Technical University Munich

Department of Mathematics

Chair of Mathematical Modelling

Autophagy in Yeast Saccharomyces Cerevisiae: Single-Cell and Multi-Cell Model Approaches

Master's Thesis

Melanie Michels

Supervisor / Advisor: Prof. Dr. Christina Kuttler
Submission Date: 19.10.2020

I hereby declare that this thesis is my own work and that no other sources have been used except those clearly indicated and referenced.

Markt Schwaben, 19.10.2020

Melanie Michels

Abstract

Autophagy is a process, where components within a cell are degraded and recycled for different purposes. This can especially be helpful when cells are starving and they have to adapt rapidly to the new conditions to ensure their survival. For studying this process, yeast has become a model organism, since autophagy is easy to control and observe within these cells. Furthermore, autophagy could be connected to different human diseases like cancer or Huntington's disease. To get a better understanding of the basic principles and interrelations, mathematical models can be helpful. Therefore, we want to study the process of autophagy from a mathematical point of view.

In this work, we first set up two single-cell models for starvation induced autophagy in yeast cells and analyse their properties mathematically. In a next step, we expand the first single-cell model to the multi-cell case and perform again a mathematical analysis. Using data on cell densities, we then fit our model parameters and examine the sensitivity of the model to its parameters. In a final step, we introduce different modifications of the multi-cell model, to improve the fit of the model to our data. We compare the resulting model fits and discuss their different properties.

Contents

Glossary

$atg2\Delta$ the Atg2 gene is nonfunctional or missing.

AIC$_c$ a second order Akaike information criterion.

alkaline phosphatase a particular type of protein enzyme.

cytoplasm components within a cell except for the cell nucleus.

eukaryote organisms whose cells possess a nucleus and a nuclear membrane around it.

lysosome a membrane-bound compartment within an animal cell, containing enzymes to break down biomolecules.

prototrophic to be able to synthesize all required metabolites by itself.

vacuole a membrane-bound compartment within a plant or fungal cell, filled with fluid and containing enzymes.

Acronyms

AIC Akaike's Information Criterion.

Atg Autophagy-related.

BIC Bayesian Information Criterion.

GFP Green Fluorescent Protein.

IVP Initial Value Problem.

ODE Ordinary Differential Equation.

PAS Pre-Autophagosomal Structure.

RSS Residual Sum of Squares.

SEM Standard Error of the Mean.

1 Introduction

"Autophagy" comes from the Greek words "auto", meaning self, and "phagy", meaning eat. So strictly speaking we look at a process where a cell is "eating itself". To be more precise: autophagy is an intracellular process, where cytoplasm is degraded and recycled for different purposes. This was first observed in mammalian cells, but here the regulation of autophagy is highly complex and hence hard to study in all its details [1, 2]. As it turned out, autophagy in yeast is much easier to control and with the discovery of the first Autophagy-related (Atg) proteins, yeast became the model organism to study this process. In addition, this pathway is highly conserved among eukaryotes. So discoveries in yeast can be used as a basis for more studies in mammalian cells [3]. Furthermore, autophagy is nowadays known to be connected to different human diseases like cancer [4] or Huntington's disease [5]. Those are only a few reasons making it interesting to study autophagy also from a mathematical point of view.

In this work we will focus on starvation-induced autophagy in yeast, as it is the model organism for this process and hence biologically well studied. We will propose two single-cell models based on the biological findings described below. After analyzing them mathematically, we will expand our first model to the multi-cell case, examining again its mathematical properties. Next, we will try to fit our model to a data set and test different model modifications to improve the fit. We will also look at the sensitivities of the different models to their parameter sets. In a final step, we will discuss our results and give an outlook for further possible work on this topic.
All code used for this thesis can be found on `https://github.com/Melanie757/Thesis-Code`.

1.1 Biological Background

Under stress conditions like starvation, cells have to adapt rapidly to ensure their survival. To this end they can use autophagy to degrade and recycle cytoplasmic components. The recycled components can then be used, for example, to build new components, adapt to the new environmental conditions, or to provide energy [1].

Autophagy can be subclassified into three groups - macroautophagy, microautophagy and chaperone-mediated autophagy, whereas the last one only appears in mammalian cells [1]. For our models we will focus on macroautophagy, the most prevalent form, and hence just refer to it as autophagy.

1.1.1 Autophagy

The process of Autophagy As a first step of the process cytoplasmic components are sequestered by a membrane, called the phagophore or isolation membrane. When this membrane closes about the cellular material, it forms a double membrane vesicle, the autophagosome. In the next step the outer membrane of this vesicle fuses with the

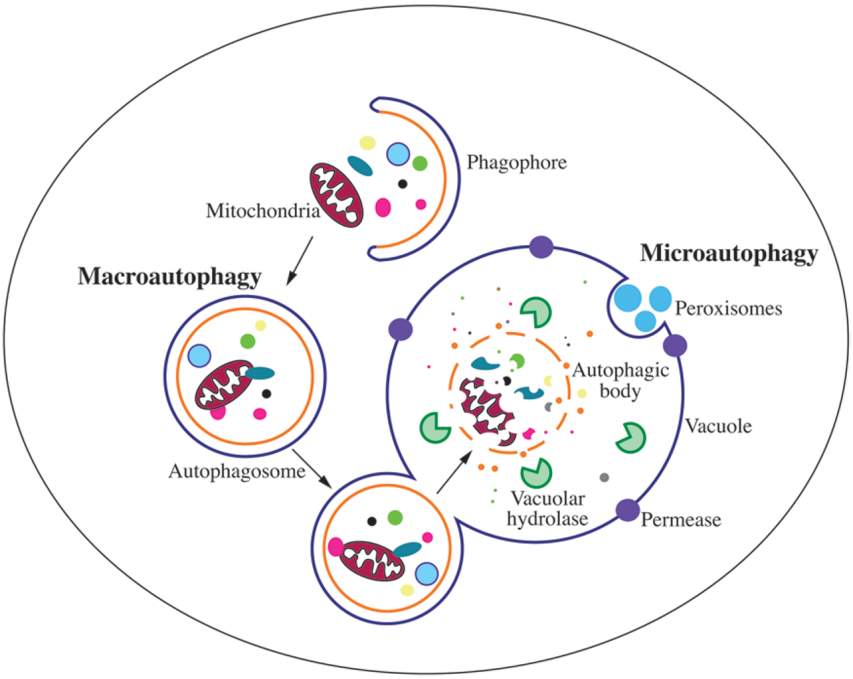

Figure 1.1: The process of autophagy in yeast cells. Cytoplasmic contents are sequestered by the phagophore. Next the phagophore closes to build the autophagosome. At last, the autophagosome fuses with the vacuole and the contents are released to be degraded. Taken from [1].

vacuole (in yeast or plants) or the lysosome (in mammals). The inner part and its contents can now be degraded and the catabolites can be recycled [3]. A visualization of this process is shown in fig. 1.1.

Selective and Bulk Autophagy A further differentiation of autophagy is selective and bulk autophagy. The process described above is the one of bulk autophagy. They both share most of the core machinery [3] and we therefore do not go into detail regarding the process of selective autophagy, as this is not relevant for our later models.

Selective autophagy can degrade unnecessary, damaged or even harmful cytoplasmic contents and thus helps the cell to stay in a stable and healthy state. This type of autophagy happens at a basal level regardless of starvation or other stress [3]. A common example in yeast is the Cytoplasm-to-vacuole targeting pathway, the so called Cvt pathway. Here only the hydrolase aminopeptidase I is selectively transported to the vacuole [6]. Other examples are the selective degradation of mitochondria (mitophagy) or of endoplasmic reticula (ER-phagy).

Bulk autophagy, in contrast, has no specific target for sequestration. In yeast, it is inhibited in nutrient-rich conditions and can easily be induced in two differente ways. One way is to shift the cells to a starvation medium. The second way uses an enzyme called TORC1 (target of rapamycin complex 1), which is crucial for the cell to sense nutrient availability. If this enzyme is inhibited by a rapamycin treatment, bulk autophagy is induced [3].

1.1.2 Monitoring Autophagy

There are different methods to analyse autophagy in yeast cells. In [3], the authors give a good overview of such approaches together with their advantages and limitations. We will not go into biological details here, so we refer the interested reader to [3] for more information on the methods.

The last steps in autophagy are the delivery of material to the vacuole and the following degradation. Methods trying to monitor this are called end-point measurements. One possibility to monitor bulk autophagy is to observe activated $Pho8\Delta60$ in the vacuole. $Pho8\Delta60$ is the alkaline phosphatase $Pho8$, where amino acids 1-60 were deleted. Outside the vacuole it appears in its inactive form. When it gets transported inside the vacuole via bulk autophagy, it is being activated. This activity can then be measured. Another possibility is to measure the delivery of $Atg8$ (Atg protein 8) to the vacuole. This can be done if $Atg8$ is tagged with Green Fluorescent Protein (GFP). These are the two most commonly used methods to assay bulk autophagy in yeast.

To measure selective autophagy, the above methods can be altered in the following ways. $Pho8\Delta60$ can be fused to a cargo protein or signal sequence to target a certain cargo organelle. After this, the activity in the vacuole can again be measured. For the second method, instead of $Atg8$, a cargo-specific protein can be tagged with GFP and measured inside the vacuole. Other approaches are the use of antibodies, to recognize proteins that are part of the cargo, or the isolation of vacuoles and probing for the cargo of interest. Another, quite straight forward, approach is to directly visualize the contents of the autophagosomes by electron microscopy.

One advantage of such measurements is, that they collect sensitive data about the function of the whole pathways. However, they offer no information on single steps in the pathways.

But there are also assays to obtain information on certain steps of the autophagic process. Before the phagophore can form, Atg proteins assemble at the vacuolar membrane to form the Pre-Autophagosomal Structure (PAS). These assemblies can then be monitored via fluorescence microscopy, to see whether the autophagic pathway starts. When the autophagosomes have formed, one can visualize them with electron microscopy. Many autophagosomes, which accumulate in the cytosplasm, could for example indicate that they fail to fuse with the vacuole. If this step is functional, the autophagosomes within the vacuole can be visualized by bright-field microscopy. One can see here as well, if there is an accumulation, indicating a disturbance of the degradation process. Another method to measure the number of autophagosomes within the vacuole is again using a GFP marker.

For a mathematical model, one could now take such measurements to quantify autophagy. In chapter 3 and chapter 4, we will set up models describing autophagy in yeast. Those models have no underlying data but use a qualitative level of autophagy between 0 and 1. In this way, they do not depend on a certain type of measurement but different types of data can be adapted to it. For example, if one has measurements of the number of autophagosomes in the vacuole, the highest number of autophagosomes could be set as a level of 1. Scaling the measurements accordingly, one can then do a parameter fit.

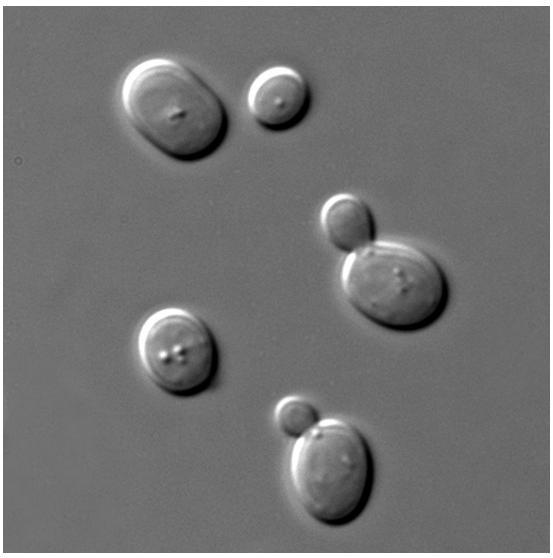

Figure 1.2: *Saccharomyces cerevisiae* cells in DIC microscopy. Taken from [8].

1.1.3 Yeast Saccharomyces Cerevisiae

For brewing and baking, the yeast *Saccharomyces cerevisiae* (fig. 1.2) has already been known for over ten thousand years. But for over a century, it is also part of laboratory research [7].

There are multiple reasons making *S. cerevisiae* a useful model organism for eukaryotes. First of all, it is a single-cell organism, hence small, and can be grown on defined media. This makes it easy to control its whole environment [9]. It has a short doubling time of about 100 minutes at 30°C [10], so effects of an experiment can be seen quite early and many generations can be studied. Furthermore, its lifecycle is ideally suited to conduct genetic analysis and efficient techniques were developed to study the impacts of mutations. Those are only some of the reasons, why *S. cerevisiae* became the first eukaryote, whose genome was completely sequenced in 1996 [9]. Leading to the setup of comprehensive gene databases like `https://www.yeastgenome.org`, this enabled researchers to create extensive genetic interaction networks for yeast with thousands of interactions [11]. For such a well-studied organism, comparison with other organism is then much easier.

Many studies have been made, investigating the connection of gene function in yeast and humans. For example, Kachroo et al. replaced essential yeast genes by their human orthologs, i.e. genes, that originated from the last common ancestor. In that way they could identify genes, which essentially perform the same role in both, yeast and human. Whether the substitution worked, was also strongly dependent on the pathway, in which the respective proteins took part [12]. As background for our work, it is particularly interesting that the basic machinery of autophagy is the same in yeast and in mammalian cells [13]. Therefore, many studies about autophagy focus on yeast, like [1, 3, 6, 13].

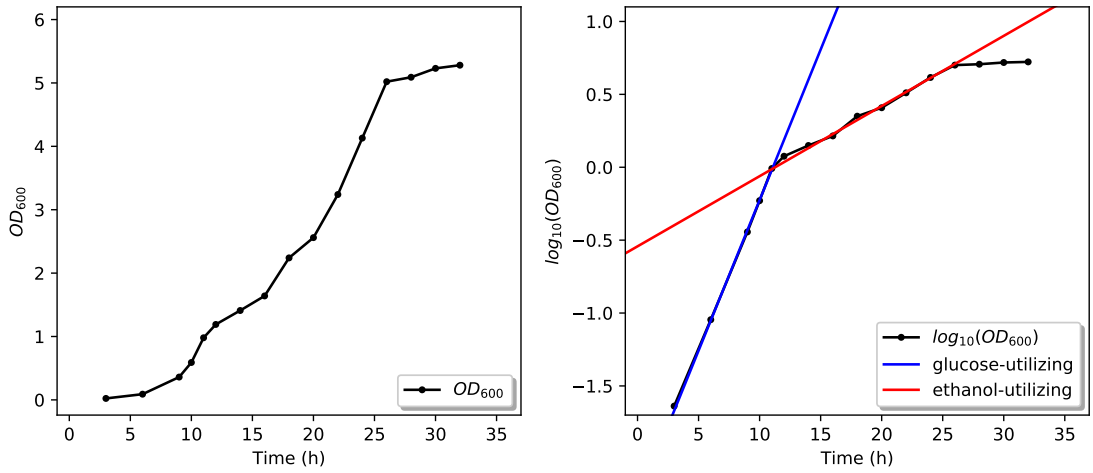

Figure 1.3: Left: Measurements of yeast cell density during growth on glucose (3-13 h) and then ethanol (13-26 h). Right: yeast cell density on a $\log_{10}$-scale and regression lines for the glucose-utilizing phase (using data from 6-11 h) and the ethanol-utilizing phase (using data from 16-24 h). Adapted from [15].

1.1.4 Diauxic Growth of Yeast

Diauxie comes from the Greek words "dis", meaning "twice", and "auxo" or "auxano", meaning "to grow". Hence, it is quite easy to see why this term should describe a growth with two different phases.

Yeast cells prefer to grow on carbon sources like glucose if they are available. This means, that if they are cultured in a medium containing glucose, the glucose is metabolized first and ethanol is produced simultaneously. Once the glucose is depleted, the yeast cells adapt their metabolism and switch to grow on the released ethanol instead. This adaption takes a short time, in which the growth rate is decreased. This is called the diauxic shift. After the depletion of ethanol, the cells go into a stationary phase. [14]

In section 4.4 we will try to fit our multi-cell model from section 4.2 to data taken from [15]. The authors of this paper wanted to investigate the diauxic growth of yeast cells and the starvation induced autophagy. Hence, this data provides a good example for this growth pattern of yeast. In the left plot of fig. 1.3 one can see the measured densities of the yeast culture throughout the experiment. Between 3 h and 13 h, the yeast cells grew on the available glucose. After this, between 13 h and 26 h, they used the produced ethanol to grow. That the growth changed between those two phases can be seen in the right plot. Here, the regression lines for the two growth phases are shown, using a logarithmic scale for the cell density.

1.2 Further Biological Findings

Autophagy is especially known as a response to starvation or nutritional stress. But it could also be linked to many other processes, like the suppression of tumors. In 1999, autophagy was connected to human breast cancer via the protein *beclin 1*, whose en-

coding gene is structurally similar to an autophagy gene in yeast. It could be shown, that *beclin 1* is less expressed in breast carcinoma cells than in normal breast epithelial cells. Enforced expression of this autophagy gene in carcinoma cells led to an inhibition of their tumour-forming potential [4].
A further process, where autophagy is known to take part in, is the intracellular defense against bacteria. Here a certain form of selective autophagy, xenophagy, targets invading bacteria [16]. Targeted are for example *S. typhimurium*, involved in human gastroenteritis [17], or *M. tuberculosis*, responsible for tuberculosis [18].
As already mentioned above, autophagy can also be connected to several neurodegenerative diseases like Alzheimer's, Parkinson's or Huntington's disease [5]. All are associated with an aggregation of misfolded proteins [19]. Ravikumar et al. studied the pathways mediating the degradation of such proteins for Huntington's disease. They inhibited or stimulated autophagy, resulting in an accumulation or enhanced clearance of the proteins [5].

2 Mathematical Prerequisites

One goal of mathematical modeling is to see the changes of a dynamical system in time or space. Therefore, a first step are discrete models. They consider just particular time steps or single locations in space. But in many cases this is not enough and a continuous model is needed. A handy way of describing a system continuously, for example in time, are differential equations. In this chapter we first give a short introduction to ordinary differential equations and related theorems needed in this work. After that we address a few common growth models and important concepts such as stationary states and their stability, for the analysis of our models.

2.1 Ordinary Differential Equations

The following is based on [20] and [21].

An Ordinary Differential Equation (ODE) is the relation of an unknown function $y(t)$ with its derivatives $\dot{y} = \frac{dy}{dt}(t), \ddot{y}, ..., y^n$. In contrast to Partial Differential Equations, an ODE depends only on derivatives in one variable, here time t. Together with an initial value we get a so-called Initial Value Problem (IVP)

$$\begin{aligned} \frac{dy}{dt}(t) &= f(t, y(t)) \\ y(t_0) &= y_0 \end{aligned} \tag{2.1}$$

where $f : D \to \mathbb{R}^n$ and $D \subseteq \mathbb{R} \times \mathbb{R}^n$ is nonempty, open and connected. In the context of this work we often set $t_0 = 0$ and $t \geq t_0$.
If the function f does only depend on $y(t)$ and not on t itself, the ODE is called autonomous.
The order of an ODE is the order of the highest derivative included in the ODE, so the above IVP is of first order.
A solution to the IVP is a function $y : I \to \mathbb{R}^n$, where $I \subseteq \mathbb{R}$ is a nonempty interval containing t_0 and y satisfies the following properties:

(i) $\dot{y}(t)$ exists for all $t \in I$,

(ii) $(t, y(t)) \in D$ for all $t \in I$,

(iii) $\dot{y}(t) = f(t, y(t))$ for all $t \in I$,

(iv) $y(t_0) = y_0$.

But how do we know whether an ODE, resp. IVP, has a solution at all? We therefore look at the following theorem.

Theorem 2.1 (Peano's Existence Theorem)**.** *Let $f \in C(D, \mathbb{R}^n)$ with $(t_0, y_0) \in D$. Then there exists $\epsilon > 0$ such that the IVP (2.1) has at least one solution in $[t_0 - \epsilon, t_0 + \epsilon]$.*

A proof can be found in [20].
Another important point is the uniqueness of a solution and the domain, where it exists. We will therefore use the following theorems and definitions from [21, pp. 7-8].

Theorem 2.2 (Picard-Lindelöf)**.** *Let f be continuous with respect to t and locally Lipschitz with respect to y on D. Then for each $(t_0, y_0) \in D$ there exists an $\epsilon > 0$ such that the IVP* (2.1) *has a unique solution $y : [t_0 - \epsilon, t_0 + \epsilon] \to \mathbb{R}$.*

This theorem gives us the existence of a unique solution on a certain interval, but not what happens with two different solutions which have intersecting domains. What we would like to have in this situation is, that the solutions on the intersection coincide. Then we would also have a unique solution on intersecting intervals. The following theorem states exactly this.

Theorem 2.3 (Uniqueness of solution)**.** *Let f be as in theorem 2.2 and $(t_0, y_0) \in D$. Let y_1 be the solution to IVP* (2.1) *on interval I_1 and y_2 on interval I_2. Then for all $t \in I_1 \cap I_2$*

$$y_1(t) \equiv y_2(t) \tag{2.2}$$

Now we also want to know the maximal interval, where we have a unique solution.

Definition 2.4 (Continuability)**.** *Consider $(t_0, y_0) \in D$ and let $y(t)$ be the solution to IVP* (2.1) *on interval I. It is said that solution $y(t)$*

(i) can be continued on the right, if there exists another solution $y_1(t)$ on interval I_1, such that $I \subset I_1$ and $\sup I$ belongs to the interior of I_1.

(ii) can be continued on the left, if there exists another solution $y_2(t)$ on interval I_2, such that $I \subset I_2$ and $\inf I$ belongs to the interior of I_2.

(iii) is continuable, if it can be continued on the right or on the left, or both.

Definition 2.5 (Maximal Solution)**.** *A solution to the IVP* (2.1) *is called a maximal solution, if it is not continuable.*

Theorem 2.6 (Existence and Uniqueness of a Maximal Solution)**.** *Let $D \subseteq \mathbb{R} \times \mathbb{R}^n$ be open, nonempty and connected and let f be as in theorem 2.2. Then, for each $(t_0, y_0) \in D$, there exists a unique maximal solution to the IVP* (2.1)*. Moreover, the interval of definition of such a maximal solution, denoted by $I_{max} = I_{max}(t_0, y_0)$, is open.*

Theorem 2.7. *Let (a, b) denote the maximal interval of existence for the IVP* (2.1)*. If $|b| < \infty$, then $lim_{t \to b^-} |y(t)| = \infty$. Similarly for a.*

Remark 2.8. *We often want to study the long term behaviour of a system and would like to have $b = +\infty$ in the maximal interval of existence (a, b). Using theorem 2.7, we can achieve this if we have a bounded solution. This holds, since for $|b| < \infty$ the solution would be unbounded for the right boundary of the interval. So, if our solution is bounded, $|b|$ cannot be finite but has to be $+\infty$, as $t_0 \geq 0$ lies in the interval.*

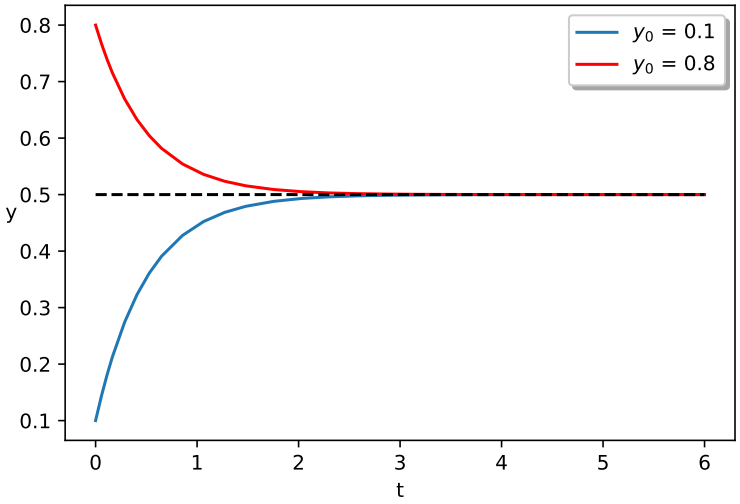

Figure 2.1: Example for model (2.3) with $a = 2$.

2.2 Growth Models

To model the increase and decrease of autophagy in cells or whole cell populations, we need some basic growth models. Since we will not find a negative population of cells or negative level of autophagy happening, we from now on assume $y(t) \geq 0$. As mentioned above, we also set $t_0 = 0$ for our models.

2.2.1 Verhulst: Saturation

One of the simplest examples of a limited growth model is

$$\dot{y}(t) = 1 - ay(t) \tag{2.3}$$

where $a > 0$ is the death rate. The solution $y(t)$ will always converge to $\frac{1}{a}$. This is quite intuitive, as $\dot{y}$ will be negative for $y > \frac{1}{a}$ and positive for $y < \frac{1}{a}$. An example can be found in fig. 2.1. Later we will use this in parts of our models for the increase of autophagy.
One advantage of this model, besides the boundedness, is its simplicity. So it is easy to handle. But a main disadvantage for population modeling is the steep increase at the beginning. If there are initially only a few individuals, the population will first probably not grow that fast. Instead, the growth will be slower and then increase with the number of individuals.

2.2.2 Verhulst: Logistic Growth

To deal with the mentioned problem of the last model, we next introduce the well known logistic growth model. It can be described in the following way:

$$\dot{y}(t) = ay(t)\left(1 - \frac{y(t)}{K}\right) \tag{2.4}$$

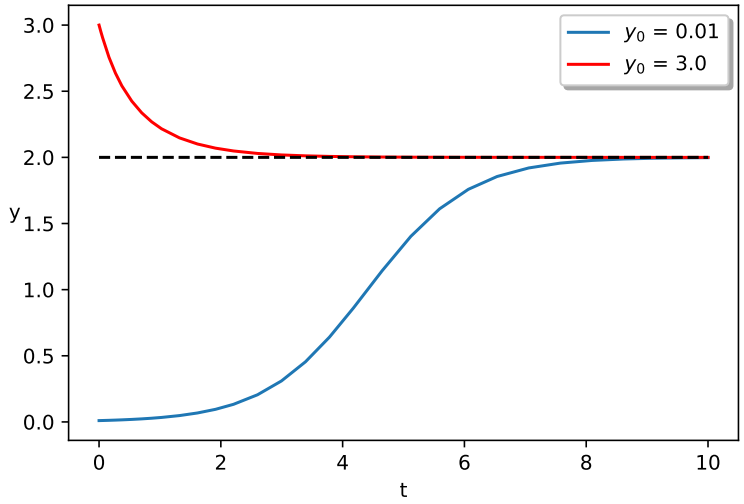

Figure 2.2: Example for model (2.4) with $a = 1.2$ and $K = 2$.

where $a > 0$ is the so-called rate constant and $K > 0$ the carrying capacity. An example can be found in fig. 2.2. We assume both a and K to be constant, but of course it would also be possible to choose time dependent versions $a(t)$ and $K(t)$.

For small values of y we have an approximately exponential growth with exponent a, but with increasing values of y the growth is extenuated proportional to the remaining capacity [22]. Here, $y(t)$ will converge to K if $y(0) > 0$ and stay 0 if we start with $y(0) = 0$. The solution to this ODE is given as

$$y(t) = \frac{y_0 K}{e^{-at}(K - y_0) + y_0} \tag{2.5}$$

A proof can be found in appendix 1.1, respectively [22, pp. 176-177]. Using this solution, we can furthermore show that $y(t)$ is point symmetric around $\frac{K}{2}$, when using a starting value between 0 and K (see appendix 1.1).

2.2.3 Generalized Logistic Growth

One can see in fig. 2.2, that the logistic growth model shows a symmetric S-shape around $\frac{K}{2}$. But what if we needed a similar but not necessarily symmetric characteristic for our model? Then the generalized logistic growth model [23] can be useful. The corresponding ODE can be formulated in the following way:

$$\dot{y}(t) = \frac{a}{n} y(t) \left(1 - \left(\frac{y(t)}{K}\right)^n\right) \tag{2.6}$$

where a, K and n are positive. If we choose $n > 1$ for example, as in fig. 2.3, we can get a slower increase for small values of y and a steeper increase when getting near K, due to the n-th power included. Hence, in contrast to the logistic growth from section 2.2.2, we get a non-symmetric solution curve. Of course, when setting $n = 1$, we arrive again at the (normal) logistic growth model.

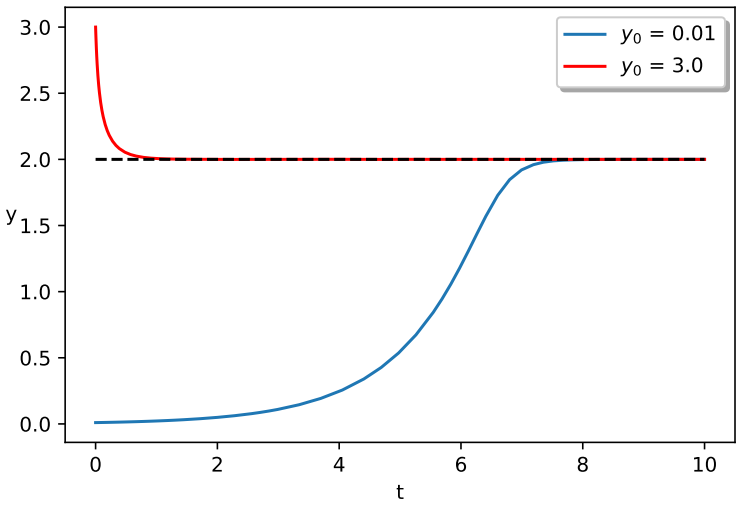

Figure 2.3: Example for model (2.6) with $a = 4$, $K = 2$ and $n = 5$.

2.3 Tools for Model Analysis

After proposing a model, we also want to find different properties to characterize its behaviour. For example, whether the model exhibits stationary states, like all of the models above, and whether this states are stable. We mostly follow here the book of Müller and Kuttler [22].

Definition 2.9. *A stationary state (also called equilibrium or stationary point) of an IVP, as stated in* (2.1)*, is a value of y that does not change in time anymore.*
Using the ODE formulation, this means $\dot{y}(t) = f(y(t)) = 0$ for such a value of y. We often denote a stationary state by y^*.

Example 2.10. *For the saturation model in section 2.2.1, we obtain the following stationary state y^*: $\dot{y}(t) = 0 \Leftrightarrow 1 - ay^* = 0 \Leftrightarrow y^* = \frac{1}{a}$.*
This coincides with the findings from the example graphs.

The next question is, whether a stationary state is stable. Roughly speaking, we want to know: if we are near the stationary state, do we reach it or at least stay near it? For example, consider the logistic growth in section 2.2.2. There we have two stationary states, 0 and K. If we start above K or below K, but bigger than 0, we always reach K in the end. So K is stable. But this is different for the 0. If we start with any $y_0 > 0$, we will never reach 0 again but converge to K. So 0 is not a stable stationary state.
The following definitions are taken from [22, Definition 2.7 and 2.8]

Definition 2.11 (Stability of a Stationary State)**.** *A stationary state y^* is called stable, if for all $\epsilon > 0$ there exists a $\delta > 0$ such that for all $y(t_0) = y_0$ with $|y_0 - y^*| < \delta$ it is $|y(t) - y^*| < \epsilon$ for all $t > t_0$.*

Remark 2.12. *If a stationary state fulfills the above definition, we know $y(t)$ will stay near it if we start close enough. But it does not guarantee us that we will eventually reach it. Therefore, we look at the next definition.*

Definition 2.13 (Asymptotic Stability of a Stationary State)**.** *A stationary state y^* is called asymptotically stable, if it is stable and there exists a $\delta_0 > 0$ such that for all y_0 with $|y_0 - y^*| < \delta_0$ it holds*

$$\lim_{t\to\infty} |y(t) - y^*| = 0 \tag{2.7}$$

Now we want to study some criteria to determine the stability of a stationary state. Let us first consider the linear case $\dot{y}(t) = Ay(t)$, where $A \in \mathbb{C}^{n\times n}$. Then the following proposition from [22, Proposition 1.60] holds.

Proposition 2.14. *Let $\sigma(A)$ be the spectrum of A. Then:*

(i) *0 is asymptotically stable $\Leftrightarrow$ $Re\,\sigma(A) < 0$*

(ii) *0 is stable $\Leftrightarrow$ $Re\,\sigma(A) \leq 0$ and all eigenvalues λ with $Re\,\lambda = 0$ are semi-simple*

(iii) *If there is a $\lambda \in \sigma(A)$ with $Re\,\lambda > 0$, then 0 is unstable. (The reverse direction may be wrong!)*

Since we often have nonlinear models, we also need some criterion for this case. The simplest way is to go back again to the linear case, by using a linearisation (at least locally). Assume we have an autonomous ODE $\dot{y} = f(y)$ with $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$, a stationary state $y^* \in \mathbb{R}^n$ (thus $f(y^*) = 0$) and some small perturbation $p(t)$. We are interested in solutions $y(t)$ in the neighbourhood of y^*, so $p(t)$ has to be small there. Using Taylor expansion for such a $y(t) = y^* + p(t)$, this yields

$$\dot{p}(t) = \dot{y}(t) = f(y(t)) = f(y^* + p(t)) = \underbrace{f(y^*)}_{=0} + \dot{f}(y^*)p(t) + o(||p||) \tag{2.8}$$

As $p(t)$ is assumed to be small close to y^*, we arrive at the corresponding linearised system

$$\dot{p}(t) = Ap(t) \tag{2.9}$$

where $A = \dot{f}(y^*)$ is the Jacobian matrix of f at y^*.
We can now use this linearisation for our stability analysis of y^*, using again a Proposition from [22, Proposition 2.17].

Proposition 2.15 (Linearisation, Stability, Perron, Lyapunov)**.** *If the real parts of all eigenvalues of $A = \dot{f}(y^*)$ are negative, then y^* is exponentially asymptotically stable, i.e. there are constants $\delta, C, \alpha > 0$ such that $||y(0) - y^*|| < \delta$ implies*

$$||y(t) - y^*|| < Ce^{-\alpha t} \quad for\ t \geq 0 \tag{2.10}$$

From $Re\,\sigma(A) \cap (0, \infty) \neq \emptyset$ it follows that y^* is unstable.

Remark 2.16. *For a linear, two dimensional system, one can also have a look at fig. 2.4 and table 2.1, which give a good overview.*

This is a very helpful statement, but does not tell us whether our linearisation and nonlinear model also correspond locally. Hence we need one more characterisation of y^* [22, Definition 2.18, Theorem 2.19].

Definition 2.17. *y^* is called hyperbolic, if $0 \notin Re\,\sigma(A)$.*

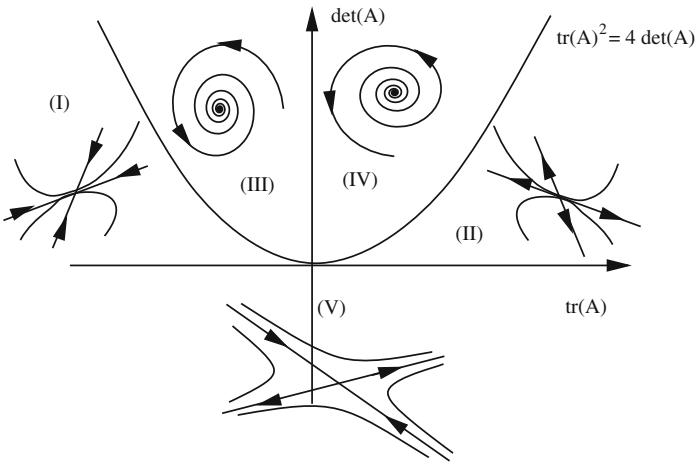

Figure 2.4: Characterization of stationary points, taken from [22].

(I)	stable node	both eigenvalues real, negative
(II)	unstable node	both eigenvalues real, positive
(III)	stable spiral	eigenvalues with a non trivial imaginary, negative real part
(IV)	unstable spiral	eigenvalues with a non trivial imaginary, positive real part
(V)	saddle	both eigenvalues real, one negative and one positive

Table 2.1: Different types of stationary points, adapted from [22].

Theorem 2.18 (Hartman and Grobman)**.** *Let y^* be hyperbolic. Then there is a neighbourhood U of y^* and a homeomorphism $H : U \to \mathbb{R}^n$ with $H(y^*) = 0$, which maps the trajectories of $\dot{y} = f(y)$ one-to-one into trajectories of $\dot{p} = Ap$, with respect to the time course.*

This means if $y(t)$ is a solution to $\dot{y} = f(y)$, then $p(t) = H(y(t))$ is a solution to $\dot{p} = Ap$ and vice versa (as long as $y(t)$ does not leave U). So, as long as y^* is hyperbolic, we can try and use proposition 2.15 for our stability analysis.

If we alter model parameters, the stationary states and their stability can change. Consider the following simple example.

Example 2.19. *Take the logistic growth from section 2.2.2 with $\dot{y}(t) = f(y(t)) = ay(t)\left(1 - \frac{y(t)}{K}\right)$, but $a \in \mathbb{R}$ instead of $a > 0$. As seen above, the logistic growth model has two stationary states, 0 and K. For $a > 0$, K is stable and 0 unstable. But if we now also take $a < 0$ into account, the stationary states change their stability properties:*

$$f'(y) = a\left(1 - \frac{2y}{K}\right) \tag{2.11}$$

so we get $f'(0) = a < 0$ and $f'(K) = a(1-2) = -a > 0$. As a consequence, 0 is stable and K unstable [22, Theorem 2.9].

This example is quite constructed, as we normally just take $a > 0$. But to get an impression of changing stability it should be enough. The phenomenon of such a change in stability depending on parameters is called bifurcation. A parameter, for which bifurcation is studied, is called bifurcation parameter. There are many types of bifurcation, hence we just mention a few examples here. See [22, pp. 223-227] for more details.

- Saddle-Node bifurcation: above a certain threshold of the bifurcation parameter, a saddle and a node appear.
- Transcritical bifurcation: two stationary states exchange stability at the parameter value, where both points meet.
- Pitchfork bifurcation: for parameter values below a threshold, e.g. $a \leq a_0$, a stationary point is stable, for $a > a_0$ two additional, stable stationary states show up and the original one becomes unstable.

Another interesting point is the positivity of solutions. Many biological systems do not exhibit negative states, e.g. when modeling cell populations. So if a model has a solution, we need to find out if it stays positive and hence biologically meaningful. Therefore, we look at the following proposition from [24, Proposition B.7.]:

Proposition 2.20 (Non-negativity of Solutions)**.** *Suppose the IVP* (2.1) *has the following properties:*

(i) solutions $y(t)$ with initial value $y_0 \geq 0$ are unique, and

(ii) whenever $y(t) \geq 0$ satisfies $y_i(t) = 0$, it holds $f_i(y) \geq 0$, for all $i \in \{1, \ldots, n\}$.

Then it follows $y(t) \geq 0$ for all $t \geq t_0$ in the domain of y, provided $y_0 \geq 0$.

3 Single-Cell Models

Here we start with our first models for autophagy in yeast. In this chapter we only consider one single cell, later we will also look at whole populations (see chapter 4). The following models involve three main components:

(i) the level of autophagy $a(t)$ from 0 to 1, where a level of 0 corresponds to no autophagy happening and a level of 1 to the maximal production of autophagosomes,

(ii) the level of cytoplasmic material $m(t)$, which can be degraded (and produced) by autophagy, from 0 to 1,

(iii) the level of catabolites $c(t)$, which are also produced through autophagy and used to build new cytoplasmic material, from 0 to 1.

Additionally, we include by a function $s(t)$ whether the cell is in a nutrient-rich environment or a starvation medium. If $s(t) = 0$, enough nutrients are available, but if $s(t) = 1$, there are no nutrients in the environment at all. So we could say, that $s(t)$ represents the current "starvation state".
Our first model disregards the catabolites and only models the levels of autophagy and the cytoplasmic material, often just called material in the following. The second model also takes the catabolites into account.
Note, that both models are dimensionless, since we only want to study levels here.

3.1 Autophagy and Cytoplasmic Material

As a first model, we chose to consider just autophagy and cytoplasmic material. The assumption hereby is, that catabolites are straight converted back to material, such that it is enough to model the material only. This is a simplification made to get a most simple model, which is easy to analyse for a first impression of the dynamics. The second assumption made is that $s(t) \in \{0, 1\}$. This is a quite reasonable simplification, as starvation stress for the cell can rapidly be induced, e.g. by shifting them to a starvation medium [3]. Otherwise, it is also easy to keep them in a "perfect" environment.

3.1.1 Model

We first consider the case with no starvation. Here we want a stable level of autophagy, the basal or selective autophagy, and therefore use the Verhulst saturation model from section 2.2.1. The mentioned disadvantages of this growth model are not that important here, as we just want the level of autophagy to converge to a stable, positive state. Furthermore, the formation of new autophagosomes presumably does not depend directly on the present ones, other than when cell populations grow by cell division. So there might be a steep increase or decrease at the beginning possible, but this is only an assumption.

Parameter	Description
α_1	rate constant of the log. growth of a
α_2	inhibition of a due to m
β	increase of m due to autophagy
γ	degradation of a
δ	normal degradation of m
M	supply of new material

Table 3.1: Description of the parameters in eq. (3.1).

We do not add a dependence on the cytoplasmic material here, as there should always be enough available in nutrient-rich conditions. For the cytoplasmic material we assume a constant supply M of new material, an increase due to autophagy and a degradation, dependent on the present level of material. Like mentioned above, we omit in this model the catabolites, which are produced through autophagy and then used for building new material. So actually, here the material is just first degraded and then produced again due to autophagy. Hence, we have a term $\beta_1 m(t)a(t) - \delta_1 m(t)a(t) = (\beta_1 - \delta_1)m(t)a(t)$ in the equation for the change of material. To simplify this term, we set $\beta := \beta_1 - \delta_1$. Furthermore, we assume $\beta > 0$ in the following, as otherwise the autophagy would in sum just degrade the level of material and would give the cell no real advantage in this model, especially in the starvation case.
When the cell is in a starvation medium, we want the autophagy level to increase until at most 1, dependent on the material available, and an inhibition of increase or even a decrease if there is insufficient material left. For modeling the increase we use the logistic growth from section 2.2.2 and for the decline an additional similar term dependent on the material level. For the level of material, the constant supply of material is omitted as there are no nutrients available in the environment.
All of the parameters are given without a unit, since we only study qualitative levels. The resulting ODE system is

$$\begin{aligned}
\dot{a}(t) &= (1 - s(t))\underbrace{(1 - \gamma a(t))}_{\text{saturation}} + s(t)(\underbrace{\alpha_1 a(t)(1 - a(t))}_{\text{log. growth}} - \underbrace{\alpha_2 a(t)(1 - m(t))}_{\text{inhibition}}) \\
\dot{m}(t) &= (1 - s(t))\underbrace{M}_{\text{supply}} + \underbrace{\beta m(t)a(t)}_{\substack{\text{growth due to} \\ \text{autophagy}}} - \underbrace{\delta m(t)}_{\text{degradation}} \\
&= (1 - s(t))M + \beta m(t)\left(a(t) - \frac{\delta}{\beta}\right)
\end{aligned} \tag{3.1}$$

where all parameters are in $\mathbb{R}_{>0}$. A description of their meaning can be found in table 3.1.

3.1.2 Mathematical Analysis

First we check if our model has a non-negative, unique solution, so whether it is meaningful for our application. Next we analyse the stationary states of the system and look at their stability. At last we examine the boundedness and domain of solutions.

Existence of a non-negative, unique solution Given our system (3.1), we want to know if there exists a solution, which is hopefully even unique. We differentiate between the cases without starvation, $s(t) = 0$, and with starvation, $s(t) = 1$.

Starting with $s(t) = 0$, our function f_n (for **n**o starvation) looks like

$$f_n \begin{pmatrix} a(t) \\ m(t) \end{pmatrix} = \begin{pmatrix} 1 - \gamma a(t) \\ M + \beta m(t) \left(a(t) - \frac{\delta}{\beta} \right) \end{pmatrix} \tag{3.2}$$

f_n is continuously differentiable with respect to a and m and hence its derivative stays bounded in $[0,1]^2$. Thus, f_n is Lipschitz continuous in $[0,1]^2$ with respect to a and m. It is also continuous with respect to t, as f_n is independent of t. Using theorem 2.2, we get that for some $\epsilon > 0$ there exists a unique solution in $[t_0 - \epsilon, t_0 + \epsilon]$. By theorem 2.6 there even exists a unique maximal solution on an open interval for initial values in $(0,1)^2$. So we have locally unique solutions, hence the first requirement of proposition 2.20 is fulfilled. Now we check the second requirement:

- Let $a(t) = 0$ and $m(t) \geq 0$, then $1 - \gamma a(t) = 1 > 0$.
- Let $m(t) = 0$ and $a(t) \geq 0$, then $M + \beta m(t) \left(a(t) - \frac{\delta}{\beta} \right) = M \geq 0$.

Consequently, a solution to our model stays non-negative, if we start with positive initial values.

Now we consider the case $s(t) = 1$:

$$f_s \begin{pmatrix} a(t) \\ m(t) \end{pmatrix} = \begin{pmatrix} \alpha_1 a(t)(1 - a(t)) - \alpha_2 a(t)(1 - m(t)) \\ \beta m(t) \left(a(t) - \frac{\delta}{\beta} \right) \end{pmatrix} \tag{3.3}$$

f_s satisfies again the requirements of theorem 2.2, using the same arguments as above. Hence, there exists a unique solution in $[t_0 - \epsilon, t_0 + \epsilon]$ for some $\epsilon > 0$ and even a unique maximal solution by theorem 2.6 on an open interval. We check the second requirement of proposition 2.20:

- Let $a(t) = 0$ and $m(t) \geq 0$, then $\alpha_1 a(t)(1 - a(t)) - \alpha_2 a(t)(1 - m(t)) = 0 \geq 0$.
- Let $m(t) = 0$ and $a(t) \geq 0$, then $\beta m(t) \left(a(t) - \frac{\delta}{\beta} \right) = 0 \geq 0$.

Thus, a solution to our model with starvation stays non-negative, if we start with positive initial values.

Stationary states Consider first $s(t) = 0$. Then we want to solve $f_n = 0$:

$$\begin{aligned} f_{n,1} = \dot{a}(t) = 0 \Leftrightarrow 0 = 1 - \gamma a^* \Leftrightarrow a^* = \frac{1}{\gamma} \\ f_{n,2} = \dot{m}(t) = 0 \Leftrightarrow 0 = M + \beta m^* \left(a^* - \frac{\delta}{\beta} \right) \Leftrightarrow m^* = \frac{M}{\delta - \beta a^*} \end{aligned} \tag{3.4}$$

So the only possible stationary state is $P = (a^*, m^*) = (\frac{1}{\gamma}, \frac{M}{\delta - \beta a^*})$. To be relevant for our model, a^* and m^* have to be in the interval $[0,1]$. Hence we get the additional constraints $\gamma \geq 1$ and $M \leq \delta - \frac{\beta}{\gamma}$.

For $s(t) = 1$, we examine $f_s = 0$:

$$\begin{aligned} f_{s,1} = \dot{a}(t) = 0 \Leftrightarrow 0 &= \alpha_1 a^*(1-a^*) - \alpha_2 a^*(1-m^*) \\ &\Leftrightarrow a^* = 0 \;\text{ or }\; a^* = 1 - \frac{\alpha_2}{\alpha_1}(1-m^*) \\ f_{s,2} = \dot{m}(t) = 0 \Leftrightarrow 0 = \beta m^* \left(a^* - \frac{\delta}{\beta}\right) &\Leftrightarrow m^* = 0 \;\text{ or }\; a^* = \frac{\delta}{\beta} \end{aligned} \tag{3.5}$$

Thus, we arrive at the following stationary states:

- $P_0 = (0, 0)$,
- $P_1 = (1 - \frac{\alpha_2}{\alpha_1}, 0)$,
- $P_2 = (\frac{\delta}{\beta}, 1 - \frac{\alpha_1}{\alpha_2}(1 - \frac{\delta}{\beta}))$, by setting equal $a^* = 1 - \frac{\alpha_2}{\alpha_1}(1 - m^*)$ and $a^* = \frac{\delta}{\beta}$, then solving for m^*.

These stationary states are meaningful for our model, if they lie between 0 and 1. This gives us some constraints for our parameters:

- for P_1: $\alpha_2 < \alpha_1$ (if $\alpha_2 = \alpha_1$, we arrive at P_0),
- for P_2: $\delta \leq \beta$ and $1 - \frac{\alpha_2}{\alpha_1} < \frac{\delta}{\beta}$ $\big($if $1 - \frac{\alpha_2}{\alpha_1} = \frac{\delta}{\beta}$, we arrive at $P_1\big)$.

Stability We start again with $s(t) = 0$.
From a biological point of view we want P to be stable, since the cell has enough nutrients available and should therefore stabilize at a positive level. For the mathematical analysis, we first determine the Jacobian matrix of system (3.1):

$$J(a, m) = \begin{pmatrix} -\gamma & 0 \\ \beta m & \beta a - \delta \end{pmatrix} \tag{3.6}$$

Now we insert our stationary state (a^*, m^*):

$$J(a^*, m^*) = J\left(\tfrac{1}{\gamma}, \tfrac{M}{\delta - \beta a^*}\right) = \begin{pmatrix} -\gamma & 0 \\ \beta \frac{M}{\delta - \beta a^*} & \beta a^* - \delta \end{pmatrix} \tag{3.7}$$

As this is a lower triangular matrix, we can instantly read off the two eigenvalues $-\gamma$ and $\beta a^* - \delta = \frac{\beta}{\gamma} - \delta$. Both are negative, if $\frac{\beta}{\gamma} - \delta < 0$. This is the case when P is biologically meaningful, since then $0 < M \leq \delta - \frac{\beta}{\gamma}$. Applying proposition 2.15 and theorem 2.18, it follows that there is a neighbourhood where (a^*, m^*) is exponentially asymptotically stable.

For $s(t) = 1$, the Jacobian matrix reads

$$J(a, m) = \begin{pmatrix} \alpha_1(1 - 2a) - \alpha_2(1 - m) & \alpha_2 a \\ \beta m & \beta a - \delta \end{pmatrix} \tag{3.8}$$

We first insert P_0:

$$J(0, 0) = \begin{pmatrix} \alpha_1 - \alpha_2 & 0 \\ 0 & -\delta \end{pmatrix} \tag{3.9}$$

This is a diagonal matrix, so the eigenvalues are just the diagonal entries. $-\delta$ is already negative and, under the assumption that $\alpha_1 < \alpha_2$, $\alpha_1 - \alpha_2$ is also negative. With proposition 2.15 and theorem 2.18, we can deduce that P_0 is stable if $\alpha_1 < \alpha_2$.
Next we insert P_1:

$$\begin{aligned} J\Big(1-\tfrac{\alpha_2}{\alpha_1},0\Big) &= \begin{pmatrix} \alpha_1(1-2(1-\frac{\alpha_2}{\alpha_1})) - \alpha_2(1-0) & \alpha_2(1-\frac{\alpha_2}{\alpha_1}) \\ \beta\cdot 0 & \beta(1-\frac{\alpha_2}{\alpha_1}) - \delta \end{pmatrix} \\ &= \begin{pmatrix} \alpha_2-\alpha_1 & \alpha_2(1-\frac{\alpha_2}{\alpha_1}) \\ 0 & \beta(1-\frac{\alpha_2}{\alpha_1}) - \delta \end{pmatrix} \end{aligned} \tag{3.10}$$

This is an upper triangular matrix, hence the eigenvalues are again on the diagonal: $\alpha_2 - \alpha_1$ and $\beta(1-\frac{\alpha_2}{\alpha_1}) - \delta$. The first eigenvalue is negative if $\alpha_2 < \alpha_1$, the same condition as for the point to be biologically meaningful. The second one is negative if $1-\frac{\alpha_2}{\alpha_1} < \frac{\delta}{\beta}$. Using again proposition 2.15 and theorem 2.18, P_1 is stable if it is relevant for us and $1-\frac{\alpha_2}{\alpha_1} < \frac{\delta}{\beta}$.
Taking the biological meaning into account, P_1 is not a stationary state we want to be stable. If there is no cytoplasmic material left, there should also be no autophagy going on as the cell is dead by then. So for our purposes we need $\alpha_1 < \alpha_2$, such that P_0 is stable and P_1 is unstable and biologically irrelevant.
Inserting P_2 gives us:

$$\begin{aligned} J\Big(\tfrac{\delta}{\beta},1-\tfrac{\alpha_1}{\alpha_2}\Big(1-\tfrac{\delta}{\beta}\Big)\Big) &= \begin{pmatrix} \alpha_1(1-2\frac{\delta}{\beta}) - \alpha_2(1-1+\frac{\alpha_1}{\alpha_2}(1-\frac{\delta}{\beta})) & \alpha_2\frac{\delta}{\beta} \\ \beta(1-\frac{\alpha_1}{\alpha_2}(1-\frac{\delta}{\beta})) & \beta\frac{\delta}{\beta}-\delta \end{pmatrix} \\ &= \begin{pmatrix} -\alpha_1\frac{\delta}{\beta} & \alpha_2\frac{\delta}{\beta} \\ \beta(1-\frac{\alpha_1}{\alpha_2}(1-\frac{\delta}{\beta})) & 0 \end{pmatrix} \end{aligned} \tag{3.11}$$

Since we cannot read off the eigenvalues of this matrix, we use a representation of the eigenvalues with trace and determinant. For details see [22, p. 47]. In the following we denote the above matrix just by J, for readability reasons.

$$\lambda_\pm = \frac{1}{2}\Big(tr(J) \pm \sqrt{tr(J)^2 - 4det(J)}\Big) \tag{3.12}$$

For the stability analysis we are only interested in the signs of the eigenvalues. First, we look at the determinant:

$$\begin{aligned} det(J) &= -\alpha_1\frac{\delta}{\beta}\cdot 0 - \alpha_2\frac{\delta}{\beta}\cdot\beta\Big(1-\frac{\alpha_1}{\alpha_2}\Big(1-\frac{\delta}{\beta}\Big)\Big) \\ &= -\underbrace{\alpha_2\delta}_{>0}\underbrace{\Big(1-\frac{\alpha_1}{\alpha_2}\Big(1-\frac{\delta}{\beta}\Big)\Big)}_{>0 \text{ if } P_2 \text{ biol. relevant}} \end{aligned} \tag{3.13}$$

So the determinant ist negative, if P_2 is biologically relevant. Resulting from this, we see that $\sqrt{tr(J)^2 - 4det(J)} > |tr(J)|$ in eq. (3.12). Thus, we have one positive and one negative eigenvalue, which gives us an unstable stationary state using proposition 2.15.

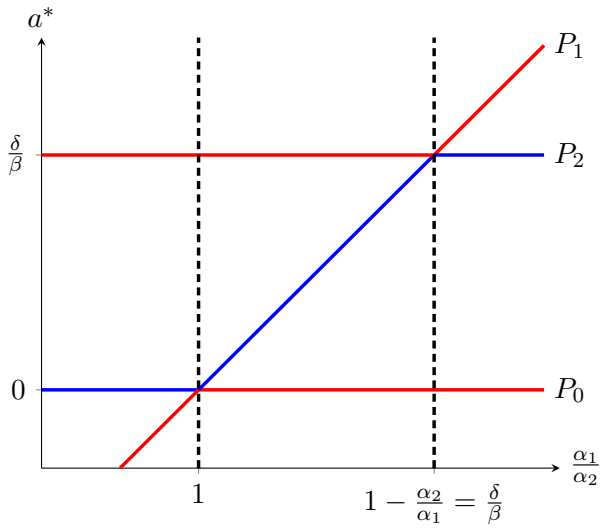

Figure 3.1: Bifurcation diagram for our ODE system (3.1) with s(t)=1, where blue shows stable and red unstable regions of the stationary points.

Bifurcation Above we characterized parameter values for our stationary states to be meaningful in our model setting. There we saw, that there exist parameter values, for which two states coincide.

First we consider P_0 and P_1. For $\alpha_1 < \alpha_2$, P_0 is stable and P_1 is unstable. In contrast, for $\alpha_1 > \alpha_2$, P_0 is unstable and P_1 is stable. For $\alpha_1 = \alpha_2$ both points coincide. So we have a transcritical bifurcation here.

We also saw that P_1 and P_2 coincide for $1 - \frac{\alpha_2}{\alpha_1} = \frac{\delta}{\beta}$. For $1 - \frac{\alpha_2}{\alpha_1} < \frac{\delta}{\beta}$, P_1 is stable and P_2 is unstable, as seen in the last paragraph. But if $1 - \frac{\alpha_2}{\alpha_1} > \frac{\delta}{\beta}$, the matrix in eq. (3.10) has one positive eigenvalue und P_1 becomes unstable. For P_2 we can see, that $det(J)$ in eq. (3.13) becomes positive. As the trace is negative, both eigenvalues become negative or have at least a negative real part. Thus, P_2 becomes stable. We again have a transcritical bifurcation.

The changes of stability for the different points are visualized in the bifurcation diagram in fig. 3.1.

Boundedness and domain Our solutions are bounded from below by 0, as we have shown that they are non-negative. It remains to show, that they are bounded from above.

For $s(t) = 0$, we get the following:

- As we consider here a simple saturation model for $a(t)$, like in section 2.2.1, we know that this will always converge to $\frac{1}{\gamma}$, no matter if we start from above or below.
  Formally, we get for $a > \frac{1}{\gamma}$ that $\dot{a} = 1 - \gamma a < 1 - 1 = 0$. Hence, $a(t)$ is decreasing as soon as it exceeds $\frac{1}{\gamma}$. From our model assumptions we know, that our initial value $a_0 := a(t_0) \leq 1$. Therefore, $a(t)$ is bounded from above by a_0, if $a_0 > \frac{1}{\gamma}$, or else by $\frac{1}{\gamma}$ itself.

- We first assume $a_0 \leq \frac{1}{\gamma}$, hence $a(t)$ does not exceed $\frac{1}{\gamma}$ as seen above. For $m > \frac{M}{\delta - \frac{\beta}{\gamma}}$, we get $\dot{m} = M + m(\beta a - \delta) < M + m(\frac{\beta}{\gamma} - \delta) < M + \frac{M}{\delta - \frac{\beta}{\gamma}}(\frac{\beta}{\gamma} - \delta) = 0$, using $\frac{\beta}{\gamma} - \delta < 0$ when P is biologically relevant and stable. So m is bounded from above by its stationary state $\frac{M}{\delta - \frac{\beta}{\gamma}}$, or its initial value.

  Now we assume $a_0 > \frac{1}{\gamma}$, hence $a(t)$ is bounded by a_0 and $\dot{m} \leq M + m(\beta a_0 - \delta)$. Further we assume $\delta \neq \beta a_0$, which is in any case biologically unlikely. We look under which condition $\dot{m}$ is negative.

$$\begin{aligned} M + m(\beta a_0 - \delta) < 0 &\Leftrightarrow m(\beta a_0 - \delta) < -M \\ &\Leftrightarrow \begin{cases} m > \frac{M}{\delta - \beta a_0} & \text{if } \beta a_0 - \delta < 0 \\ m < -\frac{M}{\beta a_0 - \delta} & \text{if } \beta a_0 - \delta > 0 \end{cases} \end{aligned} \tag{3.14}$$

  In the second case m would be negative, which is not biologically meaningful. Thus, we additionally assume $\beta a_0 - \delta < 0$. So if $m > \frac{M}{\delta - \beta a_0}$, then $\dot{m} < 0$. Hence m is in this case bounded by $\frac{M}{\delta - \beta a_0}$, or by m_0. As m should not exceed 1 in our model, we need here $M \leq \delta - \beta a_0$ as an additional constraint.

For $s(t) = 1$ we assume $a, m > 0$, as otherwise $\dot{a}$ and $\dot{m}$ are already 0.

- If $a < \frac{\delta}{\beta}$, then $\dot{m} = m(\beta a - \delta) < m(\delta - \delta) = 0$. Thus, as long as a stays bounded by $\frac{\delta}{\beta}$, also m stays bounded by its initial value.

- $\dot{a}$ is negative, if and only if

$$\begin{aligned} & \alpha_1 a(1-a) - \alpha_2 a(1-m) < 0 \\ \Leftrightarrow \quad & \alpha_1 a(1-a) < \alpha_2 a(1-m) \\ \Leftrightarrow \quad & \alpha_1 (1-a) < \alpha_2 (1-m) \\ \Leftrightarrow \quad & (1-a) < \frac{\alpha_2}{\alpha_1}(1-m) \\ \Leftrightarrow \quad & a > 1 - \frac{\alpha_2}{\alpha_1}(1-m) \end{aligned} \tag{3.15}$$

  Therefore, $a(t)$ is bounded by $1 - \frac{\alpha_2}{\alpha_1}(1 - m(t))$, or its initial value if it lies above. As $a(t)$ should be bounded by a value of at most 1, we additionally need $1 - \frac{\alpha_2}{\alpha_1}(1 - m(t)) \leq 1$. This is equivalent to $m(t) \leq 1$, what we would like to show anyway.

So for m we know it is decreasing or constant, hence definitely bounded, if $a \leq \frac{\delta}{\beta}$. Thus, if $1 - \frac{\alpha_2}{\alpha_1}(1 - m(t)) \leq \frac{\delta}{\beta}$, it would be guaranteed that a and m stay bounded. Solving for $m(t)$, we get $m(t) \leq 1 - \frac{\alpha_1}{\alpha_2}(1 - \frac{\delta}{\beta})$.
We will consider two different cases for $\frac{\delta}{\beta}$:

- Let us first assume $\frac{\delta}{\beta} \geq 1$. Then $1 - \frac{\alpha_1}{\alpha_2}(1 - \frac{\delta}{\beta}) \geq 1$ and thus $m_0 \leq 1 - \frac{\alpha_1}{\alpha_2}(1 - \frac{\delta}{\beta})$ for all $m_0 \in [0, 1]$. Clearly, $a_0 \leq \frac{\delta}{\beta}$ for every $a_0 \in [0, 1]$. Hence, we get that for any meaningful initial values, m is only decreasing or constant and a stays bounded by $\frac{\delta}{\beta}$. In particular, a stays bounded by 1 as $m \leq m_0 \leq 1$.

- Now we assume $\frac{\delta}{\beta} < 1$, leading to $1 - \frac{\alpha_1}{\alpha_2}(1 - \frac{\delta}{\beta}) < 1$. Thus, as long as $m_0 \leq 1 - \frac{\alpha_1}{\alpha_2}(1 - \frac{\delta}{\beta})$ and $a_0 \leq \frac{\delta}{\beta} \leq 1$, a is bounded by $\frac{\delta}{\beta}$ and m is decreasing or constant. If $1 - \frac{\alpha_1}{\alpha_2}(1 - \frac{\delta}{\beta}) < m_0 \leq 1$ instead, such a threshold for a_0 is dependent on the value of m_0.

Unfortunately, we did not find the exact bound in terms of our parameters and initial values. But we used the chosen parameter set from section 4.4.4, fig. 4.5, to approximate the bound for a_0 by hand in this case. This shows, that also initial values $m_0 > 1 - \frac{\alpha_1}{\alpha_2}(1 - \frac{\delta}{\beta})$ can give bounded and reasonable solutions. To find a bound for a_0, we simply changed decimal point after decimal point of a_0 and plotted the solution until time point 10000. For the plots we used the function `solve_ivp` from `scipy` [25] with the method `LSODA` [1]. Of course, the derived bound is probably not exact, but it should be a good approximation. We did only search until we found a value for a_0 where a and m converge to P_2, the limit case between the convergence to P_0 and divergence to infinity. This is the case for a small range of values, but we saw no need to specify the upper bound on a_0 more precisely. This would give no beneficial information, as we still could not derive the formula for the exact bound. Furthermore, we just solved the ODE-system numerically in python and hence the solution contains small numerical errors and is dependent on the chosen solver method. Indeed, when using `BDF` as method for the solver, $a_0 = 0.6288$ gives already an unbounded solution, whereas the solution converges to P_2 when we use the method `LSODA`.
In fig. 3.2, one can see that for $a_0 = 0.62$ both a and m converge to 0 like intended. For $a_0 = 0.6288$, in the second plot, we reach the unstable stationary point P_2 and thus still stay bounded. Note, that having an unstable stationary state does not mean it cannot be reached with certain parameter values and initial values. The simplest example is just setting the stationary state as initial value. Then the solution will stay constant. Looking at definition 2.11 and remark 2.12, stability only means that it is always possible to start close enough to the stationary state, such that the solution will stay within a certain distance to the stationary state. Hence, when a stationary state is unstable we cannot always just start close enough to stay near it, but that does not mean we never can. This also explains why we get a range of values for a_0, where the solutions tend to P_2, and not only one value. For the last plot we chose $a_0 = 0.63$. Then a crosses m between time point 150 and 175 and both increase rapidly, resulting in an unbounded solution.
Another interesting finding, which could be helpful when solutions tend to be unbounded, is that the bound for a_0 gets smaller the higher we choose m_0. An example can be found in appendix 3.1, fig. 1.

For all cases, where our solutions stay bounded, we can now use theorem 2.7. Together with remark 2.8, it follows that the maximal interval of existence for solutions to our model is (t_0, ∞).

Results We have the following constraints for our model parameters:

- If $s(t) = 0$, we need $\gamma \geq 1$ and $M \leq \delta - \frac{\beta}{\gamma}$ (meaning also $\delta - \frac{\beta}{\gamma} > 0$ to stay reasonable) for P to be stable and biologically meaningful. The first inequality additionally ensures that $a(t)$ stays bounded by 1. The second one ensures that

[1] See figures.ipynb on `https://github.com/Melanie757/Thesis-Code`

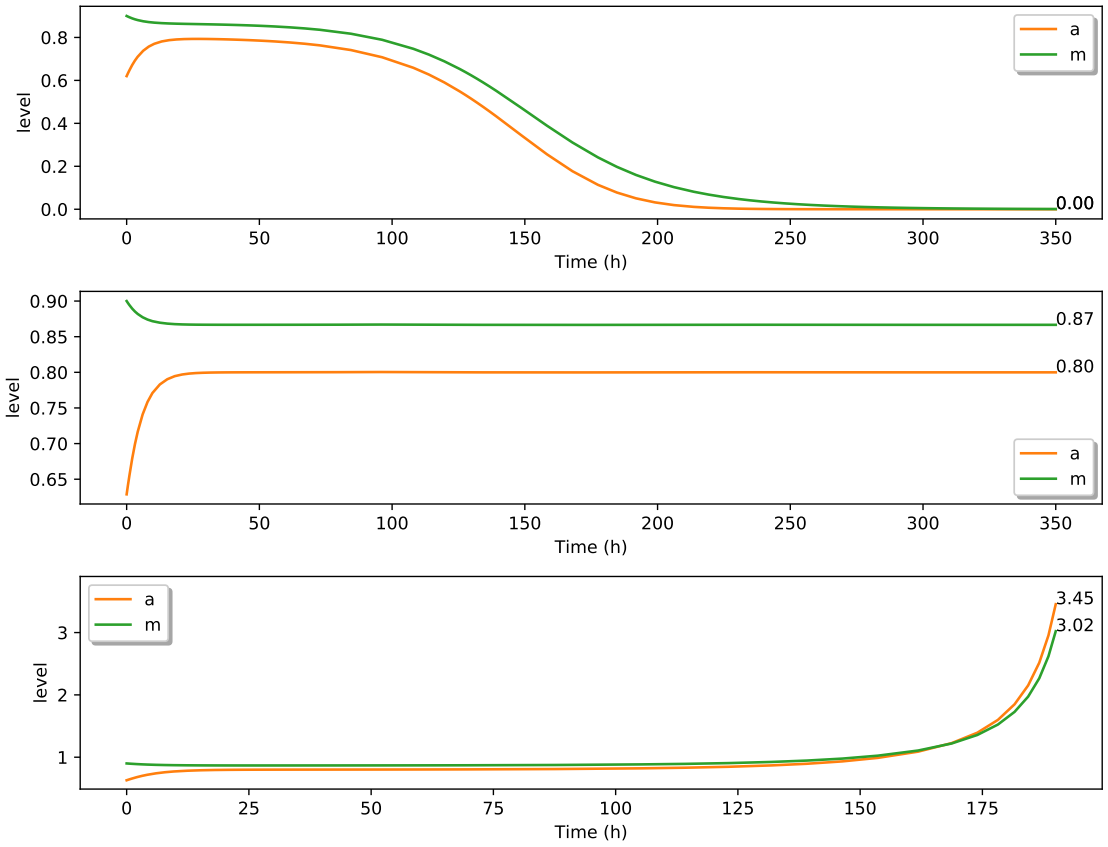

Figure 3.2: For all the plots we used the same parameter values as for fig. 4.5. The initial value of m is always set to 0.9, whereas the initial value of a changes from top to bottom: first $a_0 = 0.62$, then 0.6288 and last 0.63. We also added the values of a and m at the last time point.

$m(t)$ stays bounded by 1, as long as $a_0 \leq \frac{1}{\gamma}$. If $a_0 > \frac{1}{\gamma}$, we need $M \leq \delta - \beta a_0$ such that m stays bounded by 1.

- If $s(t) = 1$, we need $\alpha_1 < \alpha_2$ for P_0 to be stable and P_1 to be unstable. P_2 is unstable or biologically irrelevant anyway. For the boundedness of solutions, we have considered two different cases: $\frac{\delta}{\beta} \geq 1$ and $\frac{\delta}{\beta} < 1$. In the first case, all solutions stay bounded, especially by 1, as a_0 and m_0 should be smaller than 1. The second case is not that easy. Here we have bounded solutions as long as $m_0 \leq 1 - \frac{\alpha_1}{\alpha_2}(1 - \frac{\delta}{\beta})$ and $a_0 \leq \frac{\delta}{\beta}$. If $m_0 > 1 - \frac{\alpha_1}{\alpha_2}(1 - \frac{\delta}{\beta})$, we did not find appropriate constraints to ensure boundedness. But fig. 3.2 shows, that there are not only unbounded solutions in this case, but also solutions converging to P_0.

3.2 Autophagy, Catabolites and Cytoplasmic Material

Our second model takes furthermore the catabolites within the cell into account. Again we assume $s(t) \in \{0, 1\}$ for simplicity.

3.2.1 Model

We model the level of autophagy with the same approach as above. We do not add a dependence on the catabolites here, since we still assume the autophagy to just depend on the available material and not on the catabolites generated. In starvation this should be sufficient, for catabolites are used to rapidly adapt to the new environment. Hence, they are quickly turned into new material or additional energy, such that no situation with too many catabolites should appear. In nutrient-rich conditions we omit a dependence on the catabolites as there should always be enough available. What could appear, of course, are some saturation effects when the level of catabolites gets near its maximum. However, we will not investigate this further here, since we are more interested in the starvation case.
For the level of material we take additionally new built material out of catabolites into account, dependent on how much is already present. The increase is extenuated by autophagy and normal degradation happening in the cell, both dependent on the current level of material. Without starvation, a constant supply M is added.
The level of catabolites is assumed to increase dependent on the material degraded due to autophagy, and to decrease as new material is build out of catabolites and due to normal degradation processes. In nutrient-rich conditions we add a constant supply C. This assumptions are brought together in the following ODE model:

$$\begin{aligned}
\dot{a}(t) &= (1 - s(t))(1 - \gamma a(t)) + s(t)(\alpha_1 a(t)(1 - a(t)) - \alpha_2 a(t)(1 - m(t))) \\
\dot{m}(t) &= (1 - s(t))M + \underbrace{\beta_1 c(t)(1 - m(t))}_{\substack{\text{new material out} \\ \text{of catabolites}}} - \underbrace{\beta_2 a(t) m(t)}_{\substack{\text{material degraded} \\ \text{by autophagy}}} - \underbrace{\delta_1 m(t)}_{\substack{\text{other degradation} \\ \text{mechanisms}}} \\
\dot{c}(t) &= (1 - s(t))C + \underbrace{\gamma_1 a(t) m(t)}_{\substack{\text{catabolites generated} \\ \text{by autophagy}}} - \underbrace{\gamma_2 c(t)(1 - m(t))}_{\substack{\text{catabolites used for} \\ \text{new material}}} - \underbrace{\delta_2 c(t)}_{\substack{\text{other degradation} \\ \text{mechanisms}}}
\end{aligned} \tag{3.16}$$

with all parameters in $\mathbb{R}_{>0}$. A description of the parameters can be found in table 3.2.

Parameter	Description
α_1	rate constant of the log. growth of a
α_2	inhibition of a due to the level of material
β_1	increase of m due to available catabolites
β_2	degradation of m due to autophagy
γ	degradation of a
γ_1	increase of c due to autophagy and material
γ_2	decrease of c due to new built material
δ_1	normal degradation of m
δ_2	normal degradation of c
C	supply of new catabolites
M	supply of new material

Table 3.2: Description of the parameters in eq. (3.16).

3.2.2 Mathematical Analysis

We structure our analysis in the same way as in the last section. So we start with the existence of a solution.

Existence of a non-negative, unique solution In the case where $s(t) = 0$, we get

$$f_n \begin{pmatrix} a(t) \\ m(t) \\ c(t) \end{pmatrix} = \begin{pmatrix} 1 - \gamma a(t) \\ M + \beta_1 c(t)(1 - m(t)) - \beta_2 a(t) m(t) - \delta_1 m(t) \\ C + \gamma_1 a(t) m(t) - \gamma_2 c(t)(1 - m(t)) - \delta_2 c(t) \end{pmatrix} \tag{3.17}$$

for our function f_n.
This function is Lipschitz continuous in $[0, 1]^3$ with respect to a, m and c and continuous with respect to t, like in the last model. Using theorem 2.2 and theorem 2.6, there exists a unique maximal solution on an open interval for initial values in $(0, 1)^3$. Hence the first requirement of proposition 2.20 is fulfilled and we have to check the second one:

- Let $a(t) = 0$ and $m(t), c(t) \geq 0$, then $1 - \gamma a(t) = 1 \geq 0$.
- Let $m(t) = 0$ and $a(t), c(t) \geq 0$, then $M + \beta_1 c(t)(1 - m(t)) - \beta_2 a(t) m(t) - \delta_1 m(t) = M + \beta_1 c(t) \geq M > 0$.
- Let $c(t) = 0$ and $a(t), m(t) \geq 0$, then $C + \gamma_1 a(t) m(t) - \gamma_2 c(t)(1 - m(t)) - \delta_2 c(t) = C + \gamma_1 a(t) m(t) \geq C > 0$.

Therefore, a solution to our model stays non-negative for positive initial values.
If $s(t) = 1$, we have

$$f_s \begin{pmatrix} a(t) \\ m(t) \\ c(t) \end{pmatrix} = \begin{pmatrix} \alpha_1 a(t)(1 - a(t)) - \alpha_2 a(t)(1 - m(t)) \\ \beta_1 c(t)(1 - m(t)) - \beta_2 a(t) m(t) - \delta_1 m(t) \\ \gamma_1 a(t) m(t) - \gamma_2 c(t)(1 - m(t)) - \delta_2 c(t) \end{pmatrix} \tag{3.18}$$

By the same arguments as before, we get a unique maximal solution on an open interval for initial values in $(0, 1)^3$ and the first requirement of proposition 2.20 is fulfilled. It holds:

- Let $a(t) = 0$ and $m(t), c(t) \geq 0$, then $\alpha_1 a(t)(1 - a(t)) - \alpha_2 a(t)(1 - m(t)) = 0$.
- Let $m(t) = 0$ and $a(t), c(t) \geq 0$, then $\beta_1 c(t)(1 - m(t)) - \beta_2 a(t)m(t) - \delta_1 m(t) = \beta_1 c(t) \geq 0$.
- Let $c(t) = 0$ and $a(t), m(t) \geq 0$, then $\gamma_1 a(t)m(t) - \gamma_2 c(t)(1 - m(t)) - \delta_2 c(t) = \gamma_1 a(t)m(t) \geq 0$.

So we have a non-negative, unique solution for positive initial values.

Stationary states We start again with $s(t) = 0$. If we solve $f_n = 0$, we get the following:

$$\begin{aligned} f_{n,1} = \dot{a}(t) = 0 &\Leftrightarrow 0 = 1 - \gamma a^* \Leftrightarrow a^* = \frac{1}{\gamma} \\ f_{n,2} = \dot{m}(t) = 0 &\Leftrightarrow 0 = M + \beta_1 c^*(1 - m^*) - \beta_2 a^* m^* - \delta_1 m^* \\ &\Leftrightarrow m^* = \frac{M + \beta_1 c^*}{\beta_1 c^* + \beta_2 a^* + \delta_1} \\ f_{n,3} = \dot{c}(t) = 0 &\Leftrightarrow 0 = C + \gamma_1 a(t)m(t) - \gamma_2 c(t)(1 - m(t)) - \delta_2 c(t) \\ &\Leftrightarrow c^* = \frac{C + \gamma_1 a^* m^*}{\gamma_2(1 - m^*) + \delta_2} \end{aligned} \tag{3.19}$$

First, we want to notice that

$$m^* \leq 1 \Leftrightarrow M + \beta_1 c^* \leq \beta_1 c^* + \beta_2 a^* + \delta_1 \Leftrightarrow M \leq \beta_2 a^* + \delta_1 := A \tag{3.20}$$

Inserting the equation for m^* into the equation for c^* results in

$$0 = \underbrace{\beta_1 \delta_2}_{=:U} {c^*}^2 + \underbrace{(\gamma_2(A - M) + \delta_2 A - \beta_1(C + \gamma_1 a^*))}_{=:V} c^* \underbrace{-(CA + \gamma_1 a^* M)}_{=:W} \tag{3.21}$$

The solutions of this quadratic formula are

$$c_1^* = \frac{-V + \sqrt{V^2 - 4UW}}{2U} \quad \text{and} \quad c_2^* = \frac{-V - \sqrt{V^2 - 4UW}}{2U} \tag{3.22}$$

We look at $-4UW$:

$$-4UW = 4 \cdot \beta_1 \delta_2 \cdot (CA + \gamma_1 a^* M) > 0 \tag{3.23}$$

as all parameters are positive. Thus, $\sqrt{V^2 - 4UW} > |V|$ and c_2^* is negative.
c_1^* is positive and for our model we additionally need $c_1^* \leq 1$:

$$\begin{aligned} &\frac{-V + \sqrt{V^2 - 4UW}}{2U} \leq 1 \Leftrightarrow -V + \sqrt{V^2 - 4UW} \leq 2U \\ &\Leftrightarrow \sqrt{V^2 - 4UW} \leq 2U + V \Leftrightarrow V^2 - 4UW \leq (2U + V)^2 \\ &\Leftrightarrow V^2 - 4UW \leq 4U^2 + 4UV + V^2 \Leftrightarrow -UW \leq U^2 + UV \\ &\Leftrightarrow -W \leq U + V \\ &\Leftrightarrow \underbrace{CA + \gamma_1 a^* M}_{>0} \leq \underbrace{\beta_1 \delta_2 + \gamma_2(A - M) + \delta_2 A}_{>0, \text{ assuming } M \leq A \text{ (see eq. (3.20))}} - \beta_1(C + \gamma_1 a^*) \end{aligned} \tag{3.24}$$

The second last inequality holds, since $U > 0$. For the last inequality to be meaningful, $\delta_2(\beta_1 + A) + \gamma_2(A - M) > \beta_1(C + \gamma_1 a^*)$ has to hold.
So we arrive at the stationary state $(a^*, m^*, c^*) = \left(\frac{1}{\gamma}, m^*, c_1^*\right) \in (0, 1]^3$, assuming the constraints in eq. (3.20) and eq. (3.24) hold and are meaningful.
Since we assume for $s(t) = 0$ that all needed nutrients are available, we will especially consider the case where $m^* = 1$. We see from eq. (3.20), that we just need $M = A$ for this.

For $s(t) = 1$:

$$\begin{aligned} f_{n,1} = \dot{a}(t) = 0 &\Leftrightarrow 0 = \alpha_1 a(t)(1 - a(t)) - \alpha_2 a(t)(1 - m(t)) \\ &\Leftrightarrow a^* = 0 \ \text{ or } \ a^* = 1 - \frac{\alpha_2}{\alpha_1}(1 - m^*) \\ f_{n,2} = \dot{m}(t) = 0 &\Leftrightarrow 0 = \beta_1 c^*(1 - m^*) - \beta_2 a^* m^* - \delta_1 m^* \\ &\Leftrightarrow m^* = \frac{\beta_1 c^*}{\beta_1 c^* + \beta_2 a^* + \delta_1} \\ f_{n,3} = \dot{c}(t) = 0 &\Leftrightarrow 0 = \gamma_1 a(t) m(t) - \gamma_2 c(t)(1 - m(t)) - \delta_2 c(t) \\ &\Leftrightarrow c^* = \frac{\gamma_1 a^* m^*}{\gamma_2(1 - m^*) + \delta_2} \end{aligned} \tag{3.25}$$

Inserting the equation for m^* in the equation for c^*, we arrive at

$$c^* = 0 \ \text{ or } \ c^* = \frac{\beta_1 \gamma_1 a^* - A(\gamma_2 + \delta_2)}{\beta_1 \delta_2} \tag{3.26}$$

where $A := \beta_2 a^* + \delta_1$ as in eq. (3.20). For detailed computations see appendix 1.2
As a first step, we examine under which conditions each coordinate of these stationary states lies in $[0, 1]$.

- $a^* = 0$ obviously lies in $[0, 1]$. For the second case of a^*, we get the following:

$$0 < a^* \leq 1 \Leftrightarrow 0 \leq \frac{\alpha_2}{\alpha_1}(1 - m^*) < 1 \Leftrightarrow 1 - \frac{\alpha_2}{\alpha_1} < m^* \leq 1 \tag{3.27}$$

- m^* is always non-negative, as long as a^* and c^* are non-negative. For $m^* \leq 1$ we need $A \geq 0$, but this holds true for every $a^* \geq 0$.

- $c^* = 0 \in [0, 1]$, so we consider the second possibility for c^*. Here it holds

$$\begin{aligned} 0 < c^* \leq 1 &\Leftrightarrow 0 < \beta_1 \gamma_1 a^* - A(\gamma_2 + \delta_2) \leq \beta_1 \delta_2 \\ &\Leftrightarrow 0 < \beta_1 \gamma_1 a^* - (\beta_2 a^* + \delta_1)(\gamma_2 + \delta_2) \leq \beta_1 \delta_2 \\ &\Leftrightarrow 0 < a^*(\beta_1 \gamma_1 - \beta_2(\gamma_2 + \delta_2)) \leq \beta_1 \delta_2 + \delta_1(\gamma_2 + \delta_2) \\ &\Leftrightarrow \begin{cases} 0 < a^* \leq \frac{\beta_1 \delta_2 + \delta_1(\gamma_2 + \delta_2)}{\beta_1 \gamma_1 - \beta_2(\gamma_2 + \delta_2)} & \text{if } \beta_1 \gamma_1 - \beta_2(\gamma_2 + \delta_2) > 0 \\ \frac{\beta_1 \delta_2 + \delta_1(\gamma_2 + \delta_2)}{\beta_1 \gamma_1 - \beta_2(\gamma_2 + \delta_2)} \leq a^* < 0 & \text{if } \beta_1 \gamma_1 - \beta_2(\gamma_2 + \delta_2) < 0 \end{cases} \end{aligned} \tag{3.28}$$

  In the second case, a^* would be negative, hence we will assume $\beta_1 \gamma_1 - \beta_2(\gamma_2 + \delta_2) > 0$ in the following. This also implies $\beta_1 \gamma_1 - \beta_2 \gamma_2 > 0$.

Now we determine the different possibilities for the stationary states:

- If $c^* = 0$ also $m^* = 0$, using the equation for m^* in eq. (3.25). For a^* we have two possibilities, $a^* = 0$ and $a^* = 1 - \frac{\alpha_2}{\alpha_1}(1-0) = 1 - \frac{\alpha_2}{\alpha_1}$. So we arrive at the trivial stationary point $P_0 = (0,0,0)$ and $P_1 = (1 - \frac{\alpha_2}{\alpha_1}, 0, 0)$. P_1 is only in $[0,1]^3$, if $\alpha_2 \leq \alpha_1$.

- If $0 < c^* \leq 1$, we have again two possibilities for a^*. For $a^* = 0$, we get $c^* = \frac{-\delta_1(\gamma_2+\delta_2)}{\beta_1\delta_2} < 0$, hence this is not biologically meaningful. Therefore, we consider the case where $a^* = 1 - \frac{\alpha_2}{\alpha_1}(1-m^*) = \frac{\alpha_1-\alpha_2}{\alpha_1} + \frac{\alpha_2}{\alpha_1}m^*$. Fist we set the equation for c^* in the equation for m^*:

$$\begin{aligned} m^* &= \frac{\beta_1 c^*}{\beta_1 c^* + \beta_2 a^* + \delta_1} = \frac{\beta_1 c^*}{\beta_1 c^* + A} = \frac{\beta_1\gamma_1 a^* - A(\gamma_2+\delta_2)}{\beta_1\gamma_1 a^* - A(\gamma_2+\delta_2) + A\delta_2} \\ &= \frac{\beta_1\gamma_1 a^* - A(\gamma_2+\delta_2)}{\beta_1\gamma_1 a^* - A\gamma_2} \end{aligned} \tag{3.29}$$

Now we can insert a^* and solve for m^*, resulting in the following quadratic equation

$$\underbrace{\frac{\alpha_2}{\alpha_1}\zeta_1}_{:=U} m^{*2} + \underbrace{\left(\frac{\alpha_1-\alpha_2}{\alpha_1}\zeta_1 - \delta_1\gamma_2 - \frac{\alpha_2}{\alpha_1}\zeta_2\right)}_{:=V} m^* \underbrace{-\left(\frac{\alpha_1-\alpha_2}{\alpha_1}\zeta_2 - \delta_1(\gamma_2+\delta_2)\right)}_{:=W} = 0 \tag{3.30}$$

where $\zeta_1 := \beta_1\gamma_1 - \beta_2\gamma_2$ and $\zeta_2 := \beta_1\gamma_1 - \beta_2(\gamma_2+\delta_2)$ for readability reasons. From eq. (3.28) it already followed that $\zeta_1, \zeta_2 > 0$.
The solutions for m^* are

$$m_1^* = \frac{-V + \sqrt{V^2 - 4UW}}{2U} \quad \text{and} \quad m_2^* = \frac{-V - \sqrt{V^2 - 4UW}}{2U} \tag{3.31}$$

which give us two additional stationary states P_2 and P_3.
Again, we examine whether these solutions are meaningful for our model. To make the computations easier, we will assume $\alpha_1 < \alpha_2$. This is a finding from the next paragraph about stability, ensuring that P_0 is stable in the starvation setting.
With this assumption, we see that $V < 0$ and $W > 0$. Hence, both solutions are positive and real valued if $V^2 - 4UW \geq 0$. But as mentioned before, the only biologically meaningful stationary state in this setting is P_0. So we would like to rule out m_1^* and m_2^*. We tried to find constraints directly on the parameters, such that both are not relevant for our model. Unfortunately, this did not work well and gave us no meaningful conditions. Thus, we will just assume $V^2 - 4UW < 0$ here. Then m_1^* and m_2^* are complex and will not be reached if $m(t)$ stays in $[0,1]$.

Stability With $s(t) = 0$, we get the Jacobian matrix

$$J(a,m,c) = \begin{pmatrix} -\gamma & 0 & 0 \\ -\beta_2 m & -\beta_1 c - \beta_2 a - \delta_1 & \beta_1(1-m) \\ \gamma_1 m & \gamma_1 a + \gamma_2 c & -\gamma_2(1-m) - \delta_2 \end{pmatrix} \tag{3.32}$$

When we insert the stationary state $\left(\frac{1}{\gamma}, m^*, c_1^*\right)$, the eigenvalues are not that easy to compute. But as we are most interested in the case where $m^* = 1$, we only consider

$\left(\frac{1}{\gamma}, 1, c_1^*\right)$ here. Then we get again a lower triangular matrix with its eigenvalues on the diagonal:

$$J\left(\frac{1}{\gamma}, 1, c_1^*\right) = \begin{pmatrix} -\gamma & 0 & 0 \\ -\beta_2 & -\beta_1 c_1^* - \frac{\beta_2}{\gamma} - \delta_1 & 0 \\ \gamma_1 & \frac{\gamma_1}{\gamma} + \gamma_2 c_1^* & -\delta_2 \end{pmatrix} \tag{3.33}$$

Thus, the eigenvalues are $-\gamma, -\beta_1 c_1^* - \frac{\beta_2}{\gamma} - \delta_1$ and $-\delta_2$, which are all negative when c_1^* is non-negative. Using proposition 2.15 and theorem 2.18, we can conclude that $\left(\frac{1}{\gamma}, 1, c_1^*\right)$ is stable.

For $s(t) = 1$, the Jacobian looks like

$$J(a, m, c) = \begin{pmatrix} \alpha_1(1-2a) - \alpha_2(1-m) & \alpha_2 a & 0 \\ -\beta_2 m & -\beta_1 c - \beta_2 a - \delta_1 & \beta_1(1-m) \\ \gamma_1 m & \gamma_1 a + \gamma_2 c & -\gamma_2(1-m) - \delta_2 \end{pmatrix} \tag{3.34}$$

Inserting the trivial stationary state P_0, we get

$$J(0, 0, 0) = \begin{pmatrix} \alpha_1 - \alpha_2 & 0 & 0 \\ 0 & -\delta_1 & \beta_1 \\ 0 & 0 & -\gamma_2 - \delta_2 \end{pmatrix} \tag{3.35}$$

with eigenvalues $\alpha_1 - \alpha_2$, $-\delta_1$ and $-\gamma_2 - \delta_2$. The last two are clearly negative. $\alpha_1 - \alpha_2$ is negative for $\alpha_1 < \alpha_2$. As the cell should die in a starvation medium in the long run, we want P_0 to be stable and hence will assume $\alpha_1 < \alpha_2$ for our model.
This assumption makes the first component of P_1 negative and hence, P_1 cannot be reached by a solution starting with positive initial values. We also assumed P_2 and P_3 to be biologically irrelevant. For this reason we will omit the stability analysis of this three points here, as this would give us no beneficial information.

Boundedness and domain It remains to show, that our solutions are bounded from above.

We start with the case $s(t) = 0$:

- $a(t)$ always converges to $\frac{1}{\gamma}$, like in the last model. Thus, $a(t)$ is bounded by $\max\left(\frac{1}{\gamma}, a_0\right) \leq 1$.
- Assume $a_0 \leq \frac{1}{\gamma}$. Then we know $a_0 \leq a(t) \leq \frac{1}{\gamma}$. If we additionally assume $M \leq \beta_2 a_0 + \delta_1$, we get for $m > 1$:

  $$\dot{m} = M + \beta_1 c(1-m) - \beta_2 a m - \delta_1 m < M - \beta_2 a_0 - \delta_1 \leq 0. \tag{3.36}$$

  Hence, m is bounded from above by 1.
  If $a_0 > \frac{1}{\gamma}$, we know $a(t) \geq \frac{1}{\gamma}$. Assuming $m > 1$, we arrive at

  $$\dot{m} < M - \beta_2 \frac{1}{\gamma} - \delta_1 = M - A \leq 0, \tag{3.37}$$

  where the last inequality holds because of eq. (3.20). Thus, m is again bounded by 1.

- We first assume, that the requirements to ensure $0 \leq m \leq 1$ are fulfilled. Assuming furthermore $C + \gamma_1 \max\left(\frac{1}{\gamma}, a_0\right) \leq \delta_2$, we get for $c > 1$:

$$\dot{c} = C + \gamma_1 am - \gamma_2 c(1-m) - \delta_2 c < C + \gamma_1 \max\left(\frac{1}{\gamma}, a_0\right) - \delta_2 \leq 0, \tag{3.38}$$

  Therefore, also c is bounded by 1 under the above assumptions.

Next we consider $s(t) = 1$.

- Let $m > \frac{\beta_1 c}{\beta_1 c + \beta_2 a + \delta_1}$. Then

$$\begin{aligned}\dot{m} &= \beta_1 c(1-m) - \beta_2 am - \delta_1 m \\ &< \beta_1 c \frac{\beta_2 a + \delta_1}{\beta_1 c + \beta_2 a + \delta_1} - \beta_2 a \frac{\beta_1 c}{\beta_1 c + \beta_2 a + \delta_1} - \delta_1 \frac{\beta_1 c}{\beta_1 c + \beta_2 a + \delta_1} \\ &= \frac{\beta_1 c \beta_2 a + \beta_1 \delta_1 c - \beta_1 c \beta_2 a - \beta_1 \delta_1 c}{\beta_1 c + \beta_2 a + \delta_1} = 0\end{aligned} \tag{3.39}$$

  So we can deduce, that m stays bounded by $\frac{\beta_1 c}{\beta_1 c + \beta_2 a + \delta_1}$, which is in particular smaller than 1, as a and c are non-negative.

- Let $a > 1 - \frac{\alpha_2}{\alpha_1}(1-m)$. Then

$$\begin{aligned}\dot{a} &= \alpha_1 a(1-a) - \alpha_2 a(1-m) \\ &< a\left(\alpha_1\left(1 - 1 + \frac{\alpha_2}{\alpha_1}(1-m)\right) - \alpha_2(1-m)\right) \\ &= a(\alpha_2(1-m) - \alpha_2(1-m)) = 0\end{aligned} \tag{3.40}$$

  Hence, a is bounded by $1 - \frac{\alpha_2}{\alpha_1}(1-m)$. This bound is smaller than 1 if and only if $m < 1$, which we showed above.

- Let $c > \frac{\gamma_1 am}{\gamma_2(1-m) + \delta_2}$. Then

$$\begin{aligned}\dot{c} &= \gamma_1 am - \gamma_2 c \underbrace{(1-m)}_{>0 \text{ as } m<1} - \delta_2 c \\ &< \gamma_1 am - \gamma_2 \frac{\gamma_1 am}{\gamma_2(1-m) + \delta_2}(1-m) - \delta_2 \frac{\gamma_1 am}{\gamma_1(1-m) + \delta_2} \\ &= \gamma_1 am - \gamma_1 am \frac{\gamma_2(1-m) + \delta_2}{\gamma_2(1-m) + \delta_2} = 0\end{aligned} \tag{3.41}$$

  Thus, c is bounded by $\frac{\gamma_1 am}{\gamma_2(1-m) + \delta_2} < \infty$, as $a, m \in [0,1]$.
  Just using $a, m < 1$ to estimate this bound further, gives us that c stays below 1 if $\gamma_1 < \delta_2$. But this is quite a rough estimate. In fig. 3.3 [2], for example, we used parameter values with $\delta_2 < \gamma_1$. Especially the solution curve of $c(t)$ looks interesting, as it first decreases but then increases again until over 0.95.

  Using the exact bounds instead results in inequalities, which give us no easily

[2] For the corresponding python code see figures.ipynb on `https://github.com/Melanie757/Thesis-Code`.

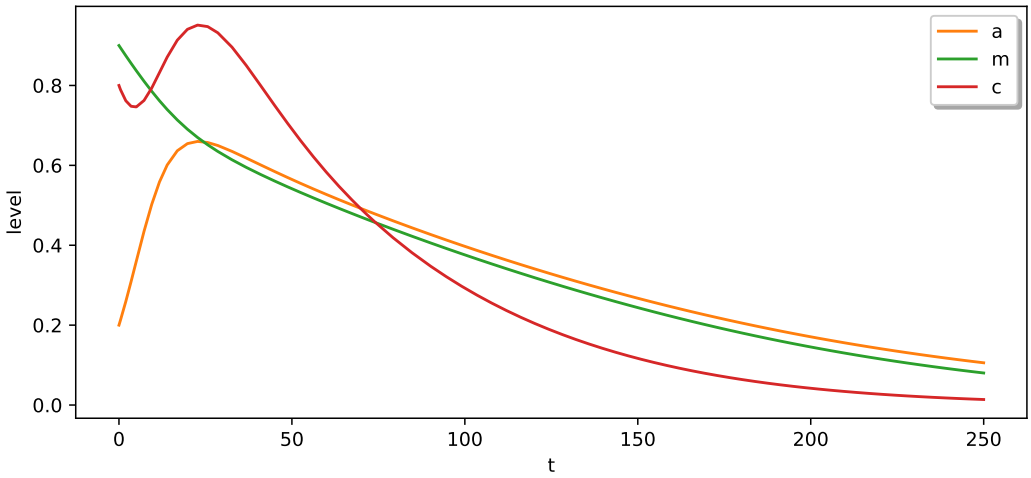

Figure 3.3: Eq. (3.16) with $s(t) = 1$, solved numerically with scipy [25] in python. The used initial and parameter values are given in table 3.3.

	Name	Value
Initial values	a_0	0.2
	m_0	0.9
	c_0	0.8
Parameters	α_1	0.2
	α_2	0.205
	β_1	0.02
	β_2	0.0065
	γ_1	0.235
	γ_2	0.1
	δ_1	0.015
	δ_2	0.075

Table 3.3: Initial values for the variables and parameter values used for fig. 3.3.

readable, hence informative, constraints on the parameters. Thus, we omit further computations here, since the boundedness by any finite value is for now sufficient for the analysis of our model. When fitting the model to real data, one can still adjust the parameters to keep c below 1.

Now that we know our solutions to stay bounded (with certain requirements to the parameters), we can use theorem 2.7. Together with remark 2.8, it follows that the maximal interval of existence for solutions to our model is (t_0, ∞).

Results

- If $s(t) = 0$, then $\gamma \geq 1$ should hold, like in the first model such that $a^* \in [0, 1]$. For $m^* \in [0, 1]$, we need $M \leq \frac{\beta_2}{\gamma} + \delta_1$ (see eq. (3.20)). For $c^* \in [0, 1]$, it is necessary that $\delta_2(\beta_1 + A) + \gamma_2(A - M) > \beta_1(C + \gamma_1 a^*)$ must hold (see eq. (3.24)).
  Assuming $m^* = 1$, the condition for $c^* \in [0, 1]$ also ensures stability of the biologically meaningful stationary state.
  If $a_0 < \frac{1}{\gamma}$, $M \leq \beta_2 a_0 + \delta_1$ is additionally needed for $m(t)$ to stay bounded by 1. For $c(t)$ to stay below 1, we need $C + \gamma_1 \max(\frac{1}{\gamma}, a_0) \leq \delta_2$.

- If $s(t) = 1$, P_0 should be stable and therefore we need $\alpha_1 < \alpha_2$. This also ensures, that P_1 is not meaningful for our model. To rule out P_3 and P_4, we want $V^2 - 4UW < 0$ with U, V and W defined like in eq. (3.30).
  a and m are already bounded by 1 without further conditions. For c we did not find a good constraint, such that it stays below 1, but at least it is bounded by a finite value we could specify.

4 From Single-Cell to Multi-Cell Model

So far we only considered models for single cells. But as cells normally form whole cell batches, the next step is to extend and adapt our first model to this new setting.

During our research for this work, we also found an interesting article by Ryo Iwama and Yoshinori Ohsumi [15]. They investigated autophagy in yeast cultures under gradual changes of nutrient supply. In the last chapter we just modeled two situations: enough nutrients available or no nutrients available at all. This is a reasonable simplification, as in laboratory research autophagy can rapidly be induced [3]. For studying the process of autophagy, inducing it at a certain time point is of course convenient. In nature, however, it is more likely that the nutrient availability changes gradually and not instantly. Hence, in the second part of this chapter, we will try to fit our model to the findings of Ryo Iwama and Yoshinori Ohsumi.

We will start by adapting our first model from section 3.1. As the growth of a cell batch can be expected to be different while the cells are starving or not, we again split the model into two parts. Furthermore, we assume every single cell to behave exactly the same, regarding the level of material and autophagy, and neglect the spacial structure of the cell batch. Again, our model is dimensionless like the single-cell model, since we only study levels.

4.1 First Adaption of the single-cell Model

When we first tried to adapt the single-cell model from section 3.1 to the multi-cell case, we had not factored out β, like we now have in all the models. So we took the equations from section 3.1 as follows

$$\begin{aligned} \dot{a}(t) &= (1 - s(t)) \cdot (1 - \gamma a(t)) + s(t) \cdot (\alpha_1 a(t)(1 - a(t)) - \alpha_2 a(t)(1 - m(t))) \\ \dot{m}(t) &= (1 - s(t)) \cdot M + m(t)(\beta a(t) - \delta) \end{aligned} \tag{4.1}$$

and adapted them in the following way by choosing a (sort of) logistic growth model for the cell number $Y(t)$:

$$\begin{aligned} \dot{Y}(t) &= (1 - s(t)) \cdot rY(t)\left(1 - \frac{Y(t)}{K}\right) + s(t) \cdot rY(t)\left(1 - \frac{Y(t)}{K}\right)(m(t) - T_m Y(t)) \\ \dot{a}(t) &= (1 - s(t)) \cdot (\alpha Y(t) - \gamma a(t)) + s(t) \cdot (\alpha_1 a(t)(Y(t) - a(t)) - \alpha_2 a(t)(Y(t) - m(t))) \\ \dot{m}(t) &= (1 - s(t)) \cdot MY(t) + m(t)(\beta a(t) - \delta) \end{aligned} \tag{4.2}$$

with all parameters in $\mathbb{R}_{>0}$. For more details see table 4.2 and table 4.3.

By replacing each 1 in the equation for $\dot{a}$ by $Y(t)$, we tried to make sure that $a(t)$ stays bounded by $Y(t)$ instead of 1. In this way $a(t)$ would already be scaled according to $Y(t)$. Thus, we did not add another scaling in the equation for $\dot{m}$, except for the constant supply M. Now a and m denote the summed levels of autophagy and material

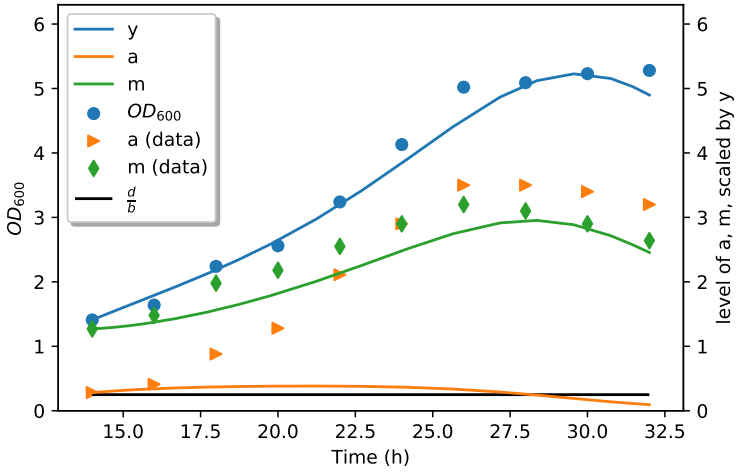

Figure 4.1: Numerical solution of eq. (4.2), computed with `scipy` [25]. The initial values and parameter values are given in table 4.1.

	Name	Value
Initial values	$a(14h)$	0.282
	$m(14h)$	1.27
	$y(14h)$	1.41
Parameters	α_1	0.186
	α_2	0.5
	β	0.6
	δ	0.15
	K	20
	r	0.3
	T_m	0.55

Table 4.1: Initial values for the variables and parameter values used for fig. 4.1.

in all cells together.
But this did not work as we wanted. What we did not realize until we factored out β was, that the sign of $m(t)(\beta a(t)-\delta)$, and hence the sign of $\dot{m}$ in the starvation case, depended only half on $Y(t)$. When writing $m(t)(\beta a(t)-\delta) = \beta m(t)\left(a(t)-\frac{\delta}{\beta}\right)$, the shortcoming becomes clear at once. Considering the difference between $a(t)$, whose value is dependent on $Y(t)$, and $\frac{\delta}{\beta}$, which does not change with $Y(t)$, is not very reasonable.

One of the best fits we could get with this model in the starvation case is shown in fig. 4.1. For a detailed description on how the fitting is done and what data is used, see section 4.4.3 and section 4.4.4. We also plotted $\frac{\delta}{\beta}$, to illustrate how m increases while $a > \frac{\delta}{\beta}$ and decreases while $a < \frac{\delta}{\beta}$. One can see, that $\frac{\delta}{\beta}$ cannot be chosen much bigger, as then m would first be decreasing while a increases until $\frac{\delta}{\beta}$. Choosing higher values for a at the beginning will also not help, as a still has to stay below Y. Hence, $a(14) \leq Y(14) = 1.41$ is needed and thus also $\frac{\delta}{\beta} \leq 1.41$. Having this in mind, $a(26)$ has

Parameter	Description
α	scaling of Y, if necessary (else set $\alpha = 1$)
β	increase of m due to autophagy
γ	degradation of a
δ	normal degradation of m
K	maximal capacity for the cell batch
M	supply of new material per cell
r	rate constant of the log. growth of Y

Table 4.2: Description of the parameters in eq. (4.3).

to be around $\frac{\delta}{\beta} \leq 1.41$ and $a(28)$ already smaller than that, to get the wanted solution curve for m. This gives us very small values for the level of autophagy after 26 h. Using the values for $Y(t)$ from table 1, we get an autophagy level of at most $\frac{1.41}{Y(26)} \approx 0.28$ per cell after 26 h. This is far away from the level we wanted to achieve.

In the next section we correct the incomplete scaling of $\dot{m}$ and then analyse the resulting model.

4.2 Adjusted Model

For readability reasons we omit $s(t)$ here and divide the model instead directly into two systems of equations. This is equivalent to choosing $s(t) \in \{0, 1\}$ as we did before.

If the cells are cultured in a nutrient-rich environment, we choose a logistic growth model. Let $Y(t)$ denote the number of yeast cells (or a similar measure). Then we get the following ODE model:

$$\begin{aligned} \dot{Y}(t) &= rY(t)\left(1 - \frac{Y(t)}{K}\right) \\ \dot{a}(t) &= \alpha Y(t) - \gamma a(t) \\ \dot{m}(t) &= M \cdot Y(t) + \beta m(t)\left(a(t) - \frac{\delta}{\beta} Y(t)\right) \end{aligned} \tag{4.3}$$

where all parameters are in $\mathbb{R}_{>0}$. A description of their meaning can be found in table 4.2. $a(t)$ and $m(t)$ are again given as levels. But as every single cell can have a level of 1, we get at most a level of $Y(t)$ for the whole batch.

Later we always set $\alpha = 1$ for the fitting of our model, as Y and a have already matching units. This has the advantage, that we have one parameter less to fit.

If the cell batch is in a starvation medium, autophagy helps the cells to turn over cytoplasmic components for new needed components and also to provide energy. In this way the cells are able to keep growing for some time, although no new nutrients are

Parameter	Description
α_1	rate constant of the log. growth of a
α_2	inhibition of a due to the level of material
T_m	ratio of material which is at least needed for Y to grow

Table 4.3: Description of the additional parameters for eq. (4.4).

available. Hence, we assume the cells can grow until the amount of material falls below a certain ratio of Y, here denoted by T_m:

$$\begin{aligned}\dot{Y}(t) &= rY(t)\left(1-\frac{Y(t)}{K}\right)(m(t)-T_mY(t))\\ \dot{a}(t) &= \alpha_1 a(t)(Y(t)-a(t)) - \alpha_2 a(t)(Y(t)-m(t))\\ \dot{m}(t) &= \beta m(t)\left(a(t)-\frac{\delta}{\beta}Y(t)\right)\end{aligned} \tag{4.4}$$

where all parameters are in $\mathbb{R}_{>0}$ and $T_m < 1$, additionally. A description of the parameters which are not already included in table 4.2 can be found in table 4.3.

4.3 Mathematical Analysis

The analysis is again structured as in the last chapter. We begin by showing, that our model has a non-negative, unique solution. Next we consider the stationary states and their stability. Last we look at the boundedness and domain of solutions.

Existence of a non-negative, unique solution Starting with the non-starvation case, we obtain

$$f_n\begin{pmatrix}Y(t)\\a(t)\\m(t)\end{pmatrix} = \begin{pmatrix} rY(t)\left(1-\frac{Y(t)}{K}\right)\\ \alpha Y(t)-\gamma a(t)\\ M\cdot Y(t)+\beta m(t)\left(a(t)-\frac{\delta}{\beta}Y(t)\right)\end{pmatrix} \tag{4.5}$$

f_n is continuously differentiable with respect to Y, a and m, hence its derivative stays bounded in $[0,K]^3$. From this it follows, that f_n is Lipschitz continuous in $[0,K]^3$ with respect to Y, a and m. Furthermore, it is continuous with respect to t, as it is independent of t. By using theorem 2.2 and theorem 2.6, we can conclude that our model has a unique maximal solution on an open interval for positive initial values. This is the first requirement for proposition 2.20. Now we check the second one:

- Let $Y(t)=0$ and $a(t),m(t)\geq 0$, then $rY(t)\left(1-\frac{Y(t)}{K}\right)=0\geq 0$.
- Let $a(t)=0$ and $Y(t),m(t)\geq 0$, then $\alpha Y(t)-\gamma a(t)=\alpha Y(t)\geq 0$.
- Let $m(t)=0$ and $Y(t),a(t)\geq 0$, then $M\cdot Y(t)+\beta m(t)\left(a(t)-\frac{\delta}{\beta}Y(t)\right)=M\cdot Y(t)\geq 0$.

Thus, solutions to our model stay non-negative for positive initial values.

Next, we consider the case where the cells are starving. Here we have

$$f_s \begin{pmatrix} Y(t) \\ a(t) \\ m(t) \end{pmatrix} = \begin{pmatrix} rY(t)\left(1 - \frac{Y(t)}{K}\right)(m(t) - T_m Y(t)) \\ \alpha_1 a(t)(Y(t) - a(t)) - \alpha_2 a(t)(Y(t) - m(t)) \\ \beta m(t)\left(a(t) - \frac{\delta}{\beta} Y(t)\right) \end{pmatrix} \tag{4.6}$$

Like above, this function is Lipschitz continuous with respect to Y, a and m and continuous with respect to t. Hence, with theorem 2.2 and theorem 2.6, it follows that our model has a unique maximal solution on an open interval for positive initial values. Looking at the second requirement of proposition 2.20:

- Let $Y(t) = 0$ and $a(t), m(t) \geq 0$, then $rY(t)\left(1 - \frac{Y(t)}{K}\right)(m(t) - T_m Y(t)) = 0$.
- Let $a(t) = 0$ and $Y(t), m(t) \geq 0$, then $\alpha_1 a(t)(Y(t) - a(t)) - \alpha_2 a(t)(Y(t) - m(t)) = 0$.
- Let $m(t) = 0$ and $Y(t), a(t) \geq 0$, then $\beta m(t)\left(a(t) - \frac{\delta}{\beta} Y(t)\right) = 0$.

So our model stays non-negative for positive initial values.

Stationary states We first consider system (4.3). Setting all equations equal to 0, we arrive at the following:

- $\dot{Y}(t) = 0 \Leftrightarrow Y^* = 0$ or $Y^* = K$.
- $\dot{a}(t) = 0 \Leftrightarrow a^* = \frac{\alpha}{\gamma} Y^*$.
- $\dot{m}(t) = 0 \Leftrightarrow M \cdot Y^* + \beta m^* \left(a^* - \frac{\delta}{\beta} Y^*\right) = 0 \Leftrightarrow m^* = \frac{M}{\delta - \frac{\alpha\beta}{\gamma}}$ or $m^* = m_n \in \mathbb{R}$, if $Y^* = 0$.

Hence, we have the following stationary states: $(0, 0, m_n)$, with $m_n \in \mathbb{R}_{\geq 0}$, as we only consider non-negative solutions, and $\left(K, \frac{\alpha}{\gamma} K, \frac{M}{\delta - \frac{\alpha\beta}{\gamma}}\right)$. The last one is the stationary state we want to be stable, as then the cell batch stabilizes at its capacity. But to be meaningful, we additionally need $\frac{\alpha}{\gamma} K \leq K$ and $0 \leq \frac{M}{\delta - \frac{\alpha\beta}{\gamma}} \leq K$. These conditions are equivalent to $\frac{\alpha}{\gamma} \leq 1$ and $M \leq K(\delta - \frac{\alpha\beta}{\gamma})$, especially $\delta > \frac{\alpha\beta}{\gamma}$ similar to the single cell case.

For system (4.4), we obtain:

- $\dot{Y}(t) = 0 \Leftrightarrow Y^* = 0$, $Y^* = K$ or $Y^* = \frac{m^*}{T_m}$.
- $\dot{a}(t) = 0 \Leftrightarrow a^* = 0$ or $a^* = Y^* - \frac{\alpha_2}{\alpha_1}(Y^* - m^*)$.
- $\dot{m}(t) = 0 \Leftrightarrow m^* = 0$, $a^* = \frac{\delta}{\beta} Y^*$ or $m^* = m_s \in \mathbb{R}$, if $Y^* = a^* = 0$.

The only biologically meaningful stationary state is $P_0 = (0, 0, 0)$, like in the single cell models from section 3.1 and section 3.2. To analyse also the stability of the other possible stationary states, we compute them anyway.
Considering all different and non-negative combinations of the stationary states found above, we get

- $P_0 = (0,0,0)$,
- $P_1 = (K,0,0) = K(1,0,0)$,
- $P_2 = (0,0,m_s)$, with $m_s \in \mathbb{R}_{>0}$,
- $P_3 = \big(K, K\big(1-\frac{\alpha_2}{\alpha_1}\big), 0\big) = K\big(1, 1-\frac{\alpha_2}{\alpha_1}, 0\big)$,
- $P_4 = \big(K, K\frac{\delta}{\beta}, K\big(1-\frac{\alpha_1}{\alpha_2}\big(1-\frac{\delta}{\beta}\big)\big)\big) = K\big(1, \frac{\delta}{\beta}, 1-\frac{\alpha_1}{\alpha_2}\big(1-\frac{\delta}{\beta}\big)\big)$.

For detailed computations see appendix 1.3.
By factoring out K in P_1, P_3 and P_4, we can see the similarity to the stationary states of the single cell version, obtained from eq. (3.5).

Stability In the non-starvation case, the Jacobian matrix is stated as follows

$$J(Y,a,m) = \begin{pmatrix} r\big(1-\frac{2Y}{K}\big) & 0 & 0 \\ \alpha & -\gamma & 0 \\ M-\delta & \beta m & \beta\big(a-\frac{\delta}{\beta}Y\big) \end{pmatrix} \tag{4.7}$$

Inserting the first stationary state:

$$J(0,0,m_n) = \begin{pmatrix} r & 0 & 0 \\ \alpha & -\gamma & 0 \\ M-\delta & \beta m_n & 0 \end{pmatrix} \tag{4.8}$$

where $m_n \in \mathbb{R}_{\geq 0}$. This matrix has the eigenvalues $r, -\gamma$ and 0. As one of the eigenvalues is 0, we cannot use theorem 2.18 this time. But one can see in eq. (4.3), that the equation for $\dot{Y}(t)$ is independent of the other two. Hence, if 0 is unstable for $\dot{Y}(t)$, the whole stationary point has to be unstable. However, the growth of Y is just described by a logistic growth, where we know that 0 is unstable.
Now we insert the second stationary state:

$$J\left(K, \frac{\alpha}{\gamma}K, \frac{M}{\delta-\frac{\alpha\beta}{\gamma}}\right) = \begin{pmatrix} -r & 0 & 0 \\ \alpha & -\gamma & 0 \\ M-\delta & \frac{\beta M}{\delta-\frac{\alpha\beta}{\gamma}} & K\big(\frac{\alpha\beta}{\gamma}-\delta\big) \end{pmatrix} \tag{4.9}$$

Here, the eigenvalues are $-r, -\gamma$ and $K\Big(\frac{\alpha\beta}{\gamma}-\delta\Big)$, which are all negative if $\delta > \frac{\alpha\beta}{\gamma}$. But this is also a premise for the point to be biologically meaningful. Using proposition 2.15 and theorem 2.18, we can deduce stability here.

No we consider the starvation-case:

$$\begin{aligned} &J(Y,a,m) = \\ &= \begin{pmatrix} r\big(1-\frac{2Y}{K}\big)(m-T_mY) - rY\big(1-\frac{Y}{K}\big)T_m & 0 & rY\big(1-\frac{Y}{K}\big) \\ \alpha_1 a - \alpha_2 a & \alpha_1(Y-2a) - \alpha_2(Y-m) & \alpha_2 a \\ -\delta m & \beta m & \beta\big(a-\frac{\delta}{\beta}Y\big) \end{pmatrix} \\ &= \begin{pmatrix} r\big(3\frac{T_m}{K}Y^2 - 2\big(\frac{m}{K}+T_m\big)Y + m\big) & 0 & rY\big(1-\frac{Y}{K}\big) \\ (\alpha_1-\alpha_2)a & \alpha_1(Y-2a) - \alpha_2(Y-m) & \alpha_2 a \\ -\delta m & \beta m & \beta\big(a-\frac{\delta}{\beta}Y\big) \end{pmatrix} \end{aligned} \tag{4.10}$$

Inserting P_0:

$$J(0,0,0) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{4.11}$$

Hence, the linearization and the eigenvalues give us no information about stability. The reason, why we get only zeroes here, is that all terms of the equations in eq. (4.4) include at least two of the variables Y, a and m or a square of one variable. Thus, the equations are just too dependent on each other, to ensure they all got to 0 in this way. For example, $Y(t)$ can only go to 0, if $m(t)$ stays below $T_m Y(t)$. But $m(t)$ can only go to 0, if $a(t)$ stays below $\frac{\delta}{\beta} Y(t)$. Therefore, it is not that easy to find suitable parameter conditions, such that $Y(t)$, $a(t)$ and $m(t)$ show a biologically meaningful course and go to 0 eventually. Later we will fit our model to data and see, that we can find reasonable parameter sets were the solution curves actually converge to 0, or at least to very small values.
Looking at the other stationary states:

$$J(P_1) = J(K,0,0) = \begin{pmatrix} rKT_m & 0 & 0 \\ 0 & (\alpha_1 - \alpha_2)K & 0 \\ 0 & 0 & -\delta K \end{pmatrix} \tag{4.12}$$

This matrix has at least one positive eigenvalue, rKT_m. Hence, P_1 is always unstable according to proposition 2.15 and theorem 2.18. Of course, it could also be the case that $\alpha_1 = \alpha_2$. Then we would have 0 as one eigenvalue and could not make use of the theorem. But this is very unlikely in a biological model and hence, we just assume $\alpha_1 \neq \alpha_2$.
For P_2 with $m_s \in \mathbb{R}_{>0}$, we get

$$J(P_2) = J(0,0,m_s) = \begin{pmatrix} rm_s & 0 & 0 \\ 0 & \alpha_2 m_s & 0 \\ -\delta m_s & \beta m_s & 0 \end{pmatrix} \tag{4.13}$$

Here we have two positive eigenvalues, rm_s and $\alpha_2 m_s$, and one that is 0. Hence, P_2 is not hyperbolic and we cannot use proposition 2.15 to deduce stability or instability for the nonlinear system. Our guess is, that P_2 is at least not asymptotically stable. Because as soon as $T_m Y(t) < m(t)$, which should happen if $Y(t)$ goes to 0 but $m(t)$ to $m_s > 0$, $\dot{Y}(t)$ would be positive again, resulting in an increase of $Y(t)$. A similar argument can be applied to $a(t)$, since $\dot{a}(t)$ includes the term $+\alpha_2 a(t)m(t)$. Thus, when $a(t)$ and $Y(t)$ get small enough, this term is bigger than the rest, causing $a(t)$ to increase again. But of course, this effects can reverse again, such that $Y(t)$ and $a(t)$ decrease until once again the above conditions hold. Hence, the point could be stable from what we see here, as this only means Y, a and m stay "close enough" P_2.
Inserting P_3:

$$\begin{aligned} J(P_3) &= J\Big(K, K\Big(1 - \tfrac{\alpha_2}{\alpha_1}\Big), 0\Big) = J\Big(K, K\tfrac{\alpha_1 - \alpha_2}{\alpha_1}, 0\Big) \\ &= \begin{pmatrix} rKT_m & 0 & 0 \\ K\frac{(\alpha_1-\alpha_2)^2}{\alpha_1} & K(\alpha_2 - \alpha_1) & K\alpha_2\frac{\alpha_1-\alpha_2}{\alpha_1} \\ 0 & 0 & \beta K\big(\frac{\alpha_1-\alpha_2}{\alpha_1} - \frac{\delta}{\beta}\big) \end{pmatrix} \end{aligned} \tag{4.14}$$

The characteristic polynomial of this matrix is

$$p(\lambda) = \big(\lambda - rKT_m\big)\big(\lambda - K\big(\alpha_2 - \alpha_1\big)\big)\left(\lambda - \beta K\left(\frac{\alpha_1 - \alpha_2}{\alpha_1} - \frac{\delta}{\beta}\right)\right) \tag{4.15}$$

One can now see, that the eigenvalues are just the diagonal entries of the matrix. Thus, we have again at least one positive eigenvalue, rKT_m, such that P_3 is also unstable. If $\alpha_1 = \alpha_2$ or $\frac{\alpha_1 - \alpha_2}{\alpha_1} = \frac{\delta}{\beta}$, one of the eigenvalues could be 0 of course and then we would not be able to deduce instability so easily. But as this is very unlikely in a biological model, we assume this does not happen.
When we insert P_4 we get:

$$\begin{aligned} J(P_4) &= J\Big(K, K\tfrac{\delta}{\beta}, K\Big(1 - \tfrac{\alpha_1}{\alpha_2}\Big(1 - \tfrac{\delta}{\beta}\Big)\Big)\Big) = \\ &= \begin{pmatrix} -r(K(T_m + \frac{\alpha_1}{\alpha_2} - 1) - \frac{\alpha_1}{\alpha_2}\frac{\delta}{\beta}) & 0 & 0 \\ (\alpha_1 - \alpha_2)\frac{\delta}{\beta} & -\alpha_1\frac{\delta}{\beta} & \alpha_2\frac{\delta}{\beta} \\ -\delta(K - \frac{\alpha_1}{\alpha_2}(K - \frac{\delta}{\beta})) & \beta(K - \frac{\alpha_1}{\alpha_2}(K - \frac{\delta}{\beta})) & 0 \end{pmatrix} \end{aligned} \tag{4.16}$$

From this matrix it is not that easy to derive the eigenvalues. But using the parameter values of the fit in fig. 4.8, we computed the resulting eigenvalues in python: 9.84, -158.09 and -0.10 (values rounded). Hence, P_4 is at least unstable for this certain parameter set, as we have one positive and two negative eigenvalues.

Boundedness and domain Since we have already shown that solutions to our model stay non-negative, it remains to show that they are bounded from above.

Starting with the nutrient-rich case:

- Let $Y > K$. Then $\dot{Y}(t) = rY(t)\Big(1 - \frac{Y(t)}{K}\Big) < rY(t)\Big(1 - \frac{K}{K}\Big) = 0$. Hence, $Y(t)$ stays bounded by K.
- Let $a > \frac{\alpha}{\gamma}Y$. Then $\dot{a}(t) = \alpha Y(t) - \gamma a(t) < \alpha Y(t) - \alpha Y(t) = 0$ and thus, $a(t)$ is bounded by $\frac{\alpha}{\gamma}Y(t)$. Especially, since Y is bounded by K, a is bounded by $\frac{\alpha}{\gamma}K \leq K$ (see paragraph about stationary states).
- Let $m > \frac{M}{\delta - \frac{\alpha\beta}{\gamma}}$. Then we get

$$\begin{aligned} \dot{m}(t) &= M \cdot Y(t) + \beta m(t)\left(a(t) - \frac{\delta}{\beta}Y(t)\right) \\ &< M \cdot Y(t) + \beta m(t)\left(\frac{\alpha}{\gamma}Y(t) - \frac{\delta}{\beta}Y(t)\right) \\ &= M \cdot Y(t) + m(t)\underbrace{\left(\frac{\alpha\beta}{\gamma} - \delta\right)}_{<0, \text{ see st. states}}Y(t) \\ &< M \cdot Y(t) + \frac{M}{\delta - \frac{\alpha\beta}{\gamma}}\left(\frac{\alpha\beta}{\gamma} - \delta\right)Y(t) \\ &= M \cdot Y(t) - M \cdot Y(t) = 0 \end{aligned} \tag{4.17}$$

  Hence, m stays bounded by $\frac{M}{\delta - \frac{\alpha\beta}{\gamma}} \leq K$ (see st. states).

Now we consider the starvation case. Note that we assume $Y, a, m > 0$, as otherwise we would have to cover all cases, where some of the derivatives are already 0.

- We first consider $Y(t)$. Note, that we always assume $Y_0 \leq K$. Looking at the last term of $\dot{Y}$, we get the following two cases:
  - $m - T_m Y > 0 \Leftrightarrow m > T_m Y$. As long as this is the case, Y stays bounded by K because of the logistic growth term $rY(1 - \frac{Y}{K})$. From a biological point of view $m > T_m Y$ should be the case at the beginning of the starvation, as the cell number should not be instantly decreasing when starvation is induced. See for example fig. 4.2, where the optical density of the cells is still increasing a bit after 26 hours, although there are almost no nutrients left.
  - $m - T_m Y < 0 \Leftrightarrow m < T_m Y$. In this case Y can only be decreasing or constant, assuming it was bounded by K before and hence $rY(1 - \frac{Y}{K}) \geq 0$. Thus, Y is still bounded by K, or more precisely by $\frac{m}{T_m}$.

  So in summary, Y stays bounded by K as long as $Y_0 \leq K$.
- Let $a > Y - \frac{\alpha_2}{\alpha_1}(Y - m)$. Then $\dot{a} < 0$, similar to the single-cell case. Hence, a is bounded by $Y - \frac{\alpha_2}{\alpha_1}(Y - m)$. We would like the bound to be smaller than Y, but this is just equivalent to m being smaller than Y, which we want to show anyway.
- $\dot{m}$ can only be negative, if $a < \frac{\delta}{\beta} Y$.

Considering the conditions on $\dot{a}$ and $\dot{m}$ to be negative, we see the similarity to the single cell case. The only difference is, that we now have Ys instead of 1 in each term. But as this is the case, we had here the same problems to find constraints for boundedness as in the single cell model. Hence, we stop here and try to fit our model without further conditions on this.

Results

- In the nutrient-rich case, we need the following constraints on our parameters: $\frac{\alpha}{\gamma} \leq 1$ and $0 < M \leq K(\delta - \frac{\alpha\beta}{\gamma})$, including the condition $\delta > \frac{\alpha\beta}{\gamma}$ of course. Then we get a biologically meaningful stable stationary state and bounded solutions.
- In the starvation case, we have two stationary states where we could neither show instability nor stability, namely P_0 and P_2. Of course we would like P_0 to be stable, as this would be the biologically meaningful stationary state. But P_2 with a very small value for m_s could also be sufficient in practice. P_1 and P_3 are unstable and P_4 is at least unstable with the fitted parameter set used for fig. 4.8.
  $Y(t)$ stays always bounded by its capacity K, like we would expect. For the boundedness of $a(t)$, we need $m(t)$ to stay below $Y(t)$. But like in the single cell case, we did not find suitable conditions to ensure this in every case.

4.4 Fitting the Model to Data

Like mentioned at the beginning of the chapter, Ryo Iwama and Yoshinori Ohsumi studied autophagy in yeast under a gradual change of nutrient availability [15]. In this section we want to try and fit our model to the data they obtained. First, we shortly

review the major points of the experimental set-up and results of the study. After this we go into detail about the parameter fitting and possible model adaptions.

4.4.1 Experimental Set-Up

As a first part, the authors of [15] tried to find a good medium for the yeast to grow. Like described in section 1.1.4, yeast shows a diauxic growth when cultured in a medium containing glucose. Because the authors also wanted to obtain more detailed information on this sort of growth behaviour, they chose glucose-containing media to try. The goal was to find a medium, where the yeast growth is only dependent on the availability of carbon sources, like glucose. To represent the growth of the yeast cell culture, they measured the optical density at a wavelength of 600 nm, namely OD_{600}. With this value the number of cells per milliliter can be approximated, when some calibration is done beforehand. But as the authors stay with the density and hence provide no factor for this conversion, we will also use the density in our fitting procedure.
In a first step they used a standard medium with glucose (SDCA) to culture prototrophic wild type cells. But this medium was not appropriate, as the cell growth almost arrested before the available ethanol was depleted. Furthermore, the cells did not grow logarithmically on ethanol, like they did on glucose. So they deduced, that the growth arrest was not caused by a too low ethanol concentration, but instead by some other factors. There are two approaches to get rid of this factors: on the one hand, to increase the concentration of other nutrients, on the other hand, to decrease the glucose concentration. In the standard medium, the cells reached quite a high cell density when glucose was depleted. Hence, the authors considered this to have a possible negative effect on the further growth of yeast on ethanol. But with this in mind, increasing the concentration of other nutrients would not produce the desired properties of the yeast growth. Thus, the authors decided to lower the glucose concentration instead, to avoid such a high cell density.
With a medium containing less glucose ($SD_{0.2}CA$), the desired diauxic growth pattern and the dependence on carbon source availability could be achieved. They divided the cell growth into three phases: glucose-utilizing (3-13 h), ethanol-utilizing (13-26 h) and ethanol-depleted phase (26-32 h). This time, the yeast growth was also logarithmic in the ethanol-utilizing phase and arrested when ethanol was depleted, see fig. 4.2.

To investigate bulk autophagy, yeast cells expressing GFP-*Atg*8 were used. As mentioned in section 1.1.2, the delivery of *Atg*8 to the vacuole can be used as an indicator for bulk autophagy. For this purpose, *Atg*8 is tagged with GFP. If this GFP-*Atg*8 is then transported inside the vacuole, *Atg*8 is degraded and free GFP remains. Using a GFP-specific antibody, GFP-*Atg*8 and free GFP can be monitored by Western blotting [3]. As can be seen in fig. 4.3 A, the authors found out that free GFP could only be detected in the ethanol-utilizing and ethanol-depleted phase. Additionally, the cells were investigated at certain time points using fluorescence microscopy and the ratio of cells having fluorescence signal in the vacuoles was computed (see fig. 4.3 B). They also used *atg*2Δ cells to compare and confirmed, that the delivery of GFP-*Atg*8 to the vacuole was almost not possible there, as shown in fig. 4.3.

To investigate selective autophagy, the same assays as above were applied, but with different proteins. For details on the used proteins see [15].

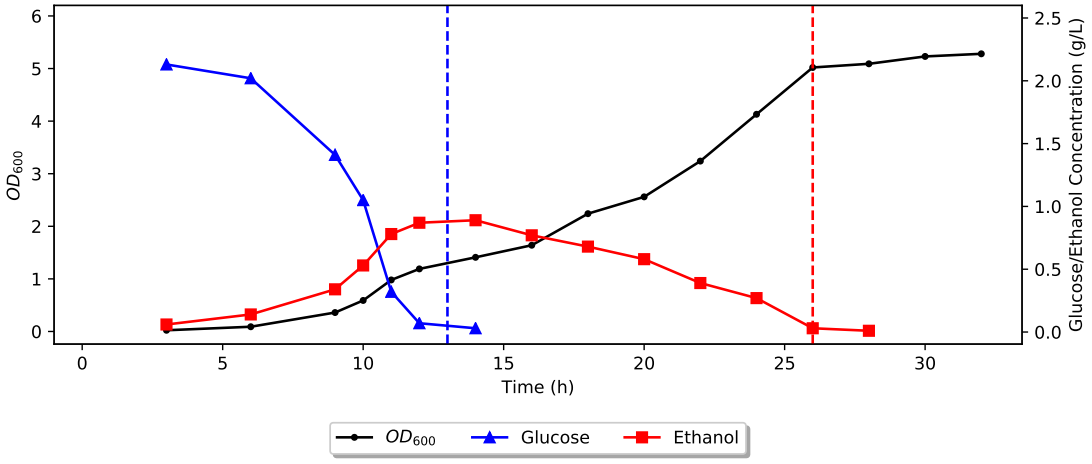

Figure 4.2: Measurements of OD_{600}, glucose and ethanol concentrations in medium with lower initial glucose concentration. The end of the glucose-utilizing phase is marked in blue (at 13 h) and the end of the ethanol-utilizing phase in red (at 26 h). Adapted from [15], datapoints derived with WebPlotDigitizer [26].

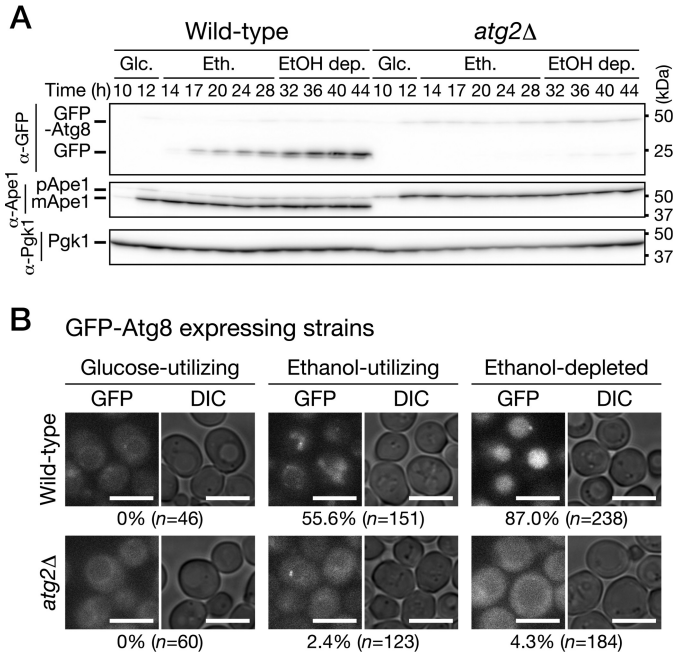

Figure 4.3: A: Western blot analysis, B: Fluorescence microscopy with ratios of cells having fluorescence signal in the vacuoles. For more information see [15]. Taken from [15].

4.4.2 Results on Autophagy

Autophagy was investigated in each of the phases with the following findings:

- Bulk autophagy and ER-phagy were first induced in the ethanol-utilizing phase. In the ethanol-depleted phase, the level of bulk autophagy increased further.
- Mitophagy only occurred when ethanol was depleted.

The authors particularly mentioned the induction of autophagy in the ethanol-utilizing phase, as the availability of ethanol was expected to be sufficient for the yeast to grow. Hence, autophagy seems to be constitutively induced when glucose is depleted and different types of autophagy are induced at different stages of nutrient availability. However, the authors gave no quantitative data we could use on autophagy for our model, but only on the amount of yeast cells and the concentration of glucose and ethanol. As shown in fig. 4.3 B, they computed the ratio of cells with a fluorescence signal in their vacuole, but this gives us no quantitative information about the level of autophagy happening within each cell. Thus, we needed to make up some additional test data to be able to fit our model in python.
Like mentioned before, we assume all cells to behave just the same. This might be a drawback when we try to fit our model to the data in [15], since fig. 4.3 B shows that the cells behave quite different. They start autophagy at different time points and hence also have different levels at each time point. We introduced a model where all cells are considered to behave identically, as this is the easiest way. But a next step could be to take this different behaviour of the cells into account.

4.4.3 Data for Fitting

As mentioned before, we had no quantitative data on the level of autophagy and cytoplasmic material we could use from the paper. However, the lack of data on the material was to be expected, as we decided on our model variables without any given datasets. We chose the material as a variable because it summarizes everything in the cell, which could get degraded through autophagy, and hence thought this to be a reasonable choice.

We originally set $Y(t)$ to be the number of yeast cells. But as the data in [15] gives OD_{600} measurements instead, $Y(t)$ will represent such optical densities in the following. Since the cell numbers and the OD_{600} values are linearly dependent, this makes no difference for our qualitative findings. According to this change, $a(t)$ and $m(t)$ are now scaled by the optical density instead of the cell number, having theoretically the same unit as $Y(t)$. In this way we can easily derive the (dimensionless) level per cell again by dividing $a(t)$, respectively $m(t)$, by $Y(t)$. But this unit for a and m makes practically not much sense, since OD_{600} measurements give cell densities and do not indicate what is happening inside the cells. Hence, a and m are of "arbitrary unit" (a.u.) here.

First, we decided to just make up some test data by hand, to see whether we could find parameter values producing an approximately, qualitatively fitting course of the solution. This data can be found in appendix 2, table 1, and was used for fig. 4.1. In fig. 4.4 we plotted the data points from table 1 and added a plot showing the level of autophagy and material per cell. In the glucose-utilizing phase we wanted a stable level of material and basal autophagy in each cell. This is in accordance to the paper, as additional bulk and selective autophagy are first induced when glucose is depleted. We

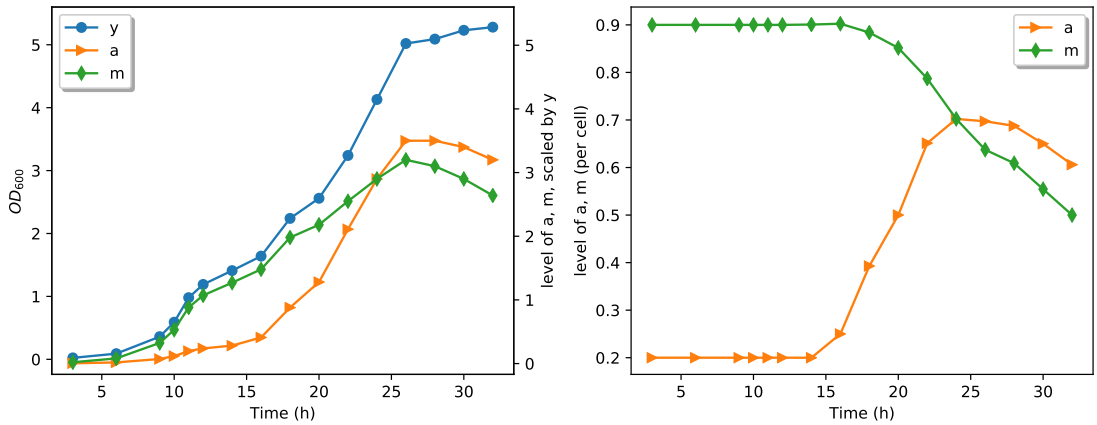

Figure 4.4: Left: y shows the OD_{600} measurements from [15], a and m are made up for fitting our model, Right: level of a and m per cell.

chose the level of material per cell in the glucose-utilizing phase to stay at 0.9. First we tried just a level of 1 per cell, as the cells should be in perfect conditions. But then many parameter sets led to unbounded solutions for a and m, as m would just have to increase a bit over Y to change the sign of $-\alpha_2(Y-m)$ in the equation for $\dot{a}$ and hence leading to a strong increase of a, m and Y. As we solved our model numerically in python, this could be also an effect of the limited numerical precision. Nevertheless, it was easier to just choose a value below 1 per cell, such that it would not matter if m increased a bit. So until 14 h, when glucose is depleted, we set $a(t)$ to be $0.2 \cdot Y$ and $m(t)$ to be $0.9 \cdot Y(t)$ in table 1. We chose a scaling of 0.2 for $a(t)$, because in this way $a(t)$ is not too high at the beginning, but is still visibly bigger 0 in the plots. As we had no data for the basal level of autophagy, this scaling seemed to be as possible as any other. One flaw in our dataset is, that the level of autophagy is decreasing in the ethanol-depleted phase instead of increasing. But as we had already great difficulty in fitting our model only approximately to this data, this made no real difference. Especially as this part of the data only comprised 4 of the 16 data points.

Next we used the first of our already derived single-cell models to generate data instead. Then we could scale this data with the cell densities given in the paper, to generate data for the whole cell batch. In this way we would hopefully be able to fit the multi-cell model more easily, as it is derived of the single-cell model.
The data used in the starvation case can be found in appendix 2, table 3, and is plotted in fig. 4.5 on the left. In the glucose-utilizing phase, the same data as before is used. First we tried parameter values resulting in a higher maximal value of $a(t)$ than plotted in fig. 4.5. But there we had great difficulty finding suitable parameter sets, as most of the time we got unbounded solution curves when $a(t)$, and then also $m(t)$, got too high. Hence, we thereafter chose parameter values resulting in a maximal value of autophagy around 0.5, since in this way we avoid problems with unbounded solution so often. Again, as we have no quantitative data to compare, we only chose in this way for the named reason.

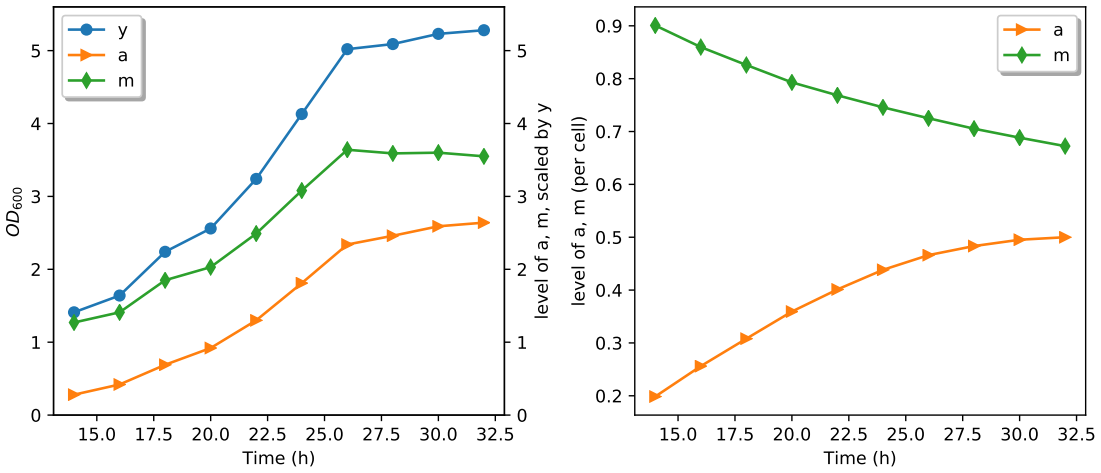

Figure 4.5: Left: data from table 3 (data plotted on the right, multiplied with y at each time point), Right: eq. (3.1) with $s(t) = 1$, solved in python using $a_0 = 0.2$, $m_0 = 0.9$, $\alpha_1 = 0.2$, $\alpha_2 = 0.3$, $\beta = 0.04$ and $\delta = 0.032$.

4.4.4 Fitting

For the fitting in python we used `symfit` [27] [1]. This is a python package providing functions and syntax to easily write fitting routines. The algorithms used are taken from another python package, `scipy`, more precisely `scipy.optimize`. A short description of the algorithms used from this package can be found in appendix 5. Additionally, we want to note, that we let choose `symfit` the objective function to be minimized. In our case, this always ended up with the default, namely least squares.
For some parts, or the fine tuning, we also adapted the parameter values by hand.

Non-starvation case For the first part of the data, in the glucose-utilizing phase, we used our equations for the non-starvation case (see eq. (4.3)) with α set to 1, such that we had one parameter less to fit. This can be done as a is already scaled according to Y and hence we do not need an additional scaling of Y.
Here it was quite easy to find good parameter values, as for $\dot{Y}(t)$ they could be fitted independently to the ones of $\dot{a}(t)$ and $\dot{m}(t)$. Hence, we first searched good values for r and K. Next we could use our knowledge of the stationary state and set $\gamma = 5$ and $M = 0.9 \cdot K(\delta - \frac{\beta}{\gamma})$ accordingly, such that $a(t)$ converges to $0.2 \cdot Y(t)$ and $m(t)$ to $0.9 \cdot Y(t)$. Thus, the only parameters left to fit were β and δ. We just tried different values by hand for them, until we found good solution curves.
When fitting only $Y(t)$, we tried different initial points. Since the first two data points are very small, the error we get when we extract the values with the WebPlotDigitizer can be comparatively high. For the point at 3 h, we first got a value of 0.04 from the WebPlotDigitizer. But when we looked at the values on a log_{10}-scale, in fig. 1.3, and compared to the paper, we saw that this value was too high. Thus, we changed it to 0.023. This already improved our first fits, which shows how sensitive the fit for the

[1] See fitting.ipynb on `https://github.com/Melanie757/Thesis-Code`.

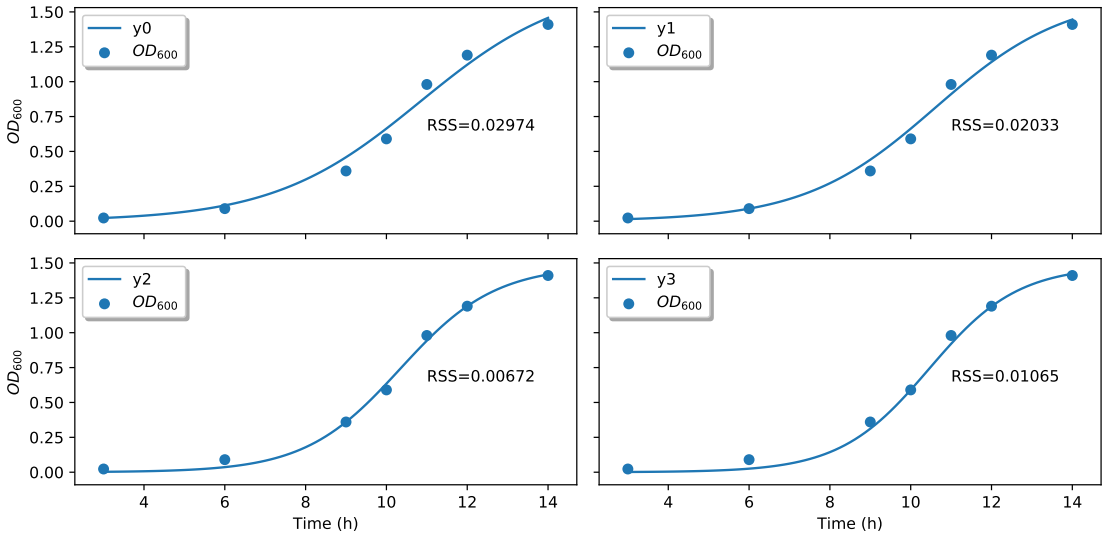

Figure 4.6: Results of fitting only the $Y(t)$ solution curve until 14 h with `symfit`, using different points as initial values. y0 corresponds to the solution when using the first point at 3 h as initial value, y1 when using the point at 6 h, y2 at 9 h and y3 at 10 h. The corresponding RSS scores are also added to the different plots.

logistic growth curve is to the values in the first, quite even part of the curve. Aside from that, we consider data from a biological experiment, where it is very unlikely that the resulting measurements follow an exact logistic growth curve. Hence, we thought it best to try the first four data points from table 1 as initial value in turn. The results can be seen in fig. 4.6. We also computed the Residual Sum of Squares (RSS), to decide which parameter set we wanted to use. We chose this value, since `symfit` does a least squares fit by default. Hence, the RSS is the objective function to be minimized. Since its value is smallest when the third point (at 9 h) is used, we chose the resulting parameter set with $r = 0.849793431417659$ and $K = 1.4801439740721747$. To check the results again we used `scipy`'s function `solve_ivp`. This is a newer function for solving IVPs than `scipy`'s `odeint`, which is used in `symfit` to compute the solutions.
One problem with `solve_ivp` was, that we needed to solve the equations also backwards, as our initial point was not our first data point in time. This worked for Y, like it worked in `symfit`, but one of the solutions for a and m were always unbounded when going back in time. This was probably due to the very small values at the beginning, where a small numerical error can have a great influence on the computed solution. Before experimenting with different values for a and m at 9 h, we decided to just read out the already computed solution for $Y(t)$ from `symfit` and use the computed value at 3 h as new initial point. The values for a and m are here only dependent on Y anyway. Then there would be no need to solve the equations back in time. Later we found out, that one could also add the initial points as parameters and let the optimization algorithm do the work. This resulted in almost the same values like we found above, so we stay with our approach for now.
Putting all together, we got the solution curves plotted in fig. 4.7 on the left. The

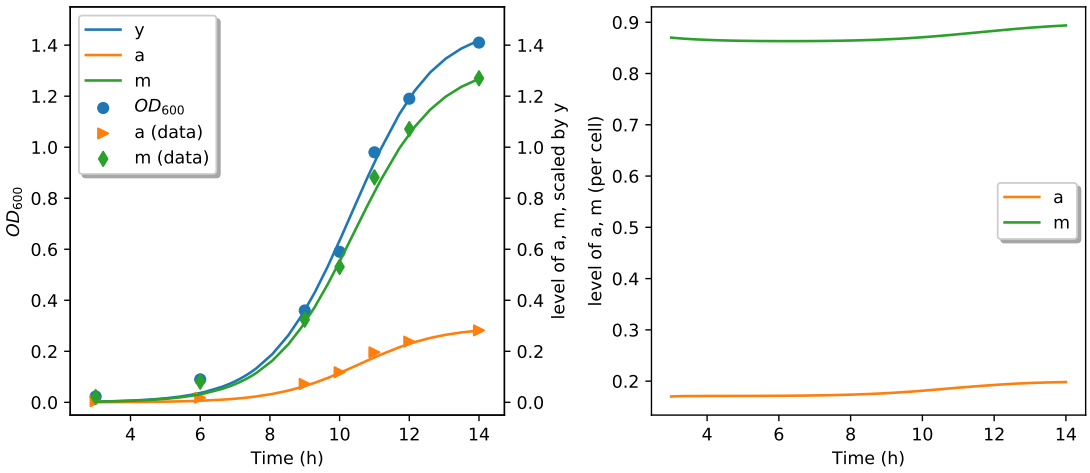

Figure 4.7: Left: Solving eq. (4.3) in `scipy` with the parameter values given in table 4.4, Right: Levels of a and m per cell.

	Name	Value
Initial values	$a(3h)$	$0.17 \cdot y(3h) \approx 0.0005$
	$m(3h)$	$0.87 \cdot y(3h) \approx 0.0025$
	$y(3h)$	0.00289816
Parameters	α	1.0
	β	0.1
	γ	5.0
	δ	0.57
	K	1.4801439740721747
	M	$0.9 \cdot K(\delta - \frac{\beta}{\gamma}) \approx 0.7327$
	r	0.849793431417659

Table 4.4: Initial values for the variables and parameter values used for fig. 4.7.

corresponding initial values and parameter values can be found in table 4.4. For a and m we chose initial values that lie slightly below the value they are supposed to reach. We did this, because in this way the convergence in the single-cell plot, fig. 4.7 on the right, is good to see. One could also just start with $a(3h) = 0.2 \cdot y(3h)$ and $m(3h) = 0.9 \cdot y(3h)$.

Starvation case When we could not find suitable parameter values with `symfit`, we tried to just solve a system of equations to get the parameter values, making use of our data in table 3 to approximate derivatives at each point by difference quotients. But using the first few points resulted in parameter sets giving far too low solution curves. When we then used points in the middle of the time span or just some points, that are distributed over the whole time span, we mostly got unbounded solution curves for a and m. Additionally, we used the computed parameter values as a starting point for the fitting in python, but this also did not work. In a further attempt we fixed one of

the curves, for example $a(t)$, by representing it by a polynome and only tried to fit the remaining parameter values. Again, this most of the time gave us unbounded solution curves for a and m. This already led us to the assumption, that we would not be able to get a "perfect" fit. Furthermore, as mentioned in section 4.4.1, the authors did not only want to investigate autophagy but also the diauxic growth of yeast. This is something we did not explicitly include in our model and hence could cause problems when trying to fit the parameter values.
Because using our data to find parameter values did not work, we again tried to fit our model with `symfit`. We used wide ranges for the parameter values, hoping to get an idea where useful parameter values could lie. But, as was to be expected, with no prior knowledge on the parameter values we had to experiment quite a while to find good parameter sets.
To make the fitting easier and faster, we reduced the number of parameters. To achieve this, we tried to find constraints on some of the parameters, such that they are only dependent on other parameters. Therefore, we used points of our dataset, where one of the derivatives should be approximately zero:

- Looking at the time course of $a(t)$ in the left plot of fig. 4.5, one can see that the level of autophagy still seems to be increasing a bit after 30 hours. This makes sense, since the authors of [15] wrote about a further increase of autophagy in the ethanol-depleted phase. However, with our generated data, this increase is only around $\frac{2.64-2.59}{2}\frac{OD_{600}}{h} = 0.025\frac{OD_{600}}{h}$. Hence, we thought it worth a try to assume $\dot{a}(32) \approx 0$, to reduce the number of parameters. This resulted in: $\dot{a}(32) \approx 0 \Leftrightarrow \alpha_1(5.28 - 2.64) \approx \alpha_2(5.28 - 3.55) \Leftrightarrow \alpha_2 \approx \frac{2.64}{1.73}\alpha_1 \approx 1.526 \cdot \alpha_1$.
- For $m(t)$ we see that its derivative is around 0 after 26 hours. Thus, we get $\dot{m}(26) \approx 0 \Leftrightarrow a(26) = \frac{\delta}{\beta}Y(26) \Leftrightarrow \beta \approx \frac{5.02}{2.34} \approx 2.145 \cdot \delta$.

The optical cell density still seems to be increasing after 30 hours, but not much anymore. As we additionally know, that glucose and ethanol are already depleted after 30 hours, we assumed $\dot{Y}(32) \approx 0$ to get an idea for the value of T_m: $\dot{Y}(32) \approx 0 \Leftrightarrow m(32) \approx T_m Y(32) \Leftrightarrow T_m \approx \frac{3.55}{5.28} \approx 0.672$. First we tried to set α_2, β and T_m fixed in this way. But then the solution curves we found did not flatten when they reached 30 h. Instead a and m just rapidly increased, resulting in an unbounded solution. So we decided to just fix α_2 and β, while T_m was fitted in a range from 0.6 to 0.7. We chose T_m, since we had at least a rough idea where its value should lie. This last approach worked best so far, as can be seen in fig. 4.8. The corresponding code can be found on `https://github.com/Melanie757/Thesis-Code` and in appendix 4 as an example for the fitting in `symfit`. Because we knew the equations of the non-starvation case would be far more easy to fit, we primarily focused on the fitting of the starvation case. Hence, we always used as initial data the values at 14 h from table 3. Later, when we had decided on the initial values and parameter values in the non-starvation case, we used the end points of this solutions as initial data for the starvation case. Instead of fitting again with this new, but only slightly different, initial data, we just adapted the values by hand. This is the reason, why some of the values have just three decimal places, while others have 16, the most python usually outputs. In this way we get a continuous solution. At first we also wanted the derivatives to be equal at this time point, to get a smooth solution. But the first problem was, that such a feature is not

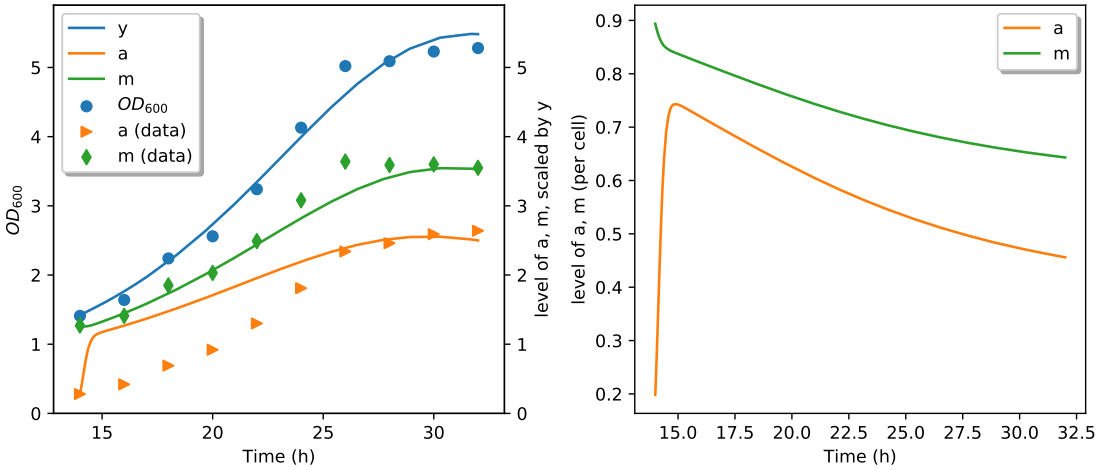

Figure 4.8: Left: Solving eq. (4.4) with the parameter values given in table 4.5, Right: Levels of a and m per cell.

	Name	Value
Initial values	$a(14h)$	0.2811824057660078
	$m(14h)$	1.2673713692691995
	$Y(14h)$	1.4180265001692174
Parameters	α_1	6.0
	α_2	$1.526 \cdot \alpha_1 \approx 10.448$
	β	$2.145 \cdot \delta_1 \approx 0.222$
	δ	0.093
	K	53
	r	0.364
	T_m	0.645

Table 4.5: Initial values for the variables and parameter values used for fig. 4.8.

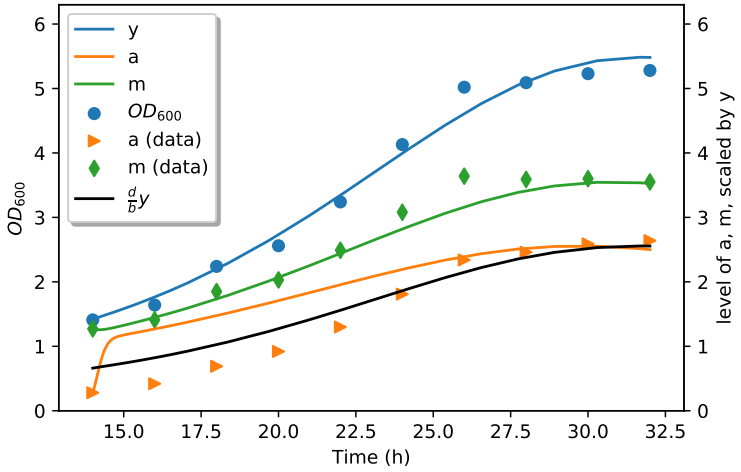

Figure 4.9: fig. 4.8 with $\frac{\delta}{\beta}Y(t)$ added.

yet implemented for ODE-systems in `symfit`. Another problem was also, that solution curves for a would then always stay very low. So we decided for now to go on without smoothness at this point.
The biggest gap by far between our data points and the numerical solution is the high increase of a between 14 and 15 hours. When adding $\frac{\delta}{\beta}Y(t)$ in the plot, like in fig. 4.9, it instantly gets clear why we need this rapid increase. As $\dot{m}$ can only be positive if $a > \frac{\delta}{\beta}Y$, a has to increase that fast to cross $\frac{\delta}{\beta}Y$. Another reason for this increase, of course, is the small value of $Y(t) - m(t)$ at the beginning. We tried lower values for $\frac{\delta}{\beta}$, but then a cannot increase as high as our data points, since otherwise it would not decrease quickly enough to get $\dot{m} < 0$ at the end of the time span. We also attempted to find values for α_1 and α_2, such that the increase would not be that steep. As visible in fig. 4.9, a could also take lower values and would cross $\frac{\delta}{\beta}Y$ all the same. But then the further increase was also very low. This resulted in a similar, albeit lower course of the solution when considering the levels for a single cell.
Looking at it from a biological point of view, we thought it not that unlikely that the cells should first rapidly increase their autophagic activity to adapt quickly to the new conditions. But the following monotone decrease of the levels per cell does not fit to the results in [15]. What does fit is the convergence of all curves to 0, or at least to a small positive value, like can be seen in appendix 3.5, fig. 6. As the cells get no further nutrients in this model, they should eventually die out.

4.5 Sensitivity Analysis

Now that we have found a set of parameter values for our system of equations (4.4), we are interested in how sensitive the solution curves respond to changes in these values. Therefore, we want to apply a sensitivity analysis. To get a first overview of different methods, we refer the interested reader to [28]. We also used this review paper to decide on the methods we wanted to cover here: the Sobol method and Morris method.

Reasons why we chose these two methods, are that both are global and show interactions between parameters. Also, applying two different methods, we have the possibility to compare the results and see, whether they agree or not. A python implementation of these two methods is provided in the library `SALib` [29].

First of all, we give a short description of the two methods and what their differences are. Next, we show and discuss the results we got with `SALib`.

Sobol Method ([30]) The Sobol method is a variance based method, meaning that it computes how much of the model variance is caused by a single parameter or interactions of multiple parameters. In this way, one can find out which parameters reduce the model variance most when they are fixed. One can also look at parameters, which lead to no significant change in the model variance, and fix them to reduce the number of free parameters.
Let $Y = f(X_1, X_2, ..., X_k)$ be the model to be analysed, where Y is a scalar. Additionally, all parameters are assumed to be independent. The variance caused by a single parameter (first order sensitivity coefficient) can now be written as

$$S_i = \frac{V_{X_i}(E_{\mathbf{X}_{\sim i}}(Y|X_i))}{V(Y)}, \tag{4.18}$$

where $\mathbf{X}_{\sim i}$ denotes a matrix of all parameters except X_i. Hence, $E_{\mathbf{X}_{\sim i}}(Y|X_i)$ is the mean of Y taken over all possible values of $\mathbf{X}_{\sim i}$, where X_i is kept fixed. Taking then the variance over all possible values of X_i and dividing by $V(Y)$, gives us a normalized coefficient (for more details see [30]).
Next, also the total effect index S_{T_i} for parameter X_i is considered. This index comprises first and higher order effects of X_i:

$$S_{T_i} = \frac{E_{\mathbf{X}_{\sim i}}(V_{X_i}(Y|\mathbf{X}_{\sim i}))}{V(Y)} = 1 - \frac{V_{\mathbf{X}_{\sim i}}(E_{X_i}(Y|\mathbf{X}_{\sim i}))}{V(Y)} \tag{4.19}$$

where the identity $V_{\mathbf{X}_{\sim i}}(E_{X_i}(Y|\mathbf{X}_{\sim i})) + E_{\mathbf{X}_{\sim i}}(V_{X_i}(Y|\mathbf{X}_{\sim i})) = V(Y)$ was used.
The authors of [30] also describe a scheme for computing the second order effects only. But as we do not need this for our analysis, as we will see later, we omit a description here.

Morris Method ([31] and [32]) In contrast to the above method, which gives us quantitative information on the influence of each parameter (including interactions), this method only gives us qualitative information. Its aim is to determine, whether the effects of each parameter are negligible, linear and additive, or nonlinear and/or involved in interactions. But it does not tell us which certain parameters are involved in interactions with each other.
Let $Y = f(X_1, X_2, ..., X_k) = f(\mathbf{X})$ be again the model we want to analyse. Furthermore, all parameters are assumed to lie in [0, 1]. For each parameter X_i, two main sensitivity measures are computed: μ_i, the overall influence of X_i on the output, and σ_i, as an estimate for all higher order effects. Both measures depend on the so-called elementary effect $d_i(\mathbf{X})$ of parameter X_i. The idea is to use a finite difference scheme for $\frac{\partial Y}{\partial X_i}|\mathbf{x}$, since this can be seen as a functional index describing the influence of X_i on Y:

$$d_i(\mathbf{X}) = \frac{Y(X_1, ..., X_{i-1}, X_i + \Delta, X_{i+1}, ..., X_k) - Y(\mathbf{X})}{\Delta} \tag{4.20}$$

To be able to compute this finite difference scheme reasonably, a region of experimentation, Ω, is defined. This corresponds to a regular k-dimensional p-level grid. In this grid, each X_i can take values from $\{0, \frac{1}{p-1}, \frac{2}{p-1}, ..., 1\}$. Now, the elementary effect of X_i can be computed for each $\mathbf{X} \in \Omega$ with $X_i \leq 1 - \Delta$, where Δ is a predetermined integer multiple of $\frac{1}{p-1}$. In [31] and [32] for example, p is chosen even and $\Delta = \frac{p}{2(p-1)}$.
To now get the desired sensitivity measures μ_i and σ_i, different values for $\mathbf{X}$ are sampled randomly from Ω to get finite distributions of the elementary effects. From these distributions, the mean μ_i and the standard deviation σ_i can be estimated.
The authors of [32] propose additionally a refined version μ_i^* of μ_i. This measure is especially effective if elementary effects can have different signs. The problem when computing μ_i in this case is, that elementary effects with opposite signs can cancel each other out. In this way μ_i may take a value near 0, although the effect of X_i is high, but can be of positive and negative sign. To avoid this problem, the absolute values of elementary effects are used to estimate μ_i^*. Then both estimates, μ_i and μ_i^*, can be used to determine, whether X_i has a significant effect and whether this effect has a positive, negative, or both signs.
In his publication [31], Morris also describes a strategy for efficiently sampling from the needed distributions. For this purpose, he constructs r trajectories consisting of $(k+1)$ points from Ω. Then, each trajectory provides one elementary effect for each parameter. This sampling method was later improved by the authors of [32]. They first sample a higher number of trajectories than needed in the end, with the original strategy. Next, they choose from these trajectories the ones with the most spread, based on a definition of distance they give in their article.

Implementation in `SALib` In `SALib`, both methods described above are implemented based on, amongst others, the sources we cited. Furthermore, the library provides sampling functions `SALib.sample.saltelli.sample` and `SALib.sample.morris.sample`. Those functions generate suitably formatted input values, such that the corresponding model outputs can then be used for the sensitivity analysis. For the Sobol method, an extension of the Sobol sequence proposed by Saltelli et al. in [30] is used. For the Morris method, the sampling strategy of Morris in [31] and the improved sampling strategy proposed in [32] are implemented. After using the sampling functions, one has to input those samples into the model to get the corresponding output values for the analysis. So, `SALib` does not directly interact with the model, but only generates the samples and then takes the model outputs again for the next step. The model outputs can be analysed with `SALib.analyze.sobol.analyze`, respectively `SALib.analyze.morris.analyze`.
The output of the Sobol analysis are the first order sensitivities S_1, total order sensitivities S_T and second order sensitivities S_2, if needed. Additionally, confidence intervals are given, per default with a confidence level of 0.95. The Morris analysis returns μ, μ^* with confidence interval and σ.

Results As input for the sensitivity analysis we chose the RSS of each solution curve with the corresponding data points. We decided on this input, since we want to use, and already used, this value to compare different fits. In this way we have three scalar outputs, one for equation $\dot{Y}$, $\dot{a}$ and $\dot{m}$. We chose to output a value for each curve instead

Parameter	Minimum	Maximum
r	0.346	0.382
K	50.35	55.65
T_m	0.613	0.645
α_1	5.700	6.100
α_2	9.000	9.614
β	0.190	0.200
δ	0.092	0.098

Table 4.6: Parameter bounds for the Sobol sensitivity analysis.

of only one value, since we also wanted to see the different influences of the parameters on every single equation.

We begin with the results of the Sobol method. Therefore, one first has to choose bounds on each parameter, such that within these bounds the samples can be generated. Our parameter bounds are given in table 4.6. We first started with bounds determined by adding or subtracting 5% of the original parameter value. But this often led to samples, which gave unbounded solutions and hence `NaN`s as RSS. To fix this problem, we printed for each parameter all the samples in a scatter plot, which led to `NaN` as output. Then we tried to restrict our bounds, such that we change the values as little as possible but get rid of as many `NaN`-producing values as possible. Of course, we always wanted the original parameter value to be still within the bounds. After three runs of this procedure, we got bounds which only rarely produced `NaN`s. Meaning that, after running the sampling function at most two or three times, we always got samples producing only valid outputs. We did not want to restrict the bounds even more, such that maybe no `NaN` values are produced at all, since we already have very tight bounds. Limiting them even more would also exclude many informative parameter sets.
For the sample size we tried different values between 100 and 1500 and found, that a sample size of 700 gives a quite stable result while running in a short time. The resulting values for the S_1 and S_T sensitivities can be found in fig. 4.10 [2]. As we did not use greek letters in our python implementations, the parameters are named a bit differently in the plots: $a1$ for α_1, $a2$ for α_2 and so on. To determine only the influence of higher order sensitivities, one can subtract the first order sensitivities from the total ones. The results are visualized in fig. 4.10 in the lower plot.
As it can be seen there, all estimated higher order sensitivities lie clearly between -0.05 and 0.05. Hence, the higher order influences are negligible compared to the first order ones. This is also the reason why we did not compute the second order sensitivities with `SALib`. Of course, negative values should not be possible. But, as the sensitivities are only estimated numerically, this can happen due to numerical errors. Additionally, we included the summed confidence intervals. They exceed 0.1 in some cases, but since the estimated values lie that near to 0, there is evidence that there are no significant higher order influences.
The plots in fig. 4.10 show, that r and K have nearly no influence on the resulting RSS (within their bounds). We also experienced this, when we tried to fit parameter values

[2] For the corresponding code see sensitivities.ipynb on `https://github.com/Melanie757/Thesis-Code`.

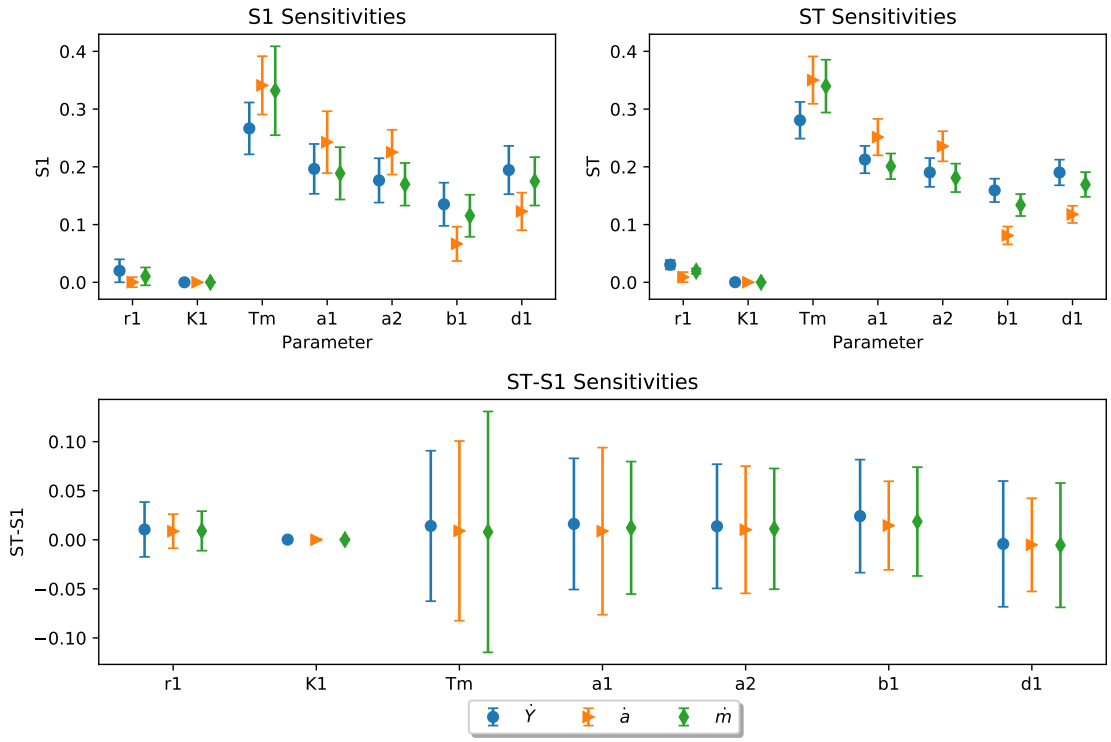

Figure 4.10: The results of the Sobol method computed with `SALib`, plotted as errorbars. Top left: S_1 sensitivities with their confidence interval for each parameter and each equation of system (4.4), Top right: the same for the S_T sensitivities, Bottom: total sensitivities minus first order sensitivities, including the resulting confidence intervals.

Parameter	Minimum	Maximum
r	0.355	0.370
K	52.50	53.50
T_m	0.610	0.647
α_1	5.800	6.030
α_2	9.100	9.610
β_1	0.1900	0.1995
δ_1	0.0927	0.1000

Table 4.7: Parameter bounds for the Morris sensitivity analysis.

by hand. β_1 seems to have only little influence as well, especially on the result for $\dot{a}$. The most influence shows T_m. Interestingly, even more on $\dot{a}$ and $\dot{m}$ than on $\dot{Y}$.

Next, we want to have a look at the results when using the Morris method. As a reminder, this method only gives us qualitative information. First, we again have to choose bounds on the parameters, which are given in table 4.7. One can see, that we often have even tighter bounds than for the Sobol sampling (table 4.6). One reason for this is the way, how the samples (respectively trajectories this time) are generated. As we already mentioned in the paragraph about the Morris method, the trajectories are generated on a regular k-dimensional p-level grid. By default, p is set to 4. We stayed with this value at the beginning, as both Morris ([31]) and the authors of [32] used this value for their different examples. As initial guess for our bounds we used the same ones as above ($\pm 5\%$). But applying now the method to adjust the bounds as before, we always have to exclude at least one of the four values, each parameter can take. Of course, changing the value of p or just trying to extend some of the bounds again could result in wider parameter ranges. But we thought it might be more reasonable to spend more time on the analysis. If the results seem inconsistent, we can still work on the bounds again, to see whether these caused problems.
Another point worth noticing is, that we used the sampling strategy of Morris instead of the improved strategy. Typically, between 10 and 50 trajectories are used. With this number the improved strategy has advantages, as it explores the parameter space more thoroughly than the original strategy with the same number of trajectories [32]. This is particularly interesting when each model evaluation takes a considerable time and thus only few but well distributed trajectories are needed. We first tried the `SALib` implementation for this strategy with 50 trajectories. Generating the samples worked well, but we got fairly different results of the analysis every time we generated new ones. Hence, we concluded that the number of trajectories is just not sufficient for the analysis of our model. But increasing the number of trajectories led to a high increase in computation time, since more and more trajectories have to be compared to find the most spread ones. However, as each of our model evaluations does not take a considerable time, we simply can use far more trajectories generated with the original strategy instead of using a few well chosen ones. This should cover the parameter space also well enough for the following analysis. We decided on a number of 1000 trajectories. This gave us far more stable results than before, but not as stable as we would like to have. Hence, we tried to increase the value of p from 4 to 8. With this we reached stable results.

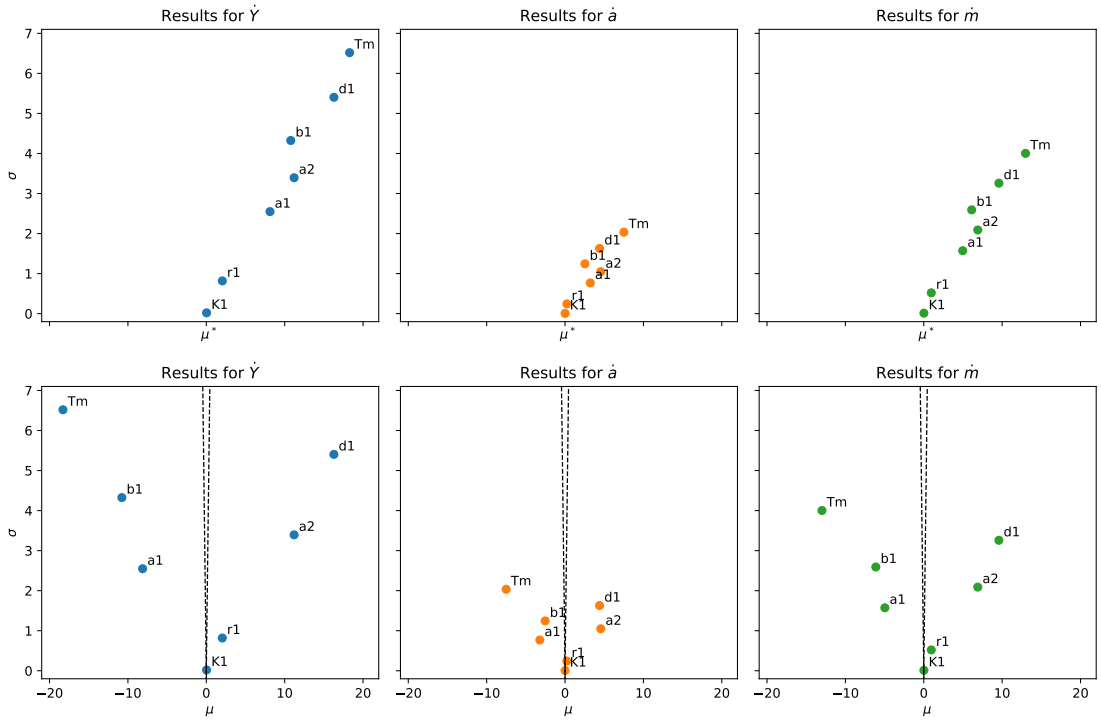

Figure 4.11: The results of the Morris method computed with `SALib`. First row: μ^* and σ plotted for all parameters and each equation we want to analyse, Second row: the same for μ and σ, together with $\mu = \pm 2 \cdot SEM$ (dashed lines).

The plots showing μ^* and σ, respectively μ and σ, can be found in fig. 4.11. The first row of plots shows μ^* and σ for every parameter and each equation. Here, one can see that the Morris method ranks the influence of T_m highest too. K and r show again only small influences. Especially the elementary effect of K has a mean and standard deviation of around 0 for every equation. The ranking of the remaining parameters is different from the one obtained with the Sobol method. This can have multiple reasons, as both methods rely on different samples and different computational schemes. But getting the same most and least influential parameters with both methods tells us, that this should be a reliable result. The second row of plots shows μ and σ, thus the mean and standard deviation for every elementary effect with its corresponding sign. Furthermore, we included the Standard Error of the Mean (SEM) in the plots like suggested by Morris in [31]. The SEM is computed as follows for the i-th parameter:

$$SEM_i = \frac{\sigma_i}{\sqrt{r}}, \tag{4.21}$$

where r denotes the number of trajectories. In his publication, Morris uses the wedge defined by $\mu = \pm 2 \cdot SEM$, to identify parameters with a non-zero elementary effect. If a point lies outside the wedge, this could give evidence that μ_i is really non-zero. Since we chose $r = 1000$ rather high, our wedge is very small. Only the points corresponding to the elementary effect of K_1 seem to lie inside, but they are near (0,0) anyway. Hence, according to Morris, all other effects have probably a non-zero mean.

When we compare both rows of plots, we do not seem to have problems with canceling signs. Comparing for example the results for T_m in the first column of plots, we can see that $\mu \approx -\mu^*$. Hence, we can use the plots with μ^* (first row) as an easy to read ranking of the parameter's influences. The second plots with μ then show us the sign of the elementary effects. This is especially interesting, as the Sobol method only tells us the proportion of model variance each parameter accounts for, but no "direction" of the influence within the parameter bounds. Using μ instead also tells us, that the elementary effect of T_m for example is probably negative. This suggests that increasing T_m within its bounds mostly leads to a decrease of our analysed function (see eq. (4.20)), which is the RSS of the solution curves. Considering the parameter bounds for T_m ([0.610, 0.647]) and the original value (0.645), this makes sense as the original value lies near the upper boundary and should provide an optimal value within these bounds.
Additionally, one has to keep in mind the standard deviation σ, which indicates higher order effects like interactions or non-linearity. In all plots of fig. 4.11, T_m has the highest standard deviation. But taking into account that the Morris method only provides qualitative results, we would in this case rely more on the result of the Sobol method, where the higher order effects of T_m were negligible. But for such cases we applied two methods, to compare and get the most information out of both.
There is one output from `SALib.analyze.morris.analyze` that we did not use here, the confidence intervals of μ^*. We tried to add them in the first row of plots in fig. 4.11, but they were just so small that they were hardly or not at all visible behind the points already plotted. Thus, we decided to omit them.

To summarize the most important results:

- Both methods suggest, that T_m has the most influence on the resulting RSS, no matter which of the solution curves (Y, a or m) we consider. Hence, the model variance can be significantly reduced when this parameter is fixed, but its value has to be chosen carefully.
  The Sobol method also gives only negligible higher order effects for each parameter.

- The influence of K is always around 0 and the influence of r at least very small compared to the rest. Hence, those two parameters can first tried to be fixed when fitting the model to data. In this way one could reduce the number of model parameters, to speed up the fitting procedure.

4.6 Modifications of the Model

In section 4.4.4 we found parameter sets for our model, where the level of autophagy per cell is first rapidly increasing and then slowly decreasing. That is not what we searched for according to the results in [15]. Thus, we tried to modify our model to get better fitting results.

4.6.1 Generalized Logistic Growth

Our first idea to improve the model was to use a generalized logistic growth for the cell number $Y(t)$ (see also section 2.2.3). This does not attenuate the rapid increase of autophagy at the beginning, but we thought it to be an interesting approach to improve

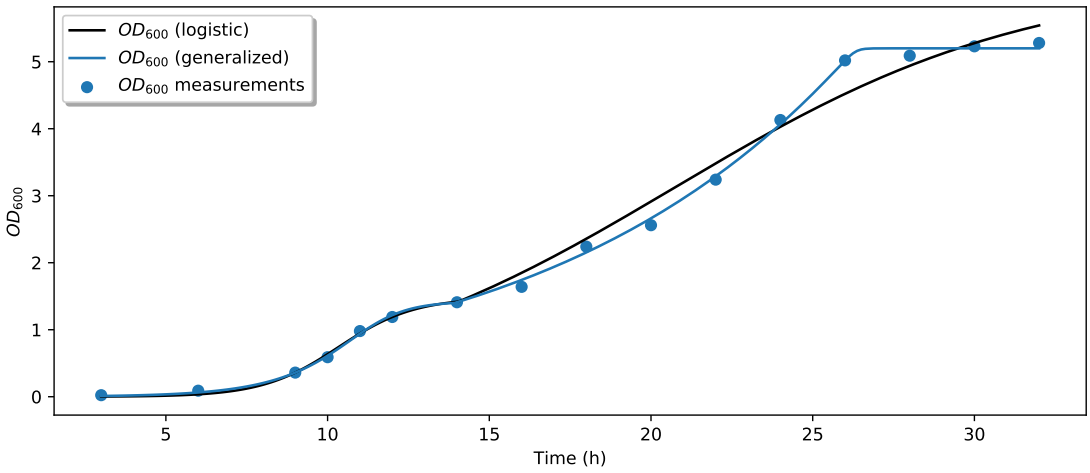

Figure 4.12: A generalized logistic and (normal) logistic growth model fitted to the OD_{600} data. In the first part, 3 - 14 h, we used the third data point as initial value for the fitting procedure. In the second part, 14 - 32 h, we just used the value at 14 h to get a continuous curve.

the fit for $Y(t)$ at least.
Unlike the (normal) logistic growth, the generalized version can also be non-symmetric, which seems to be the case for 14 - 32 h. As a short reminder, the equation for the generalized logistic growth we use is $\dot{Y} = \frac{r}{n}Y(t)\left(1 - \left(\frac{Y(t)}{K}\right)^n\right)$. To see whether this could really be beneficial, we fitted the just stated equation to the first part (3 - 14 h) and second part (14 - 32 h) of the OD_{600} data. The result can be seen in fig. 4.12.

As we already figured out in section 4.4.4, for the first part from 3 to 14 h, the third data point seems to work best as initial value for the fitting. Nevertheless, we again tested the first four data points for the generalized model. The plots with corresponding RSS and logistic growth curves can be found in appendix 3.2 fig. 2. Again, the third data point seems to give the best result. However, the fit of the logistic model was already really good and the generalized version cannot get much better.
Where we get a much better result with the generalized logistic growth is the second part from 14 to 32 h, like clearly visible in fig. 4.12. This suggests, that the assumption of a simple logistic growth of the cells in the starvation case could be improved.
However, this result will be different to what we can achieve with our whole model from eq. (4.4), since we have the additional term $(m - T_m Y)$ in our equation for $\dot{Y}$ there.

According to the above, we changed the equations for $\dot{Y}(t)$ in the following way:

- when nutrients are available: $\dot{Y}(t) = \frac{r}{n}Y(t)\left(1 - \left(\frac{Y(t)}{K}\right)^n\right)$,
- when the cells are starving: $\dot{Y}(t) = \frac{r}{n}Y(t)\left(1 - \left(\frac{Y(t)}{K}\right)^n\right)(m(t) - T_m Y(t))$.

The other equations stay the same as in eq. (4.3) and eq. (4.4), respectively.
First, the fitting was done the same way as described in section 4.4.4. But this method uses a least squares objective, hence all deviations of our solution curves from the data points are weighted equally. However, since we only aimed an improvement of $Y(t)$,

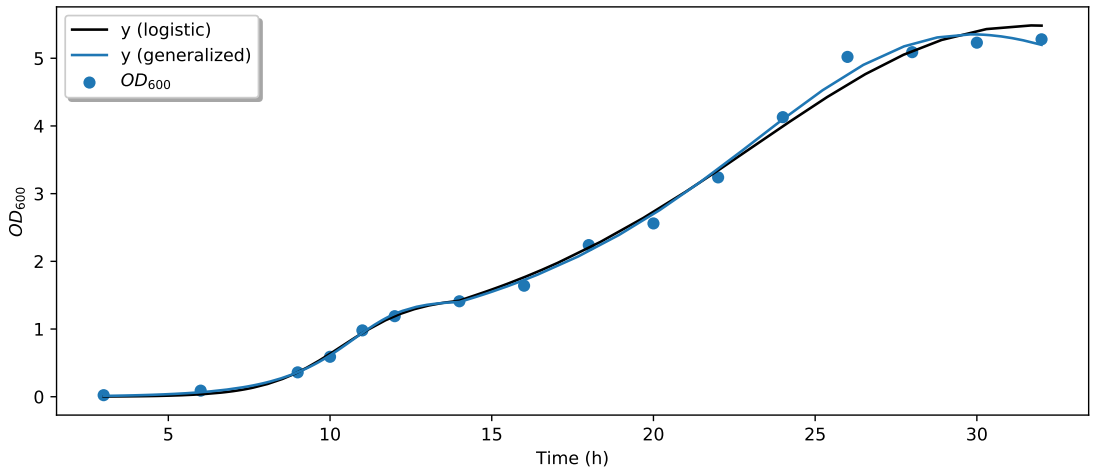

Figure 4.13: The resulting solution curves for $Y(t)$, when fitting the model with a generalized logistic growth term in $\dot{Y}(t)$. Additionally, the corresponding curves of the model with (normal) logistic growth are plotted. The initial and parameter values can be found in appendix 3.3, table 4 and table 5.

we tried our own weighted version of a least squares objective, adding also a penalty if the solution curves tend to infinity [3]. Therefore, we chose `scipy.optimize.minimize` instead of `symfit`, since the implementation of a custom objective was easier there. As starting values we used the parameter set found with `symfit` before. With this method the fits for $Y(t)$ and $m(t)$ improved compared to the results of `symfit`, when considering the corresponding RSS. For $a(t)$ we get a higher RSS value than before [4]. It was to be expected, that the RSS of either $m(t)$ or $a(t)$ would increase, as they are both not that highly weighted in the minimization as before. Compared to the original model, the RSS values of all solution curves improved (see chapter 5). The resulting plots with all solution curves and parameter values can be found in appendix 3.3. Since we tried with our modification to improve $Y(t)$, we give here only a plot in fig. 4.13 comparing this certain curves. As already described before, until 14 h it does not really make a difference whether to use a logistic growth model or the generalized version. In the second part, one can see that we get a better fit. The only drawback is, that we already have a clearly decreasing curve after 30 h, whereas the data points are still increasing slightly.

Like for the original model, we want to apply a sensitivity analysis. In this way, changes in the dependence of the model on different parameters can be identified. The visualized results for the Sobol method and applied bounds can be found in appendix 3.6, fig. 9 and table 7. One can see, that again T_m has the highest S_1 and S_T values. Although the confidence interval of $S_T - S_1$ for T_m goes up to around 0.15, the estimates for $S_T - S_1$ lie clearly between -0.05 and 0.05. Hence, we assume the higher order influences to be negligible. This time, the gap between the values for T_m and the other parameters is even higher. Furthermore, our newly added parameter n has almost no influence on the

[3] See fitting.ipynb on `https://github.com/Melanie757/Thesis-Code`.
[4] See outputs in fitting.ipynb.

variance of the RSS. This suggests, that it could be fixed early to reduce the number of parameters for the fitting. A similar result can be seen in the plots of the Morris analysis (appendix 3.6, fig. 10). The negative value of μ for T_m can again be explained by the chosen bounds on this value for the analysis.

4.6.2 First Modification of the Equation for Material

Now we want to improve the equation for $\dot{m}$ when the cells are starving, to achieve a slower increase of autophagy instead of a rapid increase. Therefore, we again look at the equation in its expanded form:

$$\dot{m}(t) = \beta m(t)\Big(a(t) - \tfrac{\delta}{\beta}Y(t)\Big) = \beta m(t)a(t) - \delta m(t)Y(t) \tag{4.22}$$

First of all, we considered changing the first term into $\beta(Y(t) - m(t))a(t)$, as then m would be bounded by $Y(t)$. Additionally, when looking at the equation in this form, multiplying with $Y(t)$ in the last term seems unnecessary. Hence, we arrive at the new equation

$$\dot{m}(t) = \underbrace{\beta(Y(t) - m(t))a(t)}_{\text{increase due to autophagy}} - \underbrace{\delta m(t)}_{\substack{\text{(normal)}\\ \text{degradation}}} \tag{4.23}$$

using the same equations for $\dot{Y}$ and $\dot{a}$ as in eq. (4.4). Thus, we have again the (normal) logistic growth term in the equation for the number of cells. We did not want to additionally change this to a generalized logistic growth, as this would be one parameter more to fit.
From a biological point of view, the first term denotes again the increase of material due to the level of autophagy and dependent on the level of material, that is already in the cells. But this time the increase gets smaller the bigger m is, accounting for the limited space within each cell. Furthermore, the degradation of material is now only dependent on the current level of material itself.
With this new system we did not find a better fit to the data. Actually, most fitted solution curves tended to approximately (K, K, K). This happened when δ was chosen too small. One stationary point of the system is $Y^* = K$, $a^* = Y^* - \frac{\alpha_2}{\alpha_1}(Y^* - m^*)$ and $m^* = \frac{\beta a^* Y^*}{\beta a^* + \delta}$. Hence, when δ is very small, m^* and a^* are approximately $Y^* = K$. But choosing a bigger δ instead resulted in monotonic decreasing solutions for $m(t)$. So we tried to find conditions on the parameters such that this point becomes unstable. Unfortunately, setting the point into the corresponding Jacobian matrix gave us no practicable constraints.
What we did find was a parameter set, which gives interesting solution curves apart from fitting the data points (see fig. 4.14 and table 4.8). As our data for autophagy and cell material are no real measurements anyway, we also want to show such findings.
One arrives at $\delta = 1.2865 \cdot \beta$, when assuming $\dot{m}(32) = 0$ and taking the corresponding values for $Y(32), a(32)$ and $m(32)$ from table 3.
Although we have again this rapid increase of autophagy at the beginning, we get a slower increase starting around 26 h, where the cells enter the ethanol-depleted phase. Furthermore, the levels of autophagy and material per cell almost stabilize between 16 and 26 h. This is not exactly what we searched for. But when considering the curves, we thought it not that far-fetched, that the cells should first strongly activate

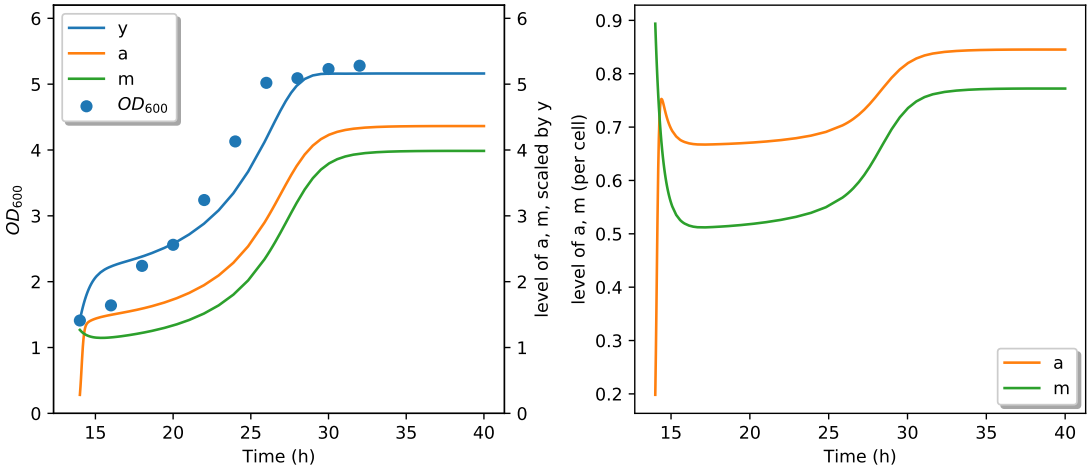

Figure 4.14: Left: Solving eq. (4.4) with our modified equation for $\dot{m}$, eq. (4.23), using the parameter values given in table 4.8, Right: Levels of a and m per cell.

	Name	Value
Initial values	$a(14h)$	0.2811824057660078
	$m(14h)$	1.2673713692691995
	$Y(14h)$	1.4180265001692174
Parameters	α_1	9.668
	α_2	6.566
	β	0.163
	δ	$1.2865 \cdot \beta \approx 0.210$
	K	5.162
	r	1.984
	T_m	0.500

Table 4.8: Initial values for the variables and parameter values used for fig. 4.14 with eq. (4.23).

autophagy, then stabilize while ethanol is available, and later increase their autophagic activity again when ethanol is depleted. However, one drawback of this solution is that all curves stabilize at a level higher than 0. Especially $Y(t)$ stabilizes at its capacity K and not at $\frac{m}{T_m}$, like we would like to have when no nutrients are available. We rather wanted the solutions to tend to 0 instead, as the cells should eventually die when no nutrients are available. On the other hand, reading in article [15], one can see in figure 8 B that the cell viability hardly changes the five days after entry into the ethanol-depleted phase. Thus, the levels could be stable for at least 120 more hours. One reason for this is probably the entry of the cells into the so-called stationary phase (for details on the stationary phase in yeast see e.g. [33]). This could suggest, that one needs an additional part in the model taking care of this effect.

Now, with $(Y(t) - m(t))$ instead of just $m(t)$ in our equation for $\dot{m}$, we got an interesting solution but no better fit to our data. Hence, we tried to finetune this approach. Therefore, we added a parameter T_y to our $\dot{m}$-equation, which should scale the value of $Y(t)$. The resulting equation looks like follows:

$$\dot{m}(t) = \underbrace{\beta\big(T_y Y(t) - m(t)\big)a(t)}_{\text{increase due to autophagy}} - \underbrace{\delta m(t)}_{\substack{\text{(normal)}\\ \text{degradation}}} \tag{4.24}$$

where $T_y \in (0, 1]$. With $T_y = 1$, we simply arrive at eq. (4.23) again. Now the stationary state of m changes to $m^* = T_y \frac{\beta a Y}{\beta a + \delta}$. Hence, we can scale the stationary state to a value smaller than K, also for small values of δ, just by choosing $T_y < 1$.
Looking at it from the biological side, we now assume that the maximal level of material gets lower when the cells are starving. Breaking it down to a single cell again, we assume that the level of material within this cell cannot exceed $T_y < 1$ anymore.
For the fitting procedure, we fixed $T_m = 0.672$ as computed in section 4.4.4. The whole list of parameter values can be found in table 4.9. This resulted in a good fit to the data points, like shown in fig. 4.15 on the left. Also the levels of a and m per cell, in fig. 4.15 on the right, show a course that fits our assumptions. The level of autophagy first increases quite quickly, then almost stabilizes, before increasing again. Only the second increase happens before ethanol is depleted, so a bit too early. At the beginning the level of autophagy per cell also decreases briefly, because of the steep increase of Y at this time. The level of material per cell first decreases quickly, but then stabilizes at lower levels, increasing slightly again when the level of autophagy increases.

Next, we again applied a sensitivity analysis. As a side note, we had this time no problems with our chosen bounds ($\pm 5\%$) and did not need to adapt them. The sensitivity analysis yielded here a quite interesting result. As it can be seen for the Sobol method in appendix 3.6, fig. 11, the only parameters with significantly high effects are T_m and T_y. Especially, when looking only at the higher order influences $S_T - S_1$, they still show values over 0.3. Because of this, we decided to let `SALib` additionally compute the S_2 sensitivity coefficients this time. This resulted in an S_2 value of around 0.43 for their interaction regarding $\dot{Y}$, 0.39 for $\dot{a}$ and 0.37 for $\dot{m}$. Thus, we have a strong interaction of T_m and T_y. The Morris analysis, which can be found in appendix 3.6, fig. 12, likewise suggests a high influence and interaction of T_m and T_y. The influences of all other parameters seem here negligible. Additionally, we see the sign of the elementary effects. Since μ is negative for T_y, an increase of T_y seems to decrease the RSS for each

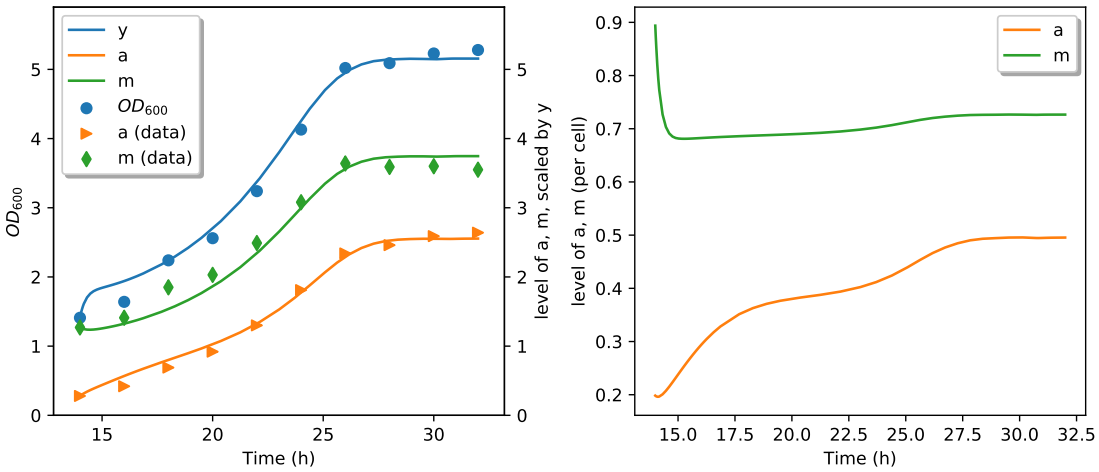

Figure 4.15: Left: Solving eq. (4.4) with our modified equation for $\dot{m}$, eq. (4.24), using the parameter values given in table 4.9, Right: Levels of a and m per cell.

	Name	Value
Initial values	$a(14h)$	0.2811824057660078
	$m(14h)$	1.2673713692691995
	$Y(14h)$	1.4180265001692174
Parameters	α_1	0.9420682080843957
	α_2	1.7377956657273594
	β	2.1860515775730747
	δ	0.024898361992930614
	K	5.155822859681016
	r	4.530153836535905
	T_m	0.672
	T_y	0.7298085885763022

Table 4.9: Initial values for the variables and parameter values used for fig. 4.15 with eq. (4.24).

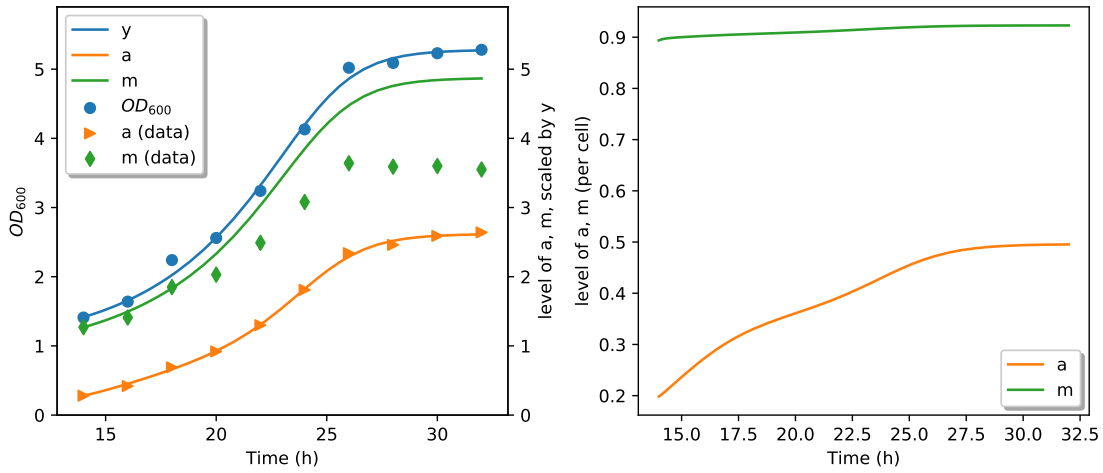

Figure 4.16: Left: Solving eq. (4.4) with our modified equation for $\dot{m}$, eq. (4.24), using the parameter values given in table 4.10, Right: Levels of a and m per cell.

equation. The sign of μ for T_m, on the other hand, is positive, suggesting that a decrease of T_m also decreases the RSS. With this information we tried to fit the values for T_m and T_y again. All other parameters were fixed to their already found values, as they seem not to have a great influence within their bounds. For T_m and T_y we chose the lower half, respectively upper half, within the bounds given in appendix 3.6, table 9. This led to only slightly smaller RSS values than before, hence we will not change our already found parameter set.

The steep increase of Y at the beginning comes from the term $m - T_m Y$ in the equation for $\dot{Y}$. Since we have no biological foundation for our chosen data for $m(t)$, and thus the value of T_m and T_y, we decided to try another fit without the data for m. This can simply be done in `symfit` by setting the data for m to `None`. But to ensure an increase of m at the beginning we took nevertheless an approximation of $\dot{m}(14)$ with the data points from table 1 and table 3: $\dot{m}(14) \approx \frac{m(16)-m(12)}{16-12} = \frac{1.41-1.071}{4} \approx 0.085$. Next we used this approximation to determine T_y dependent on the other parameters, resulting in $0.085 \approx \beta(T_y Y - m)a - \delta m \Leftrightarrow T_y \approx 0.215\frac{1}{\beta} + 3.217\frac{\delta}{\beta} + 0.901$, were the values for $Y(14)$, $a(14)$ and $m(14)$ are taken from table 3. In this way we have one parameter less to fit. The result can be seen in fig. 4.16 and the corresponding parameter values in table 4.10. Like intended, the steep increase of Y at the beginning is gone because of the higher value for T_m. The solution for m is similar but higher than before. Moreover, the steep decrease of the level of m per cell is missing. To evaluate whether this is more reasonable or not, we would need real data.

In both of the above cases $Y(t)$ converges to K instead of $\frac{m}{T_m}$, like we would rather want for the model.

	Name	Value
Initial values	$a(14h)$	0.2811824057660078
	$m(14h)$	1.2673713692691995
	$Y(14h)$	1.4180265001692174
Parameters	α_1	1.4587690241816567
	α_2	9.552509450128724
	β	9.223730441798079
	δ	0.0033533819911944897
	K	5.28029177673666
	r	2.4944097900751787
	T_m	0.870949265203856
	T_y	$0.215\frac{1}{\beta} + 3.217\frac{\delta}{\beta} + 0.901 = 0.9231398675322214$

Table 4.10: Initial values for the variables and parameter values used for fig. 4.16 with eq. (4.24).

4.6.3 Second Modification of the Equation for Material

Another modification we tried was just changing $m(t)$ to $(Y(t) - m(t))$ in the equation for $\dot{m}$, thus resulting in

$$\dot{m} = \beta\Big(Y(t) - m(t)\Big)\Big(a(t) - \tfrac{\delta}{\beta}Y(t)\Big) \tag{4.25}$$

From a biological perspective, we now assume the level of material to increase as long as the level of autophagy exceeds a certain level. But the increase is extenuated the higher the level of material gets.
Although this modification would not change the rapid increase of $a(t)$ needed at the beginning, we wanted to see whether this had a positive effect on the resulting fit. Like already assumed, it did not really improve the fit, but we had less problems with unbounded solutions. The corresponding plots and fitted parameter values can be found in appendix 3.4. One big drawback is, that the fitted curves for $Y(t)$ and $a(t)$ converge very quickly to 0 and for $m(t)$ even to a negative value, as visible in appendix 3.5, fig. 7. If $m(t)$ gets negative it is not biologically meaningful anymore. Also the rapid convergence of $Y(t)$ and $a(t)$ to 0 does not fit the information about cell viability given in [15].
The plotted results of the sensitivity analysis can be found in appendix 3.6. In both figures, fig. 13 and fig. 14, we can find three groups of parameters: r and K with very small sensitivity coefficients, T_m again with the highest values and all other parameters with similar coefficients in between. This time, all higher order effects $S_T - S_1$ are comparatively high, except for r and K. Hence, we examined the second order sensitivity coefficients S_2. However, the values all lay between -0.09 and 0.008. Thus, the high total effects must be caused by third or higher order effects, which we cannot examine in this way.

(Un)fortunately, on time we got a typo in the code and wrote $(1 - m)$ instead. Surprisingly, this gave an almost perfect fit. As m is always bigger than 1 for our data in

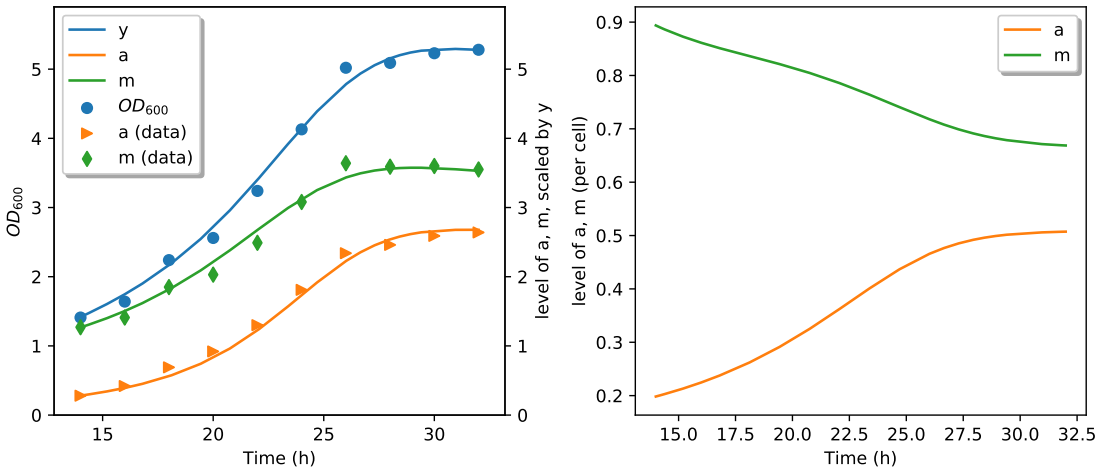

Figure 4.17: Left: Solving eq. (4.4) with the modified equation for $\dot{m}$, eq. (4.26), using the parameter values given in table 4.11, Right:Levels of a and m per cell.

	Name	Value
Initial values	$a(14h)$	0.2811824057660078
	$m(14h)$	1.2673713692691995
	$Y(14h)$	1.4180265001692174
Parameters	α_1	0.1741920909034433
	α_2	0.26027862251141193
	β	0.18705182750308452
	δ	0.0935790357918391
	K	14.93470846029889
	r	0.3610402147878471
	T_m	0.672

Table 4.11: Initial values for the variables and parameter values used for fig. 4.17 with eq. (4.26).

the starvation case, $(1-m)$ is just equivalent to a change of sign of the whole equation. Hence, we tried instead to fit the following equation for $\dot{m}$:

$$\dot{m} = \beta m(t)\left(\tfrac{\delta}{\beta}Y(t) - a(t)\right) \tag{4.26}$$

This resulted in the solution curves shown in fig. 4.17 and parameter values given in table 4.11. With this model the curves also seem to tend to 0 or at least to very low values. The corresponding plot until 10000 h can be seen in appendix 3.5.

However, when looking at the meaning of the terms, this does not match our original assumptions. In this equation, the increase of m is independent of a. Instead, it just depends on the current number of cells and level of material. Autophagy only decreases the level of material here. Thus, as $\dot{Y}$ does not directly depend on a, the level of autophagy has a negative effect on the level of material and cell number instead of a

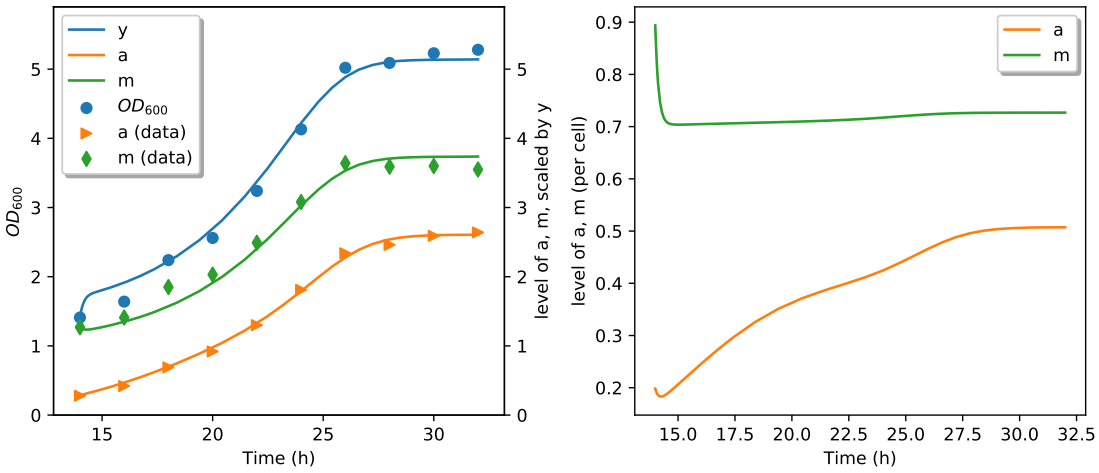

Figure 4.18: Left: Solving eq. (4.4) with the modified equation for $\dot{m}$, eq. (4.27), using the parameter values given in table 4.12, Right: Levels of a and m per cell.

positive one, as we wanted. That the model fits our data so well could be a side effect from deriving the data from our single cell model.

4.6.4 Third Modification of the Equation for Material

Instead of having the sign of $\dot{m}$ dependent on a difference of a and Y, which already made problems above, we tried an equation where it is dependent on a difference of m and Y instead. This would also help us staying with a bounded solution. This new equation for $\dot{m}$ looks like follows:

$$\dot{m}(t) = \beta a(t)\big(Y(t) - \tfrac{\delta}{\beta} m(t)\big) = \underbrace{\beta a(t) Y(t)}_{\substack{\text{increase due to} \\ \text{autophagy}}} - \underbrace{\delta a(t) m(t)}_{\substack{\text{degradation due to} \\ \text{autophagy}}} \tag{4.27}$$

The first term describes the increase of material due to autophagy and the number of cells. We did not multiply a term including m here, as the level of autophagy should decrease when the level of material gets too low anyway. The second term is the decrease due to autophagy, which is small when there is only little material left. Fitting our model with this equation for $\dot{m}$ in `symfit`, gave us the solution curves shown in fig. 4.18. The corresponding parameter values are given in table 4.12. This looks similar to our fitted curves in fig. 4.15. Only when considering the levels per cell, one can see a more uniform increase of autophagy here.
The sensitivity analysis showed a high influence of T_m, β and δ, whereas the effects of the other parameters all seem negligible. The corresponding plots can be found in appendix 3.6, fig. 15 and fig. 16. Considering only the higher order effects $S_T - S_1$ in fig. 15, we see that they still have values between 0.25 and 0.5. Thus, we let `SALib` compute the S_2 coefficients. For the interaction of T_m with β we got values lower 0.08 for each equation, hence negligible. Also the interaction of T_m with δ showed no higher values. However, the S_2 coefficients for β and δ were 0.1 for $\dot{Y}$, 0.23 for $\dot{a}$ and 0.12 for $\dot{m}$. This suggests an interaction between β and δ. The remaining parts of the total

	Name	Value
Initial values	$a(14h)$	0.2811824057660078
	$m(14h)$	1.2673713692691995
	$Y(14h)$	1.4180265001692174
Parameters	α_1	0.5157881010810613
	α_2	0.9293444104759587
	β	3.208759002703862
	δ	4.414748255442791
	K	5.143422883668531
	r	6.607682705570198
	T_m	0.69679544430603

Table 4.12: Initial values for the variables and parameter values used for fig. 4.18 with eq. (4.27).

effects then have to come from third or higher order effects.
Again, $Y(t)$ converges to K instead of $\frac{m}{T_m}$ here.

5 Discussion

In the first part of this thesis we proposed two single-cell models describing autophagy in yeast cells. Therefore, we made use of publications like [1] and [2] to understand the underlying principles and decide on the variables to use. Next, we analysed our models mathematically including the existence and non-negativity of unique solutions, their boundedness and domain, and stationary states together with their stability. We tried to find constraints on the parameters of each model to ensure desired properties like boundedness and the stability of certain stationary states. For the first model we could also conduct a bifurcation analysis, showing critical values where stationary states change their stability characteristics.

In a second step, we adapted the first single-cell model to a multi-cell setting. We analysed the resulting model in the same way as done before, trying to find meaningful constraints to ensure boundedness and stability of certain stationary states.

After this, we tried fitting the model to data we found in an article of Ryo Iwama and Yoshinori Ohsumi [15]. One drawback was, that we could only use the data on the cell density from this article, since the other measurements described quantities we did not incorporate into our model. Thus, we needed to make up test data to be able to fit our model meaningfully. Furthermore, the authors also wanted to study a second phenomenon, the diauxic growth of yeast. Our model was not designed to take this phenomenon into account, hence this could additionally be an issue when trying to find suitable parameter values. The fitting procedure was then quite complex. A major difficulty was here, that we had no prior information on the parameter values. The variables we chose were not used in other models we could find. In fact, we only found a few publications modeling autophagy in mammalian cells via molecular interactions and one publication with, probably, another ODE model about autophagy, which was not accessible. Thus, we could not take parameter values from similar approaches to get a starting point for our parameter search. Instead, we had to search a wide range of values, which was not easily possible. Moreover, good starting values are often crucial for an effective parameter fitting. We tried different means of fixing parameters, to speed up the fitting process and be able to try wider ranges for the remaining parameters in an acceptable time. Eventually, we found a parameter set, which gave a good fit (see fig. 4.8).

Especially the resulting levels of a and m per cell, however, did not match the results in [15] or our assumptions. Thus, we tried a few modifications of our original model in the starvation case to improve the fit of the solution curves. In the following, we now want to compare these different model versions. We do not consider the nutrient-rich case here, as this part was far more easy to fit and already gave very good results with our basic model. Furthermore, we omit the first version of our first modification of $\dot{m}$, since we could not get a good fit there. Also the "typo"-version of the second modification is left out, because we have no biologically meaningful explanation for it.
Since we always did a least squares fit with `symfit`, thus minimizing the RSS, we wanted

Model	RSS	AIC_c
Basic model	RSS(a) = 2.56233 RSS(m) = 0.35078 RSS(Y) = 0.31339 Sum: 3.2265	-44.03683984
Gen. logistic growth	RSS(a) = 2.43485 RSS(m) = 0.18120 RSS(Y) = 0.15248 Sum: 2.76853	-44.48642665
First modification of $\dot{m}$, adapted	RSS(a) = 0.06332 RSS(m) = 0.24739 RSS(Y) = 0.17256 Sum: 0.48327	-96.85131471
Second modification of $\dot{m}$	RSS(a) = 1.35396 RSS(m) = 0.86682 RSS(Y) = 0.59053 Sum: 2.81131	-48.16926159
Third modification of $\dot{m}$	RSS(a) = 0.01825 RSS(m) = 0.18566 RSS(Y) = 0.16345 Sum: 0.36736	-109.22116815

Table 5.1: RSS values and AIC_c for the basic model and each modification.

to compare this value for all of our models here. Additionally, we chose to compute the AIC_c, to have a second criterion also taking the number of parameters into account.
The AIC_c is a second order version of the AIC. It is bias adjusted for models, where too less samples are given compared to the number of parameters. K. Burnham and D. Anderson recommend in their book to use the AIC_c if the ratio of samples to the number of parameters is less than 40 [34]. This is the case for our models, as shown in appendix 6. There we also give more details on the computation of the AIC_c. We decided on this criterion, because it is one of the widely used and well studied model selection criteria. Another criterion, which we first considered, is the Bayesian Information Criterion (BIC). But the underlying assumption for this criterion is, that the true model is in the set of examined models, which we do not think is true in this case. For more details on the AIC, AIC_c, BIC and their comparison, see for example [34].
In table 5.1, the computed RSS values (rounded) are given for each model, together with the AIC_c. We divided the RSS into values for each single solution curve, as often only one curve shows a particularly high score. Furthermore, since our data for Y is the only measured data, we were particularly interested in the RSS value for this curve. The AIC_c was then computed with the sum of all RSS scores. Splitting also the AIC_c into three different values would not be meaningful, because also the number of parameters is taken into account. But all of our equations are dependent on each other, hence each of them is indirectly dependent on every parameter in the model.
Taking first only RSS(Y) into account, the model with generalized logistic growth fits best. This was to be expected, since this modification aimed only at a better fit for

$Y(t)$. But also the first and third modification give a good fit considering the RSS(Y). One drawback of the generalized logistic growth model is the high RSS value for $a(t)$, which also results in a high sum for the RSS.
When we consider the sum of the RSS values for each model, the third modification is the best, followed by the first modification. The RSS values for all other models are at least 5 times higher, mostly due to high RSS scores of $a(t)$. This finding also corresponds to the AIC_c scores, where the model with the lowest score should be preferred.
Following this results, one could now try in the starvation case to combine a generalized logistic growth model for $Y(t)$ with the first or third modification of the $\dot{m}$-equation.

Last, we also want to take a look at the parameter values for K in each model and the results of the sensitivity analysis. When we set up our equation for $\dot{Y}$ in the starvation case, the intention was that the cell growth should be limited by the amount of available material instead of the capacity K. Our assumption here was, that the cell batch has so much space, that the lack of nutrients hinders the cell growth before the batch reaches its maximal capacity. Otherwise we would get a mixture of the effects from the lack of nutrients and the limited space. Hence, our capacity K should be larger than 5.3, which is approximately the value where the data for Y tends to. The only fitted models satisfying this requirement are the original model, the generalized logistic growth version and the model with the second modification for $\dot{m}$. Hence, these models with their parameter values reflect our above assumptions best. Although we want to mention again, that the long term behaviour for the second modification is not biologically meaningful, since $m(t)$ converges to a negative value. In this case one has to decide, whether the solutions are only important in the time span, where measured data is given, or also in longer time periods. On the other hand, we already mentioned in section 4.6.2 that the cell viability according to [15] hardly changes for at least 120 hours after the cells enter the ethanol-depleted phase. All three models we just mentioned, however, already decrease clearly at the end of our regarded time interval. Thus, one could also interpret our $\dot{Y}$-equation in another way. Assuming that our model is only valid as long as the cells stay stable in the stationary phase, K could be a starvation dependent maximal capacity of the cell batch and the term $(m - T_m Y)$ only influences how fast the batch reaches this capacity. Then we would like the solution curves to stabilize after 32 h and presume, that we need another model when the cell batch begins to diminish due to the lack of nutrients. In this way the models with the first or third modification for $\dot{m}$ would be best, like chosen by the AIC_c.
Regarding the sensitivity analysis, we first want to note again, that this analysis strongly depends on the chosen parameter bounds. Starting at different parameter values with other bounds could change the results considerably. With this in mind, we want to shortly discuss our main results. For the basic model, we found that T_m has the highest influence on the resulting RSS. But also the other parameters, except r and K, had a discernible influence. Thus, one could adjust the solution curves by tuning multiple parameter values. This has the advantage, that we get many possibilities for different solution curves. On the other hand, having many parameters to fit is more demanding. Especially, when only small changes of some parameter values lead to unbounded or unstable solutions, like it was in our case. Then one would rather like to fix the most influential parameters, in order to be able to fit the remaining ones. We tried this by fixing T_m and setting α_2 and β at least dependent on other parameters. But this did not work, probably because the chosen value for T_m was not close enough a value allowing

a good fit. Setting only α_2 and β, dependent on α_1 and δ, then worked. The sensitivity analysis for the generalized logistic growth version showed a similar result, except that also n did not have a notable influence. For the model with our first modification of $\dot{m}$, however, the results were quite interesting. Here, only T_m and T_y show non-negligible influences. Particularly, they also showed a strong interaction. This could explain, why it was comparatively easy to fit this model when we fixed the value for T_m. Once we had suitable parameter ranges, all parameters except T_y would not have a great influence anyway, hence the fitting procedure must only find a good value for T_y. Of course, finding a good value for T_m and suitable ranges for the other parameters may not be easy, but either these ranges are very wide in our case or we were just lucky to find them so quickly. The model with our second modification for $\dot{m}$ showed again a different result. Here we had quite clearly three groups of parameters with different levels of influence. The greatest influence had again T_m. Then came α_1, α_2, β and γ, all with a similar value. Last we had r and K with only negligible influences. The higher order influences of all other parameters were still notable, so we examined the second order interactions between the parameters. These were all around 0, so we deduced that they have to come from third or even higher order interactions. This could be an interesting starting point for further analysis of this model. Especially the parameters, with which T_m has higher order interactions, could help in making the fitting of the model easier. We also find it worth noticing, that in this case both parameters included in the $\dot{a}$-equation seem to have no influence. The model with our third and last modification for $\dot{m}$ showed once again another result. In both sensitivity analysis, T_m, β and δ had a similar influence, while all other parameters were negligible. Thus, one could first try to fix these three parameter values if the model is fitted again, in order to speed up the procedure. Again, the parameters determining $\dot{a}$ seem to have no influence.

Certainly, there are many more possible modifications to adapt the model further. A next step for example could be, to use another function for $s(t)$. We only used here a step function with one jump from 0 to 1 at 14 h. But one could also try to use a continuous function. We already played around with different versions for $s(t)$ a bit, but did not get a good result yet. For example, we tried to simply approximate our previous step function with a sigmoid function of the form $\frac{1}{1+e^{-x+14}}$. But when the change from 0 to 1 was not quickly enough, we could not easily find a good parameter set.
When we searched for more background about the diauxic growth, we found an interesting article presenting an ensemble of models describing diauxic growth [35]. Since the effects of diauxic growth are explicitly included in the data we used from [15], it could also be worth trying to construct a model from this direction. Starting with a model for diauxic growth, one could try to incorporate the process of autophagy in some way and get by that an even more realistic model.

List of Figures

List of Tables

Bibliography

[1] Yuchen Feng et al. "The machinery of macroautophagy". In: *Cell Research* 24 (2014), pp. 24–41.

[2] Noboru Mizushima. "Autophagy: process and function." In: *Genes & development* 21 22 (2007), pp. 2861–73.

[3] Raffaela Torggler, Daniel Papinski, and Claudine Kraft. "Assays to Monitor Autophagy in Saccharomyces cerevisiae". In: *Cells* 6 (3 2017). ISSN: 2073-4409. DOI: 10.3390/cells6030023. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5617969/.

[4] Xiao Liang et al. "Induction of autophagy and inhibition of tumorigenesis by Beclin-1". In: *Nature* 402 (Dec. 1999), pp. 672–6. DOI: 10.1038/45257.

[5] Brinda Ravikumar, Rainer Duden, and David C. Rubinsztein. "Aggregate-prone proteins with polyglutamine and polyalanine expansions are degraded by autophagy". In: *Human Molecular Genetics* 11.9 (May 2002), pp. 1107–1117. ISSN: 0964-6906. DOI: 10.1093/hmg/11.9.1107. eprint: https://academic.oup.com/hmg/article-pdf/11/9/1107/9465005/ddf115.pdf. URL: https://doi.org/10.1093/hmg/11.9.1107.

[6] Melinda Lynch-Day and Daniel Klionsky. "The CVT pathway as a model for selective autophagy". In: *FEBS letters* 584 (Feb. 2010), pp. 1359–66. DOI: 10.1016/j.febslet.2010.02.013.

[7] Gianni Liti. "The Natural History of Model Organisms: The fascinating and secret wild life of the budding yeast *S. cerevisiae*". In: *eLife* 4 (Mar. 2015), e05835. ISSN: 2050-084X. DOI: 10.7554/eLife.05835. URL: https://doi.org/10.7554/eLife.05835.

[8] *S. cerevisiae under DIC microscopy*. 2010. URL: https://commons.wikimedia.org/w/index.php?title=File:S_cerevisiae_under_DIC_microscopy.jpg&oldid=373200746 (visited on 04/20/2020).

[9] A Goffeau et al. "Life with 6000 Genes". In: *Science (New York, N.Y.)* 274 (Nov. 1996), pp. 546, 563–7. DOI: 10.1126/science.274.5287.546.

[10] Leland H. Hartwell. "*Saccharomyces cerevisiae* cell cycle". In: *Bacteriological Reviews* 38 (2 June 1974), pp. 164–98. ISSN: 0005-3678.

[11] Amy Tong et al. "Global Mapping of the Yeast Genetic Interaction Network". In: *Science* 303 (Feb. 2004), p. 808. DOI: 10.1126/science.1091317.

[12] Aashiq Kachroo et al. "Evolution. Systematic humanization of yeast genes reveals conserved functions and genetic modularity". In: *Science (New York, N.Y.)* 348 (May 2015), pp. 921–5. DOI: 10.1126/science.aaa0769.

[13] Fulvio Reggiori and Daniel J. Klionsky. "Autophagy in the Eukaryotic Cell". In: *Eukaryotic Cell* 1.1 (2002), pp. 11–21. ISSN: 1535-9778. DOI: 10.1128/EC.01.1.11-21.2002. eprint: https://ec.asm.org/content/1/1/11.full.pdf. URL: https://ec.asm.org/content/1/1/11.

[14] Luciano Galdieri et al. "Transcriptional Regulation in Yeast during Diauxic Shift and Stationary Phase". In: *Omics : a journal of integrative biology* 14 (Dec. 2010), pp. 629–38. DOI: 10.1089/omi.2010.0069.

[15] Ryo Iwama and Yoshinori Ohsumi. "Analysis of autophagy activated during changes in carbon source availability in yeast cells". In: *Journal of Biological Chemistry* 294 (Feb. 2019), jbc.RA118.005698. DOI: 10.1074/jbc.RA118.005698.

[16] Ju Huang and John Brumell. "Bacteria-autophagy interplay: A battle for survival". In: *Nature reviews. Microbiology* 12 (Jan. 2014). DOI: 10.1038/nrmicro3160.

[17] Cheryl L. Birmingham et al. "Autophagy Controls Salmonella Infection in Response to Damage to the Salmonella-containing Vacuole". In: *Journal of Biological Chemistry* 281.16 (2006), pp. 11374–11383. DOI: 10.1074/jbc.M509157200. eprint: http://www.jbc.org/content/281/16/11374.full.pdf+html. URL: http://www.jbc.org/content/281/16/11374.abstract.

[18] Maximiliano Gutierrez et al. "Autophagy Is a Defense Mechanism Inhibiting BCG and Mycobacterium tuberculosis Survival in Infected Macrophages". In: *Cell* 119 (Jan. 2005), pp. 753–66. DOI: 10.1016/j.cell.2004.11.038.

[19] Henry Paulson. "Protein Fate in Neurodegenerative Proteinopathies: Polyglutamine Diseases Join the (Mis)Fold". In: *American journal of human genetics* 64 (Mar. 1999), pp. 339–45. DOI: 10.1086/302269.

[20] Ravi Agarwal and Donal O'Regan. "An Introduction to Ordinary Differential Equations". In: (Jan. 2008). DOI: 10.1007/978-0-387-71276-5.

[21] Tomás Caraballo and Xiaoying Han. *Applied Nonautonomous and Random Dynamical Systems: Applied Dynamical Systems*. Jan. 2016. ISBN: 978-3-319-49246-9. DOI: 10.1007/978-3-319-49247-6.

[22] Johannes Müller and Christina Kuttler. *Methods and Models in Mathematical Biology*. Jan. 2015. ISBN: 978-3-642-27250-9. DOI: 10.1007/978-3-642-27251-6.

[23] Rainer Sachs, L.R. Hlatky, and P. Hahnfeldt. "Simple ODE models of tumor growth and anti-angiogenic or radiation treatment". In: *Mathematical and Computer Modelling* 33 (June 2001), pp. 1297–1305. DOI: 10.1016/S0895-7177(00)00316-2.

[24] Hal L. Smith and Paul Waltman. *The Theory of the Chemostat: Dynamics of Microbial Competition*. Cambridge Studies in Mathematical Biology. Cambridge University Press, 1995. DOI: 10.1017/CBO9780511530043.

[25] Pauli Virtanen et al. "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". In: *Nature Methods* 17 (2020), pp. 261–272. DOI: https://doi.org/10.1038/s41592-019-0686-2.

[26] Ankit Rohatgi. *WebPlotDigitizer*. July 2020. URL: https://automeris.io/WebPlotDigitizer (visited on 06/04/2020).

[27] Martin Roelfs and Peter C Kroon. "tBuLi/symfit: symfit 0.5.2". In: (Nov. 2019). DOI: 10.5281/zenodo.3532953.

[28] George Qian and Adam Mahdi. "Sensitivity analysis methods in the biomedical sciences". In: *Mathematical Biosciences* 323 (Jan. 2020), p. 108306. DOI: 10.1016/j.mbs.2020.108306.

[29] Jon Herman and Will Usher. "SALib: An open-source Python library for Sensitivity Analysis". In: *The Journal of Open Source Software* 2 (Jan. 2017). DOI: 10.21105/joss.00097.

[30] Andrea Saltelli et al. "Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index". In: *Computer Physics Communications* 181 (Feb. 2010), pp. 259–270. DOI: 10.1016/j.cpc.2009.09.018.

[31] Max D. Morris. "Factorial Sampling Plans for Preliminary Computational Experiments". In: *Technometrics* 33.2 (1991), pp. 161–174. ISSN: 00401706. URL: http://www.jstor.org/stable/1269043.

[32] Francesca Campolongo, Jessica Cariboni, and Andrea Saltelli. "An effective screening design for sensitivity analysis of large models". In: *Environmental Modelling & Software* 22.10 (2007). Modelling, computer-assisted simulations, and mapping of dangerous phenomena for hazard assessment, pp. 1509–1518. ISSN: 1364-8152. DOI: https://doi.org/10.1016/j.envsoft.2006.10.004. URL: http://www.sciencedirect.com/science/article/pii/S1364815206002805.

[33] M Werner-Washburne et al. "Stationary phase in the yeast Saccharomyces cerevisiae." In: *Microbiology and Molecular Biology Reviews* 57.2 (1993), pp. 383–401. ISSN: 0146-0749. eprint: https://mmbr.asm.org/content/57/2/383.full.pdf. URL: https://mmbr.asm.org/content/57/2/383.

[34] Kenneth Burnham and David Anderson. "Model Selection and Multimodel Inference". In: *A Practical Information-theoretic Approach* (Jan. 2004), pp. 61–66. DOI: 10.1007/978-0-387-22456-5_5.

[35] Andreas Kremling et al. "An ensemble of mathematical models showing diauxic growth behaviour". In: *BMC Systems Biology* 12 (Dec. 2018). DOI: 10.1186/s12918-018-0604-8.

Appendix

1 Detailed Computations and Proofs

In this section we give the detailed computations or proofs of some results we only stated in the main part.

1.1 Solution and Symmetry of the Logistic Equation

First we want to show that our stated solution for $y(t)$ is the actual solution. The proof follows [22, pp. 176-177].
To solve $\dot{y}(t) = ay(t)\big(1 - \frac{y(t)}{K}\big)$, one can apply the so-called "trick of Riccati". This means, we define a new variable for which we can solve the resulting ODE more easily. In our case we take $v(t) = \frac{1}{y(t)}$, thus

$$\begin{aligned}\dot{v}(t) = -\frac{\dot{y}(t)}{y(t)^2} &= -\frac{ay(t)\big(1 - \frac{y(t)}{K}\big)}{y(t)^2} = -a\frac{y(t) - \frac{y(t)^2}{K}}{y(t)^2} \\ &= -a\left(\frac{1}{y(t)} + \frac{1}{K}\right) = -av(t) - \frac{a}{K}\end{aligned} \tag{1}$$

This ODE can now simply be solved by variation of constants. Without loss of generality we assume here $t_0 = 0$, hence $v_0 = \frac{1}{y_0}$.

$$\begin{aligned}v(t) &= v_0 e^{-A(t)} + \int_0^t e^{-(A(t)-A(s))}\tfrac{a}{K} ds \\ &= v_0 e^{-at} + \tfrac{a}{K} e^{-at}\int_0^t e^{as} ds \\ &= v_0 e^{-at} + \tfrac{a}{K} e^{-at}\tfrac{1}{a}(e^{at} - 1) \\ &= v_0 e^{-at} + \tfrac{1}{K}(1 - e^{-at})\end{aligned} \tag{2}$$

where we used $A(t) = \int_0^t a\, d\tau = at$ for the second equation, since a is constant.
Going now back to our original variable $y(t) = \frac{1}{v(t)}$, we arrive at

$$\begin{aligned}y(t) = \frac{1}{v(t)} &= \frac{1}{v_0 e^{-at} + \frac{1}{K}(1 - e^{-at})} \\ &= \frac{1}{\frac{1}{y_0} e^{-at} + \frac{1}{K}(1 - e^{-at})} \\ &= \frac{y_0 K}{K e^{-at} + y_0(1 - e^{-at})} \\ &= \frac{y_0 K}{e^{-at}(K - y_0) + y_0}\end{aligned} \tag{3}$$

Next we show that $y(t)$ is point symmetric around $\frac{K}{2}$.
That it should be $\frac{K}{2}$ can easily be seen by setting the second derivative of $y(t)$ to 0: $\ddot{y} = a\Big(1 - \frac{2y}{K}\Big) = 0 \Leftrightarrow 1 = \frac{2y}{K} \Leftrightarrow y = \frac{K}{2}$. Since we know that $\dot{y}(t) > 0$ for $y = \frac{K}{2}$, it is an inflection point. To actually show symmetry, we first need the value t_i with $y(t_i) = \frac{K}{2}$:

$$\begin{aligned} y(t) = \frac{K}{2} &\Leftrightarrow \frac{y_0 K}{e^{-at}(K - y_0) + y_0} = \frac{K}{2} \\ &\Leftrightarrow e^{-at}(K - y_0) + y_0 = 2y_0 \\ &\Leftrightarrow e^{-at} = \frac{y_0}{K - y_0} \\ &\Leftrightarrow t = -\frac{1}{a} \ln\left(\frac{y_0}{K - y_0}\right) \end{aligned} \tag{4}$$

Let $t > 0$. Then $y(t_i - t) + y(t_i + t) = K$ should hold:

$$\begin{aligned} y(t_i - t) + y(t_i + t) &= \frac{y_0 K}{e^{-a(t_i - t)}(K - y_0) + y_0} + \frac{y_0 K}{e^{-a(t_i + t)}(K - y_0) + y_0} \\ &= \frac{y_0 K}{e^{at} e^{\ln(\frac{y_0}{K - y_0})}(K - y_0) + y_0} + \frac{y_0 K}{e^{-at} e^{\ln(\frac{y_0}{K - y_0})}(K - y_0) + y_0} \\ &= \frac{y_0 K}{e^{at} \frac{y_0}{K - y_0}(K - y_0) + y_0} + \frac{y_0 K}{e^{-at} \frac{y_0}{K - y_0}(K - y_0) + y_0} \\ &= \frac{K}{e^{at} + 1} + \frac{K}{e^{-at} + 1} \\ &= K \frac{e^{at} + 1 + e^{-at} + 1}{(e^{at} + 1)(e^{-at} + 1)} \\ &= K \frac{e^{at} + 1 + e^{-at} + 1}{1 + e^{at} + e^{-at} + 1} \\ &= K \end{aligned} \tag{5}$$

Hence, we can deduce point symmetry for $y(t)$.

1.2 Stationary states 1

Inserting $m^* = \frac{\beta_1 c^*}{\beta_1 c^* + \beta_2 a^* + \delta_1}$ into $c^* = \frac{\gamma_1 a^* m^*}{\gamma_2(1 - m^*) + \delta_2}$ and using $A := \beta_2 a^* + \delta_1$:

$$\begin{aligned} c^* &= \frac{\gamma_1 a^* \frac{\beta_1 c^*}{\beta_1 c^* + A}}{\gamma_2\big(1 - \frac{\beta_1 c^*}{\beta_1 c^* + A}\big) + \delta_2} \\ &\Leftrightarrow c^* \left(\gamma_2 \frac{A}{\beta_1 c^* + A} + \delta_2\right) = \gamma_1 a^* \frac{\beta_1 c^*}{\beta_1 c^* + A} \\ &\Leftrightarrow c^* \gamma_2 A + c^* \delta_2 (\beta_1 c^* + A) = \gamma_1 a^* \beta_1 c^* \\ &\Leftrightarrow c^* = 0 \ \text{ or } \ \gamma_2 A + \delta_2(\beta_1 c^* + A) = \beta_1 \gamma_1 a^* \\ &\Leftrightarrow c^* = 0 \ \text{ or } \ \beta_1 \delta_2 c^* = \beta_1 \gamma_1 a^* - \gamma_2 A - \delta_2 A \\ &\Leftrightarrow c^* = 0 \ \text{ or } \ c^* = \frac{\beta_1 \gamma_1 a^* - A(\gamma_2 + \delta_2)}{\beta_1 \delta_2} \end{aligned} \tag{6}$$

1.3 Stationary states 2

Here we compute the stationary states given for the multi-cell model in the starvation case. We consider different cases, to make the computations more clear.

- We start by setting $Y^* = 0$. Then we have to possibilities for a^*:
  - if $a^* = 0$, then $\dot{m} = 0$, independent of the value of m. Hence, we get the stationary states $P_2 = (0, 0, m_s)$ with $m_s \in \mathbb{R}$.
  - if $a^* = Y^* - \frac{\alpha_2}{\alpha_1}(Y^* - m^*) = \frac{\alpha_2}{\alpha_1} m^*$, we have again two possibilities to achieve $\dot{m} = 0$:
    - if $m^* = 0$, we get $a^* = 0$ and thus the stationary state $P_0 = (0, 0, 0)$.
    - if $a^* = \frac{\delta}{\beta} Y^* = 0$, we set both equations for a^* equal and arrive at $m^* = 0$. Hence, we get again P_0.
- Next we set $Y^* = K$. Again, we have two possibilities for a^*:
  - if $a^* = 0$, we can get
    - $P_1 = (K, 0, 0)$, when $m^* = 0$, or
    - no valid stationary state, as $0 = a^* = \frac{\delta}{\beta} K \neq 0$ is not possible.
  - if $a^* = K - \frac{\alpha_2}{\alpha_1}(K - m^*)$
    - and $m^* = 0$, we get $a^* = K - \frac{\alpha_2}{\alpha_1} K = K\big(1 - \frac{\alpha_2}{\alpha_1}\big)$. Thus, we arrive at $P_3 = \big(K, K\big(1 - \frac{\alpha_2}{\alpha_1}\big), 0\big)$.
    - if $a^* = \frac{\delta}{\beta} K$, we can set both equations for a^* equal and solve for m^*. So we get $m^* = K\big(1 - \frac{\alpha_1}{\alpha_2}\big(1 - \frac{\delta}{\beta}\big)\big)$ and arrive at $P_4 = \big(K, K\frac{\delta}{\beta}, K\big(1 - \frac{\alpha_1}{\alpha_2}\big(1 - \frac{\delta}{\beta}\big)\big)\big)$.
- Last we can set $Y^* = \frac{m^*}{T_m}$:
  - if $a^* = 0$, we look at the two possibilities such that $\dot{m} = 0$:
    - if $m^* = 0$, also $Y^* = 0$ and we arrive at P_0.
    - if $a^* = \frac{\delta}{\beta}\frac{m^*}{T_m}$, we get $m^* = 0$ by setting both equations for a^* equal. Thus, we get again P_0.
  - if $a^* = \frac{m^*}{T_m} - \frac{\alpha_2}{\alpha_1}(\frac{m^*}{T_m} - m^*) = \frac{m^*}{T_m}(1 - \frac{\alpha_2}{\alpha_1}(1 - T_m))$:
    - if $m^* = 0$, it follows that $a^* = 0$ and $Y^* = 0$. Hence, we get again P_0.
    - if $a^* = \frac{\delta}{\beta}\frac{m^*}{T_m}$, we set both equations for a^* equal and arrive at $1 - \frac{\alpha_2}{\alpha_1}(1 - T_m) = \frac{\delta}{\beta}$. Hence, the value of m^* does not matter, but we would need this constraint on the parameters. As we look at a biological model, it is very unlikely that exactly this constraint holds, except we want it to. Therefore, we can assume this stationary state to be negligible.

2 Data Tables

In this section we give the data tables used to fit our models. The first test data can be found in table 1. The single cell data used for the second test data is given in table 2 and the resulting scaled multi-cell data in table 3.

t	3	6	9	10	11	12	14	16	18	20	22	24	26	28	30	32
Y	0.023	0.09	0.36	0.59	0.98	1.19	1.41	1.64	2.24	2.56	3.24	4.13	5.02	5.09	5.23	5.28
a	0.0046	0.018	0.072	0.118	0.196	0.238	0.282	0.41	0.88	1.28	2.11	2.90	3.50	3.50	3.40	3.20
m	0.0207	0.081	0.324	0.531	0.882	1.071	1.27	1.48	1.98	2.18	2.55	2.90	3.20	3.10	2.90	2.64

Table 1: First test data.

t	14	16	18	20	22	24	26	28	30	32
a	0.2	0.25	0.31	0.36	0.4	0.44	0.47	0.48	0.49	0.5
m	0.9	0.86	0.82	0.79	0.77	0.75	0.72	0.71	0.69	0.67

Table 2: Data for 14 - 32 h derived by solving eq. (3.1) numerically with python, using $a_0 = 0.2$, $m_0 = 0.9$, $\alpha_1 = 0.2$, $\alpha_2 = 0.3$, $\beta = 0.04$ and $\delta = 0.032$.

t	14	16	18	20	22	24	26	28	30	32
Y	1.41	1.64	2.24	2.56	3.24	4.13	5.02	5.09	5.23	5.28
a	0.28	0.42	0.69	0.92	1.3	1.81	2.34	2.46	2.59	2.64
m	1.27	1.41	1.85	2.03	2.49	3.08	3.64	3.59	3.60	3.55

Table 3: Data from table 2 multiplied with the given optical cell densities from [15].

3 Further plots

In this section we show additional plots, which can be helpful but are not necessary for the understanding in the main part.

3.1 Second Example for boundedness

In this example, we chose m_0 to be 0.95, instead of 0.9 like in fig. 3.2. Then solutions are already unbounded for $a_0 \geq 0.442$, for example, which can be seen in fig. 1. This is noticeable lower than before, where solutions stayed bounded for $a_0 \leq 0.6288$, at least.

3.2 Initial Points for the Generalized Log. Growth

Like we did for the logistic growth model for $Y(t)$, we tested the first four data points as initial value for the fitting of the generalized version. We also added the fitted logistic curves to the plots in fig. 2, to instantly see differences.

3.3 Fitted Curves for the Generalized Log. Growth Model

The plots shown in fig. 3 and fig. 4 were derived the same way as fig. 4.7 and fig. 4.8. We had to adapt the value for $m(3)$ a bit, as otherwise the level per cell would have exceeded 1. We then also adapted the value of $a(3)$, to get a smoother result. Again, some of the parameter values do not show the whole 16 decimal points of the python output. We adapted them by hand to get a slightly better result. To compare different solutions we used the RSS with our data points.

3.4 Fitted Curves for the Second Modification

Changing m to $Y - m$ in the equation for $\dot{m}$ did not really improve the fit, regarding the RSS, but it helped staying with bounded solution curve. The resulting fit of eq. (4.4) with our modification eq. (4.25) can be found in fig. 5.

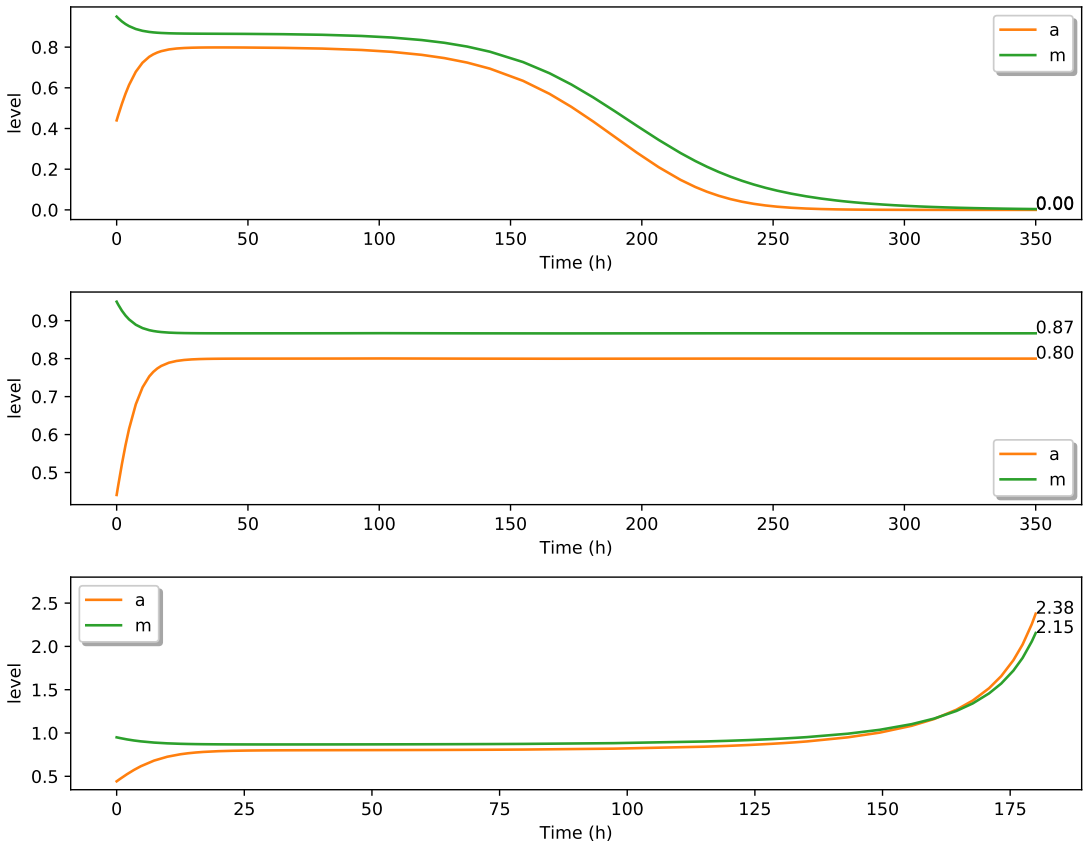

Figure 1: For all the plots we used the same parameter values as for fig. 4.5. The initial value of m is always set to 0.95, whereas the initial value of a changes from top to bottom: first $a_0 = 0.44$, then 0.44089 and last 0.442. We also added the values of a and m at the last time point.

	Name	Value
Initial values	$a(3h)$	$0.18 \cdot y(3h) \approx 0.0022$
	$m(3h)$	$0.86 \cdot y(3h) \approx 0.0104$
	$y(3h)$	0.01203952
Parameters	α	1.0
	β	0.029
	γ	5.0
	δ	0.48
	K	1.413931678779375
	M	$0.9 \cdot K(\delta - \frac{\beta}{\gamma}) \approx 0.6034$
	n	2.429362977019412
	r	1.3819066797483224

Table 4: Initial values for the variables and parameter values used for fig. 3.

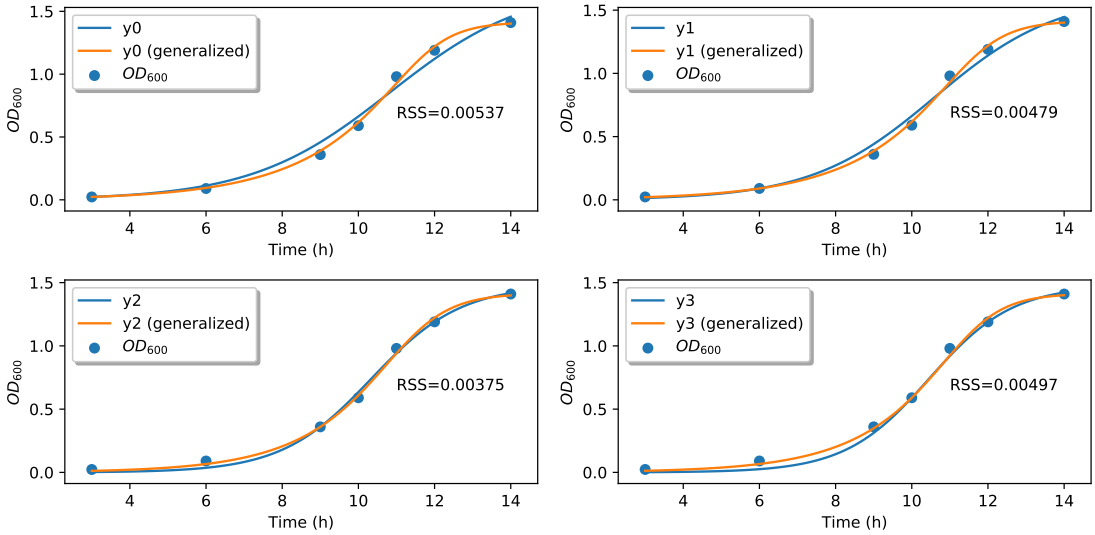

Figure 2: Results of fitting a generalized logistic growth model until 14 h with `symfit`, using different points as initial values. y0 (generalized) corresponds to the solution when using the first point at 3 h as initial value, y1 (generalized) when using the point at 6 h, y2 (generalized) at 9 h and y3 (generalized) at 10 h. The corresponding RSS are also added to the different plots. Additionally, the fits of the logistic model are plotted (see also fig. 4.6).

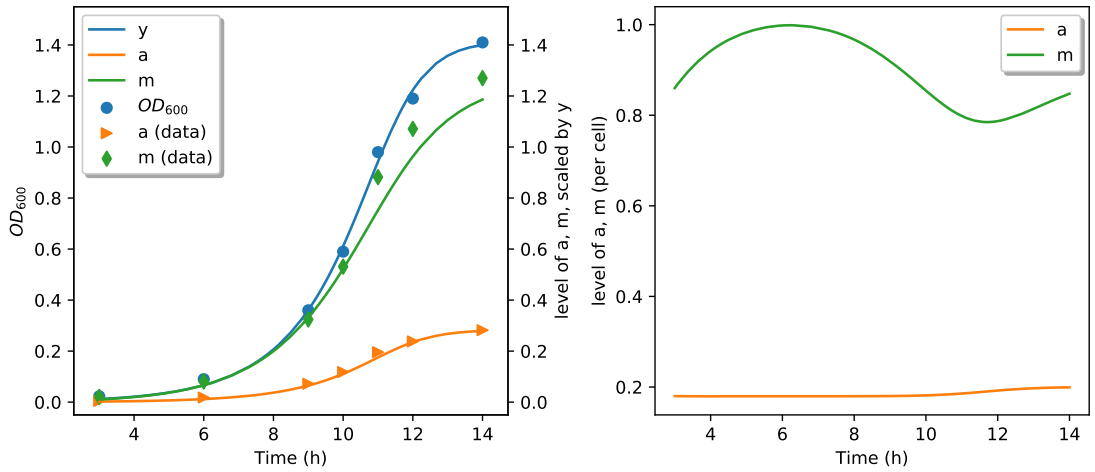

Figure 3: Left: Solving eq. (4.3), adapted with a generalized logistic growth, with the initial values and parameter values given in table 4, Right: Levels of a and m per cell.

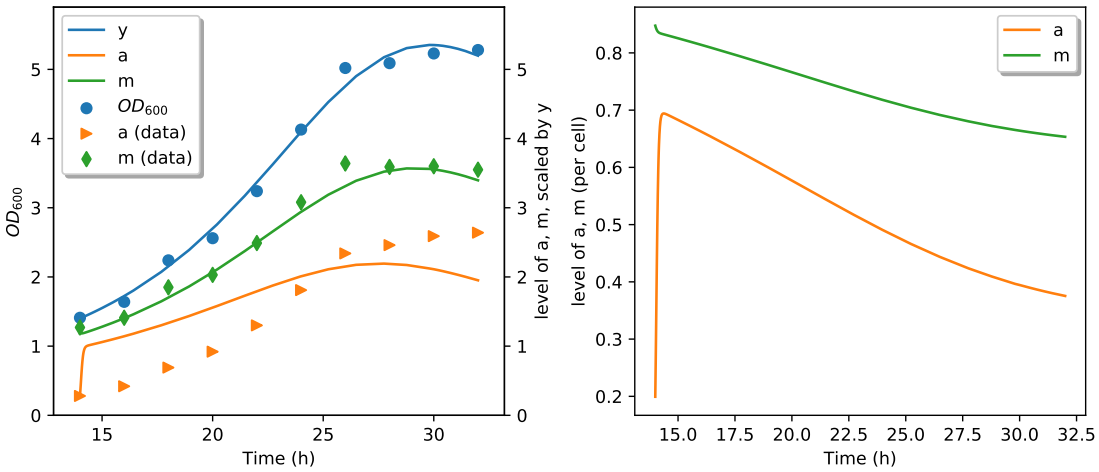

Figure 4: Left: Solving eq. (4.4), adapted with a generalized logistic growth, with the initial values and parameter values given in table 5, Right: Levels of a and m per cell.

	Name	Value
Initial values	$a(14h)$	0.2788898680560751
	$m(14h)$	1.186228590939174
	$Y(14h)$	1.3995981480119983
Parameters	α_1	19.94995054
	α_2	35.97842182
	β	0.21027698
	δ	0.0851718
	K	57.46757584
	n	1.94896191
	r	0.81570673
	T_m	0.66521046

Table 5: Initial values for the variables and parameter values used for fig. 4.

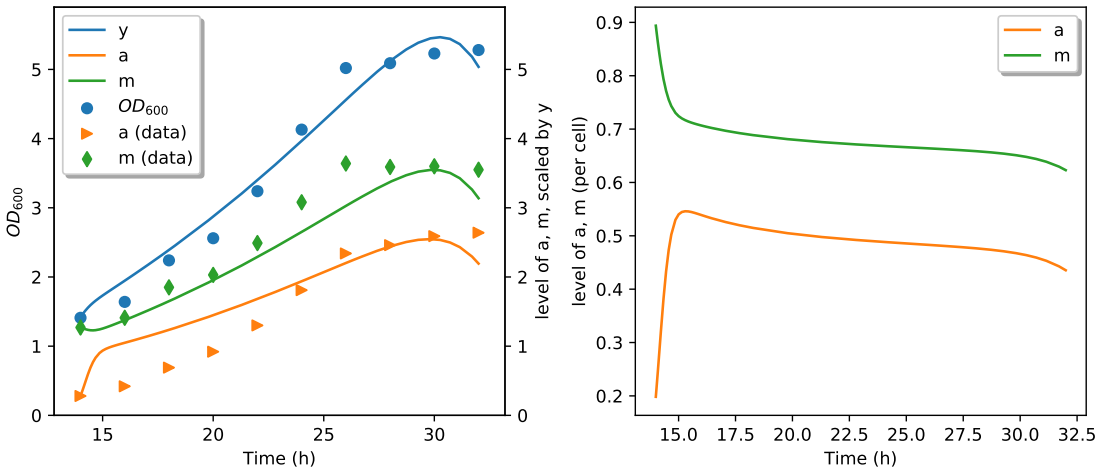

Figure 5: Solution curves of eq. (4.4) with eq. (4.25), using the parameter values given in table 6.

	Name	Value
Initial values	$a(14h)$	0.2811824057660078
	$m(14h)$	1.2673713692691995
	$Y(14h)$	1.4180265001692174
Parameters	α_1	3.2217885769458685
	α_2	$1.526 \cdot \alpha_1 = 4.9164493684193955$
	β	$2.145 \cdot \delta = 1.608284035974072$
	δ	0.7497827673538797
	K	40.80756726756603
	r	1.0087272776202856
	T_m	0.6478166024785968

Table 6: Initial values for the variables and parameter values used for fig. 5 with eq. (4.25).

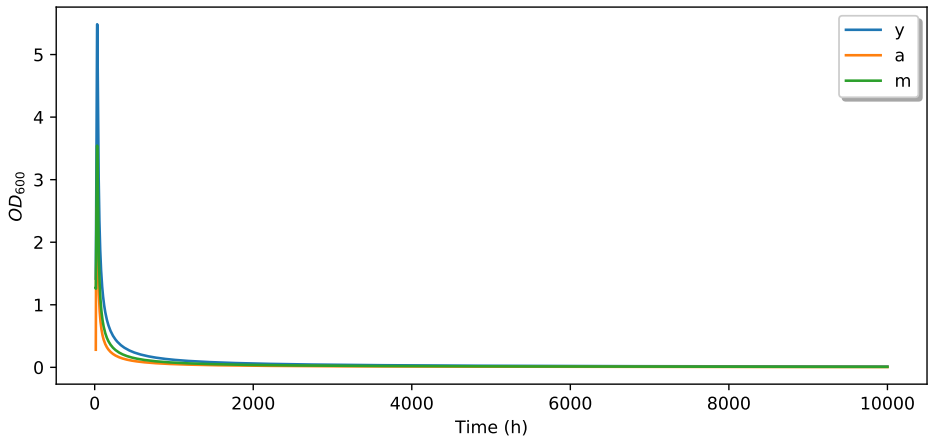

Figure 6: Solution curves of eq. (4.4) until 10000 h, using the initial and parameter values given in table 4.5.

3.5 Long Term Behaviour

Solving eq. (4.4) with the parameter values given in table 4.5 until 10000 h in python (`solve_ivp` from `scipy`) gives us the plot shown in fig. 6. We also printed the last values of the solution curves and got $Y(10000) \approx 0.012\, OD_{600}$, $a(10000) \approx 0.005\, a.u.$ and $m(10000) \approx 0.008\, a.u.$ When we computed the solutions until even higher values of t, the solution curves still got nearer 0. This suggests, that they will converge to 0.

For our second modification for $\dot{m}$ the solution curves stabilize very quickly, although m does not converge to 0 as we would like (see fig. 7). When we compute the solution in `scipy` only until 70 h, we already clearly see that $Y(t)$ and $a(t)$ go to 0, but $m(t)$ converges to a negative value. Printing the last value of the plot, we get $m(70) \approx -1.276049046521142\, a.u.$

When we let the solver compute the solutions of our model with eq. (4.26) until 10000 h, we get the plot shown in fig. 8. Printing the last values of the solution curves, we got $Y(10000) \approx 0.040\, OD_{600}$, $a(10000) \approx 0.020\, a.u.$ and $m(10000) \approx 0.026\, a.u.$ We also tried even higher values for t and the solution curves still got nearer 0, suggesting that they converge to 0.

3.6 Sensitivity Analysis for the Modified Models

Like already done for the original model, we also applied the Sobol and Morris sensitivity analysis for the model modifications. In this way one can see, whether the dependence of the model on different parameters changes.

Generalized Logistic Growth When we apply the Sobol method to our generalized logistic growth model version, we get the plots shown in fig. 9, using the bounds given in table 7. The corresponding Morris analysis and bounds are given in fig. 10 and table 8.

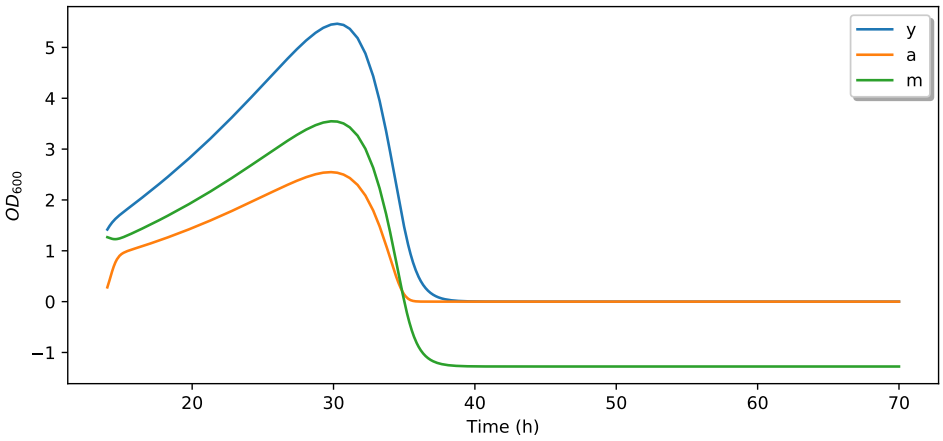

Figure 7: Solution curves of eq. (4.4) with the modified equation for $\dot{m}$, eq. (4.25), until 70 h, using the initial and parameter values given in table 6.

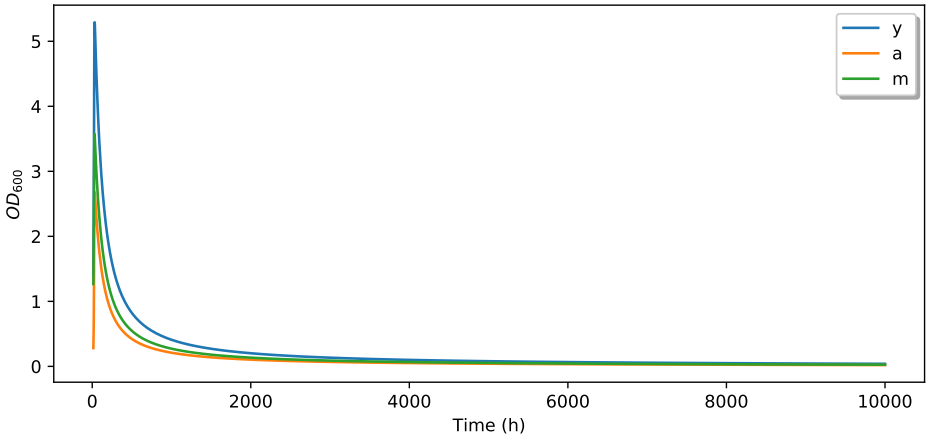

Figure 8: Solution curves of eq. (4.4) with the modified equation for $\dot{m}$, eq. (4.26), until 10000 h, using the initial and parameter values given in table 4.11.

Parameter	Minimum	Maximum
r	0.740	0.802
n	1.810	2.001
K	55.10	60.90
T_m	0.629	0.663
α_1	19.01	20.10
α_2	35.75	37.79
β	0.193	0.206
δ	0.080	0.086

Table 7: Parameter bounds used for the Sobol analysis shown in fig. 9.

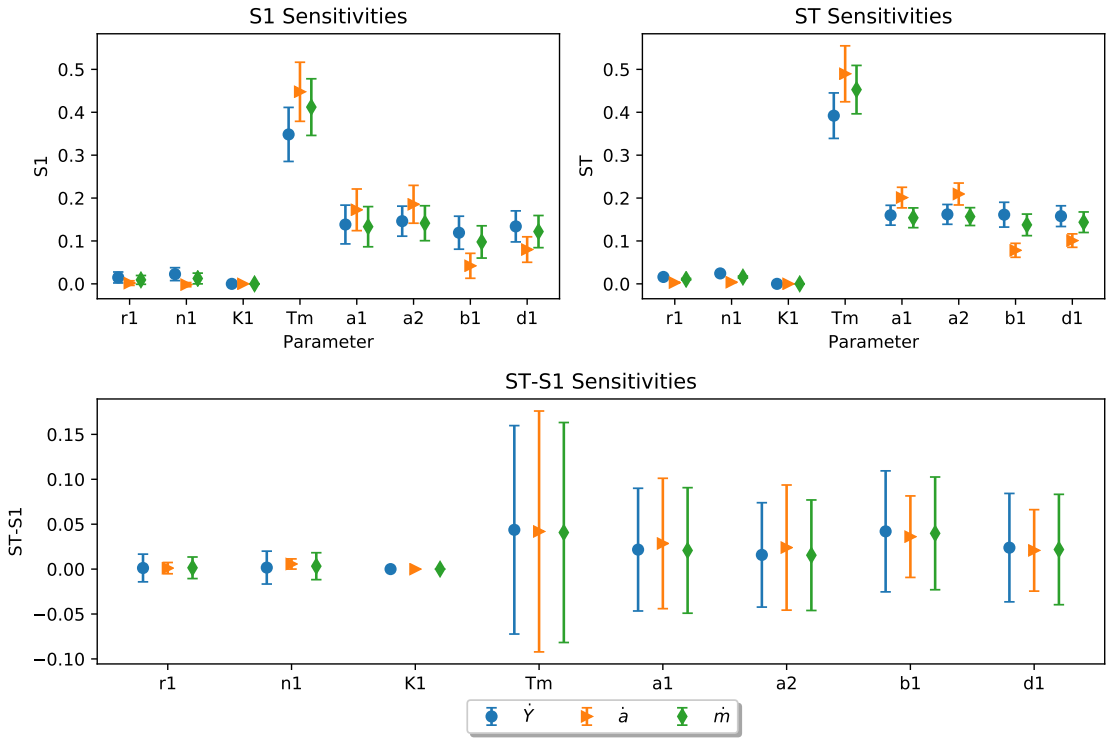

Figure 9: The results of the Sobol method computed with `SALib`, plotted as errorbars. Top left: S_1 sensitivities with their confidence interval for each parameter and each equation of eq. (4.4), where $\dot{Y}$ is changed according to section 4.6.1, Top right: the same for the S_T sensitivities, Bottom: total sensitivities minus first order sensitivities, including the resulting confidence intervals.

Parameter	Minimum	Maximum
r	0.742	0.802
n	1.810	1.935
K	55.10	60.90
T_m	0.6290	0.6619
α_1	19.01	20.10
α_2	35.50	37.79
β	0.193	0.205
δ	0.081	0.086

Table 8: Parameter bounds used for the Morris analysis shown in fig. 10.

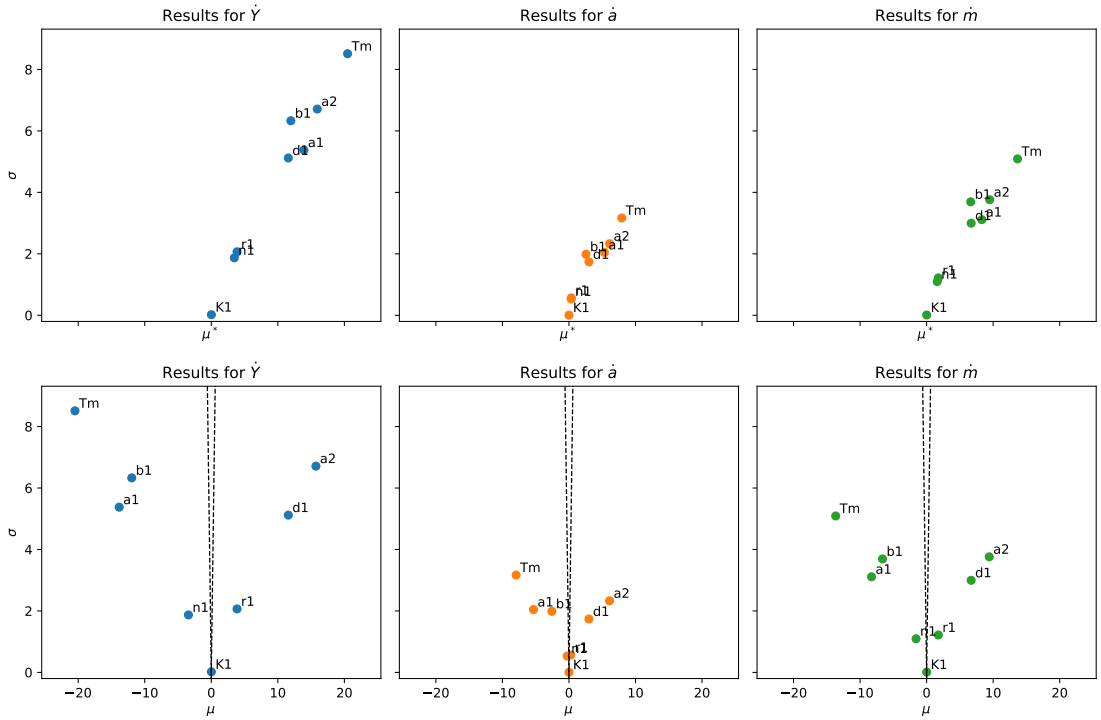

Figure 10: The results of the Morris method computed with `SALib`. First row: μ^* and σ plotted for all parameters and each equation we want to analyse (see section 4.6.1), Second row: the same for μ and σ, together with $\mu = \pm 2 \cdot SEM$ (dashed lines).

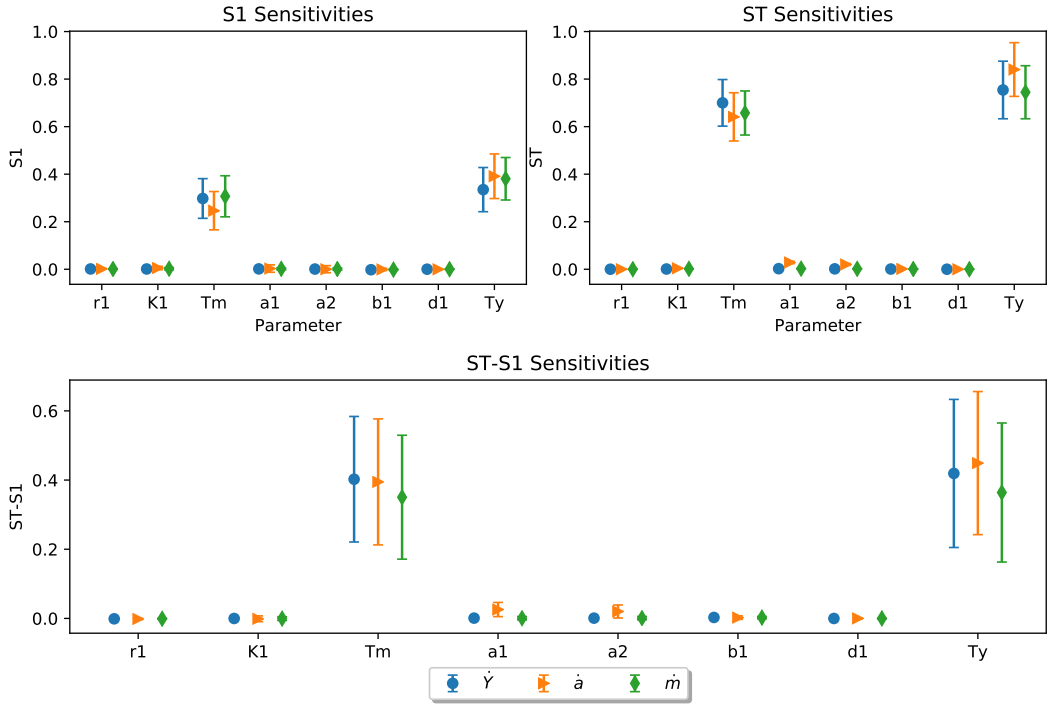

Figure 11: The results of the Sobol method computed with `SALib`, plotted as errorbars. Top left: S_1 sensitivities with their confidence interval for each parameter and each equation of eq. (4.4), where $\dot{m}$ is changed according to eq. (4.24), Top right: the same for the S_T sensitivities, Bottom: total sensitivities minus first order sensitivities, including the resulting confidence intervals.

First Modification We did the sensitivity analysis here only for the adapted modification including T_y, since we could not get a good fit for the first version. The results can be found in fig. 11 and fig. 12 and the corresponding bounds in table 9. The bounds are derived by computing $\pm 5\%$ of the original values given in table 4.9.

Second Modification Here we chose to only analyse the original modification without our typo. In the following, the visualized results of the Sobol and Morris method can be found in fig. 13 and fig. 14. The used bounds ($\pm 5\%$ of the original values in table 6) are given in table 10.

The Morris analysis results here in much larger values than for the other models. This has a simple explanation, when we look at the single terms of eq. (4.4) and eq. (4.25). Since $\dot{Y}$ includes a logistic growth term, $Y(t)$ can at most increase until K. If now $Y(t)$ becomes K, which can easily happen when we compute the solutions numerically, $\dot{Y} = 0$ and $Y(t) = K$ for the rest of the timespan. $a(t)$ and $m(t)$ can then also increase until at most K, as $\dot{a}$ and $\dot{m}$ contain $(Y - a)$ or $(Y - m)$ in each term. Thus, all solution curves are bounded by K. But this means, that we can have solution curves which almost instantly increase until K and stay there. Since they do not tend to infinity, those are valid solutions producing no `NaN` values, but their RSS score is of course very high. Computing then the elementary effect between such a solution and a good fit, results in the high values for μ and μ^* visible in fig. 14.

Parameter	Minimum	Maximum
r	4.304	4.757
K	4.898	5.414
T_m	0.638	0.706
α_1	0.895	0.989
α_2	1.651	1.825
β	2.077	2.295
δ	0.24	0.026
T_y	0.693	0.766

Table 9: Parameter bounds used for the Sobol and Morris analysis shown in fig. 11 and fig. 12.

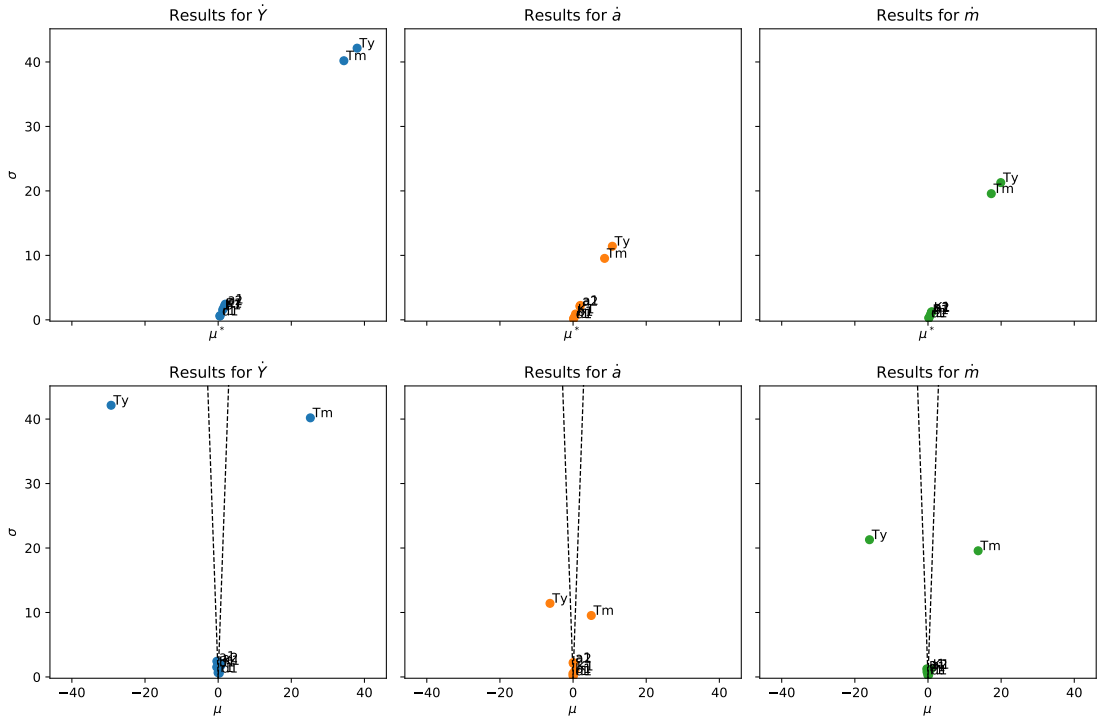

Figure 12: The results of the Morris method computed with `SALib`. First row: μ^* and σ plotted for all parameters and each equation we want to analyse (see eq. (4.4) and eq. (4.24)), Second row: the same for μ and σ, together with $\mu = \pm 2 \cdot SEM$ (dashed lines).

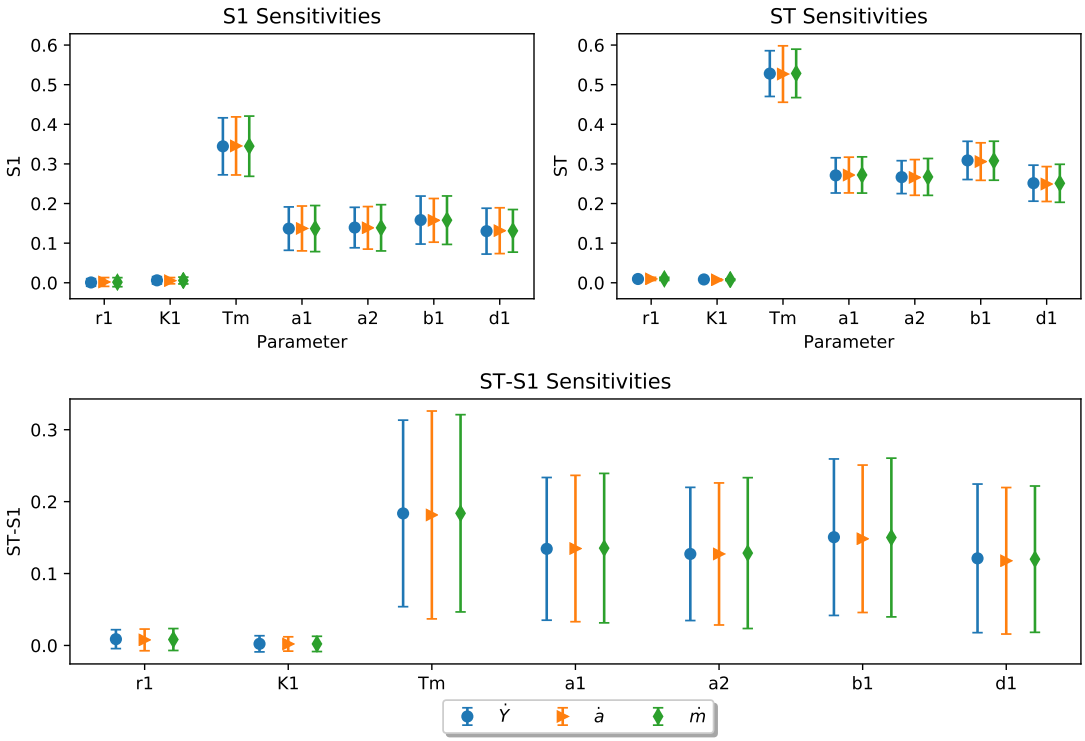

Figure 13: The results of the Sobol method computed with `SALib`, plotted as errorbars. Top left: S_1 sensitivities with their confidence interval for each parameter and each equation of eq. (4.4), where $\dot{m}$ is changed according to eq. (4.25), Top right: the same for the S_T sensitivities, Bottom: total sensitivities minus first order sensitivities, including the resulting confidence intervals.

Parameter	Minimum	Maximum
r	0.958	1.059
K	38.767	42.848
T_m	0.615	0.680
α_1	3.061	3.383
α_2	4.671	5.162
β	1.528	1.689
δ	0.712	0.787

Table 10: Parameter bounds used for the Sobol and Morris analysis shown in fig. 13 and fig. 14.

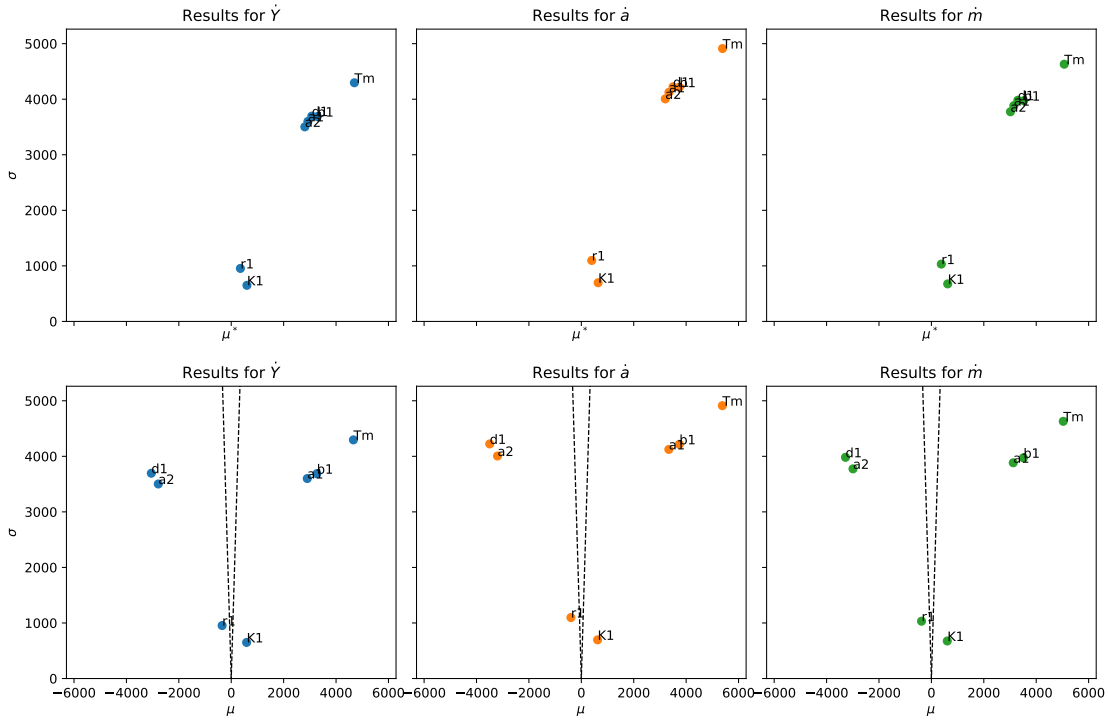

Figure 14: The results of the Morris method computed with `SALib`. First row: μ^* and σ plotted for all parameters and each equation we want to analyse (see eq. (4.4) and eq. (4.25)), Second row: the same for μ and σ, together with $\mu = \pm 2 \cdot SEM$ (dashed lines).

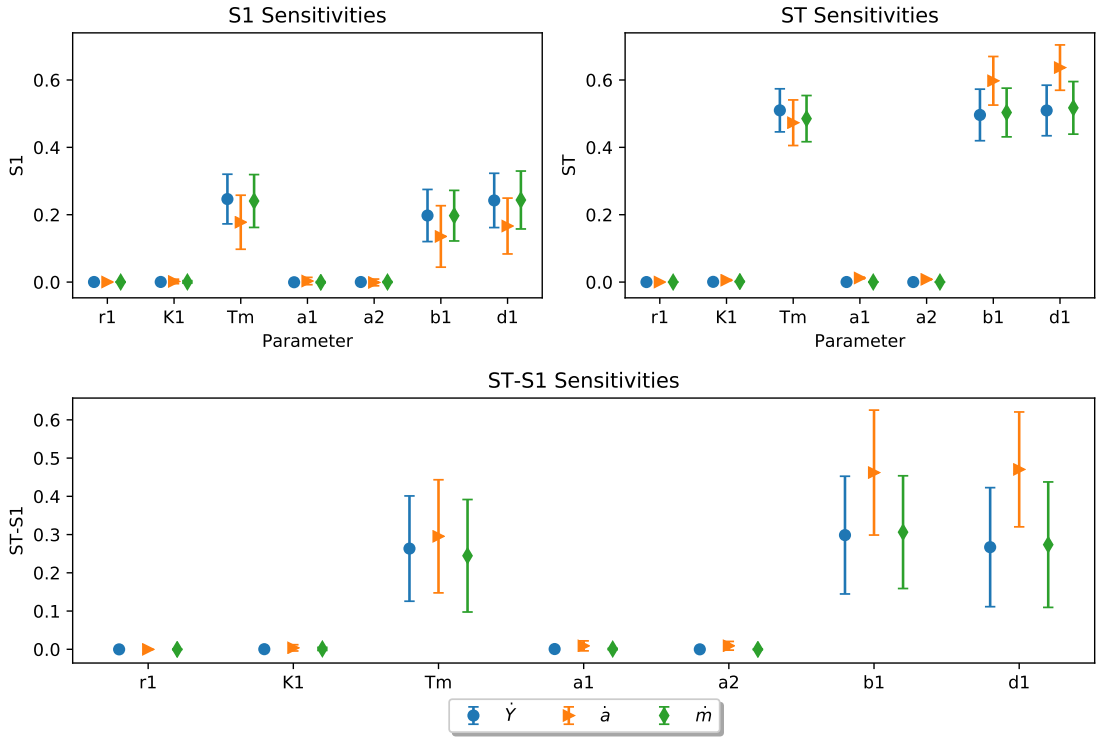

Figure 15: The results of the Sobol method computed with `SALib`, plotted as errorbars. Top left: S_1 sensitivities with their confidence interval for each parameter and each equation of eq. (4.4), where $\dot{m}$ is changed according to eq. (4.27), Top right: the same for the S_T sensitivities, Bottom: total sensitivities minus first order sensitivities, including the resulting confidence intervals.

Third Modification For our third modification of $\dot{m}$, the results of the Sobol and Morris method can be found in fig. 15 and fig. 16. Furthermore, the bounds are given in table 11, using again $\pm 5\%$ of the original values given in table 4.12.

Parameter	Minimum	Maximum
r	6.277	6.938
K	4.886	5.401
T_m	0.662	0.732
α_1	0.490	0.542
α_2	0.883	0.976
β	3.048	3.369
δ	4.194	4.635

Table 11: Parameter bounds used for the Sobol and Morris analysis shown in fig. 15 and fig. 16.

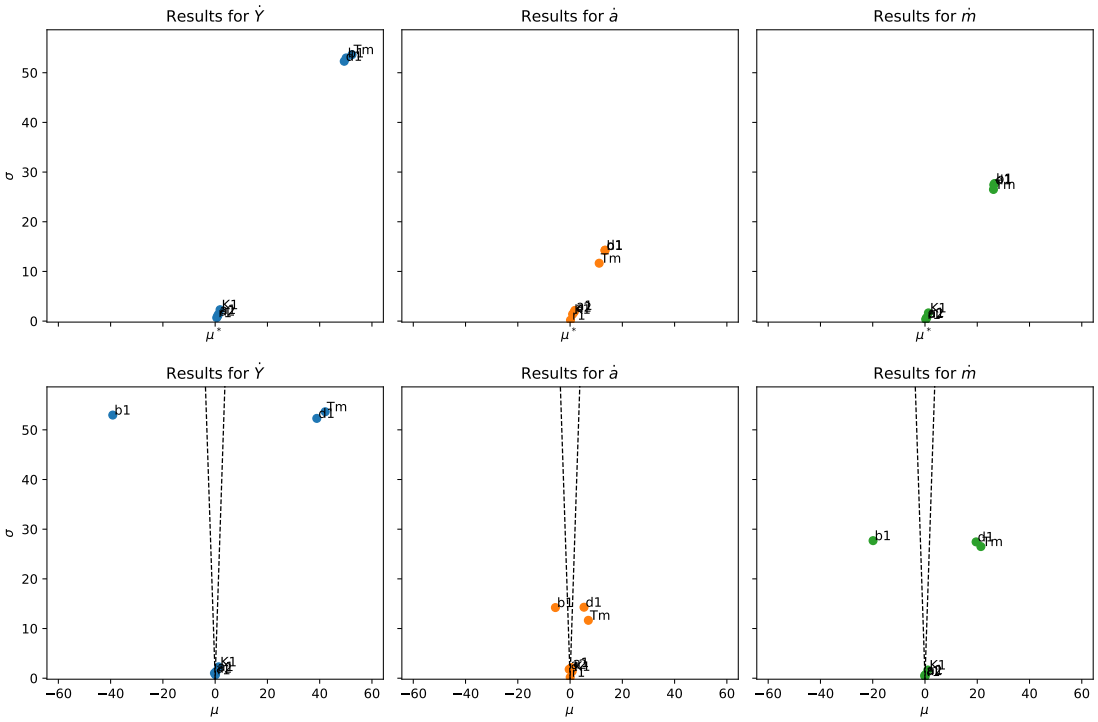

Figure 16: The results of the Morris method computed with `SALib`. First row: μ^* and σ plotted for all parameters and each equation we want to analyse (see eq. (4.4) and eq. (4.27)), Second row: the same for μ and σ, together with $\mu = \pm 2 \cdot SEM$ (dashed lines).

4 Example Code for Fitting

```
from symfit import variables, parameters, Fit, D, ODEModel
from symfit.core.minimizers import DifferentialEvolution

y, a, m, t = variables('y, a, m, t')
a1, d1, K1, r1, Tm = parameters('a1, d1, K1, r1, Tm')

a1.min, a1.max, a1.value = 6, 6.1, 6.03
d1.min, d1.max, d1.value = .08, .1, .09
K1.min, K1.max, K1.value = 44, 45, 44.8
r1.min, r1.max, r1.value = .36, .37, .364
Tm.min, Tm.max, Tm.value = .64, .65, .644

a2 = 1.526*a1
b1 = 2.145*d1

model_dict = {
    D(y, t): r1*y*(1-y/K1)*(m-Tm*y),
    D(a, t): a1*a*(y-a)-a2*a*(y-m),
    D(m, t): b1*m*(a-d1/b1*y)
}

i=0
ode_model = ODEModel(model_dict, initial={t: t1_sc[i], y: y1_sc[i], a:
    a1_sc[i] , m: m1_sc[i]})

fit = Fit(ode_model, t=t1_sc, y=y1_sc, a=a1_sc, m=m1_sc, minimizer=
    DifferentialEvolution)
fit_result = fit.execute()
print(fit_result.params)
```

Listing 1: Fitting eq. (4.4) in `symfit`.

5 Optimization Algorithms from `scipy`

In this section we give a short description of the two algorithms used in the fitting routines, based on their documentation. Both methods are global minimizers. For more details see `scipy`'s documentation on `https://docs.scipy.org/doc/scipy/reference/` and the references within.

5.1 Basin-Hopping

As a first try we often used the Basin-hopping algorithm, since it worked quite good in our first fitting attempts. This stochastic method consists of mainly two phases: a global stepping algorithm followed by a local minimization.

For a certain number of iterations, by default 100, the algorithm iteratively searches the global minimum of a function f in the following way: First, a random perturbation of the coordinates is applied, where one can also play with the used step size. Now a local minimization is applied, resulting in new coordinates. After this, the new coordinates are accepted or rejected, depending on the computed function value $f(x_{new})$. The default acceptance test is the Metropolis criterion of the standard Monte Carlo algorithm. Thus, new coordinates x_{new} are always accepted if $f(x_{new}) < f(x_{old})$. Otherwise, the acceptance probability is $e^{\frac{-(f(x_{new})-f(x_{old}))}{T}}$, where T is the so-called temperature.

5.2 Differential Evolution

When we got no good result with the Basin-hopping algorithm or wanted to apply another method on already adjusted parameter ranges, we used the Differential Evolution algorithm. This is also a stochastic method, attempting to find the global minimum. In contrast to the Basin-hopping algorithm, this method is population based.

In each generation, the algorithm mutates each candidate solution depending on a mutation constant and the chosen strategy. The resulting trial candidate then replaces the original one if its fitness is better. If it is additionally better than the best overall candidate, it replaces this one as well. For a faster convergence than in the original algorithm, the best solution vector gets updated continuously within a single iteration by default.

6 AIC and AIC_c

The following is based on [34].
The AIC in its general form is formulated as follows:

$$AIC = -2\, log(\mathcal{L}(\hat{\theta}|y)) + 2K \tag{7}$$

where K is the number of estimable parameters, $\hat{\theta}$ is an estimator for the parameter θ of the considered model, y the given data and $\mathcal{L}$ is the likelihood. Hence, $\mathcal{L}(\hat{\theta}|y)$ should be the maximum value of the likelihood function. When comparing a set of models, the model with the smallest AIC should be preferred.
Since we used a least squares estimation, we can use a special form of the above criterion, assuming normally distributed errors with constant variance σ^2:

$$\begin{aligned} &AIC = n\, log(\hat{\sigma}^2) + 2K \\ &\text{with } \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} \hat{\epsilon}_i^2 \end{aligned} \tag{8}$$

where K is the number of estimated parameters, including $\hat{\sigma}^2$, n is the number of samples and $\hat{\epsilon}_i$ are the estimated residuals of the considered model. Thus, in our case, $\hat{\sigma}^2$ can be computed in the following way:

$$\hat{\sigma}^2 = \frac{1}{30}\sum_{i=1}^{10}(Y(t_i) - Y_{t_i})^2 + (a(t_i) - a_{t_i})^2 + (m(t_i) - m_{t_i})^2 \tag{9}$$

with t_i from table 3 and Y_{t_i}, a_{t_i} and m_{t_i} the corresponding values in the same table. $Y(t_i)$, $a(t_i)$ and $m(t_i)$ denote the values of the considered model at the time point t_i. Our number of samples is 30, since we have 10 time points with 3 values each when looking only at the starvation case.
Like already mentioned in chapter 5, compared to the number of parameters we have only few data points. Our basic model has 7 parameters, like two of the modifications, while the other two modifications have 8 parameters. Thus, together with $\hat{\sigma}^2$, we have 8, respectively 9, parameters to estimate. When we look at the ratio of samples this results in $\frac{30}{8} = 3.75$, respectively $\frac{30}{9} = 3.\overline{3}$. Hence, we are far below 40, which is a suggested threshold in [34]. In this case, it is recommended to apply the AIC_c. This is a second order AIC accounting for the small sample size with a bias adjustment, resulting in

$$AIC_c = AIC + \frac{2K(K+1)}{n-K-1} \tag{10}$$

We implemented this criterion in fitting.ipynb, which can be found on `https://github.com/Melanie757/Thesis-Code`.