interval starting with t_{sim} > 11500 until end of simulation at 13000 hours simulated.

4.4.3 Yield CO2 evoked per N2 fixed

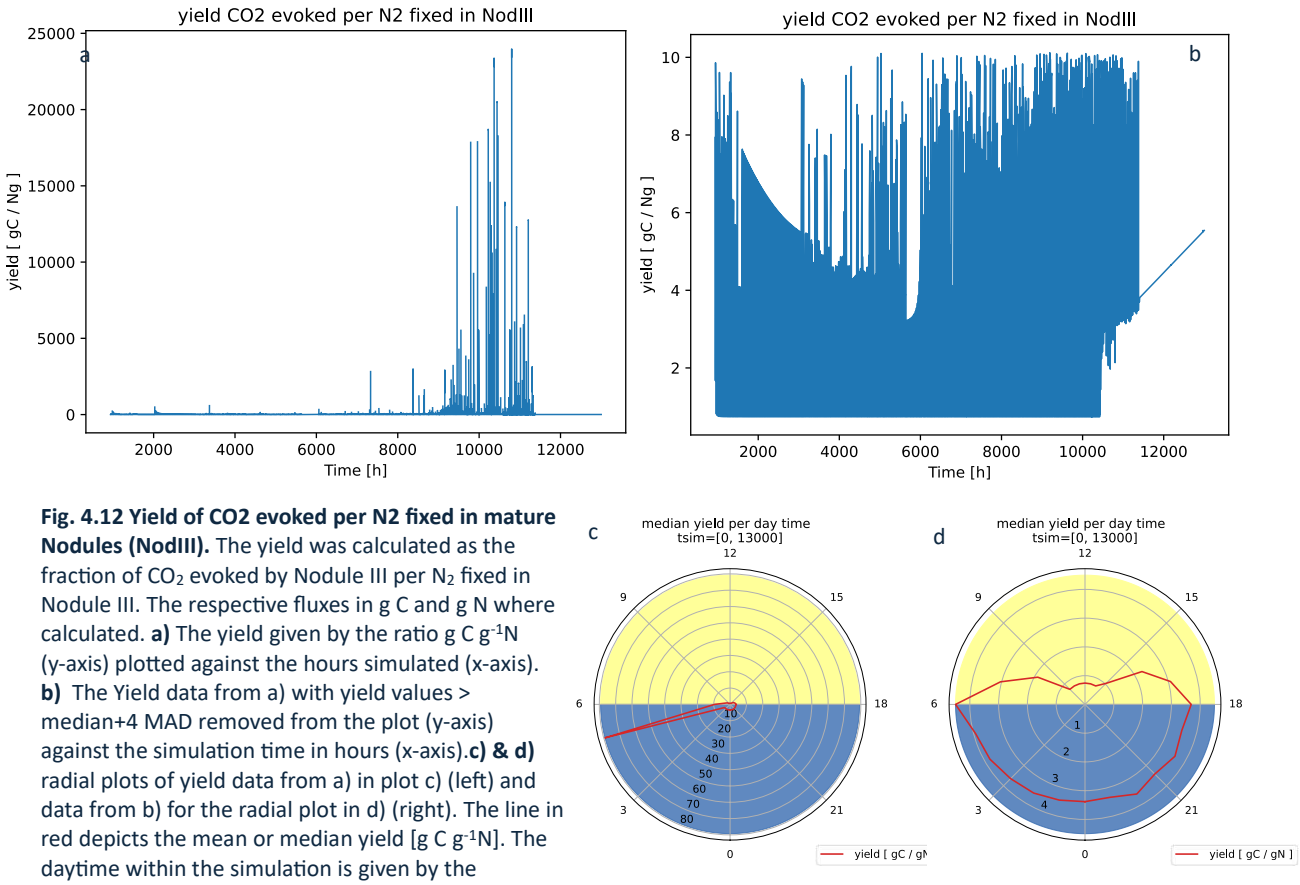

Fig. 4.12 Yield of CO2 evoked per N2 fixed in mature Nodules (NodIII). The yield was calculated as the fraction of CO_2 evoked by Nodule III per N_2 fixed in Nodule III. The respective fluxes in g C and g N where calculated. **a)** The yield given by the ratio g C g^{-1}N (y-axis) plotted against the hours simulated (x-axis). **b)** The Yield data from a) with yield values > median+4 MAD removed from the plot (y-axis) against the simulation time in hours (x-axis).**c) & d)** radial plots of yield data from a) in plot c) (left) and data from b) for the radial plot in d) (right). The line in red depicts the mean or median yield [g C g^{-1}N]. The daytime within the simulation is given by the numbers around edge of the radial plots. The blue shaded area represents nighttime, the yellow highlighted area the daytime.

(Frank R. Minchin & John F. Witty, 2005)The plotted yield in Fig. 4.12a, which represents the amount of CO_2 released by nodules per N_2 fixed over the entire simulation period, exhibits a pronounced peak around t_{sim} =[9000,11000]. In fact, the highest activity of can be observed within this time interval of the simulation. The yield displayed in this interval far exceeds the experimentally determined range for the respiratory carbon cost of N_2 fixation, which is reported as 3 – 5 [g C g^{-1} N], by (Frank R. Minchin & John F. Witty, 2005).

In Fig. 4.12b, pronounced fluctuations in the yield are observed, with a noticeable rise in the lower envelope of the oscillations around 10 500 h. This transition occurs as the model shifts form the state of biomass production to the state of starch production, due to N_2 becoming increasingly depleted. The fluctuations cease at 12 000 h, with N_2 reaching its equilibrium.

It should be noted that the initial 1,000 h are not displayed in the plots of Fig. 4.12 and Fig. 4.13. This is due to the absence of N2 fixation, concurring with the availability of NH4.

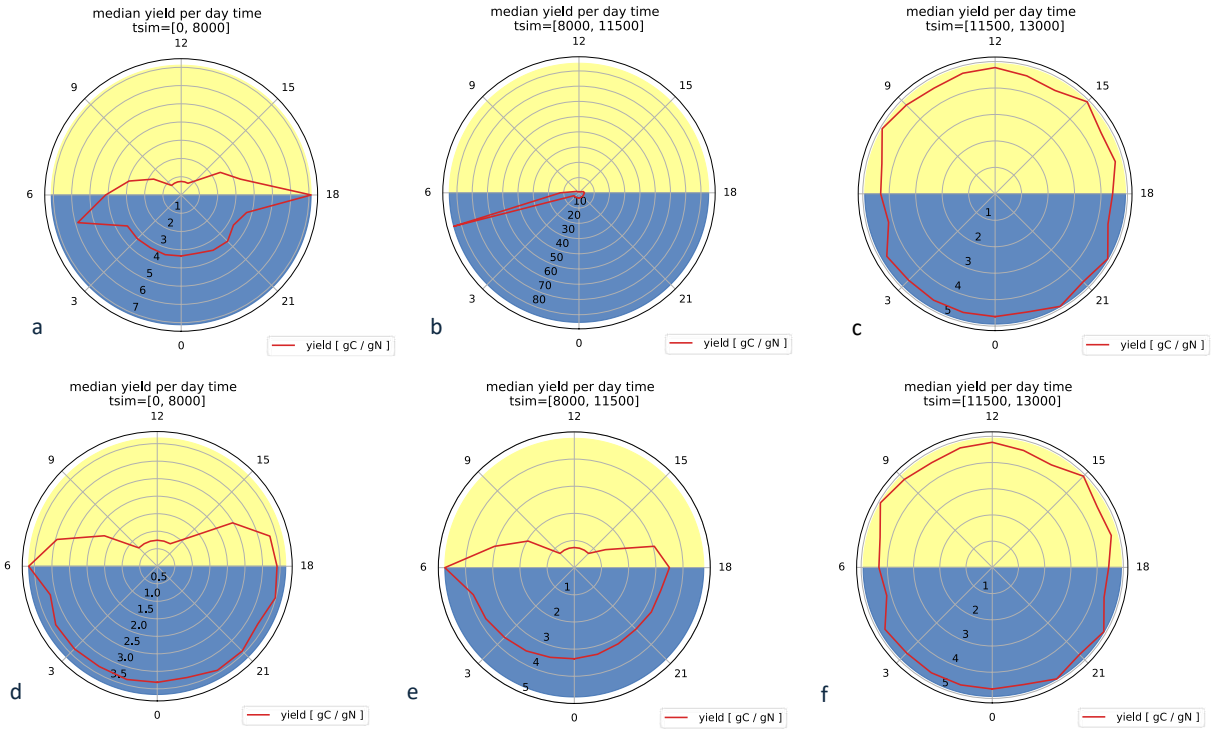

Fig. 4.13 **Radial plots of Yield of CO2 evoked per N2 fixed in mature Nodules (NodIII) grouped by daytime over different intervals of the simulation.** The line in red depicts the median yield in g C g^{-1}N. The daytime within the simulation is given by the numbers around edge of the radial plots. The blue shaded area represents nighttime, the yellow highlighted area the daytime. **a) to c)** The first row describes the Yield in the intervals as described by t_{sim} at top of the plots in hours simulated. **d) to f)** The second row of plots depict same intervals as described in first row, but data filtered with yield values > median+4 MAD removed from the plotted data.

The radial plots in Fig. 4.13 show the median yield per daytime, at different stages of the simulation. In the radial pots in the middle (Fig. 4.13b &e) we can see a sever difference between the filtered and unfiltered data, as is described in the figure caption.
It remains unclear whether the outliers visible in the middle of the simulation are genuine properties of the system's dynamics, reflect an underlying biological reality, or are merely numerical artifacts—possibly noise generated by very low levels of N_2 fixation. However, no evidence was found for significant outliers or abrupt changes in the amount of CO_2 released by the nodules. It is therefore most likely that the observed variation reflects differences in N_2 fixation.

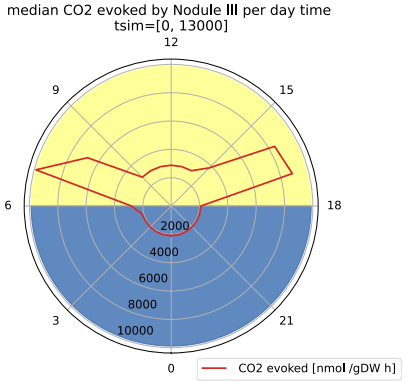

Fig. 4.14 Median of CO_2 evoked by mature Nodules per daytime. The line in red depicts the median flux of CO_2 in nmol gDW^{-1} h^{-1} grouped by daytime. The daytime within the simulation is given by the numbers around edge of the radial plots. The blue shaded area represents nighttime, the yellow highlighted area the daytime.

The plot in Fig. 4.13.c shows an equal yield at all daytimes, with a yield of ca. 4.5 g C / g N fixed. This is within a usual range, slightly higher than the reported yield for the ViNE model of 4.2 g C / g N fixed by(diCenzo et al., 2020). A decrease in CO_2 emission, could already be observed in the ViNE model outside the context of this dFBA model. That is only for the mature Nodules III, has been observed that the emission of CO_2 would increase with increasing light intensity, like in all other Nodules and the roots system. However, once photons uptake threshold of $1 \cdot 10^6$ nmol / gDWh has been reached the N-fixing Nodules III exhibit a steep decline of CO_2 emissions. This can be seen in Fig. *4.14*, where the amount of CO_2 evoked at noon drastically decreases.

Comparing the radial plots of Fig. 4.13, we can see excluding the outliers present in the period t_{sim}=[8000, 11 500], the yield at earlier phases and mid phase Fig. 4.13.b seem to agree. The exhibited range of yield also most agrees with the plots of the end phase of the simulation The slightly higher yield can be attributed to the lower N_2 fixation compared to the amount of CO_2 evoked. The most prominent difference, however, is the yield at daytime of the late stage of the simulation compared to the yield before t_{sim}= 11,500 hours. In the earlier phases we see a continuous yield consistent with the end phase, but a lower yield at daytime around the noon, a behaviour already discussed. Nevertheless, assuming the decrease in yield or CO_2 evoked is caused by the availability of a C-source would imply that we could see same behaviour in end phase as well at night as at daytime, with starch being available thus C not limiting. So, one might conclude that it is specifically relating to CO2 fixation as carbon source or the uptake of photons by the model. For, instance it could be possible that under higher irradiation the model utilizes a different pathway. But why don't we see this dent at noon not in the late pashes of the simulation? A possible an answer might be that it's the lower N_2 fixation rate due to limited N availability. In this case it would be necessary to test if the model shows the same behaviour of reduced CO_2 emissions by the nodules under restricted N_2 uptake. Unfortunately, due a technical failure all data resulting from the simulation where lost, thus it could not be investigated further.

4.4.4 Starch content

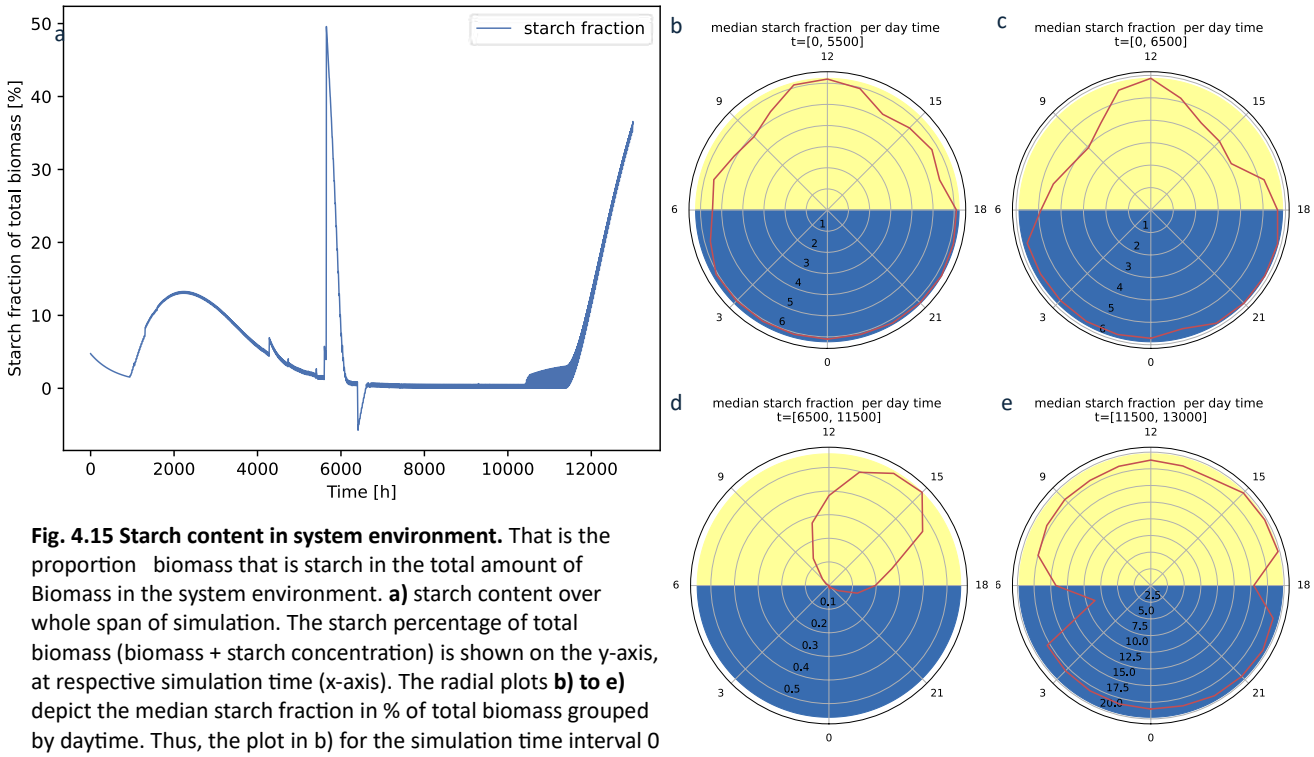

Fig. 4.15 Starch content in system environment. That is the proportion biomass that is starch in the total amount of Biomass in the system environment. **a)** starch content over whole span of simulation. The starch percentage of total biomass (biomass + starch concentration) is shown on the y-axis, at respective simulation time (x-axis). The radial plots **b) to e)** depict the median starch fraction in % of total biomass grouped by daytime. Thus, the plot in b) for the simulation time interval 0 to 5500 hour simulated**,** c) the interval 0 to 6500 h, d) the interval 6500 -11500 h and e) the interval 11500 - 13000 the end of the simulation. The line in red depicts the median starch fraction in % of Biomass. The daytime within the simulation is given by the numbers around edge of the radial plots. The blue shaded area represents nighttime, the yellow highlighted area the daytime.

In Fig. *4.15* we see the fraction of biomass that is starch in percent of total biomass in the system environment, that is the biomass plus the starch concentration in the environment of the system. Further the simulation starts with 5% starch and drops at t_{sim}=1000, where NH_4 depletes. This can be explained by the fraction of starch is only decreasing relative to biomass. Since during that time span, the model shows almost twice the growth rate then without additional NH_4. Unclear what the rise and fall of the starch fraction of biomass between t_{sim} =[1000,5500], is caused by, but in Fig. *4.10*.b we see an increase in the average daily growth rate. Thus, what we can conclude is that the starch fraction drops when growth rate is high and vice versa. At t_{sim}=[5500, 6000] h we see the outlier consistently present in all simulation results. This is followed by rather stable fluctuation of starch fraction in the day-night cycle. Until the model switches from biomass production to starch production, when N_2 approaches equilibrium. Finally with the start of starch over production at t_{sim}=11500, the starch fraction starts to raise from below 10 % up to 40% of Biomass. Experimental data on *A. Thaliana* suggest that 5% of overall pant biomass is starch (Williams et al., 2010),, however the fraction of starch in leaves is assumed to be 2.5 % (Ramonell KM, 2001). At the beginning of the simulation the avg. starch content is about in that range. However, in the mid of the simulation, compare Fig. *4.15*.c the amount of starch relative to the total biomass in the system is way below that value, at about a median of 0.5 %. In the

late phase of the simulation the starch content relative to the biomass is way too high. With higher availability of N in equilibrium of the model, we would see higher growth rate with less starch production, thus it is possible the chosen value for K_{LA} is inaccurate. This would however only delay this disproportionately high fraction of starch in the total Biomass. Therefore, this reflects a limitation of the model. First it there is only on starch pool, not differentiating between transitory starch in leaves and long-term stored starch in storage roots. Also, the produced starch is biomass but is not accounted for in the biomass reaction of the Model. What more is, the model does not constrain how much of biomass can be starch. In other words, there is no limit on how much starch can be accumulated, especially with respect to biomass.

At the end of the simulation only 40% of starch produced at daytime is being consumed at night. This is an estimate derived from *Fig. 4.7*.a and Fig. 4.8.c, the figures depict a production flux of 25 µmol gDW^{-1} h^{-1} and a consumption of around 10 µmol gDW^{-1} h^{-1} thus the production is about 2.5 times the consumption. Further in *Fig. 4.6*.b the hourly change rate of starch, shows a production of up to 100 g L^{-1} h^{-1} but only a consumption of up to 40 g L^{-1} h^{-1}. Thus, at the end of the simulation 60 % of starch remains in the pool at dusk, which would be transported to the roots for long term storage. A fraction which considerably to high given the literature indicates that 5-10% of transitory starch remains at the end of the night in *A. Thaliana* leaves (Gibon et al., 2009).

5. Discussion

5.1 Growth rate of the model

The growth rate of the model is ether limited by the availability of carbon or limited by Nitrogen. The first case is particularly true in early simulations but also in middle phase. Moreover, there is not enough Starch available at nighttime, so growth is mostly limited to daytime. The second case, where the growth rate is reduced by the availability of N, can be seen in the end phase of the simulation, where there is more starch produced then consumed in the day-night cycle. The average growth rate of 0.026 [1 /day], by the model over the whole simulation time, is about half of rate reported by (diCenzo et al., 2020), for the ViNE model of 0.044 [1 /day]. The reported rate, refers to the model with N_2 being the only N-source and light uptake limited to 1000 µmol / gDW h. The photon uptake at noon for this model is about twice as high. Consequently, we can see significantly higher growth at the start where NH_4 is available. So, for most of the simulation span µ is limited by carbon not N, thus it can be expected that mean growth rate is about half the rate of the ViNE model, or in the rage of mean growth rate ±1 std (mean = 0.026, median = 0.028, std = 0.023). However, for the end phase of the simulation the rate drops to around 0.014[1/day] due to N in equilibrium being limited by diffusion. In other words, by the chosen value of the parameter $K_L a$. This rate is low, but still in a reasonable range in under consideration of rate reported for the ViNE model. Yet, it is about one order smaller than the experimentally determined value of 0.1 [1/day] for *Medicago sativa* reported in (Lötscher et al., 2004). These measurements were conducted under a 16/8-hour day-night cycle with a light intensity of 350 µmol $m^{-2}s^{-1}$, resulting in a day light Integral (DLI) of 20.16 mol $m^{-2}day^{-1}$. To put this in comparison with the conditions of the simulation, the 12/12 hours day-night cycle, with a PFD of 400 µmol $m^{-2}s^{-1}$ at noon resulting in a DLI of 17.28 mol $m^{-2}day^{-1}$, this is difference of about 17%. Nevertheless, this difference in carbon uptake does not affect the growth rate of the end of the simulation where it is limited by N uptake.
Note the DLI values were calculated with an online calculator (Rafay, n.d.) and not calculated based on the model light function, it is thus only an estimate of the difference in light exposure. Also the light exposure can be adjust, by not only change the day length itself, but also by defining the start of the day with the I_0 parameter.

The availability of N can be adjusted in two possible ways, firstly by adjusting the parameter $K_L a$ controlling N diffusion. Secondly, by modelling regeneration of bioavailable N in the soil. This can be done be introducing a diffusion term $K_L a$ for NH_4 concentration.

With $K_L a$ being a parameter determining gas transfer from gaseous to liquid phase in Bioreactors. Thus $K_L a$ values like the one found in (Promma et al., 2024) refers to the amount of gas being transferred in a reactor tank of a certain size, shape at certain temperature, pressure and amount of agitation (Kraakman et al., 2011). This might be meaning full for simulations of bacterial batch growth in bioreactors as described in (Mahadevan et al., 2002), but doesn't necessarily translate well to N-diffusion in soil.

5.1.1 Estimation of the $K_L a$ parameter for NH_4 and N_2

The parameter $K_L a$ for NH_4, can be estimated based on the rate of mineralization occurring in soil. The flux rate could be further determined by subtracting the rate of N-immobilization through microbial community in the soil. Moreover, the mineralization and immobilization of bioavailable nitrogen depend on several factors, including temperature, soil type, pH, moisture levels, and the composition of degrading organic matter (cite) For the last mentioned, the C:N ratio in degrading soil organic matter, determines if and how much Nitrogen is immobilized by soli microbes. A C:N ratio below 40:1 favours nitrogen mineralization, whereas a higher ratio promotes nitrogen immobilization by soil microbes(Vigil & Kissel, 1991).

The net mineralization rate is difficult to determine experimentally, if even possible. Therefore, a rate describing mineralization could be obtained from theoretical models. There is a whole plethora of models, documented theoretical modelling attempts describing N-flow in the soil, reach back at least 80 years (Manzoni & Porporato, 2009). A good starting point for searching for the right model might by reviews about soil management (Kwiatkowska-Malina, 2018) or meta studies, in this case comparing 230 different models (Manzoni & Porporato, 2009). It is further of course important to choose a model at right spatial and temporal scale. At the example of the Verberne Model (Verberne et al., 1990), the decomposition rate for decomposable matter by microbial biomass is given as 0.2 day^{-1} although here it is unclear if this rate can be directly used as $K_L a$, which is given as h^{-1} when this parameter is scaled down from days to hours. What it more is, this rate doesn't account for immobilization of N. However, in the work of Verberne and colleagues, the fluxes of net mineralization are also expressed as kg N ha^{-1}, the parameters soil type (clay and sand), amount and C:N ratio of decomposable matter added to the system. The rate of diffusion in this dFBA model is given with respect to system volume (compare diffusion of N_2 in eq. (10)), as described below.

$$\frac{d\,c(N-source\,)}{dt} = K_L a \cdot \left(\, c^*(N-source) - c(N-source)\right) \quad \left[\frac{1}{h} \cdot \frac{mg}{L}\right] \tag{5.1}$$

Thus, we would need to estimate the system volume per soil area to be able the calculate the diffusion flux value, described by left hand side of eq. (5.1). With the flux estimate in the left hand side of the equation and in this case $c(N-source)\ \ being$ c(NH_4) = 0 , we can the solve for $K_L a$, to estimate a value $K_L a$ NH_4. In a similar way this can be done to estimate the $K_L a$ value for the diffusion of N_2.

5.2 Starch metabolism and starch content

The starch fraction of the total biomass is overestimated by far at the end of the simulation. With total biomass being the biomass plus the starch in the environment of the system.

The simplest yet most effective solution would be a constraint on the amount of starch that can be accumulated in the starch compartment. This could be implemented by allowing starch production only if the fraction of starch is not surpassing 5% within the total biomass Moreover 5% is the fraction of starch in the dry weight of A. Thalian plants.

One could argue that this partially comes down to the simplicity of the model, with only one starch pool. A possibility could be splitting the pool, with leaves starch pool for transitory starch and a root pool for storage starch.
Moreover, transitory starch is the surplus of Starch being produced during the day and stored in the chloroplasts of the leaves, which is then used by the plants at nighttime (Zeeman et al., 2006). Depending on implementation, splitting the Biomass reaction and the biomass of the model might not be necessary if more than one starch pool is implemented. However, the point being made here this solution might only be of relevance if the biomass is split up. Nevertheless, splitting the biomass of the different Tissues in the ViNE model would allow to address interesting questions with respect to multi tissue models and will be discussed later in section 5.2.1.

Other than the structure of the model, it could be reasoned that the chosen $V_{\max \text{ Amylase}}$,does not reflect the speed at which starch is being consumed, amylase might not be rate limiting. The mobilisation of starch in leaves is a complicated process where a whole series of enzymes is involved. Furthermore, the degradation of starch in the chloroplasts happens by network of enzymes rather than it being catalysed by a linear series of enzymes. However, the starch degradation process can be subdivided into three steps amorphization by GWD (glucan water dikinase), dephosphorylation by a Glucan phosphatase (SEX4, LSF2) and cleavage by hydrolysing enzymes like ß-amylases (BAM 1-3). On a molecular level the rate of starch synthesis is most likely regulated or heavily dependent on the activity of AGPase, for the rate of starch degradation it is not that clear. There seems to be evidence that transports out of the chloroplast like Mex1 for maltose and pGlcT for the transport of glucose play a vital role in limiting the rate of starch degradation (Lloyd et al., 2005; LU & SHARKEY, 2006)

Evidence indicates that transitory starch metabolism is governed by the circadian clock and influenced by temperature and day length, although unclear how the degradation rates of starch are controlled. However there seems to be a broad consensus in literature, that concentration of transitory starch declines linearly, so much so that the rate is controlled, in a way that almost all starch is consumed with end of night, based on the assumed day or night length of previous cycle (Graf & Smith, 2011; Stitt & Zeeman, 2012; Webb & Satake, 2015). Therefore, plants exposed to a sudden change in day length, would still adjust their starch degradation rate based on the assumed length of the night. Moreover, evidence for the regulation of the degradation rate is given by the observation that in short photoperiods and long nights the synthesis of starch happens fast and degradation slow, in short photoperiods it is reversed. For neutral photoperiods (12h/12h), as used for this simulation in this work, the rates are equal (Streb & Zeeman, 2012). Moreover, in *A. thaliana* leaves, findings indicate

that under a 12h/12h day-night cycle, the transitory starch synthesis and degradation rate is 139 µmol gDW^{-1} h^{-1} (Gibon et al., 2009). This is close to the estimated upper bound for CO_2 uptake, which is about 180 µmol $gDW^{-1}h^{-1}$, which is limiting starch production in this model.

Furthermore, varying ratios of starch degradation to synthesis, depending on day and night length, proposes that changes in starch metabolism rates may be regulated by the circadian rhythm. However, literature seem to be suggesting that circadian control is mostly orchestrated on the level of differential gene expression (Mortada et al., 2024; Venkat & Muneer, 2022). Furthermore, around 30% of *A. Thaliana* transcripts show oscillatory expression patterns in a day night cycle(Davis et al., 2015). To get to the point dFBA does not consider regulation at the level of gene expression. Therefore, such changes in starch metabolism might not be accounted for in this model.

Lastly, it should be noted that $V_{\max \text{Amylase}}$ was wrongly calculated, the corrected value is not a flux value achievable by the model. Of course, V_{max} is only a theoretical upper limit, from which the upper bound of a flux is calculated. However, the scale of this V_{max} would probably result into a nonactive upper bound. This might suggest that starch degradation in leaves is indeed not rate limited by amylase activity. Certainly, this only speculation since the correct parameter value for $V_{\max \text{Amylase}}$ has not even been tested with the model.

Addressing the starch content, in the end phase of the simulation is limited by N, it is thus questionable if higher V_{max} for amylase would lead to a lower starch fraction of the biomass. Afterall, the model can only consume as much starch at night as there is growth.

5.2.1 Splitting the biomass

Splitting the biomass reaction, thus separating the biomass of the tissues and species of the model into separate independent Biomass pools, would allow to utilize the advantages of dFBA in community modelling. Especially with respect to multi tissue plant models, where metabolism is strongly affected by a diel cycle. In the current implementation the proportion of the biomass reactions of tissues to the overall biomass reaction of the model is fixed. Simple, splitting up the biomass reaction of the model and optimising each biomass reaction independently would not be a meaningful approach. The maximization of biomass for all species in the community independently, might be meaningful in the case of microbial communities. For multi-tissue models however, a different more sophisticated objective function, would be required. Moreover, the tissues of a plant are not in competition with each other. Thus, an objective function would need to be developed that optimises the overall biomass of the system, without hardcoding the proportional contribution of different tissues to the overall biomass and without obtruding competition between the different tissues. The objective function should therefore imply a cooperative and coordinated behaviour, where each compartment has the goal to improve the overall performance or wellbeing of the organism. An example for such an objective function has been described by (de Oliveira Dal'Molin et al., 2015). The proposed objective function minimizes the energy consummation of the whole metabolic network, which is photon capture of the plant.

Furthermore, splitting up the biomass would allow to analyse behaviour like plant nodule biomass ratio under different environmental constraint. For example, does nodule biomass increase when constantly exposed to increased [CO_2] in the air, as has been discussed in (Libault, 2014).

5.3 Performance and stability of the simulation

It worth noticing, that many measures taken to improve the performance and stability of the implementation, where introduced while the model was scaled to fluxes of completely different dimension. Thus, the fluxes where wrongly assumed to be given in µmol gDW^{-1} h^{-1} instead of nmol gDW^{-1} h^{-1}. Further, most of these measures showed a bigger effect in earlier version of the model then, could be observed in the present implementation.

One of the major challenges affecting the accuracy of simulation results was numerical round-off error in the estimation of the Jacobian matrix. The Integrator function computes the Jacobian numerically using finite differences, a method that is computationally expensive and prone to numerical instability. In particular, if the step size h becomes too small ($h < \epsilon$), the method becomes numerically unstable due to round-off errors growing in proportion to the step size. When h approaches or falls below the numerical precision limit of the computer, the accuracy of the Jacobian estimation deteriorates, leading to instability. This is a limitation of the floating-point arithmetic performed by the computer (Tolsma & Barton, 1998).

An attempt to replicate the simulation described in Section 4.3 failed after approximately 10,500 hours of simulated time, which required about three days of computation. The error message indicated that the LSODA solver's step size had become too small, likely due to a failure in Jacobian matrix estimation, causing the simulation to terminate.

A potential solution to this issue is automatic differentiation (AD), which provides numerically exact and more efficient Jacobian estimation, then possible by finite differences (Tolsma & Barton, 1998). Initial tests using ODE solvers from the SciML software ecosystem (Rackauckas & Nie, 2017) showed promising results. While the SciML solver outperformed SciPy's solve_ivp function in short simulations, this advantage could not be verified for longer simulations. One challenge encountered was the incompatibility of automatic differentiation with the existing data structures of the model implementation, requiring adaptations to the model. However, SciML offers significantly more customization options for numerical methods, making it a promising direction for future improvements.

An even more promising approach to improving the performance and stability of the simulation would be, by implementing the model with a direct optimization approach (DA), as it has been proposed in(Gomez et al., 2014; Harwood et al., 2016). This approach is far more efficient and stable then the static optimization approach and the implementation of the direct approach used in this work.

The most important improvement is, that the LPs no longer need to be solved in every iteration of the solver function, when the ride hand side is evaluated. It only needs to be solved when concentrations in the environment change, more precisely if the system changes its state, which could be for instance when a metabolite depletes in the environment (De Becker et al., 2020).
This means, future optimal solutions can be predicted from initial FBA solutions, this is achieved by computing an optimal basis for the FBA solution, most likely to stay optimal during the simulation. Moreover, a basis is a collection of basic variables, where basic variables are fluxes which are non-zero. Thus, an optimal basis are the nonzero fluxes of an optimal solution of an LP problem. If a one of these variables drops and crosses zero, the basis is no longer optimal. Therefore, a new optimal basis needs to be determined, this is done by solving the LP problem, from which new a new optimal basis can be determined. Followingly, the LP problem needs to be solved at least once. Eventually it is possible to calculate the whole trajectory of the simulation, with solving just a few LP problems. (Harwood et al., 2016; The Economic Cell Collective (2025) et al., 2025, pp. 166–167)

A simulator using a direct approach has been implemented in DFBAlab by (Gomez et al., 2014), with a python implementation available by (Tourigny et al., 2020). Unfortunately, this implementation of the direct approach is not compatible with the ViNE model. For comparison, as reported by (Gomez et al., 2014) for a simulation of 24 hours with two species yeast (2191 reaction) and algae (1266 reactions) and about 5 dynamically balanced metabolites, the simulation took 30 sec. to execute. Thus, a simulation of 500 days would have taken approximately 4 hours, instead of the 4.5 days with this implementation.

6. Conclusion

This work demonstrates the successful implementation a working dFBA model. With in the scope of the limitation both inherited from the ViNE model and due to the simplicity of the model, the predictions made by the model are reasonable. The growth rate compared to the ViNE model is within a meaning full range for most of the simulation. However, the μ in equilibrium of the N concentration is whole magnitude, 10 times smaller than the experimentally determined rate. This indicates the model's limitations or degree of incompleteness. The other experimental values carbon cost for N_2 fixation and starch fraction of biomass are predicted fairly well by the model. Most of the short coming in the predictive capabilities of the model, a more detailed / nuanced modelling of N availability in the soil and an accurate estimation of the parameters would be necessary. Lastly, given the current implantation the model is of very limited practical use. The execution time is way too high to be used in practice, at least if not run on a high-performance computing cluster. Performing the simulation on a HPC cluster would, however, not solve instability issues of the simulations. Nevertheless, the most interesting question that could be addressed with the ViNE model would require a splitting of the biomass of the different tissues and organisms.

During the preparation of this work the author used ChatGPT in order to improve readability and language of this work. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the content of this work.

7. Literature

Adolfsson, L., Solymosi, K., Andersson, M. X., Keresztes, Á., Uddling, J., Schoefs, B., & Spetea, C. (2015). Mycorrhiza symbiosis increases the surface for sunlight capture in medicago truncatula for better photosynthetic production. *PLoS ONE*, *10*(1). https://doi.org/10.1371/journal.pone.0115314

Athanasiou, K., Dyson, B. C., Webster, R. E., & Johnson, G. N. (2010). Dynamic acclimation of photosynthesis increases plant fitness in changing environments. *Plant Physiology*, *152*(1), 366–373. https://doi.org/10.1104/pp.109.149351

Barker, D. G., Pfaff, T., Moreau, D., Groves, E., Ruffel, S., Lepetit, M., Whitehand, S., Maillet, F., Nair, R. M., & Journet, E.-P. (2006). *Growing M. truncatula: choice of substrates and growth conditions*.

Battino, R., Rettich, T. R., & Tominaga, T. (1984). The Solubility of Nitrogen and Air in Liquids. *Journal of Physical and Chemical Reference Data*, *13*(2), 563–600. https://doi.org/10.1063/1.555713

Berkhout, J., Bruggeman, F. J., & Teusink, B. (2012). Optimality principles in the regulation of metabolic networks. In *Metabolites* (Vol. 2, Issue 3, pp. 529–552). MDPI AG. https://doi.org/10.3390/metabo2030529

Bernhard, A. (2010). *The Nitrogen Cycle: Processes, Players, and Human Impact.* https://www.nature.com/scitable/knowledge/library/the-nitrogen-cycle-processes-players-and-human-15644632/

Bordbar, A., Monk, J. M., King, Z. A., & Palsson, B. O. (2014). Constraint-based models predict metabolic and associated cellular functions. In *Nature Reviews Genetics* (Vol. 15, Issue 2, pp. 107–120). https://doi.org/10.1038/nrg3643

Cornelissen, G. (2014). Cosinor-based rhythmometry. In *Theoretical Biology and Medical Modelling* (Vol. 11, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/1742-4682-11-16

Davis, S. J., Lu, S., Meyer, C., Yanovsky, M. J., & Romanowski, A. (2015). *Circadian rhythms and post-transcriptional regulation in higher plants*. https://doi.org/10.3389/fpls.2015.00437

De Becker, K., Bernaerts, K., & Waldherr, S. (2020). Luenberger observer design for a dynamic system with embedded linear program, applied to bioprocesses. *IFAC-PapersOnLine*, *53*(2), 15884–15891. https://doi.org/https://doi.org/10.1016/j.ifacol.2020.12.280

de Oliveira Dal'Molin, C. G., Quek, L. E., Saa, P. A., & Nielsen, L. K. (2015). A multi-tissue genome-scale metabolic modeling framework for the analysis of whole plant systems. *Frontiers in Plant Science*, *6*(JAN), 1–12. https://doi.org/10.3389/fpls.2015.00004

diCenzo, G. C., Tesi, M., Pfau, T., Mengoni, A., & Fondi, M. (2020). Genome-scale metabolic reconstruction of the symbiosis between a leguminous plant and a nitrogen-fixing bacterium. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-16484-2

Doehlert, D. C., Duyke, S. H., & Anderson, L. (1982). Beta-Amylases from Alfalfa (Medicago sativa L.) Roots'. In *Plant Physiol* (Vol. 69).

Feist, A. M., & Palsson, B. O. (2010). The biomass objective function. In *Current Opinion in Microbiology* (Vol. 13, Issue 3, pp. 344–349). https://doi.org/10.1016/j.mib.2010.03.003

Frank R. Minchin, & John F. Witty. (2005). respiratorycarbon-costs-of-symbiotic-nitrogen-fixation-in-legume. *Plant Respiration* , 195–205.

Fuller, J., An, Q., Fortunelli, A., & Goddard, W. A. (2022). Reaction Mechanisms, Kinetics, and Improved Catalysts for Ammonia Synthesis from Hierarchical High Throughput Catalyst Design. *Accounts of Chemical Research*, *55*(8), 1124–1134. https://doi.org/10.1021/acs.accounts.1c00789

Gibon, Y., Pyl, E.-T., Sulpice, R., Lunn, J. E., Höhne, M., Günther, M., & Stitt, M. (2009). *Adjustment of growth, starch turnover, protein content and central metabolism to a decrease of the carbon supply when Arabidopsis is grown in very short photoperiods*. https://doi.org/10.1111/j.1365-3040.2009.01965.x

Gomez, J. A., Höffner, K., & Barton, P. I. (2014). DFBAlab: A fast and reliable MATLAB code for dynamic flux balance analysis. *BMC Bioinformatics*, *15*(1). https://doi.org/10.1186/s12859-014-0409-8

Graf, A., & Smith, A. M. (2011). Starch and the clock: The dark side of plant productivity. In *Trends in Plant Science* (Vol. 16, Issue 3, pp. 169–175). https://doi.org/10.1016/j.tplants.2010.12.003

Gurobi Optimization LLC. (2024). *Gurobi Optimizer Reference Manual*. https://www.gurobi.com

Harcombe, W. R., Delaney, N. F., Leiby, N., Klitgord, N., & Marx, C. J. (2013). The Ability of Flux Balance Analysis to Predict Evolution of Central Metabolism Scales with the Initial Distance to the Optimum. *PLoS Computational Biology*, *9*(6). https://doi.org/10.1371/journal.pcbi.1003091

Harwood, S. M., Höffner, K., & Barton, P. I. (2016). Efficient solution of ordinary differential equations with a parametric lexicographic linear program embedded. *Numerische Mathematik*, *133*(4), 623–653. https://doi.org/10.1007/s00211-015-0760-3

Hdira, S., Haddoudi, L., Hanana, M., Romero, I., Mahjoub, A., Jouira, H. Ben, Ludidi, N., Sanchez-Ballesta, M. T., Abdelly, C., & Badri, M. (2021). Morpho-physiological, biochemical, and genetic responses to salinity in medicago truncatula. *Plants*, *10*(4). https://doi.org/10.3390/plants10040808

Henson, C. A., Duke, S. H., & Koukkari, W. L. (1986). Rhythmic Oscillations in Starch Concentration and Activities of Amylolytic Enzymes and Invertase in Medicago sativa Nodules. In *Plant Cell Physiol* (Vol. 27, Issue 2).

Hindmarsh, A. C. (1982). *ODEPACK, a Systematized Collection of ODE Solvers*. Lawrence Livermore National Laboratory. https://books.google.at/books?id=9XWPmwEACAAJ

Jenkins, A. (2013). The Sun's position in the sky. *European Journal of Physics, Volume 34*(Issue 3).

Kraakman, N. J. R., Rocha-Rios, J., & van Loosdrecht, M. C. M. (2011). Review of mass transfer aspects for biological gas treatment. *Applied Microbiology and Biotechnology*, *91*(4), 873–886. https://doi.org/10.1007/s00253-011-3365-5

Kwiatkowska-Malina, J. (2018). Qualitative and quantitative soil organic matter estimation for sustainable soil management. *Journal of Soils and Sediments*, *18*(8), 2801–2812. https://doi.org/10.1007/s11368-017-1891-1

Lammel, G., & Graßl, H. (1995). Greenhouse effect of NOX. *Environmental Science and Pollution Research*, *2*(1), 40–45. https://doi.org/10.1007/BF02987512

Libault, M. (2014). The carbon-nitrogen balance of the nodule and its regulation under elevated carbon dioxide concentration. In *BioMed Research International* (Vol. 2014). Hindawi Publishing Corporation. https://doi.org/10.1155/2014/507946

Lloyd, J. R., Kossmann, J., & Ritte, G. (2005). Leaf starch degradation comes out of the shadows. In *Trends in Plant Science* (Vol. 10, Issue 3, pp. 130–137). Elsevier Ltd. https://doi.org/10.1016/j.tplants.2005.01.001

Lötscher, M., Klumpp, K., & Schnyder, H. (2004). Growth and maintenance respiration for individual plants in hierarchically structured canopies of Medicago sativa and Helianthus annuus: The contribution of current and old assimilates. *New Phytologist*, *164*(2), 305–316. https://doi.org/10.1111/j.1469-8137.2004.01170.x

LU, Y. A. N., & SHARKEY, T. D. (2006). The importance of maltose in transitory starch breakdown. *Plant, Cell & Environment*, *29*(3), 353–366. https://doi.org/https://doi.org/10.1111/j.1365-3040.2005.01480.x

Lu, Y., Kronzucker, H. J., Yu, M., Shabala, S., & Shi, W. (2024). Nitrogen-loss and carbon-footprint reduction by plant-rhizosphere exudates Plant Science. *Trends in Plant Science*, *29*(4). https://doi.org/10.1016/j.tplants.2023.09.007

Ma, F., Jazmin, L. J., Young, J. D., & Allen, D. K. (2014). Isotopically nonstationary 13C flux analysis of changes in Arabidopsis thaliana leaf metabolism due to high light acclimation. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(47), 16967–16972. https://doi.org/10.1073/PNAS.1319485111/-/DCSUPPLEMENTAL/PNAS.1319485111.SD01.XLS

Mahadevan, R., Edwards, J. S., & Doyle, F. J. (2002). *Dynamic Flux Balance Analysis of Diauxic Growth in Escherichia coli*.

Manzoni, S., & Porporato, A. (2009). Soil carbon and nitrogen mineralization: Theory and models across scales. In *Soil Biology and Biochemistry* (Vol. 41, Issue 7, pp. 1355–1379). https://doi.org/10.1016/j.soilbio.2009.02.031

Mayali, X., Bruggeman, F. J., Navid, A., Perez-Garcia, O., Lear, G., & Singhal, N. (2016). Metabolic Network Modeling of Microbial Interactions in Natural and Engineered Environmental Systems. *Frontiers in Microbiology | Www.Frontiersin.Org*, *1*, 673. https://doi.org/10.3389/fmicb.2016.00673

Moreau, D., Burstin, J., Aubert, G., Huguet, T., Ben, C., Prosperi, J. M., Salon, C., & Munier-Jolain, N. (2012). Using a physiological framework for improving the detection of quantitative trait loci related to nitrogen nutrition in Medicago truncatula. *Theoretical and Applied Genetics*, *124*(4), 755–768. https://doi.org/10.1007/s00122-011-1744-z

Mortada, M., Xiong, L., & Mas, P. (2024). Dissecting the complexity of local and systemic circadian communication in plants. *Npj Biological Timing and Sleep*, *1*(1), 1. https://doi.org/10.1038/s44323-024-00003-3

Orth, J. D., Thiele, I., & Palsson, B. O. (2010). What is flux balance analysis? In *Nature Biotechnology* (Vol. 28, Issue 3, pp. 245–248). https://doi.org/10.1038/nbt.1614

Pfau, T., Christian, N., Masakapalli, S. K., Sweetlove, L. J., Poolman, M. G., & Ebenhöh, O. (2018). The intertwined metabolism during symbiotic nitrogen fixation elucidated by metabolic modelling. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-30884-x

Pinton, R., Tomasi, N., & Zanin, L. (2016). Molecular and physiological interactions of urea and nitrate uptake in plants. In *Plant Signaling and Behavior* (Vol. 11, Issue 1). Taylor and Francis Inc. https://doi.org/10.1080/15592324.2015.1076603

Poolman1, M. G., Fell, D. A., & Thomas, S. (2000). Modelling photosynthesis and its control. In *Journal of Experimental Botany* (Vol. 51).

Promma, I., Aucoin, M. G., Abukhdeir, N. M., & Budman, H. (2024). A coupled metabolic flux/compartmental hydrodynamic model for large-scale aerated bioreactors. *Computers and Chemical Engineering*, *189*. https://doi.org/10.1016/j.compchemeng.2024.108806

Rackauckas, C., & Nie, Q. (2017). Differentialequations.jl–a performant and feature-rich ecosystem for solving differential equations in julia. *Journal of Open Research Software*, *5*(1), 15.

Rafay, K. (n.d.). *Daily Light Integral Calculator*. Omni Calculator. https://www.omnicalculator.com/biology/daily-light-integral-dli

Ramonell KM, K. A. P. D. C. M. X. Y. M. G. M. M. (2001). *microscopy study of leaf development during long-term plant growth in subambient oxygen*. https://doi.org/10.1046/j.1365-3040.2001.00691.x

Rizwan, M., Hurain Tanveer, ·, Muhammad, ·, Ali, H., Sanaullah, · Muhammad, & Wakeel, · Abdul. (2024). Environmental Science and Pollution Research Role of reactive nitrogen species in changing climate and future concerns of environmental sustainability. *Environmental Science and Pollution Research*. https://doi.org/10.1007/s11356-024-34647-2

Rogers, A., Ainsworth, E. A., & Leakey, A. D. B. (2009). Update on Legumes and Elevated CO 2 Will Elevated Carbon Dioxide Concentration Amplify the Benefits of Nitrogen Fixation in Legumes? 1 The majority of species capable of forming a symbi-otic relationship with N 2. *Plant Physiology*, *151*, 1009–1016. https://doi.org/10.1104/pp.109.144113

Rohatgi, A. (2021). *WebPlotDigitizer* (5.2). Automeris. https://automeris.io

Sala, O. E., Chapin, F. S., Armesto, J. J., Berlow, E., Bloomfield, J., Dirzo, R., Huber-Sanwald, E., Huenneke, L. F., Jackson, R. B., Kinzig, A., Leemans, R., Lodge, D. M., Mooney, H. A., Oesterheld, M., Poff, N. L. R., Sykes, M. T., Walker, B. H., Walker, M., & Wall, D. H. (2000). Global biodiversity scenarios for the year 2100. In *Science* (Vol. 287, Issue 5459, pp. 1770–1774). https://doi.org/10.1126/science.287.5459.1770

Schuetz, R., Kuepfer, L., & Sauer, U. (2007). Systematic evaluation of objective functions for predicting intracellular fluxes in Escherichia coli. *Molecular Systems Biology*. https://doi.org/10.1038/msb4100162

Städter, P., Schälte, Y., Schmiester, L., Hasenauer, J., & Stapor, P. L. (2021). Benchmarking of numerical integration methods for ODE models of biological systems. *Scientific Reports*, *11*(1). https://doi.org/10.1038/s41598-021-82196-2

Stitt, M., & Zeeman, S. C. (2012). Starch turnover: Pathways, regulation and role in growth. In *Current Opinion in Plant Biology* (Vol. 15, Issue 3, pp. 282–292). https://doi.org/10.1016/j.pbi.2012.03.016

Streb, S., & Zeeman, S. C. (2012). Starch Metabolism in Arabidopsis. *The Arabidopsis Book*, *10*, e0160. https://doi.org/10.1199/tab.0160

The Economic Cell Collective (2025), De Martino, A., Weiße, A., Kremling, A., Goelzer, A., Mauroy, B., Goupil, C., Karamaoun, C., De Groot, D., Giannari, D., Lacoste, D., Tourigny, D., Széliová, D., Oyarzún, D. A., Noor, E., Garcia, E. P., Herbert, É., Scott, F., Noël, F., …

Liebermeister, W. (2025). *Economic Principles in Cell Biology About this book*. https://doi.org/https://doi.org/10.5281/zenodo.12592398

Tolsma, J. E., & Barton, P. I. (1998). On computational differentiation. In *Computers chem. Engng* (Vol. 22, Issue 5).

Tourigny, D., Muriel, J., & Beber, M. (2020). dfba: Software for efficient simulation of dynamic flux-balance analysis models in Python. *Journal of Open Source Software*, *5*(52), 2342. https://doi.org/10.21105/joss.02342

Udvardi, M., & Poole, P. S. (2013). Transport and metabolism in legume-rhizobia symbioses. In *Annual Review of Plant Biology* (Vol. 64, pp. 781–805). https://doi.org/10.1146/annurev-arplant-050312-120235

Venkat, A., & Muneer, S. (2022). Role of Circadian Rhythms in Major Plant Metabolic and Signaling Pathways. *Frontiers in Plant Science*, *13*. https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2022.836244

Verberne, E. L. J., Hassink, J., De Willigen, P., Groot, J. J. R., & Van Veen, J. A. (1990). Modelling organic matter dynamics in different soils. In *Netherlands Journal of Agricultural Science* (Vol. 38).

Vigil, M. F., & Kissel, D. E. (1991). *Equations for Estimating the Amount of Nitrogen Mineralized from Crop Residues*.

Walworth, J. (2013). *NitrogeN iN Soil aNd the eNviroNmeNt*.

Webb, A. A. R., & Satake, A. (2015). Understanding Circadian Regulation of Carbohydrate Metabolism in Arabidopsis Using Mathematical Models. *Plant and Cell Physiology*, *56*(4), 586–593. https://doi.org/10.1093/pcp/pcv033

Williams, T. C. R., Poolman, M. G., Howden, A. J. M., Schwarzlander, M., Fell, D. A., Ratcliffe, R. G., & Sweetlove, L. J. (2010). A Genome-scale metabolic model accurately predicts fluxes in central carbon metabolism under stress conditions. *Plant Physiology*, *154*(1), 311–323. https://doi.org/10.1104/pp.110.158535

Yousfi, N., Slama, I., Ghnaya, T., Savouré, A., & Abdelly, C. (2010). Effects of water deficit stress on growth, water relations and osmolyte accumulation in Medicago truncatula and M. laciniata populations. *Comptes Rendus - Biologies*, *333*(3), 205–213. https://doi.org/10.1016/j.crvi.2009.12.010

Zakhartsev, M., Medvedeva, I., Orlov, Y., Akberdin, I., Krebs, O., & Schulze, W. X. (2016). *Metabolic model of central carbon and energy metabolisms of growing Arabidopsis thaliana in relation to sucrose translocation*. https://doi.org/10.1186/s12870-016-0868-3

Zamani, M., diCenzo, G. C., Milunovic, B., & Finan, T. M. (2017). A putative 3-hydroxyisobutyryl-CoA hydrolase is required for efficient symbiotic nitrogen fixation in

Sinorhizobium meliloti and Sinorhizobium fredii NGR234. *Environmental Microbiology*, *19*(1), 218–236. https://doi.org/https://doi.org/10.1111/1462-2920.13570

Zeeman, S. C., Smith, S. M., & Smith, A. M. (2006). The diurnal metabolism of leaf starch. *Biochemical Journal*, *401*(1), 13–28. https://doi.org/10.1042/BJ20061393