Tumor Model Fitting Using Neural Network Estimator

Lilla Kisbenedek[1,2,*], Dániel András Drexler[1], and Levente Kovács[1]

[1]Physiological Controls Research Center, University Research and Innovation Center, Obuda University, Bécsi út 96/b, 1034 Budapest, Hungary,
{kisbenedek.lilla, drexler.daniel, kovacs}@uni-obuda.hu
[2]Applied Informatics and Applied Mathematics Doctoral School, Obuda University, Bécsi út 96/b, 1034 Budapest, Hungary
[*]Corresponding: kisbenedek.lilla@uni-obuda.hu

Abstract: Personalized therapy aims to generate treatments based on individual patient characteristics, maximizing therapeutic efficacy while minimizing adverse effects. A promising strategy involves modeling tumor progression and drug response through systems of ordinary differential equations (ODEs), where patient- and tumor-specific physiological properties are encoded in the model parameters. We present a novel framework that utilizes neural network techniques to estimate the patient-specific physiological parameters within a tumor model. While numerical integration is typically used to solve such models, conventional ODE solvers lack support for automatic differentiation. This makes them unsuitable for gradient-based training in neural network architectures. To overcome this limitation, we incorporated a differentiable ODE solver. This enables an end-to-end training process through gradient-based optimization and allows accurate parameter identification directly from tumor volume measurements and dosing schedules. The approach was validated on simulated datasets with varying noise levels, showing that it can reliably identify model parameters that reproduce the underlying tumor dynamics. Accurate identification of patient-specific parameters is essential for designing adaptive chemotherapy protocols, thereby improving future treatment outcomes.

Keywords: parameter identification; machine learning; autoencoder; tumor model; physiological parameters

1 Introduction

Cancer continues to be a major global health challenge, with nearly 20 million new cases and 9.7 million deaths reported in 2022 [1]. The most frequently diagnosed malignancies included lung, breast, colorectal, prostate and stomach cancers. These types also contributed most significantly to cancer mortality. These statistics are particularly concerning in low- and middle-income countries, where higher cancer mortality rates are often linked to insufficient access to early detection and treatment services [1]. For many years, cancer treatment options for patients were limited to surgery, radiation therapy, and chemotherapy, used separately or in combination [2].

While innovative methods, including targeted drug treatments and personalized medicine, are now being extensively researched and applied, chemotherapy is often still the best option for many patients [3, 4]. It is cost-effective compared to newer treatments, making it accessible for a broader range of patients [5]. Nevertheless, traditional chemotherapy typically employs a one-size-fits-all approach, resulting in diverse responses and a list of side effects among patients due to individual variability [6]. To solve this problem, personalized dosing strategies – such as metronomic or reduced-intensity chemotherapy – have emerged as promising alternatives, enabling personalized treatment plans that account for patient-specific differences and often maintain efficacy while reducing toxicity [7–9].

In recent years, numerous research papers have been published questioning widely applied high dosages, advocating for the use of lower amounts [7, 10–12]. Additionally, several studies propose control algorithms and mathematical models to simulate tumor behavior and guide decision-making processes regarding drug dosing [13–18]. Our research aims to optimize chemotherapy based on differential equations that describe the tumor dynamics and the drug effect [19]. Previously, a mathematical model was developed to describe both the natural growth of the tumor and the pharmacodynamics and pharmacokinetics of the drug. By utilizing this model, it is possible to determine an optimal dosing regimen based on measured, temporal tumor data [20]. However, accurate estimation of model parameters is essential to establish the optimal dosage regimen, as these parameters govern the dynamics of the tumor model equations.

Accurate parameter estimation in biological systems is usually challenging due to inherent complexities and nonlinear dynamics [21–24]. These complexities are particularly present in physiological model fitting, where inter-individual variability further complicates the modeling process. Physiological models often have a large number of parameters, increasing computational cost and making optimization more difficult. Another challenge is that clinical and experimental data are often sparse, irregularly sampled, and noisy, making model fitting more challenging.

To address some of these limitations, the integration of neural networks with systems of ordinary differential equations (ODEs) has gained significant attention in the past few years [24–27]. One approach is a physics-informed neural network, which utilizes the similarity between neural networks and differential equations. The method incorporates the structure of differential equations directly into the learning process, ensuring that machine learning models respect known physical laws [28–31]. Another prominent framework is DeepONet, which is a neural network-based architecture designed to learn mappings between function spaces, making it highly effective for solving differential equations and parameter identification. It uses a combination of two networks, namely branch and trunk networks, to efficiently predict solutions or identify parameters in complex systems. However, DeepONet generalizes well, but requires high-quality data and significant computational resources for training [27]. An alternative significant development in this area is Neural Ordinary Differential Equations (NODE). The architecture of a NODE treats the transformation of hidden states as a continuous process rather than using discrete layers. They use an ODE solver to compute the evolution of these states, allowing adaptive computation and efficient memory use. It is particularly useful for modeling time-series data and systems with underlying continuous dynamics [32–34]. These advancements in integrating neural networks with ordinary differential equations highlight the growing potential for leveraging machine learning in scientific and engineering applications. By embedding physical laws directly into the learning process, these approaches ensure that predictions remain consistent with real-world phenomena.

In this study, we propose a solution for parameter identification, which is an essential phase in a research pipeline aimed at optimizing cancer therapy. In earlier work, feed-forward neural networks were trained to predict model parameters, however, this approach performed primarily as an initial estimate for subsequent optimization methods [35]. Later, a framework combining two neural networks was introduced to predict tumor model parameters [36]. Since conventional ODE solvers do not support backpropagation, a second neural network was employed to simulate tumor dynamics. However, this design increased model complexity and reduced training accuracy, introducing additional sources of error and necessitating pre-training with fixed dosing schedules. In this work, we address these challenges by introducing a new architecture that replaces the second neural network with a differentiable ODE solver, enabling a more flexible design and a more efficient and accurate training process.

The tumor model utilized in this study is described in detail in Section 2. We incorporated this model into our proposed framework to improve the accuracy of the tumor model solution. The corresponding neural network architecture, also presented in Section 2, functions as a parameter estimator. During evaluation, rather than using a preset training dataset, we directly optimized the network on each tumor volume time series. In other words, the weights of the network were tuned for each individual case (each set of temporal tumor measurements) to find the best-fit parameters. We validated our proposed framework using *in silico* input-output pairs, applying different noise levels on the input tumor volume measurements. Furthermore, run-time performance evaluation is discussed in Section 3.

2 Materials and Methods

2.1 The Architecture of the Algorithm

The proposed architecture is illustrated in Figure 1. It takes two input time series of identical length L: the tumor volume measurements and the chemotherapy dosing schedule. Tumor volumes are sampled specifically on Monday, Wednesday, and Friday of each week, reflecting typical preclinical measurement intervals. Chemotherapy doses are administered two or three times weekly, selected randomly among weekdays, with each dose amount independently drawn from a uniform

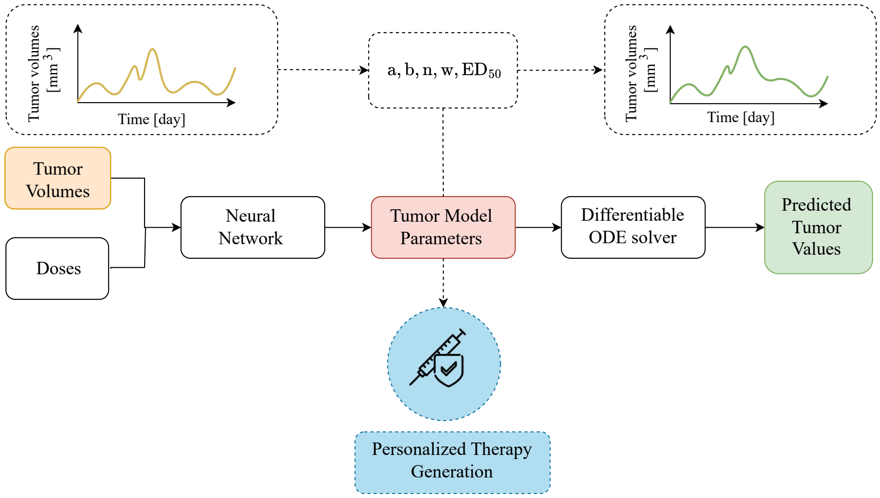

Figure 1
The architecture of the developed algorithms. The algorithm is designed to estimate tumor model parameters from tumor volume and dosage data. Based on the identified parameters, optimized therapeutic strategies can be formulated for more personalized treatment planning.

distribution ranging between 0 and 8 mg/kg. Both input time series have a length of L, however, the tumor volume measurements are further refined through interpolation at a fixed temporal resolution, introducing a predefined number of intermediate points between each measurement. The improved temporal resolution ensures finer granularity, which is essential for the accurate numerical integration of the tumor model equations and is required by deep learning frameworks to maintain differentiability throughout the training process.

The inputs are fed into the neural network, which then predicts the model parameters. Subsequently, these parameters are mapped to a physiologically relevant range using the transformation defined below. The output of the neural network is p, the vector of the five predicted parameters. On the i-th parameter value p_i, the following transformation is applied:

$$\tilde{p}_i = p_{i,\text{min}} + (p_{i,\text{max}} - p_{i,\text{min}}) \times \sigma(p_i), \tag{1}$$

where the sigmoid function $\sigma(p_i)$ is defined as:

$$\sigma(p_i) = \frac{1}{1+e^{-p_i}}. \tag{2}$$

In (1), we apply the sigmoid function to each element of p, which is a column vector with five elements, containing the predicted parameters. It transforms them to values between 0 and 1, and then scales and shifts them to the intervals limited by p_{min} and p_{max} values. Each element is constrained within its corresponding limits. The sigmoid function provides a smooth function, which is necessary for differentiability during training.

The p_{min} and p_{max} values of the parameters are listed in Table 1. This mapping ensures that the parameter values remain within biologically meaningful bounds. Given that the parameters typically have units of day^{-1} or $\text{mg} \cdot \text{kg}^{-1}$, their upper and lower limits are constrained by biological considerations.

Following the transformation of parameters to their respective ranges, a differentiable ODE solver was employed to reconstruct the tumor volume time series based on the given doses and parameter values. For simulated tumor volumes, the loss function quantifies the error as the difference between the input tumor volumes (indicated by the orange markers in the figure) and the predicted tumor volumes (indicated by the green markers in the figure). This error is then backpropagated through the neural network to update its weights. In the subsequent iteration, the neural network predicts a new set of tumor model parameters that are more likely to capture the original tumor dynamics. This process repeats until the discrepancy between the original and predicted tumor volumes is minimized. At the end of the training, the optimal parameter values are obtained. These values can be used to characterize patient-specific properties, enabling the design of personalized treatment strategies. This approach allows us to identify the most effective dosing regimen for individual cases, as illustrated with the blue markers in Figure 1.

2.2 The Neural Network and the Training Hyperparameters

The neural network used in this work is designed to predict the parameters of an ODE system representing tumor growth dynamics. The network consists of three fully connected layers with 128, 64, and a latent dimension of 5 corresponding to the number of estimated parameters (a, b, n, w, ED_{50}). It uses ReLU activation functions and a constant initialization for the final layer weights and biases. Since the parameters predicted by this network are used in an ODE system, any large or random initial outputs could cause the ODE solver to behave unpredictably or even fail. Constant initialization ensures that the initial predictions are centered within the parameter intervals after mapping, providing the network with a better starting point for parameter estimation.

The training was carried out using the Adam (Adaptive Moment Estimation) optimizer with a learning rate of 0.0001. The loss function calculated the mean squared error of the predicted tumor volumes and the input tumor volumes. The training process continues for a maximum of 500 epochs, or until the loss falls below a threshold of 0.001, or shows no improvement for 10 consecutive iterations (early stopping criterion). The model is saved, and the parameters are considered optimal when the loss is minimal during the whole training iteration process. Regularization was applied to the model parameters, and gradient clipping was used to prevent exploding gradients. For longer time series measurements, we employed alternative configurations due to the tendency of the model to converge to local minima. This was particularly evident for time series exceeding 49 days, where the exponential growth characteristic of tumor dynamics could lead to increased error. In these cases, we utilized `torch.nn.SmoothL1Loss` instead of `torch.nn.MSELoss`. The `torch.nn.SmoothL1Loss`, also known as Huber Loss, is based on applying different penalties depending on the error magnitude, thereby mitigating the impact of large deviations and improving robustness in the presence of rapidly growing tumor volumes.

2.3 The Tumor Model and the Differentiable ODE Solver

In order to describe the tumor dynamics, we utilized the tumor growth model presented in [19]. This model is defined by a set of ordinary differential equations, with four state variables, as detailed in (3)-(6). The first state variable, x_1 [mm^3], represents the living tumor volume, while the second state variable, x_2 [mm^3], captures the dead tumor volume in time. The state variables x_3 [mg·kg^{-1}] and x_4 [mg·kg^{-1}] describe the pharmacokinetics of the chemotherapeutic agent.

$$\dot{x}_1 = (a-n)x_1 - b\frac{x_1 x_3}{ED_{50}+x_3}, \tag{3}$$

$$\dot{x}_2 = nx_1 + b\frac{x_1 x_3}{ED_{50}+x_3} - wx_2, \tag{4}$$

$$\dot{x}_3 = -(c+k_1)x_3 + k_2 x_4 + u, \tag{5}$$

$$\dot{x}_4 = k_1 x_3 - k_2 x_4, \tag{6}$$

The variable u denotes the impulsive input of the model. The injection was considered over an infinitesimally short duration and injected directly into the blood. The variable x_3 describes the drug concentration in the first compartment (blood), while x_4 represents the drug concentration in the second compartment (tissue).

The model is validated using experimental data of mice, where tumor dimensions (width and length) are measured with a caliper, and the total tumor volume (y) is estimated using an approximation formula [37]. In the model, the total tumor volume is represented as the sum of the state variables x_1 and x_2.

Since the analytical solution of this system of ordinary differential equations is not known, we use numerical solvers. Specifically, a fourth-order Runge-Kutta solver is employed to solve the differential equations over the defined time interval with given initial conditions. The model parameters, summarized in Table 1, include a, b, n, w, ED_{50}, c, k_1, and k_2. The lower and upper bounds for these parameters are based on previous mouse experiments, where nonlinear mixed-effects modeling was used to estimate the parameter values of mice. The pharmacokinetic parameters c, k_1, and k_2 were previously determined and therefore set to constant during training.

To solve the differential equations of the tumor model a special ODE solver was needed. The differentiability of the ODE solver was crucial because the model parameters are optimized via gradient-based methods. Without differentiability, gradients of the loss function to the parameters cannot be computed, hindering effective optimization. We implemented the introduced model in Python language. We used the `torchdiffeq` package functions to solve the tumor model, as it

Table 1
The name, applied minimum and maximum values, and unit of measure of the model parameters. The aim of the presented algorithm is to determine these parameters.

Parameter	Name	Min.	Max.	Dimension
a	Proliferation rate coefficient	0.0603	0.8147	day^{-1}
b	Drug efficiency rate coefficient	0.002368	33	day^{-1}
n	Necrosis rate coefficient	2.5360e-05	0.0461	day^{-1}
w	Washout rate coefficient of dead tumor cells	0.0704	0.0988	day^{-1}
ED_{50}	Median effective dose of the drug	1.9256	2.3234	$\text{mg} \cdot \text{kg}^{-1}$
c	Clearance of the drug	1.8211	1.8211	day^{-1}
k_1	Flow rate coefficient of the drug from the central to peripheral compartment	14.008	14.008	day^{-1}
k_2	Flow rate coefficient of the drug from the peripheral to central compartment	136.2781	136.2781	day^{-1}

provides differentiable ODE solvers compatible with automatic differentiation of PyTorch framework [38]. It enables efficient gradient computation and parameter optimization within a framework that integrates ordinary differential equations and neural networks.

2.4 Simulation of Experimental Data for Model Validation

To evaluate the performance of the proposed architecture on time series data of tumor volumes with known physiological parameters, we first generated simulated datasets. Specifically, we created *in silico* tumor growth trajectories under chemotherapy treatment across virtual mouse populations, each characterized by distinct, predefined parameter sets.

Initial conditions for the tumor model were defined as follows: the living tumor volume, $x_1(0)$ was initialized with a random value between 0 and 200 mm^3, reflecting typical starting points for tumor measurements. The remaining state variables at $t = 0$ were initialized to zero based on biological assumptions. The dead tumor volume ($x_2(0)$) was presumed to be negligible. This is justified by the value of necrotic rate n, which has much lower values than the tumor proliferation rate a, see Table 1. Both the drug concentrations in the blood ($x_3(0)$) and tissue ($x_4(0)$) were set to zero, reflecting the fact that no drug had been administered at the initial time point.

The duration of the experiment is determined by its endpoint, which occurs either upon euthanasia of the mouse due to excessive tumor size or upon completion of the experiment. In simulations, we typically use a maximum duration of 105 days, as this is the empirically established lifespan of a mouse with cancer with standard treatments or the time of remission followed by a reasonable amount of dormant state for more advanced therapies [39].

Since excessive tumor size can also be the cause of the termination of the experiment, only simulations in which the tumor volume remained below 2000 mm^3 were utilized.

Tumor volume measurements were recorded three times weekly, reflecting the frequency used in standard preclinical protocols. Chemotherapy dosing schedules were randomly assigned to either two or three administrations per week, with individual doses sampled from a uniform distribution ranging between 0 and 8 mg/kg. Synthetic tumor volume data were generated using parameter sets sampled uniformly within the predefined ranges listed in Table 1.

To evaluate the effect of noise in realistic measurements, zero-mean Gaussian noise was added to the tumor volume measurements. Tumor volumes were sampled three times per week (Monday, Wednesday, and Friday) from the high-resolution trajectory obtained by solving the tumor model with known parameters with an ODE solver.

Noise was added only at the discrete measurement time points to reflect uncertainty arising from various sources of error. The noise values were independently drawn from a Gaussian distribution:

$$\varepsilon_i \sim \mathcal{N}(0, \sigma^2), \quad i = 1, 2, \ldots, N \tag{7}$$

where σ denotes the standard deviation of the noise, and N is the number of measurement days. The standard deviation σ was defined as a fixed absolute value, applied uniformly across all time points and samples in a given simulation. It was not scaled relative to the tumor volume magnitude at each measurement point. This modeling choice allows explicit control over the amplitude of measurement uncertainty and its impact on training. However, it may underestimate the variability of measurements at larger tumor sizes. To evaluate robustness, the noise level (σ) was systematically varied, taking values of 0%, 10%, 20%, and 40%.

After noise was added, to meet the temporal resolution requirements of the ODE solver, interpolation was applied to generate additional data points between the original measurement days. A time step of 0.1 days was used, resulting in 1050 interpolated points over a 105-day period, compared to the original 105 measurements. This smaller stepsize was necessary to avoid large value jumps between consecutive points, which could destabilize the training process. To ensure smooth transitions and differentiability, we applied Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) interpolation [40], which produces curves with continuous first derivatives. This interpolation approach led to a more stable and efficient training process compared to bigger time steps.

3 Results

To evaluate the performance of the proposed architecture, we conducted simulations of tumor volume time series under varying measurement duration and noise. Two examples are presented in Figures 2 and 4. In both cases, tumor volumes were simulated for 105 days using the same chemotherapy dosing regimen but with different underlying parameter values.

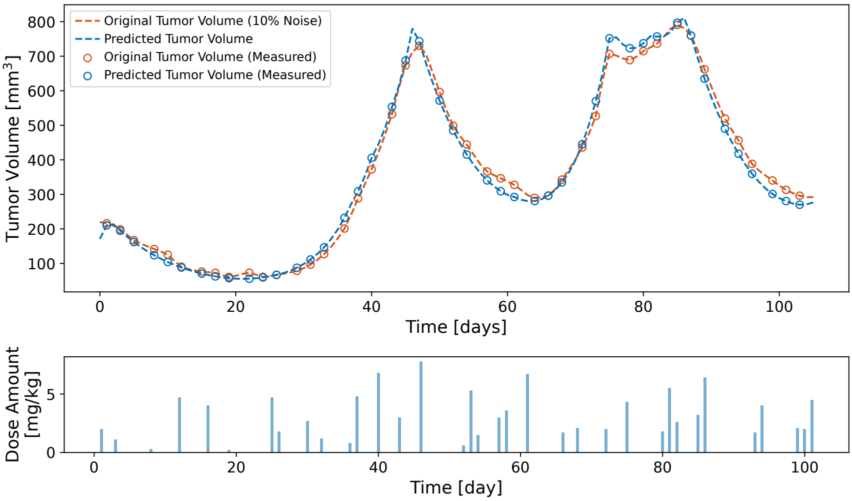

Figure 2
Example results for the 105-day tumor volume data with 10% noise are shown, with predicted tumor volumes represented in blue and the noisy, original measurements in orange. In the lower figure, the corresponding dosing regimen is visualized.

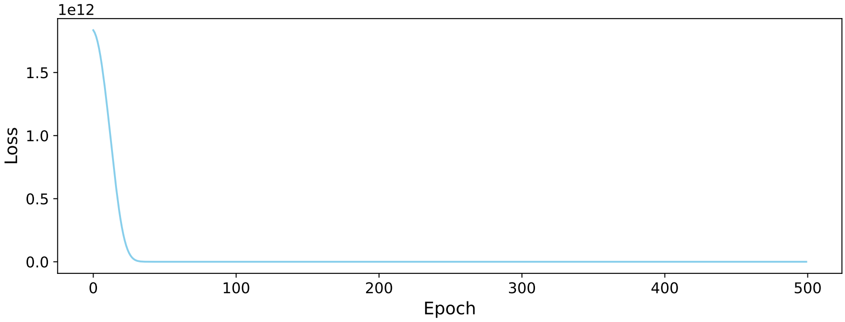

Figure 3
Training loss evolution during the fitting process for the test case shown in Figure 2. The loss is calculated as the mean squared error between predicted and original tumor volumes.

In the upper panels of both figures, the temporal evolution of tumor volumes is visualized. The original noisy measurements, sampled three times weekly (Monday, Wednesday, Friday), are marked with orange circles. A blue dashed line represents the interpolated values using PCHIP, which serves as the ground truth time series for loss computation. The predicted tumor volumes, based on the optimized model parameters, are shown in orange, with corresponding measurement-day markers. Notably, the predicted curves closely follow the trend of the measured data, suggesting that the model effectively captures the underlying tumor dynamics even under moderate noise conditions.

The bottom figures illustrate the corresponding chemotherapy dosing schedules as blue bars. These dose amounts were drawn randomly from a uniform distribution and administered on two or three randomly selected weekdays.

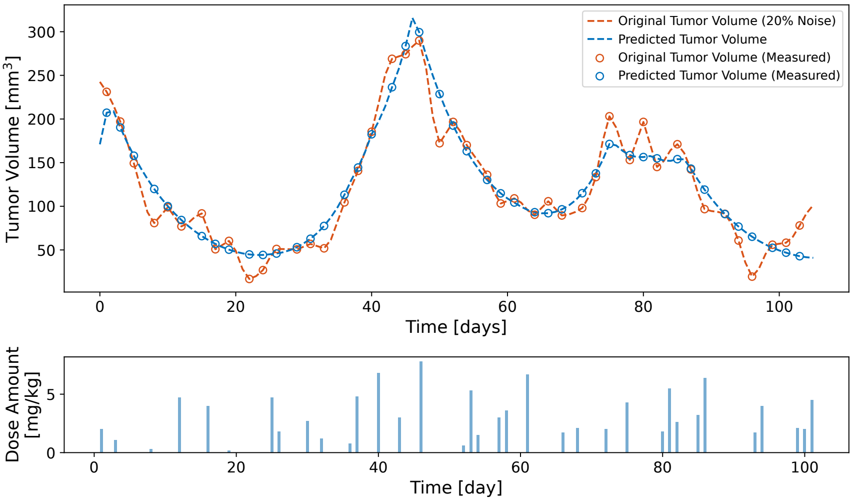

Figure 4
Example results for the 105-day tumor volume fitting are shown, with predicted tumor volumes represented in orange and the original measurements in blue. In the lower figure, the corresponding dosing regimen is visualized.

Figure 3 illustrates the trajectory of the training loss over the training epochs for the case presented in Figure 2. The loss function, defined as the mean squared error between the predicted and interpolated tumor volumes, exhibits a consistent downward trend, indicating stable convergence of the parameter estimation process. For this specific example, the final average deviation between the predicted and reference tumor volumes after the training was 44.4 mm^3, a level of accuracy that is considered acceptable. However, the noise structure *in vivo* is often more complex than the Gaussian noise assumed in this study [41]. A future direction involves integrating more realistic noise models into the simulated data generation process, thereby improving the representativeness of the training data and its alignment with experimental conditions.

To evaluate the parameter prediction capability of the model, parameter estimation was performed on 100 distinct *in silico* datasets for each noise level with known parameter values. To minimize computational time, the evaluation was conducted with a measurement length of 14 and a maximum of 500 epochs. This upper limit for the number of training iterations was determined empirically, as it was observed that training generally converged to optimal solutions beyond 500 epochs.

The accuracy of parameter prediction was evaluated under varying levels of Gaussian noise in tumor volume measurements, as described in Section 2. For each noise level, we simulated 100 test cases. We assessed our method by comparing the result tumor volume (defined by the determined parameters) to the original noiseless tumor volumes.

The differences for these test cases are shown in Figure 5. The figure shows the average differences between the predicted tumor volumes and the original tumor volumes. The errors are categorized into 10 intervals, where the group ID represents the maximum error of that group interval. The first interval shows a low error between the fitted and the original data. We can see that for error levels of 0%, 10%, and 20%, the fitting error remains below 24.5 mm^3 in at least 65% of the test cases. A significant decrease in accuracy is observed at higher noise levels (40%). Nevertheless, even in these cases, the maximum errors remain acceptable, typically staying below 100 mm^3. The median errors for noise levels of 0%, 10%, 20%, and 40% are 11.23, 11.03, 18.25, and 25.85, respectively, while the corresponding mean errors are 31.53, 32.15, 35.20, and 40.67. These values indicate that low levels of Gaussian noise have no significant impact on parameter estimation when applying this method.

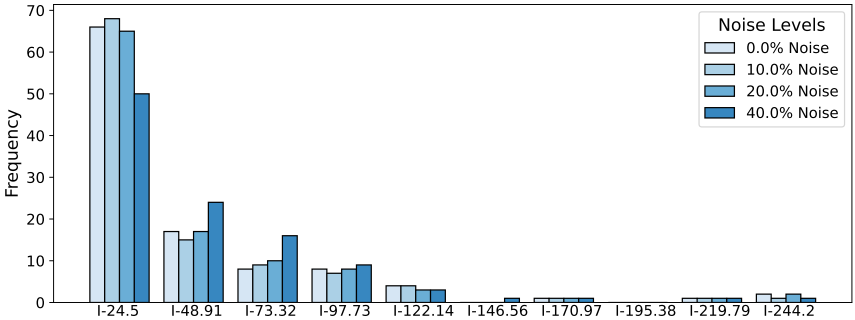

Figure 5
The histogram of the errors between the predicted and original tumor volumes. The labels indicate the error groups, with the ID including the maximum error of that interval. The label intervals are the following: I-24.5: [0.09, 24.50], I-48.91: [24.50, 48.91], I-73.32: [48.91, 73.32], I-97.73: [73.32, 97.73], I-122.14: [97.73, 122.14], I-146.56: [122.14, 146.56], I-170.97: [0.09, 170.97], I-195.38: [170.97, 195.38], I-219.79: [195.38, 219.79], I-244.20: [219.79, 244.20]. The error is determined between the predicted tumor volumes – defined by the identified parameters – and the original tumor volumes. Subsequently, the average of the absolute differences was calculated.

We also assessed the intervals of the predicted parameters. Figure 6 illustrates the distributions of the five parameters, with the original upper and lower limits indicated by grey dashed lines. The distribution of the parameters is visualized with a violin plot, which combines a box plot and a density plot, providing insight into the distribution shape of the parameters. Although a uniform distribution was used to generate the parameters, the predicted values exhibit similar to a bell curve.

This may be due to our test case restrictions, as we excluded data where the tumor volume exceeded 2000 mm^3, a threshold set to reflect realistic tumor sizes in mice. This explanation is further supported by the observed parameter values, where the proliferation rate (a) median is lower and the maximum drug effect (b) median is higher than the center of the plot. This implies that we mainly excluded test cases with significant tumor growth.

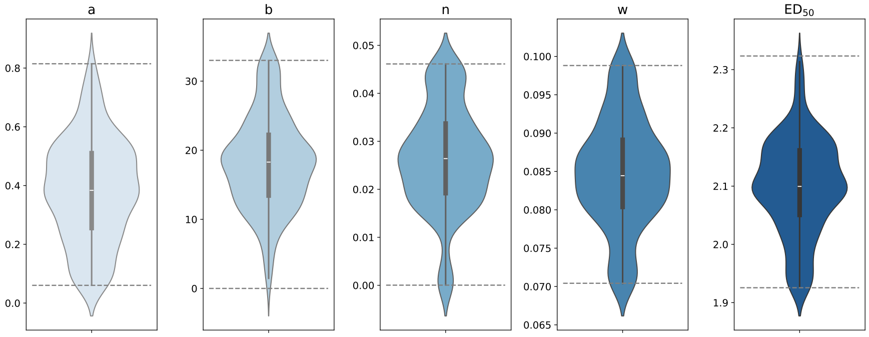

Figure 6
The distribution of the predicted parameters. The grey dashed lines indicate the predefined intervals from Table 1 used for sampling random parameter values from a uniform distribution within this specified range.

Furthermore, we conducted a runtime evaluation as a function of measurement length to assess the impact on computational efficiency. The measurement length was incrementally increased by weeks (seven days), and for each point, the evaluation was executed three times to ensure reliability. The average computation times for each measurement length were then calculated. The results, presented in Figure 7, show a linear relationship between computation time and measurement length, indicating that as the measurement length increases, the computational demands also increase proportionally. It provided helpful information for the determination of the length of the measured tumor volumes during evaluation.

In summary, the results presented in this section demonstrate that the proposed architecture can accurately estimate patient-specific model parameters across a range of Gaussian noise levels. Notably, low to moderate noise (up to 20%) does not substantially affect the performance of the algorithm. These findings highlight the potential of the method for ODE-based personalized therapy, where reliable parameter identification is necessary for effective treatment planning. Furthermore, the integration of a differentiable ODE solver enables the tumor model

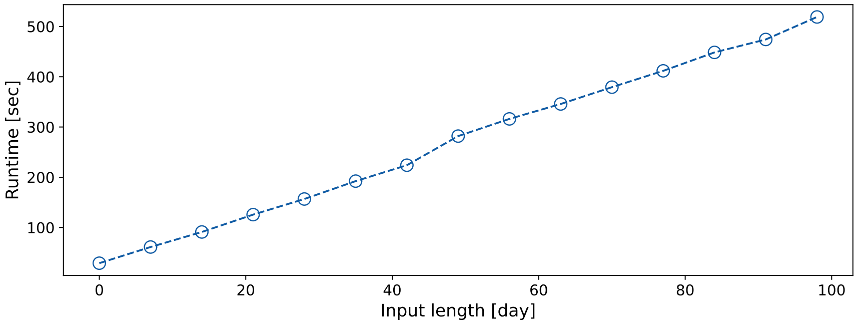

Figure 7
The results of the runtime evaluation for varying input lengths. The input size (the number of measurement days) was incremented by seven days. For each input size, runtime measurements were conducted three times and the average value was calculated and visualized in this figure.

to be incorporated into neural network frameworks that rely on gradient-based optimization, broadening its applicability across widespread machine learning architectures.

It should be emphasized, however, that our evaluation was performed exclusively on *in silico* datasets. In real experimental settings, measurement uncertainty often depends on tumor volume. For example, caliper-based tumor volume measurements exhibit volume-dependent variability: small tumors are associated with higher relative error due to skin interference and manual handling, while larger tumors tend to exhibit lower relative but higher absolute error [41]. Since our noise model assumes a constant standard deviation across all measurements, it may underestimate the true variability, particularly at larger tumor sizes. Future work should therefore include more advanced noise modeling and validation using experimental tumor volume data to assess the method's robustness under real-world measurement conditions. Additionally, benchmarking against existing parameter estimation approaches would help contextualize the relative strengths and limitations of the proposed framework.

Conclusion

In this study, we developed and evaluated a novel approach to estimate tumor model parameters using a neural network, integrated with a differentiable ordinary differential equation (ODE) solver. This framework allows for efficient and quick parameter identification, facilitating personalized cancer therapy optimization. By utilizing tumor volume data and dosing schedules, our model was able to predict parameters that govern tumor dynamics, ensuring a high level of precision in the fitted data. We evaluated our method using *in silico* tumor measurement data under both noiseless and noisy conditions. Our results indicate that low levels of Gaussian noise do not substantially affect parameter estimation performance, whereas higher noise levels lead to decline in precision. However, it is important to note that the current validation was performed only on synthetic datasets. Future work should aim to apply the framework to *in vivo* experimental data in order to assess its robustness to measurement uncertainties. Furthermore, the integration of more realistic noise models, along with comparative analyses against established parameter estimation methods, would offer deeper insights into the practical utility and relative performance of the proposed approach.

Acknowledgment

This project has been supported by the Hungarian National Research, Development and Innovation Fund of Hungary, financed under the TKP2021-NKTA-36 funding scheme. The work of Dániel András Drexler was supported by the Starting Excellence Researcher Program of Obuda University, Budapest Hungary. Lilla Kisbenedek is also with the Obuda University, Applied Informatics and Applied Mathematics Doctoral School, Budapest, Hungary. Lilla Kisbenedek acknowledges the support of the National Talent Program under the NTP-HHTDK-24-0078 project. This research was partially supported by the European Union (EU HORIZON-MSCA-2023-SE-01-01) and the Hungarian NRDI program (2020-2.1.1-ED-2024-00346) within the DSYREKI: Dynamical Systems and Reaction Kinetics Networks project.

References

[1] F. Bray, M. Laversanne, H. Sung, J. Ferlay, R. L. Siegel, I. Soerjomataram, and A. Jemal. Global cancer statistics 2022: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, 74(3):229–263, April 2024.

[2] R. Kaur, A. Bhardwaj, and S. Gupta. Cancer treatment therapies: traditional to modern approaches to combat cancers. *Molecular Biology Reports*, 50(11):9663–9676, October 2023.

[3] D. T. Debela, S. G. Muzazu, K. D. Heraro, M. T. Ndalama, B. W. Mesele, D. C. Haile, S. K. Kitui, and T. Manyazewal. New approaches and procedures for cancer treatment: Current perspectives. *SAGE Open Medicine*, 9:205031212110343, January 2021.

[4] J. C. Baltussen, N. A. de Glas, Y. van Holstein, M. van der Elst, S. Trompet, A. U. den Boogaard, W. van der Plas-Krijgsman, G. Labots, C. Holterhues, J. M. van der Bol, et al. Chemotherapy-related toxic effects and quality of life and physical functioning in older patients. *JAMA Network Open*, 6(10):e2339116–e2339116, 2023.

[5] P. Aguiar, J. J. Adashek, F. Roitberg, C. M. Noia Barreto, A. Del Giglio, and G. L. Lopes. In the era of cost-effectiveness analysis, affordability is a limiting factor for patients' access to innovative cancer treatments. *Value in Health Regional Issues*, 20:47–50, December 2019.

[6] A. Hoeben, E. A. J. Joosten, and M. H. J. van den Beuken-van Everdingen.

Personalized medicine: Recent progress in cancer therapy. *Cancers*, 13(2):242, January 2021.

[7] C. Simsek, E. Esin, and S. Yalcin. Metronomic chemotherapy: A systematic review of the literature and clinical experience. *Journal of Oncology*, 2019:1–31, March 2019.

[8] G. Ouyang, Y. Liu, J. Liu, L. Huang, F. Luo, and L. Li. Efficacy and safety of reduced-dose chemotherapy plus immunotherapy in patients with lung squamous cell carcinoma: A ¡scp¿real-world¡/scp¿ observational study. *Cancer Medicine*, 12(18):18679–18690, September 2023.

[9] L. DeRidder, D. A. Rubinson, R. Langer, and G. Traverso. The past, present, and future of chemotherapy with a focus on individualization of drug dosing. *Journal of Controlled Release*, 352:840–860, December 2022.

[10] R. K. Jain, J. J. Lee, D. Hong, M. Markman, J. Gong, A. Naing, J. Wheler, and R. Kurzrock. Phase i oncology studies: Evidence that in the era of targeted therapies patients on lower doses do not fare worse. *Clinical Cancer Research*, 16(4):1289–1297, February 2010.

[11] M. A. Banks. Challenging the high-dose paradigm for cancer drugs. *Nature*, December 2022.

[12] O. Y. Basar, S. Mohammed, M. W. Qoronfleh, and A. Acar. Optimizing cancer therapy: a review of the multifaceted effects of metronomic chemotherapy. *Frontiers in Cell and Developmental Biology*, 12, May 2024.

[13] M. Bodzioch, P. Bajger, and U. Foryś. Angiogenesis and chemotherapy resistance: optimizing chemotherapy scheduling using mathematical modeling. *Journal of Cancer Research and Clinical Oncology*, 147(8):2281–2299, May 2021.

[14] N. André, M. Carré, and E. Pasquier. Metronomics: towards personalized chemotherapy? *Nature Reviews Clinical Oncology*, 11(7):413–431, 2014.

[15] A. M. Jarrett, D. Faghihi, D. A. Hormuth, E. A. Lima, J. Virostko, G. Biros, D. Patt, and T. E. Yankeelov. Optimal control theory for personalized therapeutic regimens in oncology: Background, history, challenges, and opportunities. *Journal of clinical medicine*, 9(5):1314, 2020.

[16] B. Gergics, M. Puskás, L. Kisbenedek, M. F. Dömény, L. Kovács, and D. A. Drexler. Chemotherapy optimization and patient model parameter estimation based on noisy measurements. *Acta Polytechnica Hungarica*, 21(10):475–493, 2024.

[17] T. O. McDonald, Y.-C. Cheng, C. Graser, P. B. Nicol, D. Temko, and F. Michor. Computational approaches to modelling and optimizing cancer treatment. *Nature Reviews Bioengineering*, 1(10):695–711, 2023.

[18] U. Anand, A. Dey, A. K. S. Chandel, R. Sanyal, A. Mishra, D. K. Pandey, V. De Falco, A. Upadhyay, R. Kandimalla, A. Chaudhary, J. K. Dhanjal,

S. Dewanjee, J. Vallamkondu, and J. M. Pérez de la Lastra. Cancer chemotherapy and beyond: Current status, drug candidates, associated risks and progress in targeted therapeutics. *Genes & Diseases*, 10(4):1367–1401, July 2023.

[19] D. A. Drexler, T. Ferenci, A. Füredi, G. Szakács, and L. Kovács. Experimental data-driven tumor modeling for chemotherapy. In *Proceedings of the 21st IFAC World Congress*, pages 16466–16471, 2020.

[20] M. F. Döményi, M. Puskás, L. Kovács, and D. A. Drexler. In silico chemotherapy optimization with genetic algorithm. In *2023 IEEE 17th International Symposium on Applied Computational Intelligence and Informatics (SACI)*, pages 000097–000102, 2023.

[21] G. Lillacci and M. Khammash. Parameter estimation and model selection in computational biology. *PLoS Computational Biology*, 6(3):e1000696, March 2010.

[22] A. Gábor and J. R. Banga. Robust and efficient parameter estimation in dynamic models of biological systems. *BMC systems biology*, 9:1–25, 2015.

[23] F.-G. Wieland, A. L. Hauber, M. Rosenblatt, C. Tönsing, and J. Timmer. On structural and practical identifiability. *Current Opinion in Systems Biology*, 25:60–69, 2021.

[24] A. Reda, R. Benotsmane, A. Bouzid, and J. Vásárhelyi. A hybrid machine learning-based control strategy for autonomous driving optimization. *Acta Polytechnica Hungarica*, 20(9):165–186, 2023.

[25] A. Yazdani, L. Lu, M. Raissi, and G. E. Karniadakis. Systems biology informed deep learning for inferring parameters and hidden dynamics. *PLOS Computational Biology*, 16(11):e1007575, November 2020.

[26] I. I. Méndez-Gurrola, A. Ramírez-Reyes, and A. I. Barranco-Gutiérrez. A review and perspective on the main machine learning methods applied to physical sciences. *Acta Polytechnica Hungarica*, 19(10):205–220, 2022.

[27] L. Lu, P. Jin, G. Pang, Z. Zhang, and G. E. Karniadakis. Learning nonlinear operators via deeponet based on the universal approximation theorem of operators. *Nature Machine Intelligence*, 3(3):218–229, March 2021.

[28] M. Raissi, P. Perdikaris, and G. E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, 378:686–707, 2019.

[29] J. H. Lagergren, J. T. Nardini, R. E. Baker, M. J. Simpson, and K. B. Flores. Biologically-informed neural networks guide mechanistic modeling from sparse experimental data. *PLoS computational biology*, 16(12):e1008462, 2020.

[30] C. S. Greene and J. C. Costello. Biologically informed neural networks predict drug responses. *Cancer cell*, 38(5):613–615, 2020.

[31] K. Sel, A. Mohammadi, R. I. Pettigrew, and R. Jafari. Physics-informed neural networks for modeling physiological time series for cuffless blood pressure estimation. *npj Digital Medicine*, 6(1), June 2023.

[32] R. T. Q. Chen, Y. Rubanova, J. Bettencourt, and D. Duvenaud. Neural ordinary differential equations, 2018.

[33] M. Laurie and J. Lu. Explainable deep learning for tumor dynamic modeling and overall survival prediction using neural-ode. *npj Systems Biology and Applications*, 9(1):58, 2023.

[34] I. B. Losada and N. Terranova. Bridging pharmacology and neural networks: A deep dive into neural ordinary differential equations. *CPT Pharmacometrics Syst. Pharmacol.*, 13(8):1289–1296, August 2024.

[35] M. Puskás and D. A. Drexler. Tumor model parameter estimation for therapy optimization using artificial neural networks. In *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 1254–1259, 2021.

[36] L. Kisbenedek, M. Puskás, and D. A. Drexler. Indirect supervised fine-tuning of a tumor model parameter estimator neural network. In *IEEE 17th International Symposium on Applied Computational Intelligence and Informatics SACI 2023*, pages 109–116, 2023.

[37] J. Sápi, L. Kovács, D. A. Drexler, P. Kocsis, D. Gajári, and Z. Sápi. Tumor volume estimation and quasi-continuous administration for most effective bevacizumab therapy. *PLOS ONE*, 10(11):e0142190, November 2015.

[38] R. T. Q. Chen. torchdiffeq, 2018.

[39] L. Kovács, T. Ferenci, B. Gombos, A. Füredi, I. Rudas, G. Szakács, and D. A. Drexler. Positive impulsive control of tumor therapy—a cyber-medical approach. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 54(1):597 – 608, 2024.

[40] F. N. Fritsch and J. Butland. A method for constructing local monotone piecewise cubic interpolants. *SIAM Journal on Scientific and Statistical Computing*, 5(2):300–304, 1984.

[41] M. Puskás and D. A. Drexler. Modeling the error of caliper measurements in animal experiments. *IEEE Access*, 13:54836–54852, 2025.