An introduction to modelling through a microbial interaction application

An introduction to modelling through a microbial interaction application

Fabiana Zama [a]

[a]Department of Mathematics, University of Bologna, Italy

ARTICLE HISTORY
Compiled July 30, 2023

ABSTRACT
This paper describes a teaching experiment in a Numerical Methods course for Master of Science students. The experiment uses scientific papers to develop modelling studies in the context of wine fermentation microbial interactions. The course involves theoretical and laboratory classes that focus on implementing numerical methods using Matlab for Initial Value Problems and Boundary Value Problems. The students are asked to formalize the mathematical model and build their own experiments using the information provided in the papers. Additionally, a parameter estimation experiment is organized, which involves generating synthetic data and computing noisy data to estimate the natural death rate of sensitive yeast. The results show that data noise significantly affects the parameter estimate and that scaling the data can help reduce the impact of measurement errors. The presented results can be used to investigate other possible assignments, such as how the evaluation of the Jacobian affects the estimation performance and compare different optimization algorithms.

KEYWORDS
Numerical solution of ODEs, Parameter estimation, Graduate class assignments, Mathematical modeling, Fermentation, microbes.

1. Introduction

Mathematics courses at the Master of Science level often involve a combination of theoretical and practical components that enable students to apply mathematical principles to solve real-world problems. For example, in this teaching context, a Numerical Methods course addresses ordinary and boundary differential problems and students learn about numerical approaches to solve these problems. The course has theoretical and laboratory classes focusing on Initial Value Problems (IVPs) and Boundary Value Problems (BVPs). In the laboratory classes, students learn how to implement numerical methods for IVPs and BVPs using Matlab. Various assignments are employed in diverse contexts to enhance the practical application of the presented methods. For example, the proposed teaching experiment uses scientific papers to develop modelling studies, enriching the course content and providing a context for using numerical methods. In particular, this work focuses on modelling microbial interactions in the wine fermentation process. Wine fermentation involves various yeasts, including spoilage yeasts that can cause significant economic losses by producing negative organoleptic

CONTACT F. Z. Author. Email: fabiana.zama@unibo.it

changes. The topic is introduced by (Kuchen, Maturano, Gil, Vazquez, & Scaglia, 2022), proposing an improvement to a model discussed in a previous paper (Pommier, Strehaiano, & Delia, 2005). These articles refer to the seminal work (Ramon-Portugal, Torija, Mas, & Guillamon, 1997) that describes the kinetic study of yeasts in wine-making procedures.

Starting from (Kuchen et al., 2022), the assignment requested to formalize the mathematical model. This first part needed a minimum analysis of the referenced paper ((Pommier et al., 2005), (Ramon-Portugal et al., 1997)) to set up the mathematical model and reference parameters used in numerical simulations. Next, students were asked to use the information provided in these papers to build their own experiments, which required them to formalize the differential model that constitutes the forward problem and analyze its solution through variable steps, using Matlab Ordinary Differential Equations (ODE) solvers and implementing fixed-step solvers. Finally, the assignment requested implemening Matlab code to perform a parameter estimation experiment using a given set of nominal parameters reported in the literature. The text of the assignment, serving as an illustrative example, can be found in Appendix A.

Overall, this approach was very satisfactory, providing students with a real-world application context to apply the numerical methods they had learned in the course. In addition, using scientific papers as a basis for class assignments makes it possible to introduce students to the process of scientific inquiry and encourages them to think critically about applying mathematical principles in solving real-world problems.

The manuscript is organized as follows. In Section 2, we introduce the microbial interaction model, which serves as the basis for subsequent analysis. Then, in Section 3, we examine the numerical solution of the forward problem, specifically the initial value problem. Finally, in Section 4, we comprehensively illustrate the parameter estimation experiment.

2. Problem description

Wine fermentation involves various yeasts, including spoilage yeasts that can cause significant economic losses by producing negative organoleptic changes. Sulfur dioxide (SO_2) is commonly used to control these spoilage populations, but it harms human health. An alternative treatment known as biocontrol involves introducing new yeasts that produce a toxic protein to kill the sensitive (spoilage) ones. Mathematical models can predict the behaviour of these killer yeasts (biocontrollers) in different physico-chemical conditions and provide information on the kinetic parameters involved.

One such model, described in (Pommier et al., 2005), is a system of six ordinary differential equations that allows studying interactions between two yeasts: the spoilage yeast and the biocontroller, and it is based on the dynamic characteristics of killer/sensitive yeast mixed cultures described in the seminal work by (Ramon-Portugal et al., 1997). In this model, the biomass is segregated into four variables: the viable killer cells X_{vk} (biocontrollers), the viable sensitive cells X_{vs} (spoilage yeast), the cumulative dead killer cells X_{dk}, and the cumulative dead sensitive cells X_{ds}. The variable T stands for the toxin concentration in the culture medium, and I represents an unidentified inhibitor that is common to both killer and sensitive strains. The dynamic evolution of each of the six variables of the model is given by the system of

equations shown below

$$\begin{cases} \dot{X}_{vk} = r_{vk} - r_{dk} \\ \dot{X}_{vs} = r_{vs} - r_{ds} \\ \dot{X}_{dk} = r_{dk} \\ \dot{X}_{ds} = r_{ds} \\ \dot{T} = r_{pT} - r_{aT} \\ \dot{I} = r_I \end{cases}. \tag{1}$$

The system's dynamics are determined by seven reaction rates as shown in (2). These rates are parameterized by maximum specific growth rates, deceleration parameters, natural death rates, toxin-caused death rates, specific toxicity of the toxin, general inhibitor production rate parameter, and toxin production rate parameter.

$$\begin{bmatrix} r_{vk} & \text{killer growth rate} \\ r_{vs} & \text{sensitive growth rate} \\ r_{dk} & \text{killer death rate} \\ r_{ds} & \text{sensitive death rate} \\ r_{pT} & \text{toxin production rate} \propto r_{vk} \\ r_{aT} & \text{toxin adhesion rate} \\ r_I & \text{inibitor rate} \end{bmatrix}. \tag{2}$$

The seven rates are represented by the following equations:

$$r_{vk} = \mu_{maxk} X_{vk}(1 - A_k I), \qquad r_{dk} = \mu_{dk} X_{vk} \tag{3}$$

$$r_{vs} = \mu_{maxs} X_{vs}(1 - A_s I), \qquad r_{ds} = \mu_{ds} X_{vs} + K X_{vs} T \tag{4}$$

$$r_I = a(r_{vk} + r_{vs}), \qquad r_{aT} = W X_{vs} T, \quad r_{pT} = \alpha r_{vk}. \tag{5}$$

These equations depend on the values of several parameters, including the maximum specific growth rate of the killer/sensitive strains (μ_{maxk}, μ_{maxs}), the killer/sensitive deceleration parameters (A_k, A_s), the natural death of killer/sensitive yeast species (μ_{dk}, μ_{ds}), the toxin-caused death rate (K), the specific toxicity of the toxin (W), the general inhibitor production rate parameter (a), and the toxin production rate parameter (α). The physiochemical reactions of the model are well represented in Figure 1.

By substituting the expressions of the parameter-dependent rates, we obtain the standard form of the microbial interaction model shown in (6).

$$\begin{cases} \dot{X}_{vk} = \left((1 - A_k)\mu_{maxk} - \mu_{dk}\right) X_{vk} \\ \dot{X}_{vs} = \left(-KT + (1 - A_{s1} I)\mu_{maxs} - \mu_{ds}\right) X_{vs} \\ \dot{X}_{dk} = X_{vk}\mu_{dk} \\ \dot{X}_{ds} = (KT + \mu_{ds}) X_{vs} \\ \dot{T} = -TW X_{vs} + (1 - A_k) X_{vk} \alpha \mu_{maxk} \\ \dot{I} = -a\left(\left(-1 + A_k I\right) X_{vk}\mu_{maxk} + \left(-1 + A_{s1} I\right) X_{vs}\mu_{maxs}\right) \end{cases}, \tag{6}$$

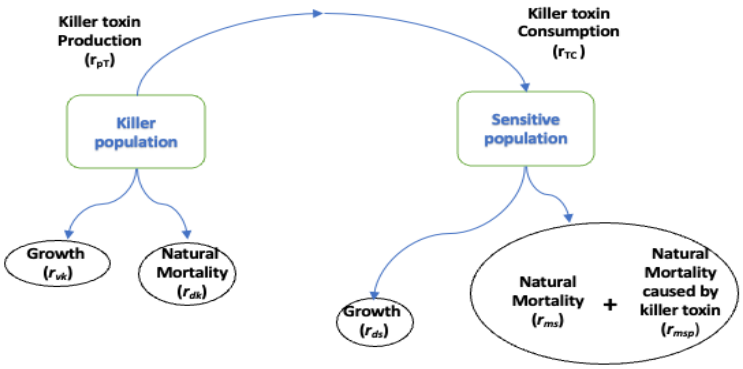


Figure 1. Reactions between a killer and a sensitive yeast in mixed culture. $r_{TC} \equiv -WX_{vs}T$ in (5), $r_{ms} \equiv \mu_{ds}X_{vs}$ in (4), $r_{msp} \equiv KX_{vs}T$ in (4).

The initial values are specified as follows:

$$X_{vs}(0) = N, \quad X_{vk}(0) = N\gamma, \quad 0 < \gamma < 1,$$

where N is the total number of sensitive cells and γ is the ratio between the viable killer and sensitive strains, and

$$X_{dk}(0) = X_{ds}(0) = T(0) = I(0) = 0.$$

Overall, mathematical models such as this one allow for the prediction of the behavior of killer yeasts in wine fermentation and provide insight into the effects of physicochemical conditions on their interactions with spoilage yeasts. The parameter γ can be utilized to evaluate the initial quantity of viable killer strains in relation to the total number of sensitive cells, thereby enabling the monitoring of their impact on population dynamics.

3. Solution of the forward problem

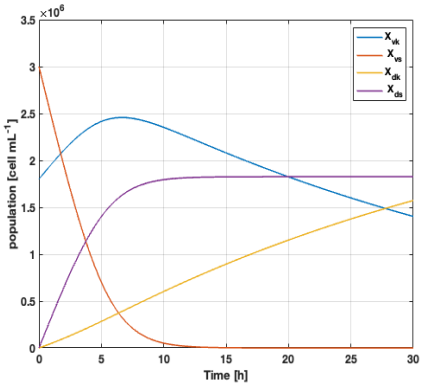
Choosing a set of nominal parameters as in Table (1) and $N = 3 \cdot 10^6$ $[cell\ mL^{-1}]$, it is possible to solve (6) numerically and observe the effect of the ratio γ on the viable sensitive population. Figures 2 and 3 depict the effect of the initial ratio $\gamma = X_{vk}/X_{vs}$

Table 1. Nominal parameters as in (Pommier et al., 2005).

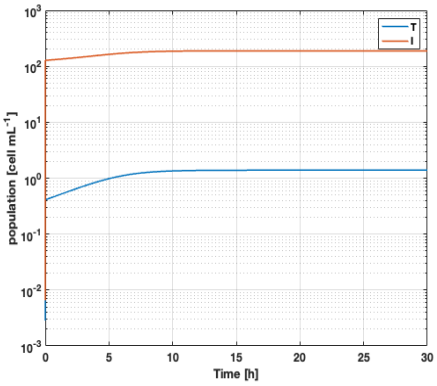
Parameter	Value	Parameter	Value
μ_{maxk}	0.358 $[h^{-1}]$	μ_{maxs}	0.278 $[h^{-1}]$
A_k	0.0053 $[lg^{-1}]$	A_s	0.0098 $[lg^{-1}]$
μ_{dk}	0.0263 $[h^{-1}]$	μ_{ds}	0.0049 $[h^{-1}]$
K	0.257 $[-]$	W	0.170 $[mL\ cells^{-1}h^{-1}]$
a	1 $[-]$	α	1 $[-]$

on population dynamics. When the initial viable killer population $X_{vk}(0)$ is approx-

imately half the initial sensitive population $X_{vs}(0)$ ($\gamma = 0.6$), the killer population reaches its maximum after about seven hours, while the sensitive population goes extinct in about 15 hours (Figure 2(a)). On the other hand, when $\gamma = 0.1$, the biocontrol



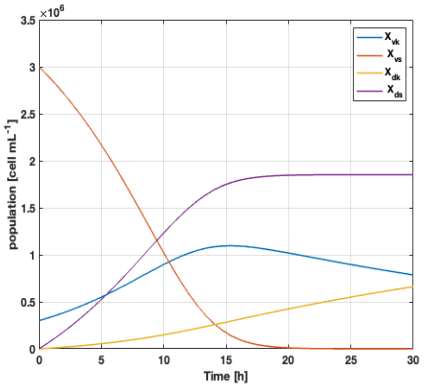
(a) Viable and Death dynamics.



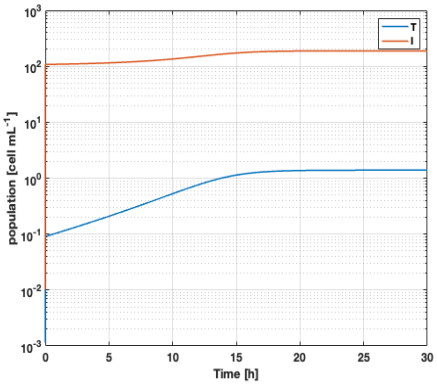
(b) Toxin and Inibitior concentration.

Figure 2. $X_{vk}(0) = X_{vs}(0) \cdot 0.6$ with $N = 3 \cdot 10^6$ $[cells\ mL^{-1}]$

process is slower, with a peak in the killer population after 15 hours, and the sensitive population going extinct in about 20 hours (Figure 3(a)). The inhibitor concentration dynamics show a rapid increase to a constant value, whereas the toxin concentration has a slower increase when γ is small (Figures 2(b) and 3(b)). The model (6) can be



(a) Viable and Death dynamics.



(b) Toxin and Inibitior concentration.

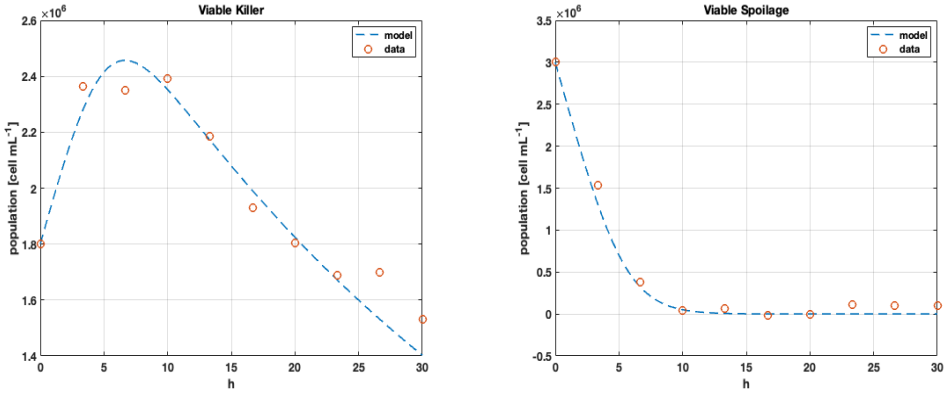
Figure 3. $X_{vk}(0) = X_{vs}(0) \cdot 0.1$ $N = 3 \cdot 10^6$ $[cells\ mL^{-1}]$

expressed in the classical initial value problem (IVP) form as:

$$\begin{cases} \dot{Y} = F(t, Y(t)) & t \in [0, 30]h \\ Y(0) = Y_0, & Y : \mathbb{R} \to \mathbb{R}^6, \end{cases} \tag{7}$$

where $Y = (X_{vk}, X_{vs}, X_{dk}, X_{ds}, T, I)^\intercal$.

The stiffness of the problem becomes evident from a posteriori observation of the computational cost reported in Table 2. Additionally, by computing the eigenvalues of the Jacobian $F_Y(0, Y_0)$, we observe that the smallest eigenvalue ($\lambda_{min} = -5.1 \cdot 10^5$)



(a) Viable killer observations

(b) Viable sensitive observations

Figure 4. Example: $m_t = 10, \delta = 10^{-2}$ Red circles represent noisy data and blue line is noiseless data.

is significantly smaller than -1 for both $\gamma = 0.1$ and $\gamma = 0.6$, confirming the IVP stiffness.

Table 2. Matlab ODE solvers

Method	Time Steps	Function Evaluations
`ode23tb`	1237	185
`ode15s`	1582	367
`ode23t`	650	187
`ode89`	2.3299 10^7	8.4973 10^5

4. Parameter Estimation Test

The final part of the assignment involves organizing a parameter estimation test to estimate the natural death rate of sensitive yeast, μ_{ds}. In this test, synthetic data are generated by extracting a set of observations from viable killer and sensitive cells, X_{vk} and X_{vs}, which were obtained by solving the forward problem with the set of nominal parameters shown in Table 1. The following is a summary of the procedure to collect synthetic measurements:

- Set a reference value of $u_{ref} \equiv \mu^*_{ds} = 0.0049$.
- Solve the forward problem $Y' = F(t, Y(t))$ with parameter $u = u_{ref}$.
- Select m_t measurement time values t_i and collect the synthetic observation data into the vector z: $z = [Y_1(t_i), Y_2(t_i)],\ i = 1, \ldots, m_t$.
- Select a value of $\delta > 0$ and a normal random vector $\boldsymbol{\eta}$ with $\|\boldsymbol{\eta}\| = 1$.
- Compute noisy data as follows: $z^\delta = z + \delta \boldsymbol{\eta} \|z\|$

Figures 4 show with red dots the noisy data in the case $m_t = 10$, and noise level of $\delta = 10^{-2}$.

The parameter μ_{vs} is estimated by solving the nonlinear least squares problem shown in (8).

$$\min_u \frac{1}{2} \|\varphi(u) - z^\delta\|_2^2 \tag{8}$$

Here, $\varphi(u)$ maps the numerical solution of the differential model (7), with parameter u, to the m_t measurement points.
The quality of the result is evaluated by computing the Parameter Relative Error (PRE) and the weighted residual norm NRes, shown in (9).

$$\mathrm{PRE}(u) = \frac{\|u - u_{ref}\|_2^2}{\|u_{ref}\|_2^2}, \qquad \mathrm{NRes}(u) = \frac{\|\varphi(u) - z^\delta\|}{\|z\|}. \tag{9}$$

The PRE(u) measures the relative error between the estimated parameter u and the reference value u_{ref}, while NRes(u) measures the residual between the observed noisy data z^δ and the computed estimates $\varphi(u)$, normalized by the reference data z.

Numerical methods based on gradient descent (Nocedal & Wright, 2006), such as Gauss-Newton or those implemented by the Matlab function `lsqnonlin`, require an initial guess u_0 and compute solution updates through the gradient of the objective function shown in (8):

$$J_u = G^\intercal(\varphi(u) - z^\delta), \quad G = \nabla_u \varphi(u).$$

Hence, one key point in the solution of the nonlinear optimization problem (8) lies in the computation of J_u.

In this example, we have one parameter, so G is obtained by aligning the two vectors $\mathbf{g_1}$ and $\mathbf{g_2}$ where

$$\mathbf{g_1} = \frac{\partial Y_1}{\partial u_1}^\intercal (t_i), \quad \mathbf{g_2} = \frac{\partial Y_2}{\partial u_1}^\intercal (t_i), \qquad i = 1, \ldots, m_t.$$

The evaluation of the components $\mathbf{g_i}$ can be performed by finite difference schemes or, more conveniently, by the sensitivity method (Chavent, 2010). In this last case, the IVP (7) is modified by adding the sensitivity components $W \equiv \frac{\partial Y}{\partial u}$ and applying the chain rule to the IVP problem. This leads to the augmented IVP problem shown in (10).

$$\begin{cases} \dot{Y} = F(Y(t), u) \\ \dot{W}_1 = F_u(Y, u) + F_Y(Y, u) W_1 \end{cases}, \qquad \begin{cases} Y(t_0) = Y_0 \\ W_1(t_0) = 0 \end{cases} \tag{10}$$

Using (10) instead of (7) inside the optimization iteration allows us to update the parameter estimation very efficiently.
With a solver available, it is possible to analyze several aspects of the parameter estimation process. For example, we can test the estimate's robustness depending on the data noise δ and the starting guess u_0. Furthermore, we can use data uncertainty to conveniently scale the data vector z^δ.
Let Z denote the given data vector, and ΔZ_j denote the uncertainty on its jth component, $j = 1, \ldots, m$. It is convenient to use this uncertainty to define data to be used in the calibration process as follows:

$$z_j = \frac{1}{\sqrt{m}} \frac{Z_j}{\Delta Z_j}$$

In the present setting, where data noise is known, it is possible to use such information

to scale the data. Therefore, the scaled problem is shown in (11),

$$\min_u \frac{1}{2}\|\varphi(u) - z^{\delta}_{scal}\|_2^2, \tag{11}$$

where the scaled data components are

$$(z^{\delta}_{scal})_i = z^{\delta}_i / \Delta Z_i, \quad \Delta Z_i = |z|\delta\boldsymbol{\eta}_i.$$

As an example, we present some results in Table 3 with noise levels of $\delta = 10^{-3}, 10^{-2}$ and initial guesses u_0 with relative distances to the true reference value of PRE0 $= 1, 10$. We used both non-scaled data z^{δ} and scaled data z^{δ}_{scal}, and the results are shown in columns $4-6$ and $7-9$, respectively.

The computation cost is evaluated in terms of the number of function evaluations (nfevals), and we used the `trust-region-reflective` solver implemented by the Matlab function `lsqnonlin` to obtain the results shown in Table 3.

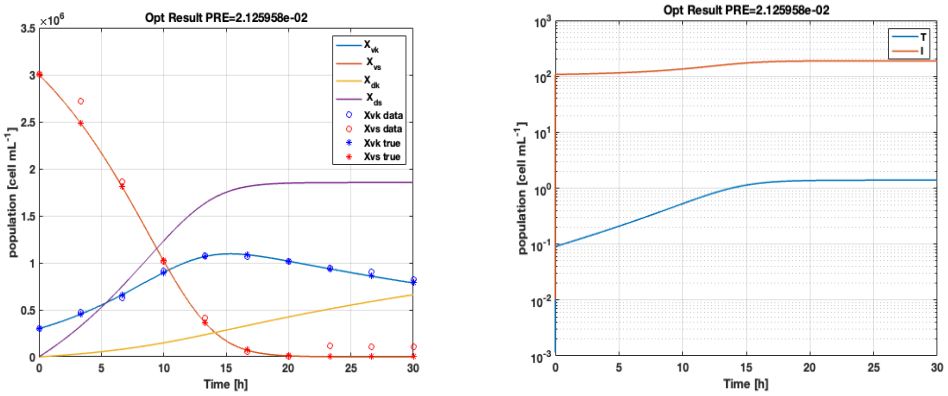
Table 3. Parameter estimation test with 10 observation times and $X_{vk}(0)/X_{vs}(0) = \rho$

ρ	δ	PRE0	No data scaling			Data scaling		
			PRE	NRes	nfevals	PRE	NRes	nfevals
0.1	0.001	10	$8.306\ 10^{-2}$	$5.732\ 10^{-3}$	12	$2.624\ 10^{-2}$	$7.651\ 10^{-4}$	20
	0.01	1	$7.813\ 10^{-1}$	$5.585\ 10^{-2}$	8	$2.511\ 10^{-2}$	$7.327\ 10^{-4}$	10
0.6	0.001	10	$3.798\ 10^{-1}$	$4.977\ 10^{-3}$	14	$5.065\ 10^{-2}$	$3.491\ 10^{-4}$	22
	0.01	0.01	3.514	$4.940\ 10^{-2}$	14	$4.621\ 10^{-2}$	$3.184\ 10^{-4}$	6

We make the following observations:

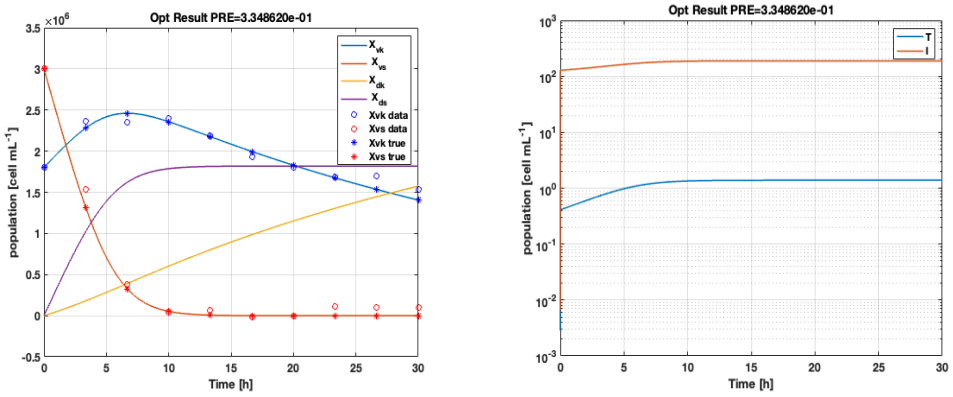
- Data noise significantly affects the parameter estimate and the Parameter Relative Error (PRE). Inaccurate measurements can lead to inaccurate parameter estimates and high PRE values. Therefore, it is essential to consider the effect of noise on the data when performing parameter estimation.
- Scaling the data is a valid strategy to improve the robustness of the estimates, especially in the presence of high noise. Scaling the data can help reduce the impact of measurement errors on the parameter estimates, resulting in more accurate and reliable results. The use of scaled data can improve the robustness of the optimization algorithm and increase the chances of obtaining accurate parameter estimates.
- The residual alone does not provide sufficient information about the precision of the parameter estimates. While the relative residual can indicate the quality of the fit between the model and the data, it is not a reliable indicator of the precision of the estimated parameters. This is evident in the fifth column of Table 3, where poorly estimated parameters exhibit similar residual values to better estimated parameters. Therefore, it is essential to consider other metrics, such as the Parameter Relative Error, or in the absence of a ground truth, compute the confidence intervals of the parameter estimates to assess their precision.

Figures 5 and 6 show two examples of data fit obtained by the estimated parameter with data scaling and initial viable killer/sensitive ratios of $\rho = 0.1$ (Figure 5) and $\rho = 0.6$ (Figure 6) for the case where the noise level is $\delta = 10^{-2}$.



(a) (b)

Figure 5. Case $\delta = 10^{-2}$, $\rho = 0.1$. (a) $X_{vk}, X_{vs}, X_{dk}, X_{ds}$ fitted (lines), X_{vk}, X_{vs} true (stars) and noisy data (circles). (b) Fitted T, I.



(a) (b)

Figure 6. Case $\delta = 10^{-2}$, $\rho = 0.6$. (a) $X_{vk}, X_{vs}, X_{dk}, X_{ds}$ fitted (lines), X_{vk}, X_{vs} true (stars) and noisy data (circles). (b) Fitted T, I.

Estimation of the initial value Expanding on the previous example, we can extend the methodology to cases where the initial data are unknown and even to scenarios involving more than one parameter. In this scenario, synthetic measurements z are obtained by evaluating the viable killer population $Y = X_{vk}$ at $m_t + 1$ measurement time values t_i, $i = 0, \dots, m_t$, and adding normalized random noise. In the current example, we consider the unknown parameters $u_1 = \mu_{vs}$ and $u_2 = X_{vk}(0)$. Therefore, we have the corresponding sensitivity components:

$$W_1 \equiv \frac{\partial Y}{\partial u_1}, \qquad W_2 \equiv \frac{\partial Y}{\partial u_2}.$$

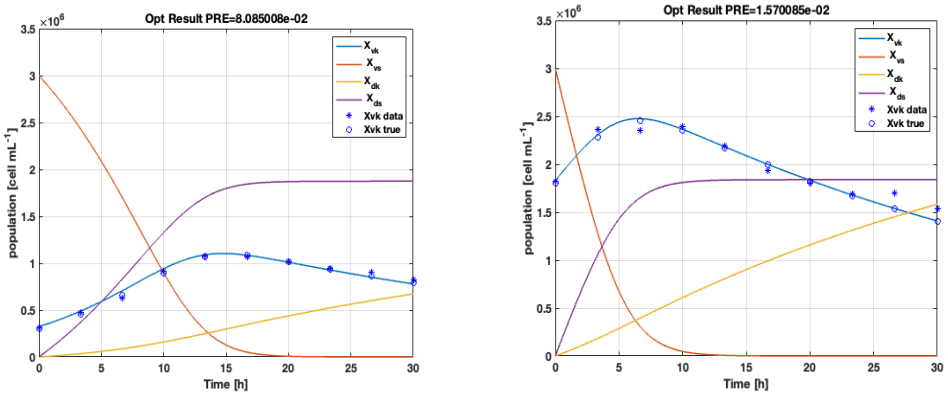
This leads us to the following augmented IVP problem:

$$\begin{cases} \dot{Y} = F(Y(t), \mathbf{u}) \\ \dot{W}_1 = F_{u_1}(Y, \mathbf{u}) + F_Y(Y, \mathbf{u}) W_1, \\ \dot{W}_2 = F_Y(Y, \mathbf{u}) W_2 \end{cases} \qquad \begin{cases} Y(t_0) = [u_2, Y_0(2), \dots, Y_0(6)]^\intercal \\ W_1(t_0) = \mathbf{0} \\ W_2(t_0) = [1, 0, 0, 0, 0, 0]^\intercal \end{cases} \tag{12}$$

Here, $\mathbf{u} = [u_1, u_2]^\top$ and $F_{u_2}(Y, \mathbf{u}) = \mathbf{0}$ since F does not depend on the initial vector Y_0. With minimal coding effort, it is possible to estimate both μ_{vs} and the initial viable killer population $X_{vk}(0)$. The estimation procedure was carried out with 10 observation times and unknown parameters μ_{vk} and $X_{vs}(0)$. The data noise level was set at $\delta = 0.01$ and data scaling was applied. The results of the estimation are presented in Table 4, which includes the performance measures PRE0, PRE, NRes, and nfevals. In Figures 7(a) and 7(b), the results are visualized for the case $\delta = 10^{-2}$. The lines depict the fitted values of $X_{vk}, X_{vs}, X_{dk}, X_{ds}$, while the stars represent the noisy data for X_{vk}. Figure 7(a) corresponds to $\rho = 0.1$ with an initial PRE0 value of 0.1, and Figure 7(b) corresponds to $\rho = 0.6$ with an initial PRE0 value of 0.01.

Table 4. Estimation test with 10 observation times and unknown parameters $\mu_{vk}, X_{vs}(0)$. Data noise $\delta = 0.01$ and data scaling.

ρ	PRE0	PRE	NRes	nfevals
0.1	0.1	$8.085\ 10^{-2}$	$3.030\ 10^{-2}$	45
0.6	0.01	$1.570\ 10^{-2}$	$7.796\ 10^{-3}$	45



(a) (b)

Figure 7. Case $\delta = 10^{-2}$. (lines) $X_{vk}, X_{vs}, X_{dk}, X_{ds}$ fitted, (stars)X_{vk}, X_{vs} noisy data. (a) $\rho = 0.1$, $PRE0 = 0.1$ (b) $\rho = 0.6$, $PRE0 = 0.01$.

Parameter Estimation: Uncertainty Analysis The inclusion of synthetic data with added noise serves as a valuable approach to assess the effectiveness of the model fitting procedure. However, it is important to acknowledge that computing the parameter relative error (PRE) becomes impractical when fitting models to real experimental data, as the true parameters are generally unknown. To overcome this limitation, the Standard Error provides useful information about the quality of the parameter estimates and enables the computation of confidence intervals (Smith (2013)). The procedure to evaluate the 95% confidence interval for each parameter is outlined in (13), where $\mathbf{u} \in \mathbb{R}^p$ represents the vector of estimated parameters (e.g., in the current

example, $p = 2$), and R is the residual vector.

$$
\begin{aligned}
&(1) \quad \sigma^2 = \frac{1}{n-p} R^\intercal R \quad \text{(Error variance estimate)} \\
&(2) \quad \mathcal{X}_{i,k} = \frac{\partial Y(t_i)}{\partial u_k} \quad \text{(Sensitivity Matrix)}, \ i = 1, \ldots, n, \ k = 1, \ldots, p \\
&(3) \quad M = \left[\mathcal{X}^\top \mathcal{X}\right]^{-1}, \quad V = \sigma^2 M \quad \text{(Covariance Matrix estimate)} \\
&(4) \quad k = 1, \ldots, p \\
&\qquad (4.1) \quad \delta_k = M_{k,k}, \quad SE = \sigma\sqrt{\delta_k}, \qquad \text{(Standard Error)} \\
&\qquad (4.2) \quad \left[u_k - t_{n-p,1-\alpha/2} SE, u_k + t_{n-p,1-\alpha/2} SE\right] \quad \text{(Confidence intervals)}
\end{aligned}
\tag{13}
$$

Here, $t_{n-p,1-\alpha/2}$ refers to the Student's t distribution with $\alpha = 0.1$, obtained using the Matlab function `tinv`. It is important to note that the sensitivity matrix $\mathcal{X}$ is obtained by evaluating the components W_1 and W_2 using the augmented system (12) when $\mathbf{u}$ represents the estimated parameter. The parameter estimates, together with the confidence intervals, are presented in Table 5. The standard error (SE) indicates the uncertainty associated with each parameter estimate, and the confidence intervals provide a range of plausible values within which the true parameter value is likely to fall. For the first test, with $\rho = 0.1$, the standard error (SE) associated with u_1

Table 5. Parameters estimates (Par), Standard error (SE) and confidence intervals for the tests in Table 4, with Student's t distribution $t_S = t_{8,0.95}$

ρ	Parameter	SE	$Par - t_S SE$	Par	$Par + t_S SE$
0.1	Par $= u_1, (\mu_{vs})$	$8.209\ 10^{-3}$	$-7.757\ 10^{-3}$	$7.509\ 10^{-3}$	$2.277\ 10^{-2}$
	Par $= u_2, (Y_{vk}(0))$	$4.801\ 10^{4}$	$2.350\ 10^{5}$	$3.243\ 10^{5}$	$4.135\ 10^{5}$
0.6	Par $= u_1, (\mu_s)$	$4.159\ 10^{-2}$	$-6.915\ 10^{-2}$	$8.191\ 10^{-3}$	$8.553\ 10^{-2}$
	Par $= u_2, (Y_{vk}(0))$	$1.689\ 10^{5}$	$1.514\ 10^{6}$	$1.828\ 10^{6}$	$2.142\ 10^{6}$

is $8.209\ 10^{-3}$, indicating a relatively low level of uncertainty in the estimate. The confidence interval ranges from $-7.757\ 10^{-3}$ to $2.277\ 10^{-2}$ for u_1, indicating a relatively low level of uncertainty in the estimate. For u_2, the SE is $4.801\ 10^4$, indicating a relatively large uncertainty. The confidence interval spans from $2.350\ 10^5$ to $4.135\ 10^5$. This wide interval suggests that there is substantial uncertainty about the true value of u_2. In the second test, with $\rho = 0.6$, the SE for u_1 is $4.159\ 10^{-2}$, indicating a relatively higher level of uncertainty compared to the first test, as confirmed by the larger confidence interval ranging from $-6.915\ 10^{-2}$ to $8.553\ 10^{-2}$. For u_2, we have a similar situation as in the case of $\rho = 0.1$. The SE is $1.689\ 10^5$, indicating a relatively large uncertainty. The confidence interval for u_2 ranges from $1.514\ 10^6$ to $2.142\ 10^6$. Again, this wide interval indicates considerable uncertainty in the estimate of u_2. The standard errors and confidence intervals provide valuable insights into the uncertainty associated with the parameter estimates. They allow us to assess the reliability and range of plausible values for each parameter, providing a measure of confidence in the estimation results.

Final Remarks In an ideal progression, starting from the initial test with the identification of a single parameter, one could proceed to include the initial data and eventually extend to all parameters.

In the real-world context, the model's validity is assessed by utilizing the assumed model and fitting parameters obtained from a specific data set, commonly referred to as the training data set. Subsequently, the next step involves evaluating how well these parameters replicate or predict the behavior observed in an independent data set, known as the test or validation set.

5. Conclusion

The results obtained provide a foundation for various potential assignments. These include exploring the impact of measurement points on estimation accuracy, evaluating the influence of Jacobian calculation on performance, comparing optimization algorithms, and addressing uncertainties in nonlinear models. From the students' perspective, engaging with practical scientific papers enhances their understanding of mathematical concepts and enables the creation of validation test problems. From the teachers' viewpoint, although the approach involves higher costs and greater oversight, it offers greater satisfaction and promotes meaningful discussions. Students often propose personal solutions to practical problems and the teacher has to be able to understand and discuss. In conclusion, assignments based on scientific papers contribute to students' preparedness for real-world applications and deepen their comprehension of the subject matter.

Appendix A. Assignment Text

A.1. Microbial interactions

Papers Kuchen et al. (2022) and Pommier et al. (2005) present advancements in the mathematical modeling of the fermentation process from grape juice to wine. From these references, select one of the proposed models and implement it in your Matlab script, using the provided parameters.

- Discuss the problem of stiffness and evaluate the efficiency of your Matlab ODE solvers compared to the built-in solvers.
- Generate a dataset using your solver and conduct a parameter estimation experiment by introducing random noise from a normal distribution. Analyze the accuracy and efficiency of your results.

Disclosure statement

No potential conflict of interest was reported by the author.

References

Chavent, G. (2010). *Nonlinear least squares for inverse problems: theoretical foundations and step-by-step guide for applications.* Springer Science & Business Media.

Kuchen, B., Maturano, Y., Gil, R., Vazquez, F., & Scaglia, G. (2022). Kinetics and mathematical model of killer/sensitive interaction under different physicochemical conditions of must/wine: a study from a biological point of view. *Letters in Applied Microbiology*, *74*(5), 718–728.

Nocedal, J., & Wright, S. (2006). Springer series in operations research and financial engineering. In *Numerical optimization.* Springer New York.

Pommier, S., Strehaiano, P., & Delia, M.-L. (2005). Modelling the growth dynamics of interacting mixed cultures: a case of amensalism. *International journal of food microbiology*, *100*(1-3), 283–293.

Ramon-Portugal, F., Torija, M.-J., Mas, A., & Guillamon, J.-M. (1997). Kinetic study and mathematical modelling of the killer and sensitive s. cerevisiae strains growing in mixed culture. *Bioprocess Engineering*, *17*, 361–366.

Smith, R. C. (2013). *Uncertainty quantification: theory, implementation, and applications* (Vol. 12). Siam.