

# Can the Kuznetsov Model Replicate and Predict Cancer Growth in Humans?

Mohammad El Wajeh<sup>1</sup>, Falco Jung<sup>1</sup>, Dominik Bongartz<sup>1</sup>, Chrysoula Dimitra Kappatou<sup>2</sup>,  
Narmin Ghaffari Laleh<sup>3</sup>, Alexander Mitsos<sup>4,1,5,\*</sup>, and Jakob Nikolas Kather<sup>3,6,\*</sup>

<sup>1</sup>*Process Systems Engineering (AVT.SVT), RWTH Aachen University, 52074 Aachen, Germany*

<sup>2</sup>*Faculty of Engineering, Department of Computing, Imperial College London, London SW7 2AZ, UK*

<sup>3</sup>*Department of Medicine III, University Hospital RWTH Aachen, 52074 Aachen, Germany*

<sup>4</sup>*JARA-CSD, 52056 Aachen, Germany*

<sup>5</sup>*Energy Systems Engineering (IEK-10), Forschungszentrum Jülich, 52425 Jülich, Germany*

<sup>6</sup>*Medical Oncology, National Center for Tumor Diseases, University Hospital Heidelberg, 69120 Heidelberg, Germany*

\* *Shared last authorship and corresponding authors: [amitsos@alum.mit.edu](mailto:amitsos@alum.mit.edu), [jkather@ukaachen.de](mailto:jkather@ukaachen.de)*

## Abstract

Several mathematical models to predict tumor growth over time have been developed in the last decades. A central aspect of such models is the interaction of tumor cells with immune effector cells. The Kuznetsov model (Kuznetsov et al. (1994), *Bulletin of Mathematical Biology*, vol. 56, no. 2, pp. 295–321) is the most prominent of these models and has been used as a basis for many other related models and theoretical studies. However, none of these models have been validated with large-scale real-world data of human patients treated with cancer immunotherapy. In addition, parameter estimation of these models remains a major bottleneck on the way to model-based and data-driven medical treatment. In this study, we quantitatively fit Kuznetsov’s model to a large dataset of 1472 patients, of which 210 patients have more than six data points, by estimating the model parameters of each patient individually. We also conduct a global practical identifiability analysis for the estimated parameters. We thus demonstrate that several combinations of parameter values could lead to accurate data fitting. This opens the potential for global parameter estimation of the model, in which the values of all parameters are fixed for all patients. Furthermore, by omitting the last two or three data points, we show that the model can be extrapolated and predict future tumor dynamics. This paves the way for a more clinically relevant application of mathematical tumor modeling, in which the treatment strategy could be adjusted in advance according to the model’s future predictions.

**Keywords**— Mathematical oncology, tumor growth modelling, tumor growth prediction, parameter estimation, parameter identifiability analysis

## 1. Introduction

Cancer immunotherapy with immune checkpoint inhibitors has revolutionized the treatment of patients with solid tumors in the last ten years. In addition to chemotherapy and molecularly-targeted therapy, immunotherapy provides a new set of tools for the oncology toolkit [1]. In several tumor types such as melanoma, non-small cell lung cancer (NSCLC), and genito-urinary tumors, immunotherapy has markedly improved the average life expectancy of patients with advanced disease. Both laboratory and clinical experiments have verified the importance of the immune system in fighting cancer [2]–[4]. Patients who suffer from acquired immunodeficiency syndrome (AIDS) are very susceptible to having some rare forms of cancer [2], [5]. This also shows the significant role the immune system plays against cancer.

One of the fundamental problems in treating patients with cancer immunotherapy is the lack of predictive biomarkers. Ideally, before the treatment begins, patients could be selected for immunotherapy, but existing biomarkers fail to deliver a high predictive value in most tumor types [6]. In addition, most patients who initially respond to immunotherapy progress experience a relapse: the tumor later on develops immune escape mechanisms due to evolutionary pressure. Forecasting the time of relapse or treatment resistance is of high practical relevance [7], [8]. However, predictions of such changes in the tumor behavior are currently not possible in clinical routine. The main problem is that most biomarkers such as tumor mutational burden (TMB) are static, i.e. they are measured at a given time point but are not dynamically updated as the tumor evolves.

In other complex systems such as financial markets [9], climate systems [10] or complex industrial processes [11], differential equation models can provide a prediction of the behavior of the system over time. By analogy, in oncology, a number of mathematical models to predict tumor growth over time have been developed in the last decades [12]. Most notably, multiple of these models explicitly include the interactions of tumors with the immune system and are therefore in principle suited to model response and resistance to cancer immunotherapy [13]. Boer and Hogeweg (1986) modeled the cellular immune reaction to tumors. They demonstrated that small doses of antigens lead to tumor dormancy [2], [14]. Kirschner and Panetta (1998) linked the oscillations in the tumor size and the long-term tumor regression to the dynamics among immune cells, tumor cells, and Interleukin-2 [2], [15]. The most prominent of these models was presented by Kuznetsov et al. [16] in 1994. Kuznetsov’s model has served as a blueprint for many other related models [2], [17], [18] and has been investigated in several theoretical studies [2], [19]–[21].

However, none of these established oncological models are currently being used in the clinic. What is more, very few of these models have been systematically fitted to actual clinical data. While some studies have fitted models to murine tumors on a small scale [2], [22], [23], the pronounced differences between mice and humans preclude the transfer of such insights to real-world cancer patients [24].

The structures of these mathematical models are well defined [16], [25]–[27]. However, in the complex biological environment of cells, little is known about the associated parameters and kinetic constants. The parameter values are essential for quantitative modeling and prediction of cancer progression. In mechanistic models, one can integrate the data from various experimental procedures and sources, and design in-silico experiments to generate hypotheses for underlying mechanisms [28], [29]. By fitting the model to the experimental data, we reverse-engineer the parameters of the system. Parameter estimation of mathematical cancer models remains a major bottleneck on the way to model-based and data-driven medical treatment of the future.

In this study, we use a mathematical model based on Kuznetsov’s model to characterize the interactions between the growing tumor and the immune system, and aim to fill this conceptual gap in the literature. We use a large dataset of thousands of cancer patients who underwent cancer immunotherapy as part of clinical trials. We then investigate how well the model can represent the actual tumor volume changes over time in these patients. After estimating the parameters of the model, we conduct an identifiability analysis to examine the uniqueness of the estimated parameters (i.e., whether we have over-fitting). Finally, we investigate if the model can be used to forecast treatment response or relapse under immunotherapy.

The remainder of the article is structured as follows. First, we briefly present the acquisition of patients data and its pre-treatment in Section 2. In Section 3, we then provide the model and all the methods used to estimate model parameters and fit the data, conduct parameter identifiability analysis and extrapolate the model for tumor growth prediction. Finally, before drawing conclusions in Section 5, we present and discuss the obtained results in Section 4.

## 2. Data Acquisition

We briefly provide here the declaration and sources of the experimental data. For more details, please refer to [30].

### 2.1. Declarations and Data Sharing

We followed the Declaration of Helsinki and International Ethical Guidelines for Biomedical Research Involving Human Subjects developed by the Council for International Organizations of Medical Sciences (CIOMS). In this study, we used a publicly available set of anonymized patient data shared by [30], which is originally derived from five large clinical trials, as we describe below. Patients gave their informed consent for data analyses as part of the original clinical trials. No specific ethical approval was sought or required for this retrospective analysis of a publicly available dataset.

### 2.2. Data Measurement and Pre-treatment

The original data for this study is from five clinical trials which were designed to evaluate the efficiency of Atezolizumab (an immune checkpoint inhibitor). Table 1 shows the original number of patients and their treatment arms. Four out of these five trials evaluated the effect of Atezolizumab on NSCLC and cohort GO29293 reported this efficiency on bladder cancer. In two of the cohorts (GO28753, GO28915), patients' responses to Atezolizumab treatment were compared to the outcome of the second treatment arm who received Docetaxel (a chemotherapy drug) as treatment. In all the trials, the longest and shortest diameter (LD and SD) of the target and non-target lesions (measured manually based on the CT scans) alongside the time intervals are reported. In this study, we used the anonymized and publicly available subset of data from [30]. This subset is created by selecting the patients with three or more measurement points. For each of these patients, only the LD measurement for one target lesion has been selected. For this reason, the total number of original patients of clinical trials has been decreased from 2693 to 1472. All the pre-processing details for the data generation has been described in more details in [30]. Moreover, before using the data we pre-processed it by first removing repetitive and null inputs. Following [30]–[32], we converted the measured LD in mm to the number of tumor cells (TC) by considering  $4/3 \times (8 \times 10^{-6}) \text{ mm}^3/\text{TC}$ . We also considered only patients with more than six net measurements. Figure 1 shows the number of patients per study and arm before and after data pre-processing. Finally, the total number of patients has been reduced from 1472 to 210.

Table 1: Description of the original data.

Study ID	Cancer Type	No. of Patients	Treatment
GO28625 [33]	NSCLC	138	Atezolizumab
GO28753 [34]	NSCLC	287	Atezolizumab/Docetaxel
GO28754 [35]	NSCLC	657	Atezolizumab
GO28915 [36]	NSCLC	1182	Atezolizumab/Docetaxel
GO29293 [37]	Bladder Cancer	429	Atezolizumab

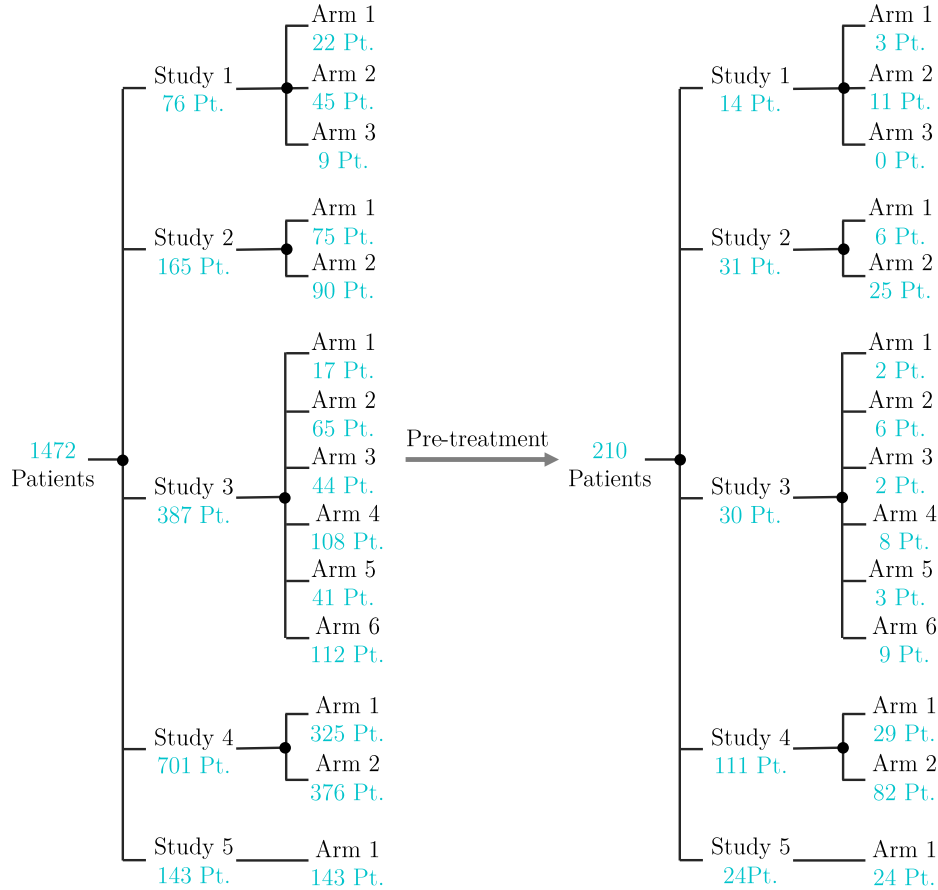


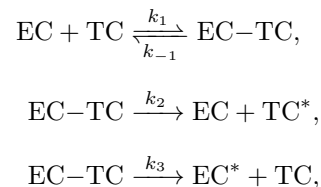
Figure 1: Number of patients (Pt.) considered per studies and arms, before and after data pre-processing.

### 3. Methods

In this section, we introduce the mathematical model of Kuznetsov [16] with slight modifications. Afterwards, we show the optimization formulations for: the estimation of the model parameters by fitting the clinical data and for the parameter identifiability analysis. We finally investigate the extrapolation capabilities of the model.

#### 3.1. Mathematical Tumor Model

To predict the frequently observed phenomena in clinics like tumor dormancy and tumor size oscillation, the tumor mathematical model has to include terms related to the response of the immune system. The inclusion of the entire immune system in the mathematical model can be very difficult [38]. The anti-tumor immune response has highly nonlinear dynamics which are complicated and not well understood. Therefore, models that describe the immune system response to tumor presence should necessarily focus on those elements of the immune system that have the highest effects on tumor dynamics [2]. Kuznetsov's model [16] describes the response of the cytotoxic T lymphocyte (CTL) to the growth of an immunogenic tumor. Usually, a cell-mediated immune response to a tumor takes place. The cytotoxic T lymphocytes and natural killer (NK) cells play the main role. The model considers immunogenic TC that are attacked by cytotoxic effector cells (EC). The EC can be, for example, CTL or NK cells. The model takes into account the possibility of EC inactivation as well as the infiltration of the TC by EC. TC and EC interaction is described through the following reactions:



where EC–TC denotes conjugates of effector and tumor cells; and  $EC^*$  and  $TC^*$  are the inactivated effector and lithally-hit tumor cells, respectively. We define  $E$ ,  $T$ ,  $C$ ,  $E^*$ ,  $T^*$  as the number of EC, TC, EC–TC conjugates,  $EC^*$ , and  $TC^*$ , respectively. The non-negative kinetic parameters,  $k_1$ ,  $k_{-1}$ ,  $k_2$  and  $k_3$ , describe the rates of the interactions. EC–TC conjugates can reversibly decompose without damaging the cells with the kinetic rate  $k_{-1}$ . However, they can also irreversibly result in  $EC^*$  or  $TC^*$  with respective kinetic rates  $k_2$  and  $k_3$ . The following system of nonlinear differential-algebraic equations describes those interactions, which is a slightly simplified version of the model of Kuznetsov [16].

$$\left. \frac{dE}{dt} \right|_t = s + F(C(t), T(t)) - hE(t) - k_1E(t)T(t) + (k_{-1} + k_2)C(t), \quad (1a)$$

$$\left. \frac{dT}{dt} \right|_t = aT(t) - k_1E(t)T(t) + (k_{-1} + k_3)C(t), \quad (1b)$$

$$\left. \frac{dC}{dt} \right|_t = k_1E(t)T(t) - (k_{-1} + k_2 + k_3)C(t), \quad (1c)$$

$$F(C(t), T(t)) = \frac{fC(t)}{g + T(t)}. \quad (1d)$$

The equations describing the rate of change of  $E^*$  and  $T^*$  are not included in the system because they are irreversibly formed and thus have no effect on the other variables, and our target is to model  $T$  and  $E$  only. In [16], the model includes a sink term in the rate of change equation of  $T$  that represents TC growth limitation due to biological environment conditions. It considers, for example, resources competition like oxygen and substrates. We do not consider this term because growth limitations of even high initial  $T$  are associated with high rates of cytotoxic EC accumulation as well as the absence of their activity suppression by TC [39].

The rate of flow of mature EC to TC localization area is characterized by the generation term  $s$ . This rate is unaffected by the presence of TC. The destruction or migration of EC from the localization region of TC is represented by the elimination rate  $h$ . The model does not take into account any TC or EC–TC conjugates migration. Both multiplication and death of TC are included in parameter  $a$  that characterizes the maximum growth rate of TC population. The function  $F(C, T)$  represents the accumulation rate of the cytotoxic EC in the TC localization area due to tumor existence (stimulated accumulation), where  $f$  and  $g$  are positive constants. The EC accumulation,  $F(C, T)$ , is due to signals, like released cytokines, generated by the EC in EC-TC conjugates. Thus this stimulated accumulation has some maximum value when  $T$  becomes large.

Following the suggestion of Kuznetsov [16], we consider a quasi-steady-state assumption for (1c), i.e.,  $dC/dt|_t \approx 0$ . Thus,  $C \approx KET$ , where  $K = k_1/(k_2 + k_3 - k_{-1})$ . As a result, (1a) and (1b) become:

$$\left. \frac{dE}{dt} \right|_t = s + \frac{fKE(t)T(t)}{g + T(t)} - hE(t) - Kk_3E(t)T(t), \quad (2a)$$

$$\left. \frac{dT}{dt} \right|_t = aT(t) - Kk_2E(t)T(t). \quad (2b)$$

For further analysis and use of the model in parameter estimation and identifiability analysis, we use the same strategy of Kuznetsov [16] for non-dimensionalizing model equations. We non-dimensionalize (2a) and (2b) by considering concentration scales  $E_0 = 10^7$  cells and  $T_0 = 10^9$  cells for EC and TC, respectively [16]. We non-dimensionalize  $t$  by relating it to the deactivation rate of TC and introducing  $\tau = k_2KT_0t/100$ . The final model formulation is:

$$\left. \frac{dx}{d\tau} \right|_\tau = \sigma + \frac{\rho x(\tau)y(\tau)}{\eta + y(\tau)} - \delta x(\tau) - \mu x(\tau)y(\tau), \quad x|_{\tau=\tau_1} = x_1, \quad (3a)$$

$$\left. \frac{dy}{d\tau} \right|_\tau = \alpha y(\tau) - \frac{E_0}{T_0} x(\tau)y(\tau), \quad y|_{\tau=\tau_1} = y_1, \quad (3b)$$

where

$$\begin{aligned} x(\tau) &= \frac{E(t)}{E_0}, & y(\tau) &= \frac{T(t)}{T_0}, & \sigma &= \frac{s}{k_2KE_0T_0}, & \rho &= \frac{f}{k_2T_0}, \\ \eta &= \frac{g}{T_0}, & \mu &= \frac{k_3}{k_2}, & \delta &= \frac{h}{k_2KT_0}, & \text{and } \alpha &= \frac{a}{k_2KT_0}. \end{aligned}$$

Thus, the final model is composed of two ordinary differential equations (ODEs) with two variables,  $x$  and  $y$ , with their respective initial values,  $x_1$  and  $y_1$ , at the initial normalized time,  $\tau_1$ , and six unknown parameters,  $\sigma$ ,  $\rho$ ,  $\eta$ ,  $\mu$ ,  $\delta$ , and  $\alpha$ .

### 3.2. Data Fitting and Parameter Estimation

To determine the parameter values for the nonlinear system (3a) and (3b) that best describe the experimental data, we conduct a regression analysis in the nonlinear least-squares sense by minimizing the sum of the squared residuals. The considered residuals are the differences between the measured values of tumor lesion longest diameters (converted to tumor number of cells as previously discussed) and the ones calculated from the model. In the present contribution, we do not seek global values of parameters, i.e, the same values for all patients. Instead, we solve the optimization problem for each patient individually to identify the parameter values that best describe the data of that patient. This constitutes a first step to check whether the model can describe the experimental data at all and analyze the ranges of parameter values. We also estimate the initial normalized value of EC number,  $x_1$ , because it is unknown. In contrast, the initial normalized value of TC number,  $y_1$ , is provided experimentally and does not need to be estimated. The nonlinear least square problem for each patient  $j \in \mathbb{J} = \{1, 2, \dots, J\}$  is expressed as follows:

$$\begin{aligned} & \underset{\mathbf{p}_j, x_{1_j}}{\text{minimize}} && \sum_{i \in \mathbb{I}_j} y_{i_j}(\tau_{i_j}, \mathbf{p}_j, x_{1_j}) - \tilde{y}_{i_j}, \quad \forall i \in \mathbb{I}_j, \\ & \text{subject to} && \text{model equations (3a) and (3b),} \\ & && \mathbf{p}_j^l \leq \mathbf{p}_j \leq \mathbf{p}_j^u, \\ & && x_{1_j}^l \leq x_{1_j} \leq x_{1_j}^u, \end{aligned} \tag{OP1}$$

where  $J$  is the total number of patients considered. After data pre-processing,  $J = 210$  patients. The index  $i \in \mathbb{I}_j$  is for the observed experimental values, where  $\mathbb{I}_j = \{1, 2, \dots, N_j\}$  with  $N_j$  being the total number of observed values for patient  $j$ . The non-dimensionalized model-predicted and observed values of TC number at the normalized time  $\tau_{i_j}$  are  $y_{i_j}(\tau_{i_j}, \mathbf{p}_j, x_{1_j})$  and  $\tilde{y}_{i_j}$ , respectively. The initial value of the non-dimensionalized EC number at  $\tau_{1_j}$ ,  $x_{1_j}$ , has lower and upper bounds  $x_{1_j}^l$  and  $x_{1_j}^u$ , respectively. The vector  $\mathbf{p}_j$  contains the non-dimensionalized parameter values of the model equations, (3a) and (3b), for patient  $j$ ,  $\mathbf{p}_j = [\sigma_j, \mu_j, \delta_j, \alpha_j, \rho_j, \eta_j]$ . The lower and upper bounds of the components of  $\mathbf{p}_j$  are the components of vectors  $\mathbf{p}_j^l$  and  $\mathbf{p}_j^u$ , respectively. The decision variables of the optimization problem are thus the components of  $\mathbf{p}_j$  and  $x_{1_j}$ . The dynamic optimization problem (OP1) is nonconvex and nonlinear, and can thus have multiple (suboptimal) local solutions. Hence, global optimization techniques are required to guarantee the global optimal solution,  $\mathbf{p}_j^{\text{opt}}$  and  $x_{1_j}^{\text{opt}}$ .

### 3.3. Parameter Identifiability Analysis

Parameter identifiability analysis determines if model parameters can be uniquely estimated [40]. Different definitions of identifiability analysis are available in the literature. Miao et al. [41] reviewed several methods of parameter identifiability analysis for nonlinear ODE models and distinguished between different methodologies including structural and practical identifiability analyses. In the former analysis, one determines if a given structure of a model allows the realization of unique parameters when certain measured variables are provided [40]. However, it only provides necessary conditions for identifiability because it does not take into consideration parameters precision [42], [43]. In contrast, practical identifiability aims to predict confidence intervals for the estimated parameters [42], [44]. It can be conducted locally (in the neighborhood of the estimated parameter values) or globally over the entire range of values. We here carry out the latter analysis and evaluate it globally to improve the confidence in the parameter values that are determined by solving (OP1).

We conduct the global practical identifiability analysis by determining the smallest box that contains the so-called feasible parameter set  $P_{e_j}$  for each patient  $j$ , as suggested in [42]. This set includes all values of parameters in  $\mathbf{p}_j$  for which the differences between model predictions  $y_{i_j}(\tau_{i_j}, \mathbf{p}_j, x_{1_j}^{\text{opt}})$  and optimal model predictions  $y_{i_j}^{\text{opt}}$  (determined by solving (OP1)) fall within certain defined bounds  $\forall i \in \mathbb{I}_j$ , that is

$$P_{e_j} = \{\mathbf{p}_j \in P_j \mid -\varepsilon y_{i_j}^{\text{opt}} \leq y_{i_j}(\tau_{i_j}, \mathbf{p}_j, x_{1_j}^{\text{opt}}) - y_{i_j}^{\text{opt}} \leq \varepsilon y_{i_j}^{\text{opt}}\}, \quad \forall j \in \mathbb{J}, \tag{4}$$

where  $y_{i_j}^{\text{opt}} = y_{i_j}(\tau_{i_j}, \mathbf{p}_j^{\text{opt}}, x_{1_j}^{\text{opt}})$ ,  $\varepsilon$  is the percentage of deviation, and  $P_j$  is the set of parameter values in  $\mathbf{p}_j$  bounded by the components of  $\mathbf{p}_j^l$  and  $\mathbf{p}_j^u$  defined in (OP1). The set  $P_{e_j}$  is depicted in dark gray as

shown in Figure 2 for the case of a two-dimensional vector  $\mathbf{p}_j$ . We approximate the nonconvex set  $P_{e_j}$  by a rectangular box (light gray color), whose edges are formed by the extreme values of the elements of  $\mathbf{p}_j$  ( $p_{j_1}^{\min}$ ,  $p_{j_1}^{\max}$ ,  $p_{j_2}^{\min}$ ,  $p_{j_2}^{\max}$ ). We determine these extreme values by solving a series of constrained dynamic optimization problems. For a vector  $\mathbf{p}_j$  that consists of  $K$  parameters, the optimization problem for parameter number  $k$  is formulated  $\forall j \in \mathbb{J}$ , and  $\forall k \in \mathbb{K}$  as follows [42], [45]:

$$\begin{aligned}
 & \underset{\mathbf{p}_j^{\min}, \mathbf{p}_j^{\max}}{\text{maximize}} && p_{j_k}^{\max} - p_{j_k}^{\min}, \\
 & \text{subject to} && -\varepsilon y_{i_j}^{\text{opt}} \leq y_{i_j}(\tau_{i_j}, \mathbf{p}_j^{\max}, x_{1_j}^{\text{opt}}) - y_{i_j}^{\text{opt}} \leq \varepsilon y_{i_j}^{\text{opt}}, \forall i \in \mathbb{I}_j, \\
 & && -\varepsilon y_{i_j}^{\text{opt}} \leq y_{i_j}(\tau_{i_j}, \mathbf{p}_j^{\min}, x_{1_j}^{\text{opt}}) - y_{i_j}^{\text{opt}} \leq \varepsilon y_{i_j}^{\text{opt}}, \forall i \in \mathbb{I}_j, \\
 & && \text{model equations (3a) and (3b),} \\
 & && \mathbf{p}_j^{\text{opt}} \leq \mathbf{p}_j^{\max} \leq \mathbf{p}_j^{\text{u}}, \\
 & && \mathbf{p}_j^{\text{l}} \leq \mathbf{p}_j^{\min} \leq \mathbf{p}_j^{\text{opt}},
 \end{aligned} \tag{OP2}$$

with  $\mathbb{K} = \{1, 2, \dots, K\}$  and  $K = 6$ . We set  $\varepsilon$  to 20%, which means that the differences between model predictions for  $\mathbf{p}_j = \mathbf{p}_j^{\max}$  and  $\mathbf{p}_j = \mathbf{p}_j^{\min}$ , and optimal model predictions for  $\mathbf{p}_j = \mathbf{p}_j^{\text{opt}}$  fall within 20% of the optimal prediction values  $\forall i \in \mathbb{I}_j$ . For each parameter  $p_{j_k}$  in  $\mathbf{p}_j$ , its minimum value  $p_{j_k}^{\min}$  in  $\mathbf{p}_j^{\min}$  and its maximum value  $p_{j_k}^{\max}$  in  $\mathbf{p}_j^{\max}$  are determined by solving (OP2). Therefore, (OP2) is solved  $K$  times for each patient  $j \in \mathbb{J}$ . As a result, the approximation of  $P_{e_j}$  is determined, and its edges are the elements of  $\mathbf{p}_j^{\min}$  and  $\mathbf{p}_j^{\max}$ . The set  $P_{e_j}$  allows for the determination of confidence regions of the estimated parameter values in  $\mathbf{p}_j^{\text{opt}}$ . When  $P_{e_j}$  covers a large space in the direction of  $p_{j_k}$ , the estimated parameter  $p_{j_k}^{\text{opt}}$  is not identifiable, in the sense that there is a large range of values for  $p_{j_k}$  that could lead to a good fit. When the set covers a small space,  $p_{j_k}^{\text{opt}}$  is identifiable because the parameter is determined to a sufficient accuracy. To decide on identifiability a certain threshold is thus needed. We here do not define a cutoff, we rather analyze the identifiability qualitatively.

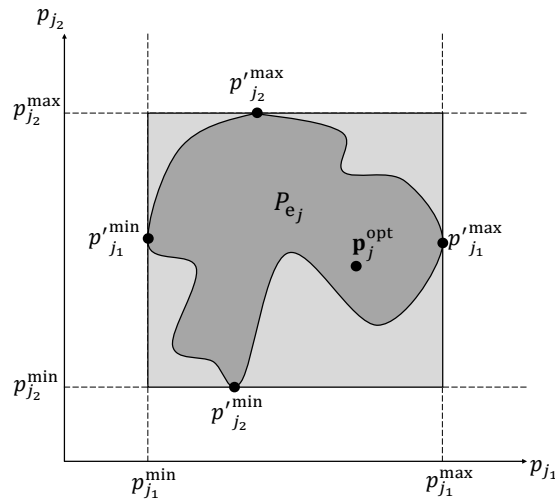


Figure 2: Illustration of parameter identifiability via global confidence intervals (based on [42]). The dark gray region indicates the feasible set  $P_{e_j}$  that includes all values of parameters in the two-dimensional vector  $\mathbf{p}_j$  for which the differences between model predictions  $y_{i_j}(\tau_{i_j}, \mathbf{p}_j, x_{1_j}^{\text{opt}})$  and optimal model predictions  $y_{i_j}^{\text{opt}}$  fall within certain defined bounds. The light gray region shows the rectangular box that conservatively approximates this feasible set, where the box edges are the extreme values of the elements of  $\mathbf{p}_j$  ( $p_{j_1}^{\min}$ ,  $p_{j_1}^{\max}$ ,  $p_{j_2}^{\min}$ ,  $p_{j_2}^{\max}$ ).



### 3.4. Tumor Growth Prediction

Beyond being able to reproduce experimental data a posteriori, a more clinically relevant application of mathematical tumor modeling would be if the model was able to predict tumor growth. This could lead to model-based tumor treatment, as the supplied doses to patients could be adjusted in advance according to model predictions. In Subsection 3.2, we fitted the model to all experimental data points when estimating the parameters. In order to compare model extrapolation capabilities and future predictions to the clinical data, we now do not include the last  $\zeta$  points of the data when fitting the model and solving (OP1). For each patient  $j$ , we solve an optimization problem similar to (OP1), but using only the data points in the set  $\mathbb{I}_j^{\text{ext}}$  instead of  $\mathbb{I}_j$ , where  $\mathbb{I}_j^{\text{ext}} = \{1, 2, \dots, N_j - \zeta\}$ . The optimal values of the decision variables obtained from this problem are called  $\mathbf{p}_j^{\text{ext}}$  and  $x_{1_j}^{\text{ext}}$ . We then integrate the model equations for  $\mathbf{p}_j = \mathbf{p}_j^{\text{ext}}$  and  $x_{1_j} = x_{1_j}^{\text{ext}}$  from  $\tau_{1_j}$  to  $\tau_{(N_j)_j}$ , thus extrapolating beyond the data used for fitting. The results of the integrated  $y_{i_j}^{\text{ext}} = y_{i_j}(\tau_{i_j}, \mathbf{p}_j^{\text{ext}}, x_{1_j}^{\text{ext}})$  between  $\tau_{(N_j-\zeta)_j}$  and  $\tau_{(N_j)_j}$  are the extrapolated part of the model that can be compared with the remaining  $\zeta$  data points to gauge the extrapolation capabilities. We consider two model extrapolation cases in which  $\zeta$  is equal to two and three.

Moreover, we formulate another optimization problem to investigate how far model extrapolation could deviate from the actual values. We aim to find two “extreme-case” lines that are designed to be as far away from each other at the final time point ( $\tau_{(N_j)_j}$ ) while both being within some  $\theta$  tolerance of the found optimal fit ( $y_{i_j}^{\text{ext}}$ ) for the fitted time before extrapolation starts. For this, we solve the following optimization problem  $\forall j \in \mathbb{J}$ :

$$\begin{aligned}
 & \underset{\mathbf{p}_j^{\text{upp}}, \mathbf{p}_j^{\text{low}}, x_{1_j}^{\text{upp}}, x_{1_j}^{\text{low}}}{\text{maximize}} && y_{(N_j)_j}(\tau_{(N_j)_j}, \mathbf{p}_j^{\text{upp}}, x_{1_j}^{\text{upp}}) - y_{(N_j)_j}(\tau_{(N_j)_j}, \mathbf{p}_j^{\text{low}}, x_{1_j}^{\text{low}}), \\
 & \text{subject to} && -\theta y_{i_j}^{\text{ext}} \leq y_{i_j}(\tau_{i_j}, \mathbf{p}_j^{\text{upp}}, x_{1_j}^{\text{upp}}) - y_{i_j}^{\text{ext}} \leq \theta y_{i_j}^{\text{ext}}, \quad \forall i \in \mathbb{I}_j^{\text{ext}}, \\
 & && -\theta y_{i_j}^{\text{ext}} \leq y_{i_j}(\tau_{i_j}, \mathbf{p}_j^{\text{low}}, x_{1_j}^{\text{low}}) - y_{i_j}^{\text{ext}} \leq \theta y_{i_j}^{\text{ext}}, \quad \forall i \in \mathbb{I}_j^{\text{ext}}, \\
 & && \text{model equations (3a) and (3b),} \\
 & && \mathbf{p}_j^{\text{l}} \leq \mathbf{p}_j^{\text{upp}} \leq \mathbf{p}_j^{\text{u}}, \\
 & && \mathbf{p}_j^{\text{l}} \leq \mathbf{p}_j^{\text{low}} \leq \mathbf{p}_j^{\text{u}}, \\
 & && x_{1_j}^{\text{l}} \leq x_{1_j}^{\text{upp}} \leq x_{1_j}^{\text{u}}, \\
 & && x_{1_j}^{\text{l}} \leq x_{1_j}^{\text{low}} \leq x_{1_j}^{\text{u}},
 \end{aligned} \tag{OP3}$$

where  $\mathbf{p}_j^{\text{upp}}$  and  $\mathbf{p}_j^{\text{low}}$  are the vectors that contain the parameter values for the upper and lower “extreme-case” model extrapolation deviations, respectively. The initial values of the non-dimensionalized EC number at  $\tau_{1_j}$  for these upper and lower deviation cases are  $x_{1_j}^{\text{upp}}$  and  $x_{1_j}^{\text{low}}$ , respectively. We set  $\theta$  to 10%. By integrating (3a) and (3b) from  $\tau_{1_j}$  to  $\tau_{(N_j)_j}$ , we get  $y_{i_j}^{\text{upp}} = y_{i_j}(\tau_{i_j}, \mathbf{p}_j^{\text{upp}}, x_{1_j}^{\text{upp}})$  and  $y_{i_j}^{\text{low}} = y_{i_j}(\tau_{i_j}, \mathbf{p}_j^{\text{low}}, x_{1_j}^{\text{low}})$ , which allow the comparison of these two “extreme-case” extrapolation deviations with the remaining  $\zeta$  data points after  $\tau_{(N_j-\zeta)_j}$  (start of extrapolation).

### 3.5. Implementation

We implement the model, (3a) and (3b), in MATLAB R2019b [46]. All optimization problems are solved in the MATLAB version of the global optimization toolbox MEIGO using the enhanced scatter search metaheuristic (eSS) method [47]. The eSS is stochastic and employs some elements of the scatter search and path re-linking methodologies [48]. Thus, the solution depends on the initial conditions and the global optimum is not guaranteed. We set the maximum number of function evaluations, the maximum CPU time and the maximum absolute violation of the constraints to  $10^5$ , 100 s and  $10^{-5}$ , respectively. We use a 50% probability of biasing the search towards bounds and the dynamic hill climbing (DHC) [49] as a local search method. For all aforementioned optimization problems, we set  $x_{1_j}^{\text{l}}$  and  $x_{1_j}^{\text{u}}$  to  $10^{-2}$  and  $10^2$ , respectively. All elements of  $\mathbf{p}_j^{\text{l}}$  and  $\mathbf{p}_j^{\text{u}}$  are set to  $10^{-2}$  and  $10^2$ , respectively.

The ODE (3a) and (3b) are solved using the variable-step, variable-order (VSVO) solver based on the numerical differentiation formulas (NDFs) of orders one to five (ode15s) [50]. We set the relative and absolute error tolerances to  $10^{-3}$  and  $10^{-6}$ , respectively. The solution refinement factor is one and the maximum step size is  $0.1(\tau_{(N_j)_j} - \tau_{1_j})$ .



## 4. Results and Discussion

We now show the results of estimation, identifiability, and predictions. We show the results of six selected patients in details and give performance measures for all 210 patients. We selected those six patients in a way to provide the different profiles of TC dynamics. Data fitting and growth prediction results of all 210 patients are provided in the supplementary material.

### 4.1. Data Fitting and Parameter Estimation

We fitted the parameters of a modified Kuznetsov model on a dataset of solid tumors in human patients under immunotherapy treatment. The model predicted the different tumor growth profiles represented by a selection of representative patients, as well as in the total (unselected) cohort. As shown in Figure 3 and Table 2, model prediction and experimental data profiles are qualitatively and quantitatively very close for these patients. The mean absolute error (MAE), the root-mean-square error (RMSE), and the coefficient of determination ( $R^2$ ) for the selected six patients are given in Table 2. We found that the model gave very high goodness of fit as measured by  $R^2$ . Across 210 patients in all studies, an average  $R^2$  of 0.784 was achieved. Moreover, Table 4 provides the number of patients and the  $R^2$  values per study and per arm. Study 1 and Study 5 have higher  $R^2$  than the remaining studies. Arm 1 has the highest  $R^2$  in all studies except for Study 2. Although direct comparison of this performance with the previous work in [30] is not possible, comparing the MAE of the selected six patients with the reported average MAE in the previous study indicates good fitting performance of the developed Kuznetsov model. Furthermore, by analyzing the goodness of fit in individual patients, we found that the modified Kuznetsov model was able to fit clinically interesting patterns. In particular, the modified Kuznetsov model was able to predict relapse after initial tumor response (patient #207 in Figure 3) and other types of fluctuating behavior, solving a key limitation of previously used simpler models as in [30].

Table 2: Quantification of the goodness of fit of the model as shown in Figure 3. We compare the values of all data points to model results when calculating MAE, RMSE and  $R^2$ .

Patient #	MAE	RMSE	$R^2$
22	0.178	0.226	0.931
83	0.065	0.103	0.996
163	0.084	0.108	0.963
186	1.171	1.993	0.949
203	1.638	2.039	0.974
207	0.115	0.154	0.987
All Patients (1 → 210)	—	—	0.784

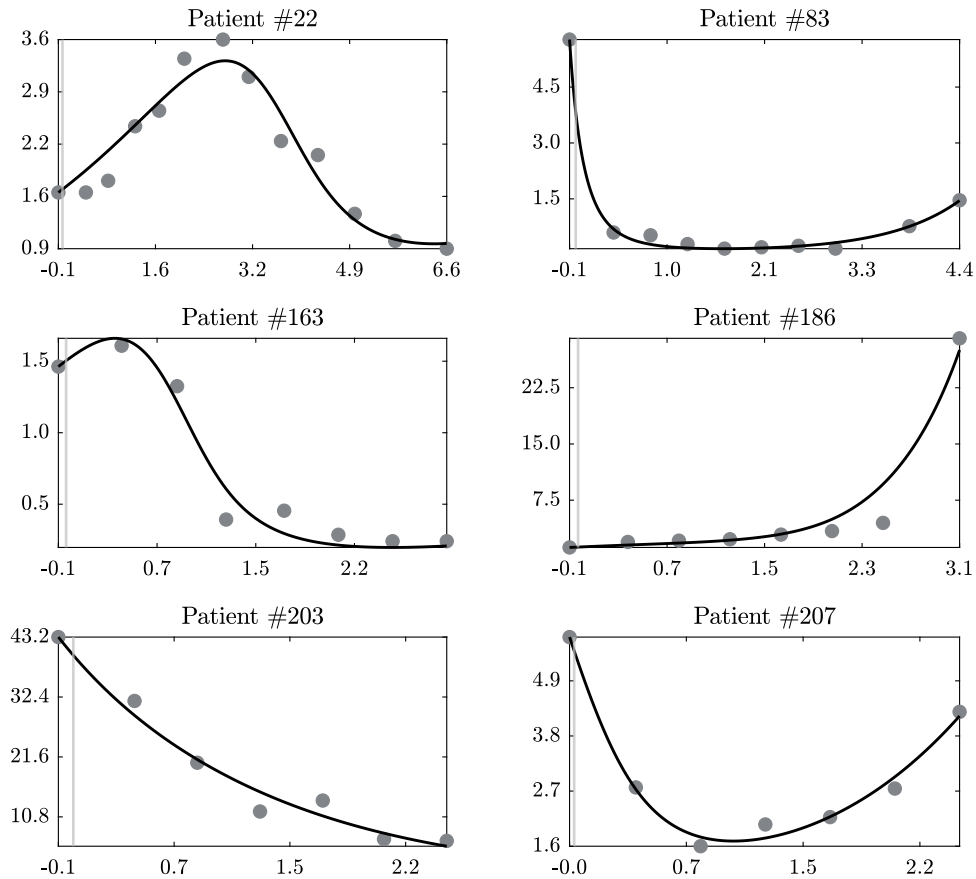


Figure 3: Data fitting results of TC number of the selected patients. The solid (black) line shows model results, where all data points are used when estimating the parameters. The points represent the measured data. Ordinates: normalized number of tumor cells. Abscissas: normalized treatment time, where negative values indicate time before the start of treatment. The model can fit experimental data with different qualitative trends (e.g., up, down and “U”-curve).

We fitted the parameters of the modified Kuznetsov model to a large clinical dataset obtained from five clinical trials. In general, the distributions of the resulting parameter values were similar between the studies (Figure 4). These ranges can be useful for further studies because they enable other researchers to determine plausible ranges and optimal boundaries when fitting the same model to other datasets, thereby simplifying the optimization procedure. A global parameter estimation, however, was not performed in this study and could be attempted in future studies.

Figure 4 shows the estimated values of model parameters of all 210 patients. All parameter values vary per patient. The values are distributed all over the bounds, except for  $\alpha$ , which has a maximum of 6.331. The parameter  $\alpha$  is the normalized parameter for  $a$ , which represents the maximum growth rate of TC population. In addition, most of the values of  $\mu$  and  $\rho$  are close to the upper bound. Although parameter values are quite distributed between the bounds, the aforementioned findings can help in narrowing the expected ranges of values of parameters when global parameter estimation is targeted. Moreover, the distribution of parameter values is compared among the considered five studies. As we can see in Figure 4, the distribution densities between the bounds of parameter values are the same for all studies.

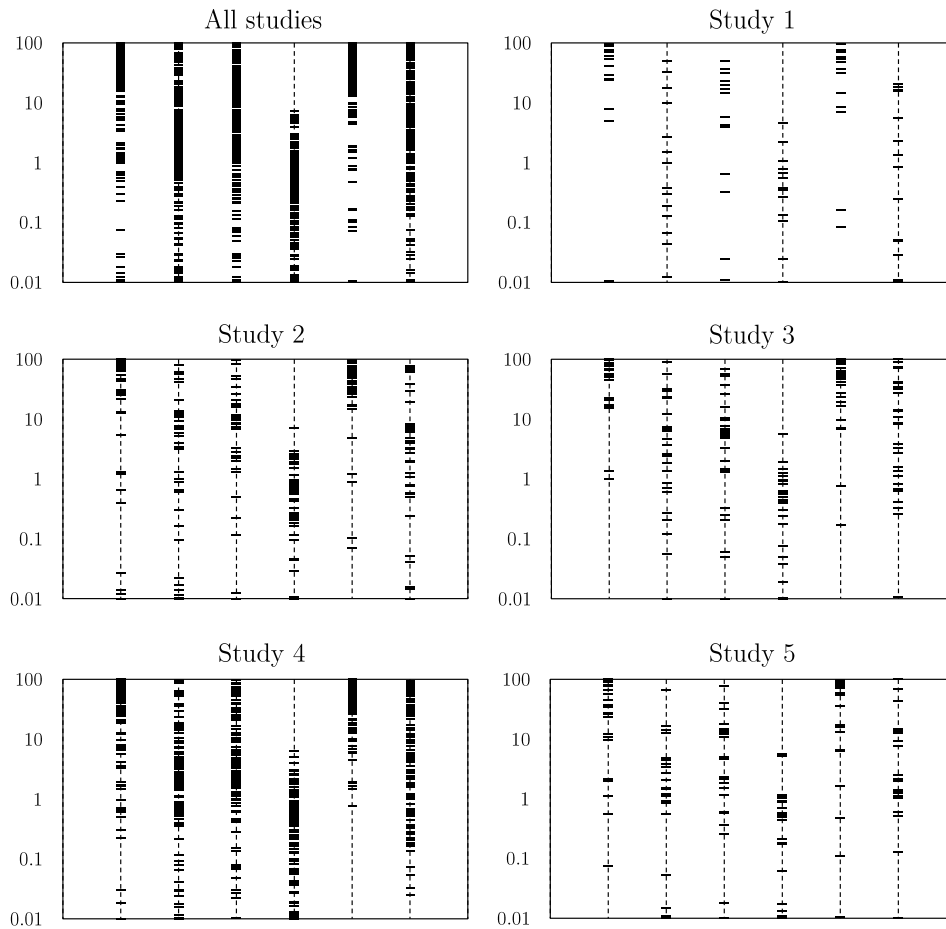


Figure 4: Estimated values of the parameters of (3a) and (3b) model for all 210 patients (after data pre-processing) individually. Ordinates: parameter values. Abscissas: non-dimensionalized model parameters, placed in the following order from left to right:  $\sigma$ ,  $\mu$ ,  $\delta$ ,  $\alpha$ ,  $\rho$ ,  $\eta$ . The values are scattered all over the bounds' ranges, but  $\alpha$  values, which have a maximum of 6.331. Moreover, the distribution densities of parameter values are very close to each other among the studies.

#### 4.2. Parameter Identifiability Analysis

Figure 5 provides the identifiability analysis results of the estimated model parameter values of the selected patients. Depending on the patient and the parameter, the estimated parameter values can be unique or take other values. For patient #22, the approximated feasible parameter set covers a small space in all parameters directions. Their estimated parameter values are close to being unique and thus identifiable. On the other hand, the approximated feasible parameter set for patient #203 covers a large space in all parameter directions. Therefore, the found values are not unique and other combinations of values could lead to good fits and predictions. For the remaining selected patients, the feasible parameter sets can be small or large depending on the patient and the parameter.

When the target is to estimate parameter values of the tumor dynamics model of a certain patient, identifiable values are what we need. However, when the aim is to find global parameters values for all patients, large spaces of the feasible parameter sets can be desirable for finding those global values, in which they are independent of a considered patient. In summary, the results of Figure 5 show that for several combinations of parameter values, good data fitting and model predictions can be achieved. This opens the potential for global parameter estimation, in which the estimated values of model parameters are the same for all patients.

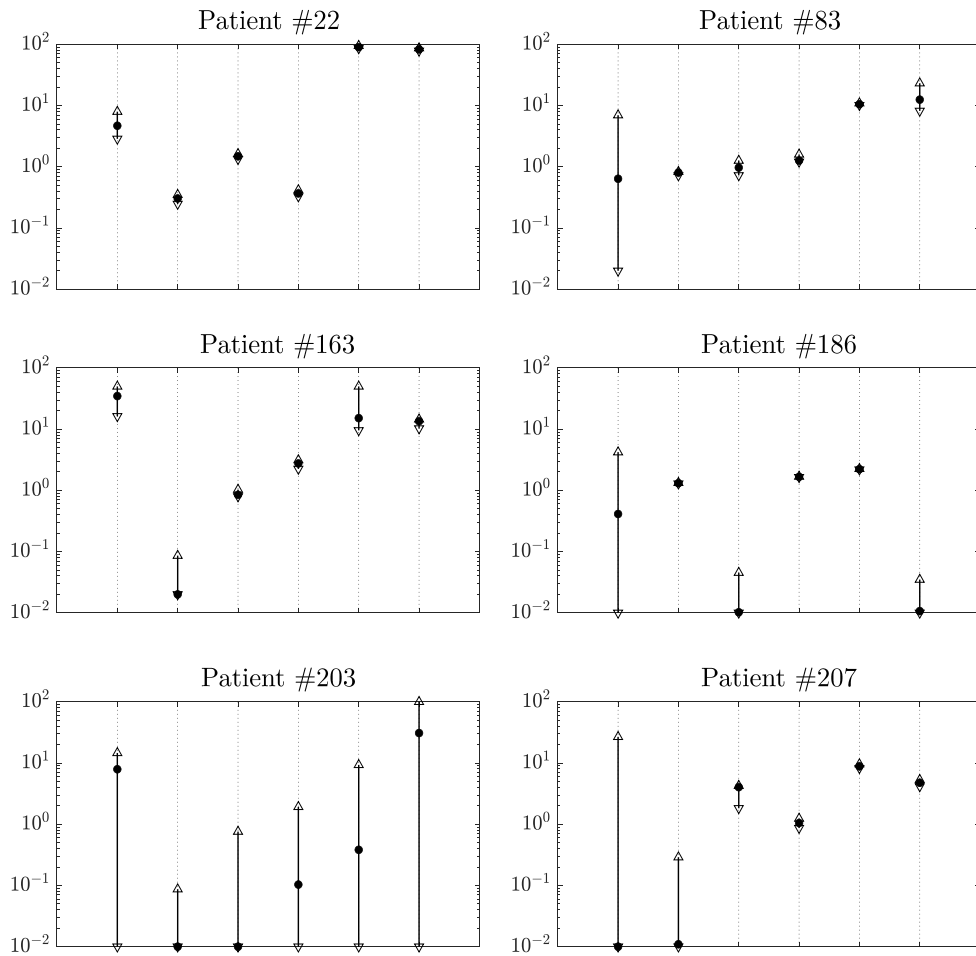


Figure 5: Identifiability analysis results of estimated values of model parameters of the selected patients. Ordinates: parameter values. Abscissas: non-dimensionalized model parameters, placed in the following order from left to right:  $\sigma$ ,  $\mu$ ,  $\delta$ ,  $\alpha$ ,  $\rho$ ,  $\eta$ . The points are the estimated values. The arrowheads are the maximum and minimum values found by the analysis. The results show that there are several combinations of parameter values at which the model can fit the experimental data.

### 4.3. Tumor Growth Prediction

In clinical decision making, a possible role of mechanistic models is forecasting tumor growth during treatment, potentially enabling physicians to adjust the treatment strategy earlier. We found that the modified Kuznetsov model indeed was able to extrapolate beyond the initial time points when the last two or three data points are not included when fitting the model. The solid black lines in Figure 6 show model extrapolation results of the selected patients when the last two data points are not considered for fitting. For the six patients, the model quantitatively forecasts tumor dynamics, except for patient #186. The last data point of patient #186 is almost impossible to forecast because it suddenly shifts upward after a mild and constant increase of tumor growth. However, the model can still qualitatively predict the growth. For the other patients, the predictions are very close to the experimental data. In Table 3, the first sub-table “optimal Extra.” provides the MAE, RMSE, and  $R^2$  of the selected patients, as well as the average deviation for the extrapolated part, represented as “dev”. The values of  $R^2$  for the six patients are close to one except for patient #186 due to the aforementioned explanation. Compared to the previous work in [30], model extrapolation results have here higher  $R^2$  values, specifically, an  $R^2$  of 0.979 was reached for patient #207. The average values of  $R^2$  and dev of all 210 patients are also provided in the table. Model extrapolation results when omitting the last two data points for all 210 patients are provided in the supplementary material. Also, we provide model extrapolation results while omitting the last three data points for the selected six patients in the supplementary material. For the latter results, the model could also predict the growth very well.

In addition, Table 4 provides the  $R^2$  and dev values per study and per arm for model extrapolation when omitting the last two data points. In general, and as expected, the calculated average  $R^2$  for the extrapolation experiment is lower than the  $R^2$  when using all the data points in all the studies. Similar to data fitting,

Study 1 and Study 5 have higher  $R^2$  than the remaining studies. In Study 3, Arm 2 and Arm 3 have the lowest  $R^2$ . Interestingly, in Study 2 and Study 4, the extrapolation performance of the model is markedly better for the Docetaxel group (Arm 2) than for the Atezolizumab group (Arm 1). This indicates that patients receiving Docetaxel have a more predictable response trajectory compared to patients receiving immunotherapy in whom unexpected patterns of tumor response can occur even at later time points.

Moreover, Figure 6 shows how far model extrapolation could deviate from the actual values. The results of the two “extreme-case” (dash-dotted blue and the dashed magenta) lines are designed to be as far away from each other at the final time point ( $\tau_{(N_j)_j}$ ) while both being within some 10% tolerance of the found optimal fit for the fitted time before extrapolation starts. For some patients, the “extreme-case” lines can significantly deviate from the actual values, especially for the upper case (dash-dotted blue lines), as for patients #22 and #83. In contrast, the “extreme-case” lines for patients #163 and #186 are very close to the optimal extrapolated ones. Sub-tables “Upper Extra.” and “Lower Extra.” in Table 3 provide the MAE, RMSE,  $R^2$  and dev for the selected patients for the two “extreme-case” extrapolations.

To sum up, the model can forecast tumor dynamics of the patients and the “extreme-case” extrapolation scenarios were conducted to check the worst model predictions. For some patients the “extreme-case” extrapolation results deviate from the actual values, for others they remain close to the optimal extrapolation results.

Table 3: Quantification of the goodness of fit and prediction of the model as shown in Figure 6. Here the last two data points are not considered for parameter estimation. We compare the values of all data points to model results when calculating the errors and  $R^2$ . The average deviation of model prediction results from the measured values during extrapolation time (shaded region in Figure 6) is represented by dev. The Optimal Extra. sub-table shows the values for the optimal model prediction results (the solid (black) line in Figure 6). The Upper Extra. sub-table relates values to the upper “extreme-case” extrapolation deviation results (the dash-dotted (blue) line in Figure 6). The Lower Extra. sub-table relates values to the lower “extreme-case” extrapolation deviation results (the dashed (magenta) line in Figure 6).

	Patient #	MAE	RMSE	$R^2$	dev
Optimal Extra.	22	0.179	0.229	0.929	0.049
	83	0.171	0.279	0.971	0.474
	163	0.086	0.117	0.956	0.079
	186	3.020	8.194	0.134	0.414
	203	1.745	2.044	0.974	0.323
	207	0.152	0.196	0.979	0.068
	All Patients (1 → 210)	—	—	0.419	0.768
Upper Extra.	22	0.394	0.566	0.568	1.295
	83	0.484	1.302	0.363	1.503
	163	0.106	0.140	0.938	0.330
	186	2.983	8.035	0.168	0.403
	203	3.561	5.719	0.798	1.559
	207	0.625	0.834	0.618	0.428
	Lower Extra.	22	0.301	0.372	0.813
83		0.224	0.409	0.937	0.721
163		0.122	0.151	0.928	0.475
186		3.258	8.499	0.069	0.485
203		3.538	4.145	0.894	0.828
207		0.465	0.811	0.639	0.348

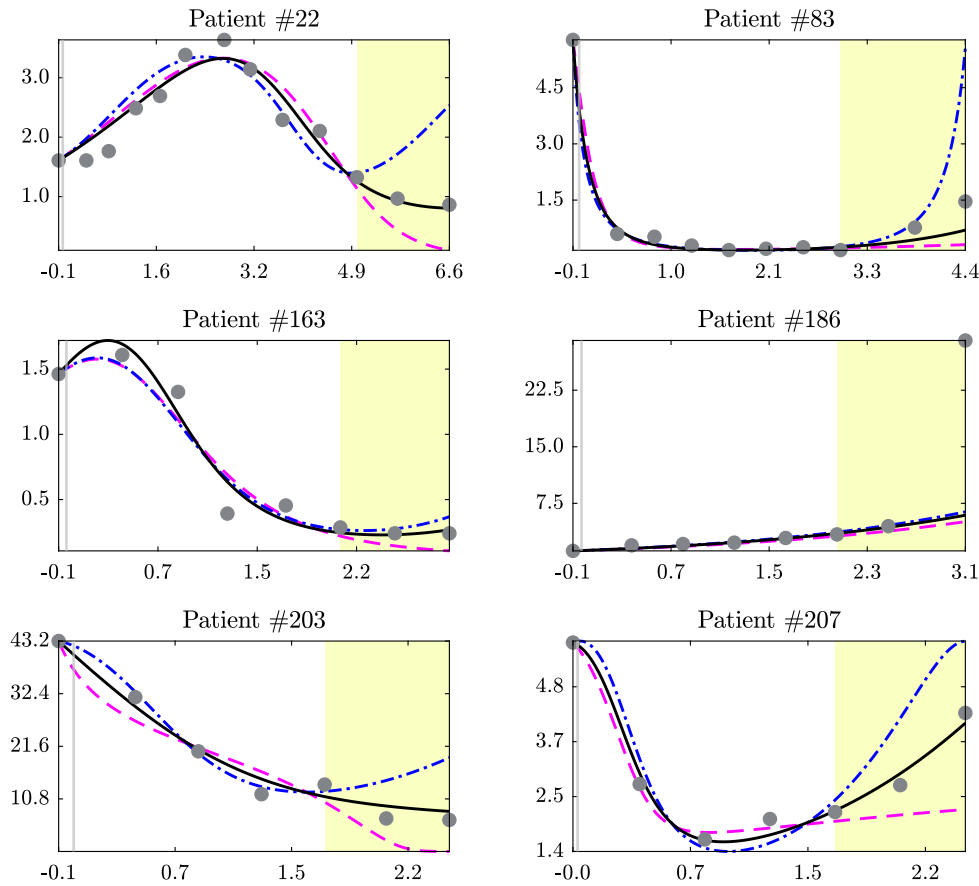


Figure 6: Model extrapolation results of the selected patients. The solid (black) line represents the optimal model prediction. The dash-dotted (blue) and the dashed (magenta) lines show the upper and the lower “extreme-case” model extrapolation deviation results, respectively. Here the last two data points are not considered for parameter estimation. The points show the measured data. Shaded areas highlight regions of model extrapolation. Ordinates: normalized number of tumor cells. Abscissas: normalized treatment time, negative values indicate time before the start of treatment. The model is capable of forecasting tumor dynamics qualitatively and sometimes quantitatively.



Table 4: Quantification of the goodness of fit and extrapolation of the model per the conducted studies and arms. We compare the values of all data points to model results when calculating the  $R^2$ . For model extrapolation, the average deviation of model prediction results from the measured values during extrapolation time (shaded region in Figure 6) is represented by dev. A detailed description of the included studies is available in [30]. Briefly, study arms with “Atezolizumab” or “MPDL” are immunotherapy. “Docetaxel” is the main chemotherapy in these trials.

Study No. (Name)	Arm No. (Name)	No. of Patients	Data Fitting	Extrapolation	
			$R^2$	$R^2$	dev
1 (FIR)	1 (MPDL3280A-1)	3	0.869	0.582	1.256
	2 (MPDL3280A-2)	11	0.847	0.535	0.406
	3 (MPDL3280A-3)	0	—	—	—
	Total	14	0.852	0.545	0.588
2 (POPLAR)	1 (Atezolizumab)	6	0.581	0.241	0.943
	2 (Docetaxel)	25	0.747	0.474	0.761
	Total	31	0.715	0.429	0.796
3 (BIRCH)	1 (MPDL3280A-1a)	2	0.964	0.980	1.983
	2 (MPDL3280A-2a)	6	0.587	-0.077	1.143
	3 (MPDL3280A-3a)	2	0.243	-0.036	0.317
	4 (MPDL3280A-1b)	8	0.864	0.522	0.525
	5 (MPDL3280A-2b)	3	0.749	0.226	0.394
	6 (MPDL3280A-3b)	9	0.884	0.575	0.670
Total	30	0.769	0.382	0.762	
4 (OAK)	1 (Atezolizumab)	29	0.811	0.218	0.393
	2 (Docetaxel)	82	0.783	0.455	0.869
	Total	111	0.790	0.393	0.745
5 (IMvigor 210)	1 (Atezolizumab)	24	0.825	0.501	0.953
Total	210	0.784	0.419	0.768	

## 5. Conclusion

In this study, we show that quantitative mathematical models can be used to describe and forecast the behavior of cancer. Previous studies have used the same datasets to fit very simple ODE models to the tumor volume measurements over time [30]. However, in Ghaffari Laleh et al. [30], it was observed that all established ODE models were not able to fit “U”-shaped trajectories well. In clinical terms, patients who relapsed after an initial response, or patients who showed a delayed response, were not adequately represented in these previous models. Compared to this, the present study evaluates a more complex model which has the benefit of being able to describe a larger variety of real-world time series. This specific model is a slight simplification of the Kuznetsov model [16], which has not been linked with or validated in large amounts of quantitative real-world human data. Specifically, it could quantitatively fit the Kuznetsov model to a large dataset of 1472 patients. Data are collected from patients undergoing immunotherapy or chemotherapy treatments.

In the parameter estimation for each patient, we found that some parameters for some patients are not unique (identifiability analysis). This means that many combinations of parameter values could lead to good fitting and predictions. This opens the potential for global parameter estimation, in which parameter values are the same for all patients. Still, the model could predict tumor growth (2-3 omitted measurements) which could indicate practical usefulness as a predictive biomarker. Specifically, the model fitting and prediction could potentially describe and forecast the behavior of cancer, improve the understanding of underlying biological mechanisms, and provide model approaches for cancer treatments. Future studies should attempt a global parameter estimation, use deterministic and less computationally demanding solution methods. Another possibility is the reduction of the number of model parameters in order to have less number of parameters to estimate and thus enhance the possibility to reach global values.

The model complexity is mostly limited by the availability of suitable data. Therefore, if more data become available, it would be possible to include other terms in the model, e.g., explicitly representing different types of immune cell populations or multiple cancer cell clones.

## 6. Additional Information

### 6.1. Author Contributions

MEW performed the numerical experiments; NGL and JNK selected the publicly available data and provided clinical insight; MEW, FJ, DB, CK, AM analyzed the data; MEW, NGL and JNK wrote the initial version of the manuscript; all authors contributed to the interpretation of the results; all authors edited the manuscript and made the decision to submit for publication.

### 6.2. Competing Interests

None.

### 6.3. Acknowledgments

None.

## 7. References

- [1] C. J. Wheeler, A. Das, G. Liu, J. S. Yu, and K. L. Black, “Clinical responsiveness of Glioblastoma multiforme to chemotherapy after vaccination,” *Clinical Cancer Research*, vol. 10, no. 16, p. 5316, 2004.
- [2] L. G. de Pillis, W. Gu, and A. E. Radunskaya, “Mixed immunotherapy and chemotherapy of tumors: Modeling, applications and biological interpretations,” *Journal of Theoretical Biology*, vol. 238, no. 4, pp. 841–862, 2006.
- [3] J. D. Farrar, K. H. Katz, J. Windsor, G. Thrush, R. H. Scheuermann, J. W. Uhr, and N. E. Street, “Cancer dormancy. VII. A regulatory role for CD8+ T cells and IFN-gamma in establishing and maintaining the tumor-dormant state,” *Journal of Immunology (Baltimore, Md. : 1950)*, vol. 162, no. 5, pp. 2842–2849, 1999.
- [4] K. J. O’Byrne, A. G. Dalglish, M. J. Browning, W. P. Steward, and A. L. Harris, “The relationship between angiogenesis and the immune response in carcinogenesis and the progression of malignant disease,” *European Journal of Cancer (Oxford, England : 1990)*, vol. 36, no. 2, pp. 151–169, 2000.
- [5] A. G. Dalglish and K. J. O’Byrne, “Chronic immune activation and inflammation in the pathogenesis of AIDS and cancer,” *Advances in Cancer Research*, vol. 84, pp. 231–276, 2002.
- [6] S. K. Chatterjee and B. R. Zetter, “Cancer biomarkers: Knowing the present and predicting the future,” *Future Oncology*, vol. 1, no. 1, pp. 37–50, 2005.
- [7] A. R. A. Anderson and V. Quaranta, “Integrative mathematical oncology,” *Nat Rev Cancer* 8, vol. 3, pp. 227–234, 2008.
- [8] R. C. Rockne and J. G. Scott, “Introduction to mathematical oncology,” *JCO clinical cancer informatics*, vol. 3, pp. 1–4, 2019.
- [9] Olivier Ledoit, Pedro Santa Clara, and Michael Wolf, *Flexible multivariate GARCH modeling with an application to international stock markets*, 2001.
- [10] S. Manabe, “Carbon Dioxide and climatic change,” in *Advances in Geophysics : Theory of Climate*, B. Saltzman, Ed., vol. 25, Elsevier, 1983, pp. 39–82.
- [11] M. L. Thompson and M. A. Kramer, “Modeling chemical processes using prior knowledge and neural networks,” *AIChE Journal*, vol. 40, no. 8, pp. 1328–1340, 1994.
- [12] L. Norton, R. Simon, H. D. Brereton, and A. E. Bogden, “Predicting the course of Gompertzian growth,” *Nature*, vol. 264, no. 5586, pp. 542–545, 1976.
- [13] Y. Kogan, K. Halevi-Tobias, M. Elishmereni, S. Vuk-Pavlović, and Z. Agur, “Reconsidering the paradigm of cancer immunotherapy by computationally aided real-time personalization,” *Cancer Research*, vol. 72, no. 9, pp. 2218–2227, 2012.
- [14] R. J. de Boer and P. Hogeweg, “Interactions between macrophages and T-lymphocytes: Tumor sneaking through intrinsic to helper T cell dynamics,” *Journal of Theoretical Biology*, vol. 120, no. 3, pp. 331–351, 1986.
- [15] D. Kirschner and J. C. Panetta, “Modeling immunotherapy of the tumor-immune interaction,” *Journal of Mathematical Biology*, vol. 37, no. 3, pp. 235–252, 1998.
- [16] V. A. Kuznetsov, I. A. Makalkin, M. A. Taylor, and A. S. Perelson, “Nonlinear dynamics of immunogenic tumors: Parameter estimation and global bifurcation analysis,” *Bulletin of Mathematical Biology*, vol. 56, no. 2, pp. 295–321, 1994.
- [17] A. Rhodes and T. Hillen, “Implications of immune-mediated metastatic growth on metastatic dormancy, blow-up, early detection, and treatment,” *BioRxiv*, 2019.
- [18] A. M. Makhlof, L. El-Shennawy, H. A. Elkaranshawy, and D. Diller, “Mathematical modelling for the role of CD4+T cells in tumor-immune interactions,” *Computational and Mathematical Methods in Medicine*, 2020.
- [19] N. Bellomo and L. Preziosi, “Modelling and mathematical problems related to tumor evolution and its interaction with the immune system,” *Mathematical and Computer Modelling*, vol. 32, no. 3, pp. 413–452, 2000.
- [20] M. Kolev, “Mathematical modelling of the competition between tumors and immune system considering the role of the antibodies,” *Mathematical and Computer Modelling*, vol. 37, no. 11, pp. 1143–1152, 2003.
- [21] K. Owens and I. Bozic, “Modeling CAR T-cell therapy with patient preconditioning,” *Bulletin of Mathematical Biology*, vol. 83, no. 5, p. 42, 2021.
- [22] S. Benzekry, C. Lamont, A. Beheshti, A. Tracz, J. M. L. Ebos, L. Hlatky, and P. Hahnfeldt, “Classical mathematical models for description and prediction of experimental tumor growth,” *PLOS Computational Biology*, vol. 10, no. 8, 2014.
- [23] C. Vaghi, A. Rodallec, R. Fanciullino, J. Ciccolini, J. P. Mochel, M. Matri, C. Poinard, J. M. L. Ebos, and S. Benzekry, “Population modeling of tumor growth curves and the reduced Gompertz

- model improve prediction of the age of experimental tumors,” *PLOS Computational Biology*, vol. 16, no. 2, 2020.
- [24] B. A. Ruggeri, F. Camp, and S. Miknyoczki, “Animal models of disease: Pre-clinical animal models of cancer and their applications and utility in drug discovery,” *Biochemical Pharmacology*, vol. 87, no. 1, pp. 150–161, 2014.
- [25] W. Kolch, M. Halasz, M. Granovskaya, and B. N. Kholodenko, “The dynamic control of signal transduction networks in cancer cells,” *Nature Reviews. Cancer*, vol. 15, no. 9, pp. 515–527, 2015.
- [26] J. J. Tyson, W. T. Baumann, C. Chen, A. Verdugo, I. Tavassoly, Y. Wang, L. M. Weiner, and R. Clarke, “Dynamic modelling of oestrogen signalling and cell fate in breast cancer cells,” *Nature Reviews. Cancer*, vol. 11, no. 7, pp. 523–532, 2011.
- [27] F. Fröhlich, T. Kessler, D. Weindl, A. Shadrin, L. Schmiester, H. Hache, A. Muradyan, M. Schütte, J.-H. Lim, M. Heinig, F. J. Theis, H. Lehrach, C. Wierling, B. Lange, and J. Hasenauer, “Efficient parameter estimation enables the prediction of drug response using a mechanistic Pan-cancer pathway model,” *Cell Systems*, vol. 7, no. 6, pp. 567–579, 2018.
- [28] L. E. Clegg and F. Mac Gabhann, “Molecular mechanism matters: Benefits of mechanistic computational models for drug development,” *Pharmacological Research*, vol. 99, pp. 149–154, 2015.
- [29] R. E. Baker, J.-M. Peña, J. Jayamohan, and A. Jérusalem, “Mechanistic models versus machine learning, a fight worth fighting for the biological community?” *Biology Letters*, vol. 14, no. 5, 2018.
- [30] N. G. Laleh, C. M. Lavinia Loeffler, J. Grajek, K. Staňková, A. T. Pearson, H. S. Muti, C. Trautwein, H. Enderling, J. Poleszczuk, and J. N. Kather, “Classical mathematical models for prediction of response to chemotherapy and immunotherapy,” *BioRxiv*, 2021.
- [31] A. Faustino-Rocha, P. A. Oliveira, J. Pinho-Oliveira, C. Teixeira-Guedes, R. Soares-Maia, R. G. da Costa, B. Colaço, M. J. Pires, J. Colaço, R. Ferreira, and M. Ginja, “Estimation of rat mammary tumor volume using caliper and ultrasonography measurements,” *Lab Animal*, vol. 42, no. 6, pp. 217–224, 2013.
- [32] M. Shevtsov, E. Pitkin, A. Ischenko, S. Stangl, W. Khachatryan, O. Galibin, S. Edmond, D. Lobinger, and G. Multhoff, “Ex vivo Hsp70-activated NK cells in combination with PD-1 inhibition significantly increase overall survival in preclinical models of glioblastoma and lung cancer,” *Frontiers in Immunology*, vol. 10, p. 454, 2019.
- [33] D. R. Spigel, J. E. Chaft, S. Gettinger, B. H. Chao, L. Dirix, P. Schmid, L. Q. Chow, R. J. Hicks, L. Leon, J. Fredrickson, *et al.*, “Fir: Efficacy, safety, and biomarker analysis of a phase ii open-label study of atezolizumab in pd-11–selected patients with nsccl,” *Journal of Thoracic Oncology*, vol. 13, no. 11, pp. 1733–1742, 2018.
- [34] L. Fehrenbacher, A. Spira, M. Ballinger, M. Kowantetz, J. Vansteenkiste, J. Mazieres, K. Park, D. Smith, A. Artal-Cortes, C. Lewanski, *et al.*, “Atezolizumab versus docetaxel for patients with previously treated non-small-cell lung cancer (poplar): A multicentre, open-label, phase 2 randomised controlled trial,” *The Lancet*, vol. 387, no. 10030, pp. 1837–1846, 2016.
- [35] S. Peters, S. Gettinger, M. L. Johnson, P. A. Jänne, M. C. Garassino, D. Christoph, C. K. Toh, N. A. Rizvi, J. E. Chaft, E. C. Costa, *et al.*, “Phase ii trial of atezolizumab as first-line or subsequent therapy for patients with programmed death-ligand 1–selected advanced non-small-cell lung cancer (birch),” *Journal of Clinical Oncology*, vol. 35, no. 24, p. 2781, 2017.
- [36] A. Rittmeyer, F. Barlesi, D. Waterkamp, K. Park, F. Ciardiello, J. Von Pawel, S. M. Gadgeel, T. Hida, D. M. Kowalski, M. C. Dols, *et al.*, “Atezolizumab versus docetaxel in patients with previously treated non-small-cell lung cancer (oak): A phase 3, open-label, multicentre randomised controlled trial,” *The Lancet*, vol. 389, no. 10066, pp. 255–265, 2017.
- [37] A. V. Balar, M. D. Galsky, J. E. Rosenberg, T. Powles, D. P. Petrylak, J. Bellmunt, Y. Loriot, A. Necchi, J. Hoffman-Censits, J. L. Perez-Gracia, *et al.*, “Atezolizumab as first-line treatment in cisplatin-ineligible patients with locally advanced and metastatic urothelial carcinoma: A single-arm, multicentre, phase 2 trial,” *The Lancet*, vol. 389, no. 10064, pp. 67–76, 2017.
- [38] A. Perelson and G. Weisbuch, “Immunology for physicists,” *Reviews of Modern Physics*, vol. 69, 1997.
- [39] R. J. de Boer and M. C. Boerlijst, “Diversity and virulence thresholds in AIDS,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 91, no. 2, pp. 544–548, 1994.
- [40] E. Walter and L. Pronzato, *Identification of parametric models from experimental data*. Berlin, New York, and Paris: Springer and Masson, 1997.
- [41] H. Miao, X. Xia, A. S. Perelson, and H. Wu, “On identifiability of nonlinear ODE models and applications in viral dynamics,” *SIAM Review*, vol. 53, no. 1, pp. 3–39, 2011.
- [42] F. Jung, F. A. L. Janssen, A. Ksiazkiewicz, A. Caspari, A. Mhamdi, A. Pich, and A. Mitsos, “Identifiability analysis and parameter estimation of microgel synthesis: A set-membership approach,” *Industrial & Engineering Chemistry Research*, vol. 58, no. 30, pp. 13 675–13 685, 2019.

- [43] A. Raue, C. Kreutz, T. Maiwald, U. Klingmuller, and J. Timmer, “Addressing parameter identifiability by model-based experimentation,” *IET Systems Biology*, vol. 5, no. 2, pp. 120–130, 2011.
- [44] A. Gábor, A. F. Villaverde, and J. R. Banga, “Parameter identifiability analysis and visualization in large-scale kinetic models of biosystems,” *BMC Systems Biology*, vol. 11, no. 1, p. 54, 2017.
- [45] R. Paulen, M. E. Villanueva, and B. Chachuat, “Guaranteed parameter estimation of non-linear dynamic systems using high-order bounding techniques with domain and CPU-time reduction strategies,” *IMA Journal of Mathematical Control and Information*, vol. 33, no. 3, pp. 563–587, 2016.
- [46] MATLAB, *9.7.0.1190202 (R2019b)*. Natick, Massachusetts: The MathWorks Inc, 2019.
- [47] J. A. Egea, D. Henriques, T. Cokelaer, A. F. Villaverde, A. MacNamara, D.-P. Danciu, J. R. Banga, and J. Saez-Rodriguez, “MEIGO: An open-source software suite based on metaheuristics for global optimization in systems biology and bioinformatics,” *BMC Bioinformatics*, vol. 15, no. 1, p. 136, 2014.
- [48] J. A. Egea, R. Mart, and J. R. Banga, “An evolutionary method for complex-process optimization,” *Comput. Oper. Res.*, vol. 37, no. 2, pp. 315–324, 2010.
- [49] D. Yuret and M. de La Maza, “Dynamic Hill Climbing: Overcoming the limitations of optimization techniques,” in *In The Second Turkish Symposium on Artificial Intelligence and Neural Networks*, 1993, pp. 208–212.
- [50] L. F. Shampine and M. W. Reichelt, “The MATLAB ODE suite,” *SIAM Journal on Scientific Computing*, vol. 18, no. 1, pp. 1–22, 1997.