Optimal Therapy Design Using Artificial Neural Network Surrogate Models of Fluid and Solute Transport in Tumors

by

Samuel Thomas Degnan-Morgenstern

Submitted to the Department of Chemical & Biomolecular Engineering
in partial fulfillment of the requirements for the degree of

Honors Scholar, Bachelor of Science in Chemical Engineering

at the

University of Connecticut

May 2022

Author . . . .
Department of Chemical & Biomolecular Engineering
April 29, 2022

Certified by . . . .
Dr. Matthew D. Stuber
Assistant Professor
Thesis Supervisor

Accepted by . . . .
Dr. Kelly A. Burke
Honors Advisor

Contents

Abstract

The application of quantitative and formal methods in cancer research can enable a fundamental understanding of the limitations of conventional therapies, uncover novel treatment strategies, and help optimize therapy for individual patient outcomes. Particularly, normalization of the tumor microenvironment (TME) has been proposed to improve the delivery of oxygen, antibodies, and nanomedicines to solid tumors and in turn treatment efficacy. A physiological transport framework based on a mechanistic model has been established to simulate the fluid and macromolecule transport in a tumor (Baxter et al). This model was used to solve parameter estimation problems to fit the model against in vivo experimental data using deterministic global optimization. However, it was found that the parameter estimation problems are computationally expensive to achieve global optimality using currently developed bounding routines for the mechanistic model. In this work, we propose to use artificial neural networks (ANN) to establish surrogate models for solving the parameter estimation problems to global optimality. Furthermore, we also propose to use the ANN surrogates to enable the solution to a simultaneous dosage and nanocarrier design problem for better patient outcomes. The use of ANN models has been confirmed to provide accurate solutions for solving challenging global optimization problems with substantially accelerated speed. The experimentally validated computational modeling approach enables fast, multidimensional TME-normalizing therapy design and anticancer drug design. Furthermore, this work provides insight into how in silico modeling approaches can aid in predicting dose-response in preclinical studies.

Introduction

Solid tumors are characterized by a heterogeneous tumor microenvironment (TME) consisting of a compressed extracellular matrix (ECM) and leaky, underdeveloped blood vessels [1]. These irregularities are significant biophysical barriers to the delivery of anticancer drugs during treatment, thus harming patient outcomes. In recent years, researchers have studied the TME using an *in silico* modeling approach with a first principles model of fluid and solute transport in the tumor [2]. In particular, Martin et al. [3] studied TME normalization both experimentally and computationally using a mechanistic model of fluid and solute transport in tumors as visualized in Figure 1-1. Using the glucosteroid dexamethasone (DEX), the authors found that TME normalization therapies administered before treatment will enable more effective drug delivery by decompressing the extracellular matrix and repairing the tumor vasculature [3]. Within this study, the numerical experiments utilized local optimization to estimate the underlying physiological parameters and quantify the effects of the studied treatments. Wang et al. [4] built upon this work by developing a rigorous computational framework for studying and designing TME normalizing therapies using deterministic global optimization. In this thesis, we explore the role of machine learning based surrogate models to reduce the computational burden of this workflow and expand the possibilities of therapy design.
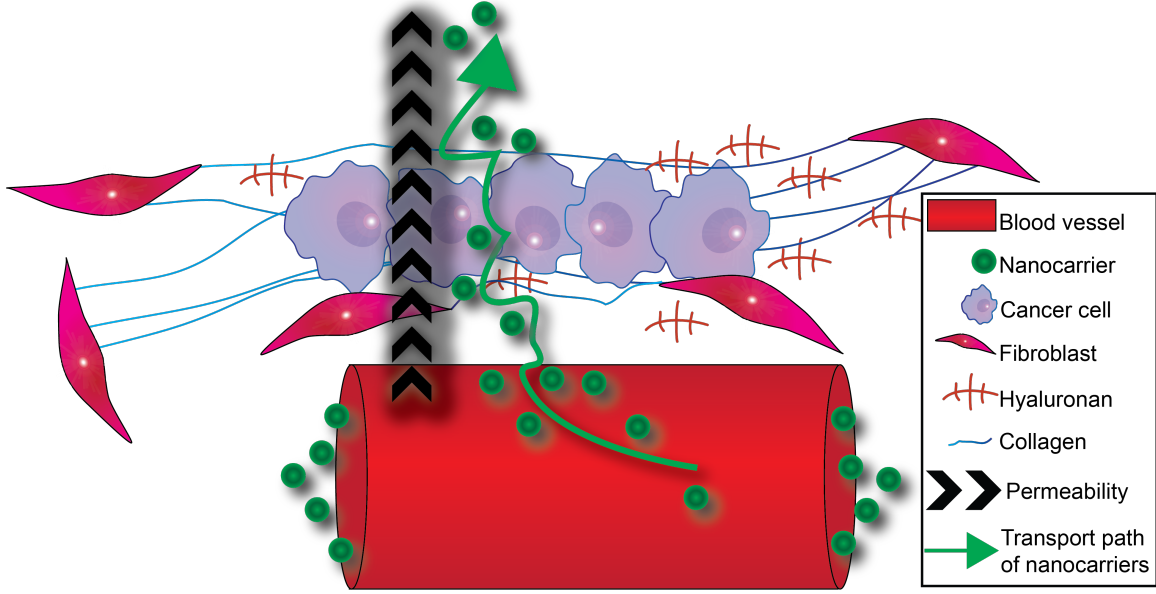
Figure 1-1: A schematic representing fluid and solute transport in the TME [3]. Nanocarriers flow through the porous blood vessels into the tumor interstitium through a combination of convective and diffusive transport.

1.1 Modeling and Simulation for Cancer

Using a first-principles approach, Baxter and Jain developed a comprehensive mathematical framework to analyze fluid and mass transport in the tumor microenvironment[5, 6, 7, 8]. Specifically, they established the first model that captured the complex transport phenomena in the porous vasculature and interstitium of the tumor. The development of this model allowed for a two-pronged approach to cancer research that coupled experimental and numerical methods. Specifically, researchers used computational models to guide experimental analysis, and conversely they used experimental data to develop more representative models.

Baxter and Jain utilized a one-dimensional model in a spherical coordinate system to simplify their simulations [5]; however, further studies have developed and applied models of higher complexity. Baish et al. [9] and Chauhan et al. [10] established a two dimensional fluid and solute transport model that accounts for the coupling between the tumor vascular and interstitial space. Sweeney et al. [11] expanded on this by developing a three dimensional model of fluid transport in tumors that fully accounted for the spatial heterogeneity. Regardless of model dimensionality, these studies of TME normalization generally recommended the establishment of a pressure gradient in the tumor to increase convective mass transport. Subsequently,

this study will rely on the more computationally efficient one-dimensional model.

In the past, mathematical oncology studies have relied heavily on the sensitivity analysis of physiological model parameters to draw biological conclusions [12]. Furthermore, many of these works do not employ experimental validation to ensure agreement between *in silico* and *in vivo* studies [12]. Using a systems engineering approach, Martin et al. [3] applied local optimization methods to fit their mechanistic model against experimental data on the physiological response to DEX pretreatment. This study effectively established a workflow for formally validating the mathematical model against experimental data to conclude the efficacy of different treatment strategies. Due to the nondeterministic nature of local optimization, Wang et al. [4] expanded this approach by developing a systematic approach to model TME normalization therapy *in silico* using deterministic global optimization. In this workflow, a parameter estimation problem is first solved to find the model parameters based on experimental data tracking solute accumulation. This phase of the workflow enables specific experimental treatments to be analyzed to find their biological responses (such as changes in pore size and pressure gradients). Following the parameter estimation, the varying therapeutic responses are generalized through simple regression models. Then, this information is used to perform treatment design studies to find optimal treatment regimens. Specifically, the dose of DEX and the size of the nanocarrier (agent used to deliver the anticancer drug) are optimized to maximize the accumulation of nanocarriers in the tumor interstitium. These optimization-based studies facilitate a rigorous computational workflow for studying TME normalization to provide deeper physical insight and optimal treatment recommendations.

1.2 Machine Learning for Surrogate Modeling

The use of global optimization for model validation and optimal therapy design is critical to ensure the best patient outcomes [13]. However, a major challenge is the computational intractability associated with performing global optimization on partial differential equation (PDE) based models such as the mechanistic tumor transport model [14]. To perform these studies, custom bounding procedures are developed for the model's governing PDEs. Despite this, global optimization using the mechanistic transport model exhibits slow convergence and a limited ability to scale for multi-dimensional problems [4]. As such, this work explores the use of machine learning based models to reduce the complexity of the mechanistic tumor transport model and accelerate the use of deterministic global optimization.

Machine learning based surrogate modeling is a powerful framework for simplifying computationally expensive simulations through the creation of black-box regression models that capture input-output relationships with high accuracy [15]. In particular, artificial neural network (ANN) based deep learning approaches are extremely capable of approximating nonlinear mappings provided there is a significant amount of data [16]. The simplest form is known as a feed-forward multilayered perceptron (MLP), which is composed of an input layer, a series of one or more hidden layers, and an output layer [16] as shown in Figure 1-2. Interested readers are advised to consult the work of Svozil et al. [17] for a complete technical description of MLPs.
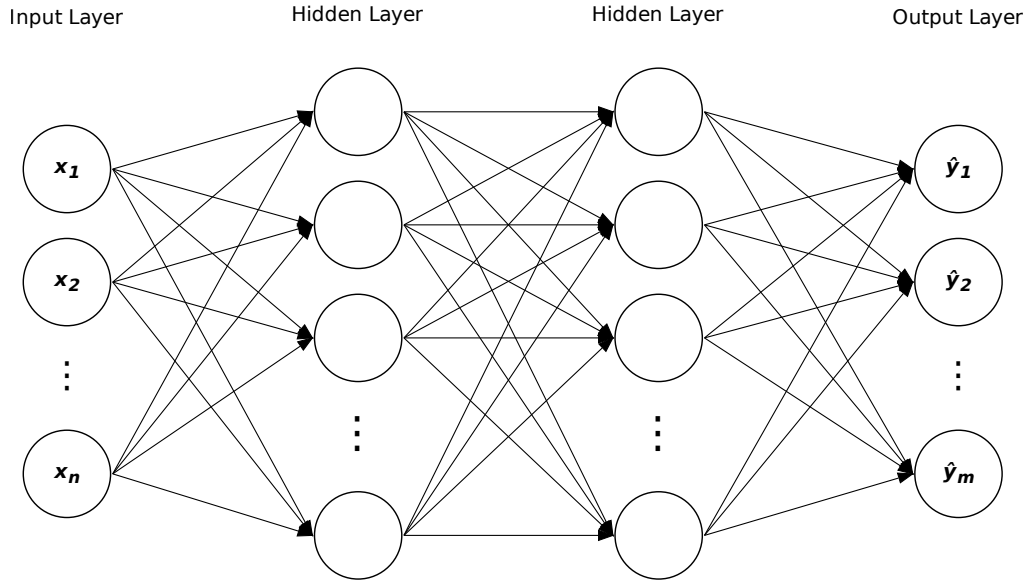

Figure 1-2: A general feed-forward multilayer perceptron model (MLP) is depicted here. MLPs are universal approximators that can represent any function to an arbitrary degree of accuracy. Here, the model takes in an input vector $\mathbf{x} \in \mathbb{R}^n$ and maps it to a target output vector $\hat{\mathbf{y}} \in \mathbb{R}^m$. Each layer is fully connected to the next layer in the network, and each connection is associated with a weight. For a given node, the node's value is computed by first taking a linear combination of all the previous layers nodes and the associated weights, then adding a bias term, and then passing the value through an activation function. This procedure is repeated at every node in all hidden layers and the output layer. The model designer has a multitude of hyperparameter choices including: the number of hidden layers, the number of nodes in each hidden layer, and the activation function for a given layer.

Previously, ANN surrogate models have been used to reduce the computational

cost of solving non-deterministic optimization problems [18, 19]. However, deterministic global optimization with ANN surrogates still remains an emerging area of research [20]. Current developments have demonstrated it can be a viable and effective solution methodology in chemical engineering applications [20, 21, 22]. Furthermore, our in house global optimizer, EAGO.jl [23, 24] provides integrated support for global optimization with ANN based surrogate models. In this work, we extend these developments by successfully deploying an ANN surrogate based approach to solve challenging global optimization problems for parameter estimation and cancer therapy design.

A machine learning based modeling framework is used to develop a simplified workflow for studying TME normalization. ANN surrogate models are developed as a highly accurate, black-box representation of the underlying transport phenomena. Using the surrogate models, multidimensional parameter estimation and therapy design problems are accurately solved to global optimality with significantly reduced solution times relative to the equivalent approach with the mechanistic model. Machine learning based modeling serves a key role in establishing a practical and scalable *in silico* framework for designing optimal TME normalizing cancer treatments.

Methods

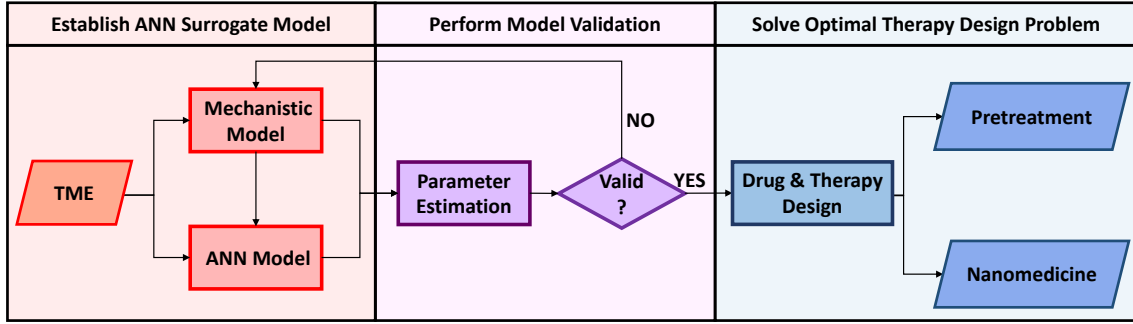Figure 2-3 illustrates a high-level overview of the workflow employed in this study.

Figure 2-3: The computational workflow for this study is illustrated in this figure. The first phase of this study utilizes physical knowledge of the TME to establish a mechanistic model of fluid and solute transport. The ANN surrogate model is developed using data generated from the mechanistic model. Using experimental data, a parameter estimation problem is solved with the surrogate model to reveal the effect of different DEX treatments on key biological parameters. The gained insight is then used to solve a drug and therapy design problem. In this problem, an optimal DEX dosage and nanocarrier size are found such that nanocarrier accumulation is maximized in the tumor interstitium.

2.1 Tumor Transport Model

In this section, we describe the 1-dimensional tumor transport model introduced by Baxter and Jain [5, 6, 7, 8] to model fluid and solute transport in tumors. Tumors are spatially heterogeneous and consist of a necrotic core, a rapidly dividing outer region, and a large vasculature to support this continuous growth. These differences in characteristics cause a spatial dependence of physiological parameters that introduces a high degree of complexity. For the purpose of this study, we make the simplifying assumption of spatial homogeneity. Furthermore, we ignore the role of lymphatic and extravascular binding because of their insignificant presence in tumors. We do not

explicitly consider the presence of microscopic structures in the tumor, such as blood vessels, cells, and the extracellular matrix (ECM), since we are focused primarily on capturing macroscopic phenomena. Spatial averaging is employed on the results of the solute transport model to provide a concise and direct description of the underlying dynamics. Finally, we assume that the sources of fluid and solute are continuously distributed throughout the tumor.

2.1.1 Fluid Transport

The interstitial fluid transport in a tumor follows Darcy's law for axisymmetric flow in spherical coordinates:

$$\mathbf{u} = -K\nabla\mathbf{p} \tag{2.1}$$

which simplifies to

$$u = -K\frac{dp}{dr}, \tag{2.2}$$

where u is the interstitial fluid velocity ($\frac{cm}{s}$), K is the hydraulic conductivity of the tumor interstitium ($\frac{cm^2}{mm\ Hg\ -\ s}$), p is the interstitial fluid pressure (mm Hg), and r is the radial position (cm). The fluid continuity equation for steady-state incompressible flow in a tumor in 1D spherical coordinates is given by:

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2 u\right) = L_p\frac{S}{V}(p_v - p), \tag{2.3}$$

where L_p is the hydraulic conductivity of the vascular wall ($\frac{cm^2}{mm\ Hg\ -\ s}$), $\frac{S}{V}$ is the ratio of the vascular surface area per unit volume (cm^{-1}), and p_v is the vascular pressure (mm Hg). Substituting (2.2) into (2.3), we obtain the steady-state fluid transport model as follows:

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dp}{dr}\right) = \frac{\alpha^2}{R^2}(p - p_{ss}), \tag{2.4}$$

where

$$\alpha = R\sqrt{\frac{S}{V}\frac{L_p}{K}}. \tag{2.5}$$

Here, α is a dimensionless parameter, R is the radius of the tumor (cm), and p_{ss} is the steady-state interstitial pressure where the efflux and influx from the vasculature are equal. For this study, p_v and p_{ss} are equivalent.

This model has the following boundary conditions:

$$\left.\frac{dp}{dr}\right|_{r=0} = 0,$$

$$p|_{r=R} = p_\infty$$

which represents the no-flux symmetry condition at the center of the tumor and a fixed tissue pressure p_∞ (mm Hg) at the tumor edge R.

The analytical solution to the fluid transport model can be derived in dimensionless form as:

$$\hat{p}(\hat{r}) = \left(1 - \frac{\sinh(\hat{r}\alpha)}{\hat{r}\sinh(\alpha)}\right), \tag{2.6}$$

where

$$\hat{r} = \frac{r}{R},$$

$$\hat{p} = \frac{p - p_\infty}{p_{ss} - p_\infty}.$$

Here, $\hat{r}$ is the dimensionless form of the radial position in the tumor and $\hat{p}$ is the dimensionless form of the interstitial fluid pressure.

2.1.2 Solute Transport

Nanocarrier solute transport in tumors is characterized by the convection-diffusion equation:

$$\frac{\partial c}{\partial t} + \nabla \cdot (\mathbf{u}c) = \nabla \cdot (D\nabla c) + \phi_s \tag{2.7}$$

which simplifies to the following in 1D spherical coordinates:

$$\frac{\partial c}{\partial t} + \frac{\partial (r^2 u c)}{\partial r} = D\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial c}{\partial r}\right) + \phi_s. \tag{2.8}$$

11

Here, c is the solute concentration in the tumor interstitium ($\frac{\text{g}}{\text{mL}}$), D is the diffusion coefficient ($\frac{\text{cm}^2}{\text{s}}$), and ϕ_s is the distributed source term based on the vessel pore model for transcapillary exchange as given by:

$$\phi_s = L_p \frac{S}{V}(p_v - p)(1 - \sigma)c_v + P\frac{S}{V}(c_v - c)\frac{Pe}{e^{Pe} - 1}. \tag{2.9}$$

Here, σ is the solute reflection coefficient, P is the vascular permeability of the solute ($\frac{\text{cm}}{\text{s}}$) through the vascular wall, and c_v is the vascular solute concentration that decays exponentially over time as represented in the following relationship:

$$c_v = c_o e^{-t/k_d}$$

where c_o is the initial concentration of blood solutes ($\frac{\text{g}}{\text{mL}}$) and k_d is the half-life circulation of nanocarriers. Additionally, Pe is the Péclet number representing the ratio of convective flux to diffusive flux across the vascular wall as given by:

$$Pe = \frac{L_p(p_v - p)(1 - \sigma)}{P}.$$

For large values of the Pe number ($Pe > 1000$), we can simplify (2.9) by recognizing that the term $\frac{Pe}{e^{Pe}-1}$ approximately goes to 0. The source term for the high-Péclet (HP) regime is given by the following equation:

$$\phi_{s,HP} = L_p \frac{S}{V}(p_v - p)(1 - \sigma)c_v. \tag{2.10}$$

It is assumed that there are initially no solutes in the tumor interstitium before injection; thus, the solute transport model has the following initial condition:

$$c(t = 0, r) = 0.$$

The solute transport model has the following boundary conditions:

$$-D\left.\frac{\partial c}{\partial r}\right|_{r=0} + uc(r = 0) = 0,$$
$$c(r = R) = c_\infty,$$

which represents the no-flux condition in the center of the tumor and the fixed continuous concentration at the edge of the tumor, as denoted by c_∞ ($\frac{\text{g}}{\text{mL}}$) representing

the concentration in the surrounding tissue.

2.1.3 Pore Theory

We model the pore theory based on the approach developed by Deen [25] and Bungay and Brenner [26]. Specifically, we assume the vascular pores to be cylindrical and we denote the transvascular hydraulic conductivity L_p, the vascular permeability P, and the solute reflection coefficient σ as follows:

$$L_p = \frac{\gamma r_o^2}{8\mu L} \tag{2.11}$$

$$P = \frac{\gamma H D_o}{L}$$

$$\sigma = 1 - W$$

where γ is the fraction of vascular surface area occupied by pores, r_o is the pore radius (nm), μ is the blood viscosity (mm Hg -s), and L is the thickness of the vascular wall (cm). D_0 is the diffusion coefficient of the nanocarrier diffusion coefficient in free solution at 37°C given by the Stokes-Einstein relationship as follows:

$$D_0 = \frac{k_B T}{6\pi\mu r_p}$$

where k_B is the Boltzmann constant, T is the absolute temperature, and r_p is the radius of the nanocarrier. H and W, respectively, denote the diffusive and convective hindrance factors and are given by:

$$H = \frac{6\pi\Phi}{K_t},$$

$$W = \frac{\Phi(2 - \Phi)K_s}{2K_t},$$

where Φ is the partition coefficient denoting the equilibrium ratio of the average intrapore concentration to the average concentration in the bulk medium. For purely steric interactions between solutes and pore wall, the partition coefficient is given by $\Phi = (1 - \lambda)^2$ and $\lambda = \frac{r_p}{r_0}$, where λ is the ratio between the particle size r_p (nm) and

the pore size r_0 (nm). Furthermore, the K_t and K_s factors are given by:

$$K_t = \frac{9}{4}\pi^2\sqrt{2}(1-\lambda)^{-5/2}[1 + \sum_{k=1}^{2}\alpha_k(1-\lambda)^k] + \sum_{k=0}^{4}\alpha_{k+3}\lambda^k$$

$$K_s = \frac{9}{4}\pi^2\sqrt{2}(1-\lambda)^{-5/2}[1 + \sum_{k=1}^{2}\beta_k(1-\lambda)^k] + \sum_{k=0}^{4}\beta_{k+3}\lambda^k$$

where α_k and β_k are the coefficients defined in Table 2.1 below.

Table 2.1: The hydrodynamic coefficients for the cylindrical pore model are given in this table. [25].

k	1	2	3	4	5	6	7
α_k	-73/60	77293/50400	-22.5083	-5.6117	-0.3363	-1.216	1.647
β_k	7/60	-2227/50400	4.0180	-3.9788	-1.9215	4.392	5.006

2.2   Globally Optimal Parameter Estimation

DEX has been identified as an effective pre-treatment adjunct agent for normalizing the vasculature and extracellular matrix in solid tumors [3]. The induced process of tumor microenvironment (TME) normalization has been shown to improve nanocarrier accumulation and thus promote tumor shrinkage [27]. In Martin et al. [3], the tumor transport model introduced in Section 2.1 was used to solve a parameter estimation problem through local optimization. In particular, this study found the optimal values of interstitial and vascular hydraulic conductivity to obtain the best fit between the mathematical model and the experimental data collected in that study.

In this section, we present the formulation of Wang et al. [4] to verify these solutions using deterministic global optimization to guarantee the most accurate fit between the model and the experimental data. We show how this challenging global optimization problem is simplified by linearizing the inequality constraints and utilizing a surrogate modeling approach.

2.2.1   Parameter Estimation using Mechanistic Transport Model

In the experiments conducted by Martin et al. [3], a control dose and DEX at 3 mg/kg and 30 mg/kg was administered into mice implanted with 4T1 tumors. After four days of treatment, the nanocarriers 70 kDa rhodamine-bound and 500 kDa FITC-bound

dextran were injected. Experimental data characterizing the distribution of dextran was collected using a confocal laser scanning microscope system. The resulting images were analyzed to quantify the effective permeability P_{eff} of the transvascular flux. Effective permeability quantifies both convective and diffusive solute transport, but it overestimates diffusive transport [5]. Using the measured effective permeability, the accumulation of nanocarriers in the tumor interstitium is calculated with the conservation equation:

$$\frac{dc_{\text{avg}}^{\text{data}}}{dt} = P_{eff}\frac{S}{V}\left(c_v - c_{\text{avg}}^{\text{data}}\right), \tag{2.12}$$

where $c_{\text{avg}}^{\text{data}}$ (g/mL) is the spatially averaged accumulation of nanocarriers. This equation is solved over a discrete time-horizon to yield an experimental concentration profile.

The work of Wang et al. sought to determine key physiological parameters in the tumor transport models to validate the models on the experimental data and verify the physiological effects of TME normalization. The parameter estimation problem seeks to minimize the sum-of-squared errors (SSE) between the dimensionless spatially averaged concentration profiles provided by the experimental measurements calculated from (2.12) and the tumor model described in Section 2.1. Both concentration profiles are calculated over a discrete time horizon of 5 minutes with 21 nodes to maximize the accuracy of the effective permeability measurement in (2.12). To limit the feasible set to physically viable solutions, inequality constraints are utilized on the dimensionless interstitial fluid pressure in the superficial region of the tumor (the periphery of the tumor). The parameter estimation problem is formulated as:

$$\min_{\boldsymbol{\pi}\in\Pi}\sum_{i=1}^{n}\left(\hat{c}_{\text{avg}}(t_i,\boldsymbol{\pi}) - \hat{c}_{\text{avg}}^{\text{data}}(t_i)\right)^2 \tag{2.13}$$

$$\text{s.t. } \hat{p}_{peri}(\boldsymbol{\pi}) \le \hat{p}_{peri,\max}$$

$$\hat{p}_{peri}(\boldsymbol{\pi}) \ge \hat{p}_{peri,\min},$$

where the spatially-averaged experimental concentration profile $\hat{c}_{\text{avg}}^{\text{data}}$ is calculated using the dimensionless form of (2.12), the model predicted spatially-averaged concentration profile $\hat{c}_{\text{avg}}$ is calculated by solving (2.8) in dimensionless form, and the dimensionless tumor pressure in the superficial region $\hat{p}_{peri}$ is calculated by evaluating (2.6) at $\hat{r} = 0.99$. Here, $\boldsymbol{\pi} \in \Pi \subset \mathbb{R}^{n_\pi}$ represents the unknown physiological parameters of interest: interstitial hydraulic conductivity K and vascular hydraulic

conductivity L_p. The experimental and model-predicted concentrations were calculated at discrete time steps t_i with $i \in \{1, \ldots, n\}$ where $n = 21$. The lower and upper bounds on peripheral tumor pressure denoted respectively by $\hat{p}_{peri,\min}$ and $\hat{p}_{peri,\max}$ were measured experimentally by Martin et al. [3] and reported in Table 2.2. The parameter estimation problem was solved for all 6 experimental cases corresponding to the administration of the 70 kDa and 500 kDa tracer molecules with pre-treatment of 0 mg/kg DEX (control), 3 mg/kg DEX, and 30 mg/kg DEX. In Wang et al. [4], the parameter estimation problem is solved to global optimality using the spatial branch-and-bound algorithm [28] as implemented in the EAGO solver [23, 24]. This algorithm requires the calculation of rigorous global bounds on the system of interest. For transient systems governed by partial differential equations (PDE), rigorous bounds are extremely computationally expensive to compute. To solve this problem, Wang et al. [4] implemented custom bounding routines for the tumor transport model using differential inequalities.

Table 2.2: The bounds on the interstitial fluid pressure in the superficial region of the tumor for the control, 3 mg/kg, and 30 mg/kg DEX treatment case are listed in this table as measured experimentally by Martin et al. [3].

Dose	Control	3 mg/kg	30 mg/kg
$\hat{p}_{peri,\min}$ (mmHg)	4.87	3.02	1.95
$\hat{p}_{peri,\max}$ (mmHg)	5.67	3.62	2.45

2.2.2 Simplification of Inequality Constraints

The parameter estimation formulated in (2.13) is subject to complex nonlinear constraints on the IFP in the periphery of the tumor. In this section, we show a procedure for reducing nonlinear constraints to simple linear constraints on the decision variables. This reformulation reduces the computational complexity of solving the parameter estimation problem.

The IFP in the superficial region of the tumor is given by:

$$\hat{p}_{peri}\left(L_p, K\right) = 1 - \frac{\sinh\left(\hat{r}_{peri}\alpha\right)}{\hat{r}_{peri}\sinh\left(\alpha\right)} \tag{2.14}$$

where $\hat{r}_{peri}$ is the dimensionless radial value of the superficial region of the tumor and α is a function of the decision variables L_p and K as given by (2.5). For this problem,

it is known $L_p, K, \alpha > 0$ due to the underlying physiology of the TME.

The peripheral IFP given by (2.14) can be fixed at an arbitrary constant IFP $\tilde{p}$ as follows:

$$\tilde{p} = 1 - \frac{\sinh(\hat{r}_{peri}\alpha)}{\hat{r}_{peri}\sinh(\alpha)}$$

and then differentiated with respect to L_p to yield:

$$\left(-\frac{\hat{r}_{peri}\sinh(\alpha)\cosh(\hat{r}_{peri}\alpha) - \sinh(\hat{r}_{peri}\alpha)\cosh(\alpha)}{\hat{r}_{peri}\sinh^2(\alpha)}\right)\left(\frac{d\alpha}{dL_p}\right) = 0.$$

Since $\alpha > 0$, the expression $\left(-\frac{\hat{r}_{peri}\sinh(\alpha)\cosh(\hat{r}_{peri}\alpha)-\sinh(\hat{r}_{peri}\alpha)\cosh(\alpha)}{\hat{r}_{peri}\sinh^2(\alpha)}\right)$ can be verified to be equal to a nonzero constant. Thus, we obtain the following result:

$$\frac{d\alpha}{dL_p} = \frac{d\left(R\sqrt{\frac{SL_p}{VK}}\right)}{dL_p} = 0. \tag{2.15}$$

Given that α is only a function of the decision variables L_p and K, in order for the expression given by (2.15) to hold, K must necessarily be a scalar multiple of L_p. Through this procedure, we proved that if (2.14) is fixed at a constant IFP then:

$$\exists\, \zeta \in \mathbb{R} \text{ such that } K = \zeta L_p. \tag{2.16}$$

To produce the nonlinear constraints in the parameter estimation problem given by (2.13), the dimensionless pressure at the periphery of the tumor is bounded by the the experimentally measured values in Table 2.2. Since the upper and lower bounds are fixed, we can apply (2.16) to yield linear constraints on the decision variables as follows:

$$K \leq \zeta_{\max} L_p, \tag{2.17}$$
$$K \geq \zeta_{\min} L_p, \tag{2.18}$$

where $\zeta_{\max}$ and $\zeta_{\min}$ are constants tabulated in Table 2.3. These constants can be

calculated by solving for ζ_j in the following equation:

$$\hat{p}_{peri,j} = 1 - \frac{\sinh\left(\hat{r}_{peri}\left(R\sqrt{\frac{S}{V}\zeta_j}\right)\right)}{\hat{r}_{peri}\sinh\left(\left(R\sqrt{\frac{S}{V}\zeta_j}\right)\right)},$$

where $\hat{p}_{peri,j}$ denotes the jth peripheral pressure entry in Table 2.2 and ζ_j denotes the corresponding constant on the inequality constraint. Note that an accelerated procedure to compute ζ_j was used in this study by the following steps:

1. Generate a set of values of L_p and K that satisfy (2.14) at a fixed pressure $\hat{p}_{peri,j}$ in Table 2.2.

2. Calculate a line of best fit by performing linear regression on the generated samples.

3. Record the slope of the trend line as the corresponding ζ_j.

Table 2.3: The coefficients of the linearized superficial pressure constraints for the control, 3 mg/kg, and 30 mg/kg DEX treatment case are listed in this table as calculated by the procedure in Subsection 2.2.2.

Dose	Control	3 mg/kg	30 mg/kg
$\zeta_{\min}$ (cm)	0.2855	0.7355	1.5898
$\zeta_{\max}$ (cm)	0.3967	1.0577	2.4447

2.2.3 Simplified Parameter Estimation Using Surrogate Model

The parameter estimation problem introduced in (2.13) can be further simplified by using an artificial neural network (ANN) surrogate model. An ANN is used as a regression model to learn the underlying mapping relationship between the decision variables L_p and K and the average accumulation profile $\hat{c}_{avg}$. The ANN surrogate model offers a reduction in computational complexity relative to the evaluation of the mechanistic tumor transport model. Furthermore, the simpler mathematical structure of the ANN allows for the more efficient construction of tight bounds of the ANN model relative to the PDE-governed tumor transport model. Furthermore, the employed spatial branch and bound optimizer, EAGO.jl [23, 24], has specialized tight

bounding routines for ANN models. The development of the employed ANN surrogate model is explored in further detail in Subsection 2.4.1.

To reduce the computational complexity of the mechanistic parameter estimation problem, we substitute the mechanistic tumor transport and nonlinear peripheral pressure constraints for the ANN surrogate model and the linearized pressure constraints. The simplified parameter estimation problem can be formulated as:

$$\min_{\boldsymbol{\pi} \in \Pi} \sum_{i=1}^{n} \left(\hat{c}_{\text{avg},i}^{\text{ANN}}(\boldsymbol{\pi}) - \hat{c}_{\text{avg}}^{\text{data}}(t_i) \right)^2 \tag{2.19}$$

$$\text{s.t. } \pi_2 \leq \zeta_{\max} \pi_1$$

$$\pi_2 \geq \zeta_{\min} \pi_1,$$

where ζ_{max} and ζ_{min} are listed in Table 2.3.

2.3 Simultaneous Therapy Design for Dose Selection and Drug Design

In this section, we propose a rigorous and accelerated workflow for simultaneously optimizing treatment and drug design for improved patient outcomes. The results of the parameter estimation problem are used to find correlations that generalize the relationship between the dosage of DEX pretreatment and the values of interstitial hydraulic conductivity and vascular hydraulic conductivity. In combination with these correlations, the tumor transport surrogate model was employed to simultaneously find an optimal dose of DEX pretreatment and nanocarrier size to maximize the accumulation of nanocarriers in the tumor interstitium.

The values of interstitial hydraulic conductivity K and vascular hydraulic conductivity L_p depend on the dose of DEX pretreatment. As such, the results of the parameter estimation problem in (2.19) are used to construct a nonlinear regression model between the DEX dosage and L_p and K. A rational model is established as follows:

$$f_{L_p}^r(x) = \frac{a_0^{L_p} x^2 + a_1^{L_p} x + a_2^{L_p}}{x + a_3^{L_p}}, \tag{2.20}$$

$$f_K^r(x) = \frac{a_0^K x^2 + a_1^K x + a_2^K}{x + a_3^K}, \tag{2.21}$$

where x is the pretreatment DEX dose (mg/kg) and $f_{L_p}^r(x)$ and $f_K^r(x)$ predict the values of L_p and K, respectively, after pretreatment with DEX. The values of the constants $a_i^{L_p}$ and a_i^K, $i \in \{0, \ldots, 3\}$ are reported in Table 2.4.

The experimental data from Martin et al. [3] is limited to three pretreatment DEX doses; however, we seek to establish a methodology that is generalizable to further experimental studies. Subsequently, we incorporate fictitious data to show the robustness of our approach to more complex dose-dependent treatment relationships. Data points are added at 10 mg/kg, 15 mg/kg, 20 mg/kg, and 25 mg/kg. The cumulative data is used to generate polynomial regression models given by the following equations:

$$f_{L_p}^p(x) = b_0^{L_p} x^5 + b_1^{L_p} x^4 + b_2^{L_p} x^3 + b_3^{L_p} x^2 + b_4^{L_p} x + b_5^{L_p}, \qquad (2.22)$$

$$f_K^p(x) = b_0^K x^5 + b_1^K x^4 + b_2^K x^3 + b_3^K x^2 + b_4^K x + b_5^K, \qquad (2.23)$$

here x is the pretreatment DEX dose (mg/kg) and $f_{L_p}^r(x)$ and $f_K^r(x)$ predict the values of L_p and K, respectively, after pretreatment with DEX. The values of the constants $b_i^{L_p}$ and b_i^K, $i \in \{0, \ldots, 5\}$ are reported in Table 2.5. A plot of these regression models are given in Figure 2-4.
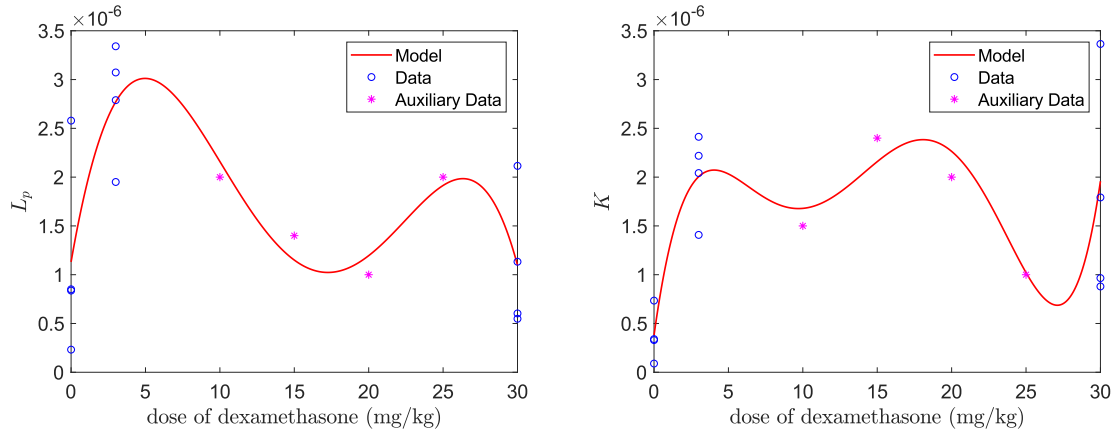

Figure 2-4: The polynomial regression models are plotted here to show the relationship between the DEX dose and the prediction of L_p and K as given by (2.22) and (2.23).

Table 2.4: The constants used in the rational regression model given by (2.20) and (2.21) are provided in this table.

i	0	1	2	3
$a_i^{L_p}$	-7.519×10^{-8}	3.355×10^{-6}	6.944×10^{-7}	0.6175
a_i^K	-2.458×10^{-8}	2.524×10^{-6}	2.916×10^{-7}	0.7816

Table 2.5: The constants used in the polynomial regression model given by (2.22) and (2.23) are provided in this table.

i	0	1	2	3	4	5
$b_i^{L_p}$	-6.2×10^{-13}	-6.0×10^{-11}	5.6×10^{-9}	-1.3×10^{-7}	8.8×10^{-7}	1.1×10^{-6}
b_i^K	1.1×10^{-11}	-8.4×10^{-10}	2.2×10^{-8}	-2.4×10^{-7}	1.1×10^{-6}	3.8×10^{-7}

In the simultaneous therapy and drug design, the nanocarrier size d_m is optimized to maximize the delivery of nanocarriers to the tumor. To perform this adjustment, we must account for the dependencies of physiological parameters on nanocarrier size. Specifically, we determine correlations for the diffusion coefficient D and the half-life circulation time k_d as a function of d_m. Based on data reported in Chauhan et al. [10], we develop a power law regression model for D and a Gaussian model for k_d as a function of d_m as follows:

$$f_D(d_m) = 1.981 \times 10^{-6} \cdot d_m^{-1.157} + 2.221 \times 10^{-8}, \tag{2.24}$$

$$f_{k_d}(d_m) = 1081 \exp\left(-\left(\frac{d_m + 16.63}{84.82}\right)^2\right) + 517.4 \exp\left(-\left(\frac{d_m - 65.61}{996.6}\right)^2\right), \tag{2.25}$$

where f_D and f_{k_d} predict the values of D and k_d, respectively, as a function of d_m.

We propose a therapy design strategy to optimize the dose of DEX and the size of the nanocarrier to maximize nanocarrier accumulation in the tumor interstitium. It is necessary to constrain the concentration at the tumor boundary in order to provide a strict guarantee of patient safety and therapy performance. This will prevent an ineffective low concentration and a toxic high concentration. The simultaneous

therapy design problem using the ANN surrogate model is formulated as follows:

$$\max_{x \in X, d_m \in Z} \hat{c}_{\text{avg}}^{\text{ANN}} \left(\left(f_{L_p}^j(x), f_K^j(x) \right), d_m \right) \tag{2.26}$$

$$\text{s.t. } \hat{c}_{peri}^{\text{ANN}} \left(\left(f_{L_p}^j(x), f_K^j(x) \right), d_m \right) \leq \lambda_1$$

$$\hat{c}_{peri}^{\text{ANN}} \left(\left(f_{L_p}^j(x), f_K^j(x) \right), d_m \right) \geq \lambda_2.$$

Here, $\hat{c}_{\text{avg}}^{\text{ANN}}$ is the tumor transport surrogate model described in Subsection 2.4.2, t_f is the final time of 5 minutes, $\hat{c}_{peri}$ is the dimensionless nanocarrier conentration at the periphery of the tumor, and $j \in \{r, p\}$ represents the usage of the polynomial and rational correlations. The safety constraint is given by $\lambda_1 = 4.5$ (g/mL), which is double the periphery conentration for the 3 mg/kg DEX treatment case. The performance constraint is given by $\lambda_2 = 3.6$ (g/mL), which is the periphery nanocarrier concentration in the control case without DEX treatment. These constraints are chosen to demonstrate that the therapy design problem can be bounded by performance and safety guidelines. As such, these values should are not physiologically significant.

2.4 Machine Learning Surrogate Model

In this work, machine learning surrogate models are proposed to reduce the computational complexity of the mechanistic tumor transport model. Specifically, ANNs are employed for their ability to capture nonlinear relationships with high accuracy. To solve the parameter estimation problem, an intermediate complexity model is established to represent the trajectory of the spatially averaged integration of the tumor transport model. To solve the simultaneous therapy design problem, a low complexity model is developed to represent the mapping between the physiological parameters and the average and peripheral accumulation of nanocarriers in the tumor.

2.4.1 Machine Learning Model for Parameter Estimation

Machine learning surrogates are used to represent the mapping between interstitial and vascular hydraulic conductivity and the spatially averaged accumulation profile at a reduced computational complexity. Using a multidimensional regression approach, the model learns the entire trajectory of the spatially averaged PDE over a time horizon of 5 minutes. In this approach, the computational burden is offloaded to the generation of data and model training instead of the optimization problem itself.

In this study, both 70 kDa and 500 kDa nanocarriers are studied, as the different nanocarrier sizes lead to different transport properties. Furthermore, the administration of dexamethasone (DEX) normalizes the TME and drastically alters the underlying transport phenomena. To account for these variations and maximize accuracy, four distinct ANN surrogates are constructed for each treatment case as follows: 70 kDa nanocarrier control case, 70 kDa nanocarrier 3 mg/kg and 30 mg/kg DEX treatment cases, 500 kDa nanocarrier control case, and 500 kDa nanocarrier 3 mg/kg and 30 mg/kg DEX treatment cases.

We constructed a dataset by varying the values of interstitial and vascular hydraulic conductivity and solving the tumor transport model at each combination of these parameters. For each case, a uniform distribution on L_p and K is defined by setting a lower and upper bound as described in Table 2.6. These bounds are set to restrict physiological parameters to physically meaningful values as expected from prior studies [3]. A Sobol [29] sequence sampling procedure is used to randomly select a million samples from each distribution of L_p and K. For each combination of the physiological parameters, the discretized fluid and solute transport models were solved using the method of lines via the stiff QNDF solver in DifferentialEquations.jl [30] over a time horizon of 5 minutes. The results of this simulation are then averaged over the spatial domain and concatenated into a 21-node output vector. As the inputs are uniformly distributed, min-max normalization is employed to scale the dataset. Furthermore, the dataset is randomly divided into training (70 %) and test (30 %) sets.

Table 2.6: The bounds on the input variables, L_p and K, used for the construction of the surrogate model are listed in this table.

Bounds	Control		Treatment	
Variable	Lower bound	Upper bound	Lower bound	Upper bound
L_p (cm/mm Hg-sec)	1.00×10^{-7}	1.75×10^{-6}	5.00×10^{-7}	3.50×10^{-6}
K (cm^2/mm Hg-sec)	1.00×10^{-7}	1.00×10^{-6}	7.00×10^{-7}	4.00×10^{-6}

In this study, a feed-forward ANN is constructed using Flux.jl[31, 32]. The following model architectures were considered: 2-4 hidden layers, 16-32 nodes per hidden layer, and different activation functions (sigmoid, tanh, gelu, and swish). Through early tuning and comparison, a two hidden layer model with 24 neurons per layer with the swish activation function was selected. This architecture is depicted in Figure 2-5.
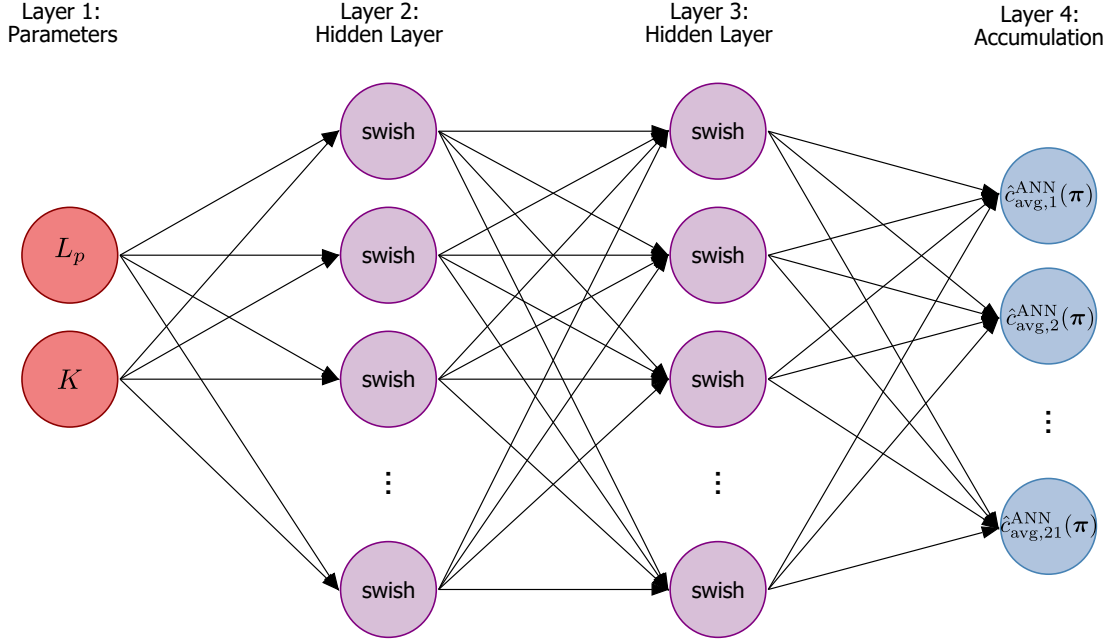
Figure 2-5: As depicted in this illustration, the surrogate model used for the simplified parameter estimation is a fully connected feed-forward multiplayer perceptron (MLP). The first layer consists of two nodes that input the physiological parameters vascular hydraulic conductivity L_p and interstitial hydraulic conductivity K. Layers 2 and 3 are fully connected hidden layers of 24 neurons per layer with the swish activation function in each layer. The final layer of the MLP represents the output of this model and consists of 21 nodes. Specifically, the output represents the temporally discretized accumulation profile over a time horizon of 5 minutes.

For this model, the mean squared error (MSE) was chosen as the loss function for this model. MSE is an appropriate choice for regression models as used in this study; it is given by the following formula:

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \tag{2.27}$$

where N is the number of samples, y_i is the true output, and $\hat{y}_i$ is the predicted output. The model was trained in Flux.jl[31, 32] using the ADAM optimizer with a mini-batch size of 10% of the training data set and a constant learning rate of 10^{-3}. The model was trained for 50 epochs using an early stopping MSE threshold of 10^{-7}. Model assessment was performed by calculating the MSE and the mean percent error on the test set. This training protocol was found to be effective as indicated by the accuracy and performance metrics as listed in Table 2.7.

Table 2.7: The performance and accuracy benchmark metrics for development time and performance of artificial neural network surrogate models of difference cases (70 kDa - control; 70 kDa - treatment; 500 kDa - control; 500 kDa - treatment) are tabulated. The data generation and training time are tracked using BenchmarkTools.jl [33]. The mean squared error (MSE) is calculated by evaluating (2.27) on the test set. Similarly, the mean percent error (MPE) is evaluated using the true and predicted outputs on the test set.

| | Time Metrics | | Performance Metrics | |
Case	Data Generation (s)	Training (s)	MSE	MPE (%)
70 kDa - Control	166	473	5.49×10^{-7}	0.339
70 kDa - Treatment	166	538	2.32×10^{-7}	0.102
500 kDa - Control	166	474	3.23×10^{-7}	0.467
500 kDa - Treatment	166	539	1.55×10^{-7}	0.096

2.4.2 Machine Learning Model for Simultaneous Therapy Design

In this section, we present the two machine learning regression models used to solve the simultaneous therapy design problem. Specifically, these models capture the mapping between interstitial hydraulic conductivity K, vascular hydraulic conductivity L_p, and the nanocarrier size d_m and the concentration of nanocarriers in the tumor. The average accumulation model maps physiological parameters to the spatially averaged nanocarrier accumulation inside the tumor interstitium $\hat{c}_{avg}^{ANN}$. The peripheral concentration model maps the physiological parameters to the nanocarrier concentration at the periphery of the tumor $\hat{c}_{peri}^{ANN}$ after 5 minutes.

These models are developed following an equivalent procedure to the one presented in Section 2.4.1. Specifically, one million samples of L_p, K, d_m were randomly sampled using a Sobol sequence [29] from a uniform distribution as described in Table 2.8. At each sample, the tumor transport model was solved to record the spatially averaged accumulation after 5 minutes or the concentration of nanocarriers at the tumor periphery after a time horizon of 5 minutes. This data was scaled with min-max normalization and divided randomly into training (70%) and test (30%) sets.

Table 2.8: The bounds of the input variables (L_p, K, d_m) used for the construction of the surrogate model are listed in this table.

Variable	Lower bound	Upper bound
L_p (cm/mm Hg-sec)	5.00×10^{-7}	5×10^{-6}
K (cm^2/mm Hg-sec)	5.00×10^{-7}	5×10^{-6}
$d_m(nm)$	10	60

The accumulation and periphery concentration models were developed using an equivalent model tuning, training, and testing process to that described in Subsection 2.4.1. An ANN model architecture consisting of two hidden layers with 12 neurons per layer and the tanh activation function was selected. The models were trained using Flux.jl [31, 32] with the ADAM optimizer [34] and a learning rate of 10^{-4}. After training, the MPE and MSE values were calculated on the test set and recorded in Table 2.9.

Table 2.9: The benchmark metrics are presented here for the ANN surrogate models of average accumulation and peripheral concentration. The data generation and training time are quantified using BenchmarkTools.jl [33]. The mean squared error (MSE) and the mean percent error (MPE) values are evaluated on the test set.

	Time Metrics		Performance Metrics	
Model	Data Generation (s)	Training (s)	MSE	MPE (%)
$\hat{c}_{avg}^{ANN}$	122	180	7.49×10^{-7}	0.172
$\hat{c}_{peri}^{ANN}$	122	66	5.22×10^{-7}	0.276

2.5 Settings for Machine Learning and Optimization

In this section, we describe the settings used for machine learning and optimization in this study. The numerical problems were solved on a high performance computer with an Intel Xeon W-2195 18-core/36-thread CPU operating at 2.3 GHz/4.30 GHz (base/turbo) and 64GB RAM running on Windows 10 Pro version 21H1. For constructing the machine learning models, data was generated using the mechanistic tumor transport model with a spatial discretization of 100 nodes and a temporal discretization over 5 minutes with 21 nodes. These simulations were performed in parallel using 36 CPU threads. The full list of physiological parameters used in these

simulations are provided in Table 2.10 below. The parameter estimation and simultaneous therapy design problems are solved to global optimality using the EAGO v0.6.1 solver [23] via JuMP v0.21.4 [35] in the Julia programming language [36]. For the parameter estimation problems, an absolute global convergence tolerance of 10^{-6} was used and a relative convergence tolerance of 10^{-1} was used. For the simultaneous therapy design problem, an absolute global convergence tolerance of 10^{-6} was used and a relative convergence tolerance of 10^{-2} was used. All machine learning experiments are conducted using Flux.jl [31, 32] on the CPU.

Table 2.10: The values of the biological parameters used in the tumor transport model are listed here. Separate values are given for parameters that vary between the 70kDa and 100kDa nanocarriers used in the *in vivo* experiments.

Parameter	Definition	Value	Reference
S/V (cm^{-1})	Vascular density	200	[37]
D (cm^2/sec)	Diffusion coefficient	2×10^{-7} (70 kDa)	[38]
		1.375×10^{-7} (500 kDa)	[38]
p_v (mm Hg)	Vascular pressure	25	[39]
k_d (min)	Blood circulation time	1480 (70 kDa)	[40]
		1278 (500 kDa)	[40]
μ (mm Hg-sec)	Blood viscosity	3×10^{-5}	[9]
L (μm)	Vessel wall thickness	5	[41]
γ	Fraction of pore area	1×10^{-3}	[10]

Results and Discussion

3.1 Results for Parameter Estimation

In this section, the results of the parameter estimation described by (2.13) and (2.19) are presented. Building on the work presented in Martin et al. [3], the parameter estimation problems are solved to global optimality to validate the tumor transport model against experimental data. Deterministic global optimization is required to provide rigorous guarantees of optimality for parameter estimation, improving confidence in the use of models for decision-making in clinical settings.

In this work, surrogate models were constructed to accelerate the speed of performing model validation using global optimization. The results of the parameter estimation problem using both the mechanistic tumor transport model in Wang et al. [4] and the machine learning surrogates are reported in Table 3.11. A further analysis is conducted in Section 3.3 to discuss the accuracy of the ANN solutions and the difference in time costs between the mechanistic and surrogate models.

Table 3.11: The global optima for the parameter estimation problems described in (2.13) and (2.19) using the mechanistic and ANN surrogate models are presented in these tables for the 70kDa and 200 kDa nanocarriers.

70 kDa Dextran Molecular Weight			
Dose	Control	3 mg/kg	30 mg/kg
P_{eff} (cm/sec)	9.60×10^{-7}	4.61×10^{-6}	2.80×10^{-6}
L_p^* - mechanistic model	8.51×10^{-7}	2.80×10^{-6}	1.12×10^{-6}
L_p^* - ANN model	8.39×10^{-7}	2.77×10^{-6}	1.13×10^{-6}
K^* - mechanistic model	3.35×10^{-7}	2.03×10^{-6}	1.80×10^{-6}
K^* - ANN model	3.32×10^{-7}	2.04×10^{-6}	1.78×10^{-6}

500 kDa Dextran Molecular Weight			
Dose	Control	3 mg/kg	30 mg/kg
P_{eff} (cm/sec)	8.18×10^{-7}	4.30×10^{-6}	1.62×10^{-6}
L_p^* - mechanistic model	8.62×10^{-7}	2.22×10^{-6}	7.50×10^{-7}
L_p^* - ANN model	8.43×10^{-7}	2.22×10^{-6}	7.58×10^{-7}
K^* - mechanistic model	3.34×10^{-7}	2.34×10^{-6}	1.21×10^{-6}
K^* - ANN model	3.29×10^{-7}	2.36×10^{-6}	1.20×10^{-6}

For a given DEX treatment and nanocarrier, interstitial and vascular hydraulic conductivity are found by validating the surrogate model against *in vivo* data. The resulting interstitial and vascular hydraulic conductivity values are a critical starting point for further analysis on the efficacy of nanocarrier transport. Specifically, interstitial and vascular hydraulic conductivity are used to calculate pore size, interstitial fluid pressure, convective transvascular flux, diffusive transvascular flux, etc. Using the results reported in Table 3.11, Wang et al. found the 3 mg/kg DEX treatment maximized nanocarrier accumulation more effectively compared to other treatments [4]. Interested readers are advised to consult Wang et al. for a full report on the physiological effects of the investigated DEX treatments.

3.2   Results for Simultaneous Therapy Design

In this section, we present the results for the simultaneous therapy design problem presented in (2.26) as well as the simultaneous design performed in Wang et al. [4]. DEX was studied for its ability to normalize the tumor vasculature and extracellular matrix. In terms of the tumor transport model, these effects are realized by changing the values of vascular and interstitial hydraulic conductivity. Furthermore, previous studies have shown that the vascular permeability of nanocarriers is size dependent [42, 10, 43]. For these reasons, the dosage of DEX and the nanocarrier size are two important decision variables when formulating a treatment strategy. Subsequently, it is important to consider the interdependent effects of varying the dose of DEX and the nanocarrier size simultaneously. Using the surrogate modeling approach, reduced complexity models are used to simultaneously optimize both the DEX dosage and nanocarrier size without a significant computational cost.

Two cases were considered corresponding to the two sets of regression models for

vascular and interstitial hydraulic conductivity with and without fictitious data. Design case 1 represents the solution to the simultaneous therapy design with regression models described in (2.20) and (2.21). Similarly, design case 2 represents the solution to the simultaneous therapy design with regression models described in (2.22) and (2.23) based on the use of fictitious data. In Wang et al. [4], the equivalent study using the mechanistic model would not converge in a reasonable time limit. For this reason, the simultaneous therapy design problem using the mechanistic model was solved with a multi-start local optimization procedure. The optimal solutions for each case with the ANN surrogate and mechanistic model are reported in Table 3.12.

Table 3.12: Optimal solutions for the simultaneous therapy design problem are reported in this table. Here x (mg/kg) represents the dosage of dexamethasone and d_m (nm) is the nanocarrier diameter.

Case Study	Case 1	Case 2
x^* - ANN model	5.32	4.38
x^* - mechanistic model	5.33	4.36
d_m^* - ANN model	16.40	12.41
d_m^* - mechanistic model	16.54	12.58

The results of the simultaneous therapy design study demonstrate the success of the proposed workflow in this study. Non-trivial values of DEX dosage and nanocarrier size were found to maximize nanocarrier accumulation in the tumor interstitium. The difference in results between case 1 and 2 reveals the importance of collecting more data on additional DEX treatment strategies to create better regression models. Specifically, more representative regression models between the dose of DEX and L_p and K will enable the discovery of truly optimal values of the dose of DEX and the nanocarrier size. Furthermore, the results of this study indicate the size dependence of nanocarrier transport in tumors. Future researchers and clinicians should incorporate optimal nanocarrier size recommendations to design more effective cancer treatment regimens. A complete analysis on the physiological implications of this optimization study is presented in Wang et al. [4].

3.3 Analysis of Accuracy and Computational Cost

In this section, we discuss the impact of the surrogate modeling approach on the accuracy and computational cost of performing model validation and simultaneous therapy design using global optimization. The viability of the surrogate modeling strategy is dependent on the accuracy of these models. From Tables 2.7 and 2.9, the mean percent error of every model falls under 0.5%; this result indicates the ANN surrogate models are predicting the same results as the mechanistic model with a high degree of accuracy. This level of accuracy provides a justification for the use of ANN surrogate models in the optimization-based numerical experiments. However, it is important to verify the results of the optimization studies using the ANN models against the results of the equivalent studies with the mechanistic model. For the parameter estimation and simultaneous therapy design problems, the percent error between the ANN solutions and the mechanistic model solutions have been quantified in Figure 3-6 based on the results in Tables 3.11 and 3.12. In these studies, the percent error between the different results is below 2.5%, thus certifying the accuracy of the ANN models. It is likely that a correlation exists between the accuracy of the ANN model against the mechanistic model and the accuracy of the optimal solutions. Further studies should investigate this phenomenon and attempt to further reduce the error of machine learning surrogates through a larger training dataset, a more meticulous hyperparameter search, or a different learning algorithm.
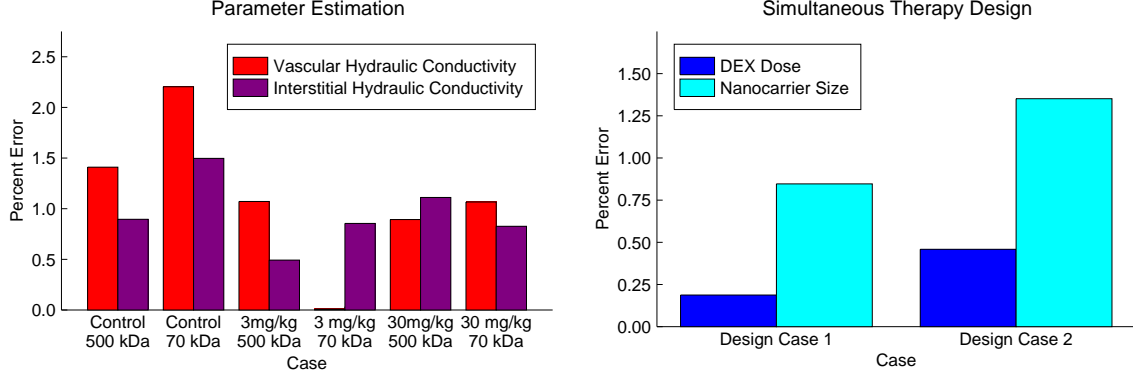
Figure 3-6: The percent error of the results of the parameter estimation studies and the simultaneous therapy design studies are plotted here. For the parameter estimation problem, the percent error is calculated for the optimal value of interstitial hydraulic conductivity and vascular hydraulic conductivity in each study. The largest error is just above 2% in the control case for the 70 kDa. The higher error in the control case studies is likely a result of the higher error of the ANN models relative to the DEX treatment models. Furthermore, many of the DEX treatment studies have between 0% and 1.5% accuracy due to the higher accuracy of the corresponding ANN models. For the simultaneous therapy design problem, the percent error is calculated for the optimal DEX dose and the optimal nanocarrier size. The percent error for these problems all fall under 1.5% indicating a high level of confidence in these results. However, it is important to note that the problem with the mechanistic model was solved using local optimization. Thus, the compared values are local optima that might not be the true global optima.

The time costs of solving the parameter estimation problems are reported in Table 3.13. To solve these problems to global optimality, the solution procedure with the mechanistic model is extremely computationally expensive. Despite the use of high performance computing, the problems required days to solve; this is a major limitation for the applicability of this work. In contrast, ANN models are quickly solved in seconds to minutes. This acceleration of the optimization demonstrates a major advantage of the surrogate modeling approach. Furthermore, the data generation and model development can be done offline prior to optimization studies. Even taking into account these offline procedures, the ANN surrogate models typically perform 2-3 orders of magnitude faster than the equivalent procedure using the mechanistic model as shown in Figure 3-7. However, one exception to this is the control cases where the mechanistic model is faster than the ANN model by an order of magnitude. This unexpected result is likely from a weaker lower bound of the ANN model leading to slower convergence than the custom bounding procedure used to solve the mech-

anistic problem. Despite this exception, global optimization using the ANN models significantly outperforms optimization with the mechanistic model in the worst case. This speed-up enables a scalable procedure that will be more accessible to future researchers.

Furthermore, the simultaneous design problem in Wang et al. [4] using the mechanistic model was unable to be solved using global optimization. Due to the nature of the multidimensional design problem, the custom bounding procedure used for global optimization exhibited weak convergence. This slow convergence is likely due to the curse of dimensionality, which amplifies the computational complexity of the branch-and-bound algorithm. On the other hand, the ANN surrogate models have a simpler mathematical structure relative to the PDE based mechanistic model. The EAGO [23, 24] branch-and-bound solver has native support for MLP models, such as the ANN models in this study. This enables the construction of tight bounds that allow for a rapid solution.
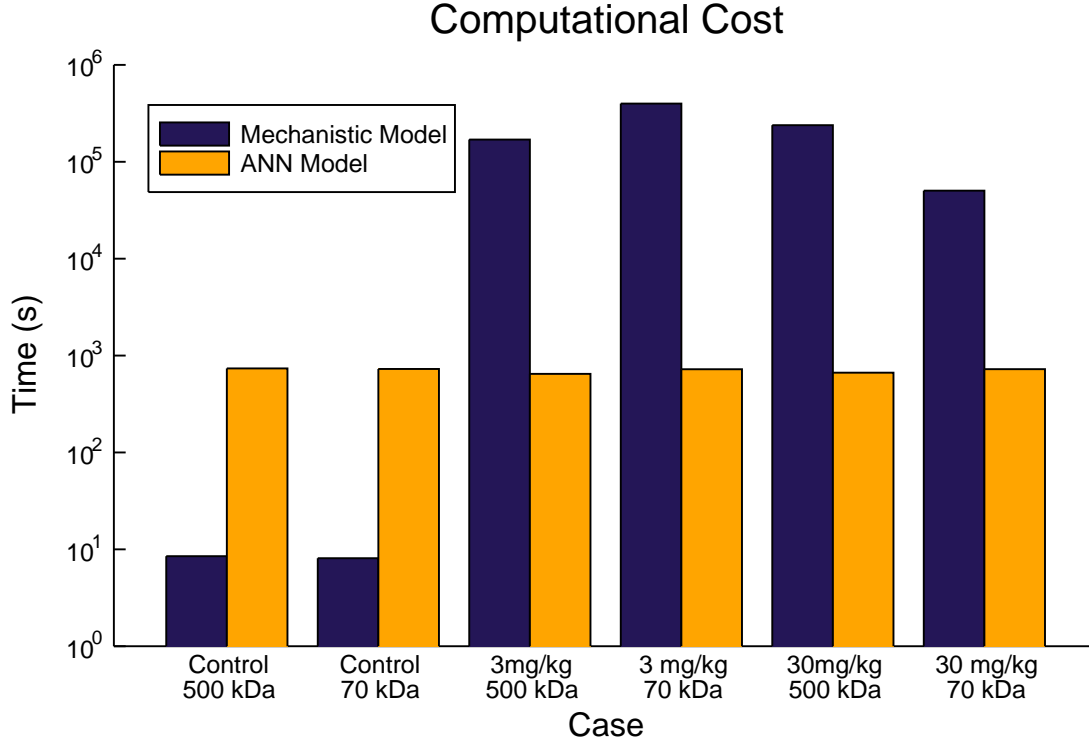
Figure 3-7: In this figure, the total computational cost of model development and solving the parameter estimation problem is plotted here on a logarithmic scale for both the mechanistic model and the ANN surrogate models. The ANN models require an offline time cost for data generation and model training that takes approximately 700 seconds. Global optimization with the ANN models is generally performed in less than a minute. As such, the ANN models have a relatively constant computational cost for the parameter estimation problems. In contrast, the cost of using the mechanistic model is primarily determined by the complexity of the parameter estimation problem. For the control case, strong lower bounds yield a quick solution for the mechanistic model. However, for the DEX treatment cases, the optimization problem takes over a day to converge. Thus in the worst case, the ANN models yield significantly accelerated solutions relative to the solutions of the mechanistic model.

Table 3.13: The benchmarks of the computational time cost for solving the parameter estimation problems using the mechanistic model and the ANN model are tabulated here for the 70kDa and 500kDa dextran bound nanocarrier. For the control case, the mechanistic model slightly outperforms the ANN models. In cases of DEX treatment, global optimization with the ANN model requires significantly less time than the mechanistic model. These time costs do not account for the computational cost of developing the ANN surrogate models.

70 kDa Dextran Molecular Weight			
Dose	Control	3 mg/kg	30 mg/kg
Mechanistic model (s)	8.5	169558.3	238732.2
ANN model (s)	97.9	7.3	25.8

500 kDa Dextran Molecular Weight			
Dose	Control	3 mg/kg	30 mg/kg
Mechanistic model (s)	8.1	398792.2	50368.1
ANN model (s)	23.2	17.6	18.8

Future Work and Conclusions

The intersection of machine learning, systems engineering, and cancer biology presents an exciting avenue of future research. Future work in this field should incorporate state-of-the-art tools from machine learning to explore more complex modeling approaches and robust design problems.

An immediate extension of the machine learning modeling approach is to employ a recurrent neural network (RNN) model. RNNs are a type of ANN designed to effectively handle sequential data by utilizing a hidden state. In Funashi and Nakamura [44], the authors proved that RNNs have the capacity to approximate dynamical systems over a continuous, finite time horizon. The model in Section 2.4.1 could be reformulated as an RNN that represents the integrator of the PDE. This formulation of the model would allow you to build an accumulation profile as a sequence of concentration values. This model could be advantageous, as it would likely improve the accuracy and enable the use of more features. However, the recurrence relationship within the RNN model may lead to weak or exploding bounds when used within a spatial branch and bound solver.

One significant hurdle to multidimensional parameter estimation and optimal treatment design is the computational expense arising from the curse of dimensionality. For the goal of optimizing a large number of physiological parameters, one could develop a novel modeling scheme using dimensionality reduction. Specifically, an autoencoder, an ANN used to learn efficient, latent representations of unlabeled data, could be used to generate a low-dimensional latent variable representation of a high-dimensional set of biological parameters. A second ANN model could be trained to map the latent representation to the state variable of interest. Global optimization can then be performed on the second model to find an optimal value of the low-dimensional latent variable. Finally, the autoencoder can be used to map the optimal latent variable back into the corresponding biological parameters.

The solute transport model used in this work is a simplified 1D spherical PDE model. However, for direct application to patients, a 3D model of fluid and solute

transport would be extremely advantageous. Such a model could be directly integrated with data from a 3D MRI. However, the corresponding 3D PDE model of fluid and solute transport would be extremely computationally complex to embed in global optimization frameworks. Furthermore, ANN surrogate models of the 3D PDE would likely be inaccurate due to the extreme complexity of the physics. A potential solution to these concerns would be to use a physics-aware machine learning model that is trained directly on experimental or clinical data. The physics-aware machine learning algorithm could utilize a physics-informed neural network [45], a neural ordinary differential equation[46], or a universal differential equation model [47].

In this work, we presented the use of machine learning for accelerated, rigorous model validation and cancer treatment design. A first principles approach was utilized to characterize fluid and solute transport within the tumor and discover the effect of TME normalization therapies. A machine learning approach was used to develop an ANN surrogate of the PDE-based tumor transport model. This model enabled the use of fast, deterministic global optimization to validate the model and design an optimal treatment regimen. Specifically, a parameter estimation problem was solved with the ANN surrogates to quantify the values of interstitial and vascular hydraulic conductivity in response to different DEX treatments. Using this information, a simultaneous therapy design problem was then solved with the ANN surrogate to find an optimal DEX dosage and nanocarrier size in order to maximize the accumulation of nanocarriers in the tumor interstitum. These solutions to these problems were found to be highly accurate relative to the equivalent approaches using the mechanistic model. The surrogate models accelerated the global optimization studies by over 2-3 orders of magnitude relative to the mechanistic model. Given the significant gains in computational efficiency, machine learning based modeling will be crucial to the development of future rigorous methods for *in silico* cancer research.

Bibliography

[1] Andrew I. Minchinton and Ian F. Tannock. "Drug penetration in solid tumours". In: *Nat Reviews Cancer* 6.8 (Aug. 2006), pp. 583–592. DOI: 10.1038/nrc1893. URL: https://doi.org/10.1038/nrc1893.

[2] Edmon Begoli, Tanmoy Bhattacharya, and Dimitri Kusnezov. "The need for uncertainty quantification in machine-assisted medical decision making". In: *Nature Machine Intelligence* 1.1 (2019), pp. 20–23. ISSN: 2522-5839. DOI: 10.1038/s42256-018-0004-1.

[3] John D. Martin et al. "Dexamethasone Increases Cisplatin-Loaded Nanocarrier Delivery and Efficacy in Metastatic Breast Cancer by Normalizing the Tumor Microenvironment". In: *ACS Nano* 13.6 (2019), pp. 6396–6408. DOI: 10.1021/acsnano.8b07865.

[4] C. Wang et al. "Optimal Therapy Design With Tumor Microenvironment Normalization". In: *AIChE Journal* (2022).

[5] Laurence T Baxter and Rakesh K Jain. "Transport of fluid and macromolecules in tumors. I. Role of interstitial pressure and convection". In: *Microvascular research* 37.1 (1989), pp. 77–104.

[6] Laurence T Baxter and Rakesh K Jain. "Transport of fluid and macromolecules in tumors. II. Role of heterogeneous perfusion and lymphatics". In: *Microvascular research* 40.2 (1990), pp. 246–263.

[7] Laurence T Baxter and Rakesh K Jain. "Transport of fluid and macromolecules in tumors: III. Role of binding and metabolism". In: *Microvascular research* 41.1 (1991), pp. 5–23.

[8] Laurence T Baxter and Rakesh K Jain. "Transport of fluid and macromolecules in tumors. IV. A microscopic model of the perivascular distribution". In: *Microvascular research* 41.2 (1991), pp. 252–272.

[9]   James W Baish, Paolo A Netti, and Rakesh K Jain. "Transmural coupling of fluid flow in microcirculatory network and interstitium in tumors". In: *Microvascular research* 53.2 (1997), pp. 128–141. DOI: `https://doi.org/10.1006/mvre.1996.2005`.

[10]   Vikash P. Chauhan et al. "Normalization of tumour blood vessels improves the delivery of nanomedicines in a size-dependent manner". In: *Nature Nanotechnology* 7.6 (2012), pp. 383–388. DOI: `10.1038/nnano.2012.45`.

[11]   Paul W. Sweeney et al. "Modelling the transport of fluid through heterogeneous, whole tumours in silico". In: *PLOS Computational Biology* 15.6 (2019). Ed. by Scott L Diamond, e1006751. DOI: `10.1371/journal.pcbi.1006751`.

[12]   Klea Panayidou et al. "GetReal in mathematical modelling: a review of studies predicting drug effectiveness in the real world". In: *Research Synthesis Methods* 7.3 (Aug. 2016), pp. 264–277. DOI: `10.1002/jrsm.1202`. URL: `https://doi.org/10.1002/jrsm.1202`.

[13]   William Crown et al. "Constrained Optimization Methods in Health Services Research—An Introduction: Report 1 of the ISPOR Optimization Methods Emerging Good Practices Task Force". In: *Value in Health* 20.3 (Mar. 2017), pp. 310–319. DOI: `10.1016/j.jval.2017.01.013`. URL: `https://doi.org/10.1016/j.jval.2017.01.013`.

[14]   Marc C Steinbach. "On PDE solution in transient optimization of gas networks". In: *Journal of computational and applied mathematics* 203.2 (2007), pp. 345–361.

[15]   Kevin McBride and Kai Sundmacher. "Overview of surrogate modeling in chemical process engineering". In: *Chemie Ingenieur Technik* 91.3 (2019), pp. 228–239.

[16]   Mohsen Pirdashti et al. "Artificial neural networks: applications in chemical engineering". In: *Reviews in Chemical Engineering* 29.4 (2013), pp. 205–239.

[17]   Daniel Svozil, Vladimir Kvasnicka, and Jiri Pospichal. "Introduction to multilayer feed-forward neural networks". In: *Chemometrics and Intelligent Laboratory Systems* 39.1 (1997), pp. 43–62. ISSN: 0169-7439. DOI: `https://doi.org/10.1016/S0169-7439(97)00061-0`. URL: `https://www.sciencedirect.com/science/article/pii/S0169743997000610`.

[18] M. Chambers and C.A. Mount-Campbell. "Process optimization via neural network metamodeling". In: *International Journal of Production Economics* 79.2 (Sept. 2002), pp. 93–100. DOI: 10.1016/s0925-5273(00)00188-2. URL: https://doi.org/10.1016/s0925-5273(00)00188-2.

[19] Cláudio Augusto Oller Nascimento, Reinaldo Giudici, and Roberto Guardani. "Neural network based approach for optimization of industrial chemical processes". In: *Computers & Chemical Engineering* 24.9-10 (Oct. 2000), pp. 2303–2314. DOI: 10.1016/s0098-1354(00)00587-1. URL: https://doi.org/10.1016/s0098-1354(00)00587-1.

[20] Sun Hye Kim and Fani Boukouvala. "Surrogate-based optimization for mixed-integer nonlinear problems". In: *Computers & Chemical Engineering* 140 (2020), p. 106847. ISSN: 0098-1354. DOI: https://doi.org/10.1016/j.compchemeng.2020.106847. URL: https://www.sciencedirect.com/science/article/pii/S0098135419306970.

[21] Artur M. Schweidtmann and Alexander Mitsos. "Deterministic Global Optimization with Artificial Neural Networks Embedded". In: *Journal of Optimization Theory and Applications* 180.3 (Oct. 2018), pp. 925–948. DOI: 10.1007/s10957-018-1396-0. URL: https://doi.org/10.1007/s10957-018-1396-0.

[22] Chenyu Wang, Matthew E. Wilhelm, and Matthew D. Stuber. "Semi-Infinite Optimization with Hybrid Models". In: *Industrial & Engineering Chemistry Research* (Apr. 2022). DOI: 10.1021/acs.iecr.2c00113. URL: https://doi.org/10.1021/acs.iecr.2c00113.

[23] M. Wilhelm and M. D. Stuber. "Easy Advanced Global Optimization (EAGO): An Open-Source Platform for Robust and Global Optimization in Julia". In: *AIChE Annual Meeting 2017*. Minneapolis, MN, Oct 31, 2017.

[24] Matthew E. Wilhelm and Matthew D. Stuber. "EAGO.jl: easy advanced global optimization in Julia". In: *Optimization Methods and Software* (2020), pp. 1–26. DOI: 10.1080/10556788.2020.1786566. URL: https://github.com/PSORLab/EAGO.jl.

[25] WM Deen. "Hindered transport of large molecules in liquid-filled pores". In: *AIChE Journal* 33.9 (1987), pp. 1409–1425.

[26] Peter M Bungay and Howard Brenner. "The motion of a closely-fitting sphere in a fluid-filled tube". In: *International Journal of Multiphase Flow* 1.1 (1973), pp. 25–56.

[27] Helen Lee et al. "64Cu-MM-302 Positron Emission Tomography Quantifies Variability of Enhanced Permeability and Retention of Nanoparticles in Relation to Treatment Response in Patients with Metastatic Breast Cancer". In: *Clinical Cancer Research* 23.15 (Aug. 2017), pp. 4190–4202. ISSN: 1078-0432. DOI: `10.1158/1078-0432.CCR-16-3193`. eprint: `https://aacrjournals.org/clincancerres/article-pdf/23/15/4190/2040086/4190.pdf`. URL: `https://doi.org/10.1158/1078-0432.CCR-16-3193`.

[28] Reiner Horst and Hoang Tuy. *Global Optimization*. Springer Berlin Heidelberg, 1996. DOI: `10.1007/978-3-662-03199-5`.

[29] I.M Sobol'. "On the distribution of points in a cube and the approximate evaluation of integrals". In: *USSR Computational Mathematics and Mathematical Physics* 7.4 (1967), pp. 86–112. DOI: `10.1016/0041-5553(67)90144-9`.

[30] Christopher Rackauckas and Qing Nie. "Differentialequations.jl–a performant and feature-rich ecosystem for solving differential equations in julia". In: *Journal of Open Research Software* 5.1 (2017).

[31] Michael Innes et al. "Fashionable Modelling with Flux". In: *CoRR* abs/1811.01457 (2018). arXiv: `1811.01457`. URL: `https://arxiv.org/abs/1811.01457`.

[32] Mike Innes. "Flux: Elegant Machine Learning with Julia". In: *Journal of Open Source Software* (2018). DOI: `10.21105/joss.00602`.

[33] Jiahao Chen and Jarrett Revels. "Robust benchmarking in noisy environments". In: *arXiv e-prints*, arXiv:1608.04295 (Aug. 2016). arXiv: `1608.04295 [cs.PF]`.

[34] Diederik P Kingma and Jimmy Ba. "Adam: A method for stochastic optimization". In: *arXiv preprint arXiv:1412.6980* (2014).

[35] Iain Dunning, Joey Huchette, and Miles Lubin. "JuMP: A Modeling Language for Mathematical Optimization". In: *SIAM Review* 59.2 (2017), pp. 295–320. DOI: `10.1137/15m1020575`.

[36] Jeff Bezanson et al. "Julia: A Fresh Approach to Numerical Computing". In: *SIAM Review* 59.1 (2017), pp. 65–98. DOI: `10.1137/141000671`.

[37] Fan Yuan et al. "Mirovascular permeability and interstitial penetration of sterically stabilized (stealth) liposomes in a human tumor xenograft". In: *Cancer research* 54.13 (1994), pp. 3352–3356.

[38]   Alain Pluen et al. "Role of tumor–host interactions in interstitial diffusion of macromolecules: cranial vs. subcutaneous tumors". In: *Proceedings of the National Academy of Sciences* 98.8 (2001), pp. 4628–4633. DOI: `https://doi.org/10.1073/pnas.081626898`.

[39]   Yves Boucher and Rakesh K Jain. "Microvascular pressure is the principal driving force for interstitial hypertension in solid tumors: implications for vascular collapse". In: *Cancer research* 52.18 (1992), pp. 5110–5114.

[40]   Zoran Popović et al. "A nanoparticle size series for in vivo fluorescence imaging". In: *Angewandte Chemie* 122.46 (2010), pp. 8831–8834. DOI: `10.1002/ange.201003142`.

[41]   Rakesh K Jain. "Transport of molecules across tumor vasculature". In: *Cancer and Metastasis Reviews* 6.4 (1987), pp. 559–593.

[42]   H. Cabral et al. "Accumulation of sub-100 nm polymeric micelles in poorly permeable tumours depends on size". In: *Nature Nanotechnology* 6.12 (2011), pp. 815–823. DOI: `10.1038/nnano.2011.166`.

[43]   Fan Yuan et al. "Vascular permeability in a human tumor xenograft: molecular size dependence and cutoff size". In: *Cancer Research* 55.17 (1995), pp. 3752–3756.

[44]   Ken-ichi Funahashi and Yuichi Nakamura. "Approximation of dynamical systems by continuous time recurrent neural networks". In: *Neural Networks* 6.6 (1993), pp. 801–806. ISSN: 0893-6080. DOI: `https://doi.org/10.1016/S0893-6080(05)80125-X`. URL: `https://www.sciencedirect.com/science/article/pii/S089360800580125X`.

[45]   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". In: *Journal of Computational Physics* 378 (Feb. 2019), pp. 686–707. DOI: `10.1016/j.jcp.2018.10.045`. URL: `https://doi.org/10.1016/j.jcp.2018.10.045`.

[46]   Ricky T. Q. Chen et al. "Neural Ordinary Differential Equations". In: *Advances in Neural Information Processing Systems*. Ed. by S. Bengio et al. Vol. 31. Curran Associates, Inc., 2018. URL: `https://proceedings.neurips.cc/paper/2018/file/69386f6bb1dfed68692a24c8686939b9-Paper.pdf`.

[47]   Christopher Rackauckas et al. "Universal differential equations for scientific machine learning". In: *arXiv preprint arXiv:2001.04385* (2020).