Model Comparison and Uncertainty Quantification in Tumor Growth

E. A. PAIXÃO[1], G. T. NAOZUKA[2*], J. V. O. SILVA[3],
M. P. C. MENEZES[4] and R. C. ALMEIDA[5]

ABSTRACT. Mathematical and computational modeling have been increasingly applied in many areas of cancer research, aiming to improve the understanding of tumorigenic mechanisms and to suggest more effective therapy protocols. The mathematical description of the tumor growth dynamics is often made using the exponential, logistic, and Gompertz models. However, recent literature has suggested that the Allee effect may play an important role in the early stages of tumor dynamics, including cancer relapse and metastasis. For a model to provide reliable predictions, it is necessary to have a rigorous evaluation of the uncertainty inherent in the modeling process. In this work, our main objective is to show how a model framework that integrates sensitivity analysis, model calibration, and model selection techniques can improve and systematically characterize model and data uncertainties. We investigate five distinct models with different complexities, which encompass the exponential, logistic, Gompertz, and weak and strong Allee effect dynamics. Using tumor growth data published in the literature, we perform a global sensitivity analysis, apply a Bayesian framework for parameter inference, evaluate the associated sensitivity matrices, and use different information criteria for model selection (First- and Second-Order Akaike Information Criteria and Bayesian Information Criterion). We show that such a wider methodology allows having a more detailed picture of each model assumption and uncertainty, calibration reliability, ultimately improving tumor mathematical description. The used *in vivo* data suggested the existence of both a competitive effect among tumor cells and a weak Allee effect in the growth dynamics. The proposed model framework highlights the need for more detailed experimental studies on the influence of the Allee effect on the analyzed cancer scenario.

Keywords: predictive oncology, inverse problem, Allee effect, logistic model, Gompertz model, exponential model.

---

*Corresponding author: Gustavo Taiji Naozuka – E-mail: naozuka@lncc.br

[1]Laboratório Nacional de Computação Científica, Av. Getúlio Vargas, 333, 25651-075, Petrópolis, RJ, Brazil – E-mail: earantes@lncc.br https://orcid.org/0000-0003-1496-5324

[2]Laboratório Nacional de Computação Científica, Av. Getúlio Vargas, 333, 25651-075, Petrópolis, RJ, Brazil – E-mail: naozuka@lncc.br https://orcid.org/0000-0002-7331-9763

[3]Laboratório Nacional de Computação Científica, Av. Getúlio Vargas, 333, 25651-075, Petrópolis, RJ, Brazil – E-mail: jvos@lncc.br https://orcid.org/0000-0002-4158-4888

[4]Laboratório Nacional de Computação Científica, Av. Getúlio Vargas, 333, 25651-075, Petrópolis, RJ, Brazil – E-mail: mpessoa@lncc.br https://orcid.org/0000-0001-7953-802X

[5]Laboratório Nacional de Computação Científica, Av. Getúlio Vargas, 333, 25651-075, Petrópolis, RJ, Brazil – E-mail: rcca@lncc.br https://orcid.org/0000-0002-9756-8131

1   INTRODUCTION

Cancer is a generic term that refers to various diseases characterized mainly by abnormal cell growth, affecting different tissues and organs of the body. It is the second leading cause of death in the world, accounting for approximately 9.6 million deaths in 2018 [16, 34]. Understanding its growth dynamics is a challenge and may contribute to a better knowledge of the mechanisms involved and new developments of more effective therapy protocols. In this context, mathematical and computational modeling has been widely applied in many cancer research areas [4]. This approach allows the assessment of hypotheses and theories, providing quantitative predictions and directions for future research in cancer biology [1]. However, for a model to provide reliable predictions, it is necessary to have a rigorous evaluation of the uncertainty inherent in the modeling process. Uncertainties are present in almost all practical problems, especially in biological processes. They are manifested in various ways including the variability associated with knowledge (or lack of knowledge) about the parameters and the model uncertainty, that is, the uncertainty associated with the representation of reality by the mathematical model. Even when small, uncertainties can have a significant effect on the model output. To account for parameter and model uncertainties, here we put forward a model framework that integrates sensitivity analysis, model calibration, and model selection methods.

The first step in the overall model analysis is the Sensitivity Analysis (SA). It is a well known strategy for identifying how uncertainties in model factors impact the quantities of interest QoIs (model outputs). There is a vast literature accumulated on the diversity of SA methods [30, 31]. Sensitivities are largely understood as derivatives of a specific output with respect to a particular input. Local SA methods evaluate sensitivities of a QoI with respect to the variation of a single input factor, while all the others are kept fixed. Global SA methods, in which all parameters are varied simultaneously, provide sensitivity measures that help the design of more parsimonious models [31]. It complements uncertainty quantification methods by indicating or quantifying how much of the variation of the desired QoI is driven by the variability of each model parameter. It has been used to guide experiments, as well as calibration and modeling processes. Specifically, by identifying critical parameters for the description of the studied phenomenon, SA can drive the experimental area towards obtaining the most appropriate data to inform the model and thus to estimate its parameters more accurately and carefully; more influential parameters may suggest directions for model improvement while those to which the QoI is insensitive may yield model simplification or may be fixed during the calibration process [28]. Overall, SA allows a better understanding of the model, pointing out limitations and capabilities. Of note, local SA methods that evaluate derivatives of the QoI at some specific points of the parametric space should be avoided since the corresponding result may be misleading [30].

The parameter values of a model are obtained by solving an inverse problem, a process that is also called model calibration. Given observational data, the idea is to estimate parameter values so that model outcomes match the available experimental data. Classical methods such as least-squares fitting provide only point estimates, a result that does not detail the influence of model and data uncertainties. In the Bayesian approach, the prior information establishing the current

knowledge about the model parameters is updated to make the model outcome consistent with the available observations, which are also accompanied by uncertainties. Model uncertainties, or lack of credibility in the model, are also included in the analysis in a very natural way, allowing us to infer the posterior distribution of the parameters that best fit the theoretical model to the data. Overall, Bayesian calibration provides not only the maximum *a posteriori* probability (MAP) estimate of each parameter but also its probability distribution, quantifying its uncertainty on the lights of the available data. Parameter uncertainties are then propagated throughout the model, a process that ultimately yields the uncertainty quantification of the model outcome.

The uncertainties should also be considered in the procedure of selecting the best model in a set of candidate models for the available experimental data. Murphy *et al.* [23] pointed out that model selection based only on goodness-of-fit criteria to data is not a guarantee of best model predictions. A recent cell invasion study [33] highlights the need to encompass model complexities and uncertainties as well as data uncertainties, among other issues, in the search for a better model.

One of the first discussions on model plausibility, calibration, and selection in a wider Bayesian framework is provided in [25]. Those issues were later systematically grouped in a single algorithm called OPAL (Occam Plausibility Algorithm) and used for model selection and validation in [22] for predicting glioma growth in murines. In OPAL, Bayesian calibration is preceded by a SA step that aims at eliminating models with parameters that do not significantly affect the selected QoI. The authors showed that such elimination may not be adequate since the model selected at the end of the analysis would have been discarded from the set of candidate models. Indeed, global SA that allows identifying the importance of parameter variability on the model outcome is an important tool for model analysis and can be instrumental in model calibration [26, 30]. A different approach for model selection in dynamical systems was proposed in [32]. Such methodology, named Approximate Bayesian Computation (ABC), does not require explicit evaluation of likelihoods for parameter inference and also allows for model selection by combining it with a sequential Monte Carlo method (ABC-SMC). This approach was successfully used in [12] for comparing tumor growth models with and without chemotherapy using hypothetical tumor cells data. Other applications in system biology can be found in [18, 21]. Although the previously mentioned methodologies have been increasingly applied in many areas, the most used model selection methods for comparing both nested and non-nested models are the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), and their variants. The AIC is based on the Kullback-Leibler divergence as a measure of information content while BIC aims to maximize the posterior model probability. Although designed for different purposes, both are unbiased estimators that have the same goodness-of-fit term but different penalty terms. The higher the number of parameters in a model, the greater the model is penalized. Since the BIC's penalty term also directly depends on the number of measurements, it imposes a greater penalty for large sample settings. In [6], the authors shed some light on the differences in design and objectives between AIC and BIC, among other methods for model selection. They mainly pointed out the appropriate properties of AIC-type criteria and recommend

their use, especially in medicine, biological, and social sciences, for selecting the most parsimonious model for a given experimental data set. More interesting, they showed that both AIC and BIC can be derived using Bayesian and non-Bayesian procedures so that the choice to use one or the other criterion should not be based on the character of the analysis (see [7] for further details).

The development of predictive mathematical cancer models to represent the growth of cancer cell populations has always been challenging. Ordinary differential equations are often used including exponential, logistic, and Gompertz models. Murphy *et al.* [23] pointed out that an appropriate choice of growth model is extremely important both for studies that directly analyze tumor growth and for derivative studies related to treatment evaluation, its efficiency, optimization, and resistance development. Their studies indicated that the model choice impacts predictions and estimates of maximum tumor size, doubling time, and the minimum amount of chemotherapy required for tumor elimination.

In addition to the classical tumor growth models mentioned before, a possible approach to cancer modeling is to consider the tumor and its microenvironment as an ecosystem [2, 19]. In this context, several ecological concepts and theories have contributed to the advancement of the development of tumor growth models, such as the Allee effect. This is a biological phenomenon in which there is a "positive density dependence", that is, a direct correlation between some individual fitness and population density [10, 11]. Some populations exhibit reproductive, eating, spreading, and general survival behaviors that need a minimum population size (denoted by Allee threshold) for them to settle and maintain themselves in a given environment, a mechanism that is usually called strong Allee effect. In the absence of an Allee threshold for population survival, the positive density dependence mechanism is denoted by weak Allee effect. Ecological studies of Allee effect-related phenomena have contributed to a better understanding of population dynamics and, consequently, impacted their conservation and management [10]. In the cancer biology approach, the investigation of Allee effect mechanism on tumor growth dynamics at low population densities shows promising results, especially regarding progression, recurrence, and metastasis, which significantly impact the study and the development of therapies [5, 14, 17, 19, 24]. What remains to be assessed, and is particularly intriguing, is the possibility of the Allee effect impacting cancer cell population at higher density population levels, which can occur in some ecological systems [10].

In this work, our main objective is to show how the integration of suitable modeling tools can improve and systematically characterize model and data uncertainties. We examined a simple scenario of the growth of breast adenocarcinoma tumor cells in nude mice, whose experimental data are available in [23, 35]. The set of candidate models to describe those data is composed of five models with different complexities. Three of these models are widely used in the mathematical description of tumor growth dynamics and two of them can characterize either the strong or the weak Allee effect. In this way, we also assess the presence or absence of the Allee effect in the considered, biologically realistic, scenario.

2   MATERIALS AND METHODS

Figure 1 shows a schematic workflow of the model framework that we put forward in this work. The first step encompasses experimental data acquisition and the definition of a set of mathematical models capable of describing the data. The next steps include sensitivity analysis, model calibration, and model selection. At the end of the process, we identify the evidence towards the model (in the candidate model set) that better supports the data, we quantify the uncertainty on estimated parameter values given the data, and how these uncertainties impact the tumor growth. All those steps are described in details in the following subsections.

2.1   Experimental Data

The experimental data used for calibration of the evaluated models were extracted by [23] from [35]. We used the *WebPlotDigitizer* tool [29] in [23] to obtain the data on breast cancer cell volume growth (GI-101A xenografts) in athymic mice, consisting of $m = 14$ tumor volume points (in mm^3), distributed over a period of 114 days. We define the times when the measurements were taken by $ti, i = 1, \ldots, m$. For completeness, this data set is presented in the Appendix.

2.2   Mathematical Models

The models selected for analysis have different complexities. The first candidate model displays the exponential growth dynamics developed by Malthus (1766-1834) which describes unlimited growth and presents a constant *per capita* population growth rate over time. Next, we consider resource-limited growth by selecting the Verhulst (1804-1849) logistic and the Gompertz (1779-1865) models, whose *per capita* population growth rates are monotonically decreasing. Moreover, due to recent experimental evidence [5, 14, 17, 24], we also consider two candidate models that include the Allee effect for which the *per capita* population growth rate reaches a maximum value at an intermediate population size. The Extended Allee models considered here may have strong or weak Allee effect combined with either exponential or logistic growth laws, respectively denoted by EGAE and LGAE. In the strong Allee effect, the population undergoes a negative growth rate at very low population sizes, and thus there is a threshold below which the population goes to extinction. On the other hand, in the weak Allee effect, the growth rate is low but always positive at small population sizes.

These models are presented in Table 1, which includes the corresponding analytical solutions when available. Otherwise, numerical solutions were obtained by applying the fourth-order Runge-Kutta method [20]. In the models, N denotes tumor volume, measured in mm^3, defined as the QoI used in the model analysis, N_0 represents its initial value, and t indicates time, measured in day. The parameter a (day^{-1}) present in all models represents the (maximum) tumor growth rate. In the logistic and Gompertz models, the parameter b (mm^3) represents the carrying capacity of the environment. In the Extended Allee models, d (mm^3) is a non-negative parameter that indicates how the *per capita* population growth rate varies with population density and c (mm^3)
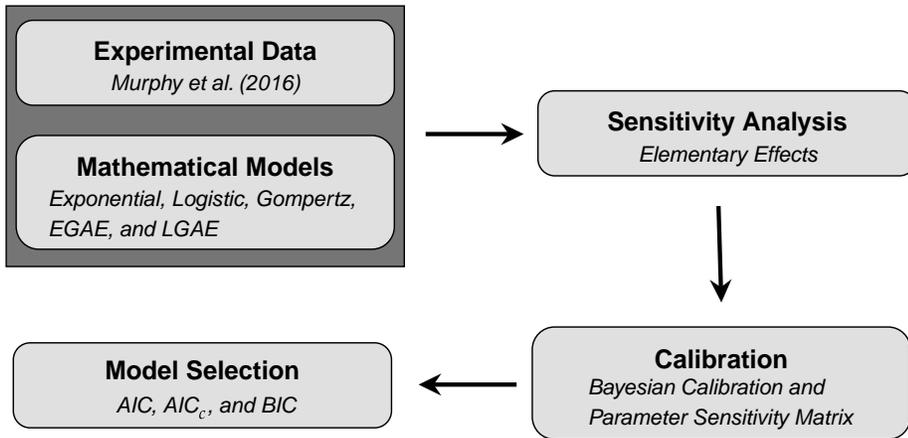
Figure 1: A schematic workflow of this study, which starts with the acquisition of the experimental data set on tumor growth and the choice of the set of candidate mathematical models for describing the data (EGAE and LGAE stand for Exponential Growth with Allee Effect and Logistic Growth with Allee Effect, respectively). Sensitivity analysis is carried out, followed by the calibration of model parameters, and the selection of the best model given the data. In each of these blocks, we indicate the associated information or used methods.

Table 1: Candidate set of mathematical models to describe the selected experimental tumor growth data with corresponding equations and analytical solutions when available. For the Extended Allee models, numerical solutions were obtained by applying the fourth-order Runge-Kutta method [20].

Model	Equation	Analytical Solution
Exponential	$\dfrac{dN}{dt} = aN$	$N(t) = N_0 e^{at}$
Logistic	$\dfrac{dN}{dt} = aN\left(1 - \dfrac{N}{b}\right)$	$N(t) = \dfrac{N_0 e^{at}}{1 - \dfrac{N_0}{b}(1 - e^{at})}$
Gompertz	$\dfrac{dN}{dt} = aN \ln\left(\dfrac{b}{N}\right)$	$N(t) = b \exp\left(\ln\left(\dfrac{N_0}{b}\right)e^{-at}\right)$
EGAE	$\dfrac{dN}{dt} = aN\left(1 - \dfrac{c+d}{N+d}\right)$	–
LGAE	$\dfrac{dN}{dt} = aN\left(1 - \dfrac{N}{b}\right)\left(1 - \dfrac{c+d}{N+d}\right)$	–

refers to the Allee threshold: when $c > 0$, there is a strong Allee effect; when $c \leq 0$ and $d > |c|$, there is a weak Allee effect; and when $d = -c$, there is no Allee effect.

Associated with each model, we define the vector $\boldsymbol{\theta} = \{\theta_1, \ldots, \theta_n\}$, which includes all model parameters and the initial condition N_0. For example, for the LGAE model, this vector is defined as $\boldsymbol{\theta} = \{a, b, c, d, N_0\}$. Thus, vector $\boldsymbol{\theta}$ and the hyperparameter σ, which accounts for both data and model uncertainties, were estimated by Bayesian calibration (see Subsection 2.4).

2.3   Sensitivity Analysis

In our model framework, we perform a global SA to identify how uncertainties in the model parameters influence the QoI using the Elementary Effects (EE) method [31]. The EE method is a simple and informative screening method, which allows ranking the parameters by their order of importance while requiring a relatively small number of model evaluations when compared to variance-based methods. The EE for a parameter θ_i is defined as:

$$EE_i = \frac{N(\theta_1, \ldots, \theta_i + \delta, \ldots, \theta_n) - N(\theta_1, \ldots, \theta_i, \ldots, \theta_n)}{\delta}, \tag{2.1}$$

in which $N(\boldsymbol{\theta})$ denotes the QoI, $\delta \in \left\{ \frac{1}{p-1}, \ldots, 1 - \frac{1}{p-1} \right\}$, and p is the number of discretization levels of a n-dimensional unit hypercube representing the parametric space. Hence, each hypercube direction is associated with a parameter whose range is mapped between 0 and 1. We compute the following sensitivity indices from EE_i, $i = 1, \ldots, n$:

$$\mu_i = \frac{1}{r} \sum_{j=1}^{r} EE_i^j; \qquad \mu_i^* = \frac{1}{r} \sum_{j=1}^{r} \left| EE_i^j \right|; \qquad \sigma_i = \sqrt{\frac{1}{r-1} \sum_{j=1}^{r} \left(EE_i^j - \mu_i \right)^2}, \tag{2.2}$$

in which r is the number of trajectories through the parametric space. The global sensitivity indices μ_i^* and σ_i indicate, respectively, the influence and nonlinear importance of the ith parameter on the QoI. The more influential the parameter, small variations in its value will have a major impact on the QoI estimation, highlighting the importance of accurate calibration. In this paper, we performed the SA at some time points along the experimental time frame to capture the changes of parameter interplay at different moments of the tumor growth dynamics. For all models, we adopted $p = 4$ and $\delta = \frac{2}{3}$, the recommended choices for an appropriate screening of the parametric space [8, 31], and set $r = 20$ that yields convergent results.

2.4   Bayesian Calibration

We apply a Bayesian approach for estimating the model parameters. To present it in a general context, let $\boldsymbol{y}$ be the vector of m experimental data and $\tilde{\boldsymbol{y}}$ the corresponding vector of the values obtained by simulating the model using $\boldsymbol{\theta}$. With these definitions, Bayes' theorem [3] states: given an initial knowledge of the parameters, defined by a prior probability distribution $p(\boldsymbol{\theta})$, and a likelihood function $\mathscr{L}(\boldsymbol{\theta}|\boldsymbol{y})$, the knowledge about the parameters can be improved by evaluating a posterior probability distribution $p(\boldsymbol{\theta}|\boldsymbol{y})$ from:

$$p(\boldsymbol{\theta}|\boldsymbol{y}) \propto \mathscr{L}(\boldsymbol{\theta}|\boldsymbol{y}) \, p(\boldsymbol{\theta}). \tag{2.3}$$

The likelihood function $\mathscr{L}(\boldsymbol{\theta}|\boldsymbol{y})$ is assumed here as Gaussian:

$$\mathscr{L}(\boldsymbol{\theta}|\boldsymbol{y}) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\sum_{i=1}^{m}\left(\frac{y(ti)-\tilde{y}(ti)}{\sigma}\right)^2\right). \tag{2.4}$$

It defines the likelihood of getting the model dynamics represented by the experimental data using model simulations with the vector of parameters $\boldsymbol{\theta}$. Ultimately, the likelihood function updates the prior knowledge about $\boldsymbol{\theta}$ by considering the information on $\boldsymbol{\theta}$ that comes from the empirical data. We also assume that the parameters are independent and uniformly distributed because of limited information about them. Notwithstanding, different definitions for the likelihood function and prior probability distributions may be used. We used Markov Chain Monte Carlo (MCMC) sampling procedure to numerically obtain the posterior distribution. Specifically, parameter inference of each model is performed using the Multilevel Monte Carlo method implemented in the open-source library QUESO (Quantification of Uncertainty for Estimation, Simulation and Optimization) [27], using 40,000 samples. Remark that once the joint posterior distribution $p(\boldsymbol{\theta}|\boldsymbol{y})$ is determined, the posterior distributions of the model parameters are the associated marginals. The MAP estimates of marginal distributions are denoted by $\hat{\boldsymbol{\theta}}$. Model simulations are then obtained propagating (drawing samples from) the joint posterior distribution, determining the tumor dynamics along time, with the corresponding 95% credible interval [15].

2.5   Parameter Sensitivity Matrix

The model's predictions can significantly deteriorate due to uncertainties in the parameter values. To systematically assess the effects of the inference made in the calibration step, we may calculate a local measure of the sensitivity of the QoI around the MAP estimates. Defining $\mathbf{N}(\boldsymbol{\theta}) = [N_{t1}(\boldsymbol{\theta}),\ldots,N_{tm}(\boldsymbol{\theta})]^T$ as the vector of simulated QoI values at the time of the experimental measurements, we want to evaluate the sensitivity of $\mathbf{N}(\boldsymbol{\theta})$ to small variations around $\hat{\boldsymbol{\theta}}$. It is not desirable for the behavior of $\mathbf{N}(\hat{\boldsymbol{\theta}})$ to differ considerably from $\mathbf{N}(\hat{\boldsymbol{\theta}}+\varepsilon\hat{\boldsymbol{\theta}})$, with ε a small constant, which would indicate high uncertainty on QoI around MAP estimates. In these cases, alternatives to reduce parameter uncertainties must be pursued.

A possible procedure to quantify the calibration reliability is by computing the condition number of the Jacobian (or sensitivity) matrix associated with each model at $\hat{\boldsymbol{\theta}}$ [9]. By definition, the coefficients of the $m \times n$-dimensional sensitivity matrix, denoted by $J(\hat{\boldsymbol{\theta}})$, are measures of the sensitivity of the QoI at time $ti, i = 1,\ldots,m$, with respect to variations in the jth parameter, $j = 1,\ldots,n$. The coefficients $\left[J(\hat{\boldsymbol{\theta}})\right]_{ij}$ can be numerically evaluated using centered differences by:

$$\left[J(\hat{\boldsymbol{\theta}})\right]_{ij} = \frac{\partial N_{ti}}{\partial \theta_j}\bigg|_{\hat{\boldsymbol{\theta}}} \approx \frac{N_{ti}(\hat{\boldsymbol{\theta}}+\varepsilon\hat{\theta}_j\mathbf{e}_j)-N_{ti}(\hat{\boldsymbol{\theta}}-\varepsilon\hat{\theta}_j\mathbf{e}_j)}{2\varepsilon\hat{\theta}_j}, \tag{2.5}$$

in which $\mathbf{e}_j$ is the standard unit vector in the jth direction. Notice that a small value of $\left[J(\hat{\boldsymbol{\theta}})\right]_{ij}$ indicates that perturbations in $\hat{\theta}_j$ yield small changes in N_{ti}. In this case, basically the same value for N_{ti} would be obtained for a wide range of values of $\hat{\theta}_j$. It is desirable to have the sensitivity

matrix $J(\hat{\boldsymbol{\theta}})$ with a small condition number, which would mean that the QoI does not vary significantly to small changes in the estimated parameters. In this work, the condition number of $J(\hat{\boldsymbol{\theta}})$ is measured in the 2-norm (the ratio between the maximum and minimum singular values of $J(\hat{\boldsymbol{\theta}})$) and we assumed $\varepsilon = 10^{-5}$ for calculating $\left[J(\hat{\boldsymbol{\theta}})\right]_{ij}$.

2.6 Model Selection

We use the Akaike (AIC), the second-order Akaike (AIC$_c$), and the Bayesian (BIC) information criteria [6] to select the best model that fits the experimental data from a set of candidate models. These criteria have a goodness-of-fit term, that depends on the log-likelihood function at its maximum point, denoted by $\log(\mathcal{L}(\hat{\boldsymbol{\theta}}|\boldsymbol{y}))$. They also have a bias correction term that depends on the number of estimated parameters in the model (n). Both AIC and AIC$_c$ are based on the Kullback-Leibler (K-L) information (or divergence) and they are built to asymptotically select the best model in the candidate set. In other words, the model with the best trade-off between goodness-of-fit and model complexity is considered the most adequate. They are mathematically defined as:

$$\text{AIC} = -2\log(\mathcal{L}(\hat{\boldsymbol{\theta}}|\boldsymbol{y})) + 2n \quad \text{and} \quad \text{AIC}_c = \text{AIC} + \frac{2n(n+1)}{m-n-1}. \tag{2.6}$$

The AIC may perform poorly if there are too many parameters when compared to the data size m. The bias correction term of AIC$_c$ overcomes this difficulty for cases where the ratio m/n is small. Burnham and Anderson [6] suggest the use of AIC$_c$ when this ratio m/n is less than 40.

Based on the assumption that a "true model" belongs to the candidate set and has a small dimension, BIC has the same goodness-of-fit of AIC and a more stringent bias term, so it tends to favor smaller models than AIC. It is determined by:

$$\text{BIC} = -2\log(\mathcal{L}(\hat{\boldsymbol{\theta}}|\boldsymbol{y})) + n \cdot \log(m). \tag{2.7}$$

A quick and useful way to rank the candidate models is using the criterion differences in which the criterion values are rescaled based on the minimum value for each criterion. The differences are given by:

$$\Delta \text{AIC}_i = \text{AIC}_i - \text{AIC}_{\min}; \quad \Delta \text{AIC}_{c_i} = \text{AIC}_{c_i} - \text{AIC}_{c_{\min}}; \quad \Delta \text{BIC}_i = \text{BIC}_i - \text{BIC}_{\min}, \tag{2.8}$$

with $i = 1, \ldots, R$, where R is the number of candidate models. AIC$_{\min}$, AIC$_{c_{\min}}$, and BIC$_{\min}$ correspond to the models with the lowest AIC, AIC$_c$, and BIC, respectively. Models for which the differences are smaller than 2 are considered to have empirical support [6]. The bigger the differences, the smaller the empirical support so that models with differences greater than 10 should be dismissed. Denoting (2.8) generically by Δ_i, it is also useful to evaluate $\exp\left(-\frac{1}{2}\Delta_i\right)$ that represents the likelihood of the ith model given the data. Using

$$w_i = \frac{\exp\left(-\frac{1}{2}\Delta_i\right)}{\sum\limits_{r=1}^{R} \exp\left(-\frac{1}{2}\Delta_r\right)} \tag{2.9}$$

leads to the weight of the ith model being the best model in the candidate set. Thus, w_i represents the probability of the ith model being the best given the data and the model set. In the context of Akaike criteria, it is possible to relate models i and j by using the evidence ratios, defined as w_i/w_j. These ratios express the evidence about the models which is best in the sense of K-L information. In particular, we are interested in the ratio w_1/w_j, considering that the probabilities of the models are descending, i.e., from best to worst model.

3   RESULTS AND DISCUSSIONS

The prior distributions used in the Bayesian calibration process are shown in Table 2. When applicable, we used the knowledge about the parameters available in [23]. For the Extended Allee models, we assumed a wide range for the possible values of the parameters, taking into account the biological constraints. Also, we assumed $N_0 \sim \mathcal{U}(10, 460)$ mm^3 and $\sigma \sim \mathcal{U}(0, 348)$ mm^3 for all models. Of note, each prior knowledge for the parameter values was considered as parameter uncertainties in the SA step.

Table 2: Prior distributions of the parameters associated with each model used in the Bayesian calibration process. We also assumed $N_0 \sim \mathcal{U}(10, 460)$ mm^3 and $\sigma \sim \mathcal{U}(0, 348)$ mm^3 for all models.

Model	a (day^{-1})	b (mm^3)	c (mm^3)	d (mm^3)
Exponential	$\mathcal{U}(0.017, 0.032)$	–	–	–
Logistic	$\mathcal{U}(0.021, 0.038)$	$\mathcal{U}(3480.000, 15000.000)$	–	–
Gompertz	$\mathcal{U}(0.001, 0.184)$	$\mathcal{U}(3480.000, 15000.000)$	–	–
EGAE	$\mathcal{U}(0.001, 0.100)$	–	$\mathcal{U}(-10000.000, 10000.000)$	$\mathcal{U}(1.000, 20000.000)$
LGAE	$\mathcal{U}(0.001, 0.100)$	$\mathcal{U}(3480.000, 15000.000)$	$\mathcal{U}(-10000.000, 10000.000)$	$\mathcal{U}(1.000, 20000.000)$

Assuming that parametric uncertainties are described by the defined prior distributions (Table 2), we first performed the SA of each model. To get an overall view of the global behavior of the first (μ^*) and second (σ) order sensitivity indices, SA was performed at six experimental points representative of the dynamics (20, 43, 65, 82, 98, and 114 days). The corresponding normalized SA results are presented in Table 3. Considering μ^*, the influence of the initial condition (N_0) decreases over time for all models, except for the LGAE model, for which it slightly grows. For the exponential model, N_0 is the most influential parameter throughout the simulation although the importance of the tumor growth rate (a) significantly increases over time, from around 10% up to 40% at the end of the simulation. For the logistic model, the decrease in the importance of N_0 is accompanied by the increasing influence of the parameter a and the carrying capacity (b), although the rank of importance among them does not change over time. At early times, a is more influential than b, which can be explained by the fact that the resource limitation is not a determinant issue. However, the competition for resources begins to play an increasing role over time, so that b becomes more and more influential. This is also observed for the Gompertz model,

for which b becomes the most influential parameter at the intermediate times of the dynamics. The growth rate a becomes the second most influential parameter while the role of the initial condition N_0 has shown no significance in comparison with the other parameters. As in the exponential model, the parameter a of the EGAE model is very influential, with its importance on the QoI increasing in time. In contrast, the parameter d, that modulates how the *per capita* population growth rate varies, is the most influential parameter at early times but switches position with a over time. Although much smaller than that of parameters a and d, the importance of the Allee threshold (c) increases over time. For the LGAE model, the parameter c is the most influential at early times of the dynamics followed by a, N_0, and d. The carrying capacity b does not appear at the plot since its first order sensitivity index is at least three orders of magnitude smaller than those of the other parameters, although it increases over time as in the logistic model. As the dynamics evolve, the importance of d increases, so that it is the third most influential parameter of the LGAE model at the end of simulation time. Regarding the nonlinear roles of the parameters measured by the second order sensitivity index (σ), we remark the similarities with the results obtained for the first order sensitivity index.

Overall, considering the parameter ranges used in the SA, the analysis demonstrates that the present parameter uncertainties significantly impact the tumor growth estimates for the considered models. Indeed, it is essential to calibrate them more accurately to make robust predictions, which requires more observational data.

The estimation of the parameters was performed applying the Bayesian technique using the 14 tumor volume points over the 114 days presented in [23]. Table 4 contains the samples of the posterior distribution of the parameters N_0, a, b, c, d, and σ. Visual inspection of the posterior distributions allows identifying the quality of the performed calibrations. In general, we observe that there was a substantial improvement of the knowledge about the parameters with the application of Bayesian inference. Of note, the marginal posterior distributions of the carrying capacity for the logistic and Gompertz models present high uncertainty and looks truncated at the upper limit of the range. This keeps happening even increasing the upper limit of the corresponding prior distributions. It is also worth observing the spreading of the samples of the posterior distributions of some parameters of the EGAE and LGAE models. We remark that: (i) the experimental data do not show saturation behavior which compromises the calibrations of the carrying capacity; (ii) likewise, estimates of parameters associated with models considering the Allee effect present high uncertainty. Given the previous SA, such uncertainties are expected to have a significant impact on model predictions. For better estimates, new experimental data are required.

Table 5 presents the MAP estimates of the calibrated parameters, including the ones obtained for σ. We observe that the parameter values for the exponential, logistic, and Gompertz models are similar to those obtained in [23]. Note that the posterior distributions presented in Table 4 indicate that the computed estimates are accompanied by uncertainties, about which there was no information provided in [23]. The LGAE model displays the least MAP estimate for the hyper-

Table 3: Normalized global sensitivity indices, μ_i^* and σ_i, calculated for each parameter in the respective models, at times $t = 20, 43, 65, 82, 98,$ and 114 days.
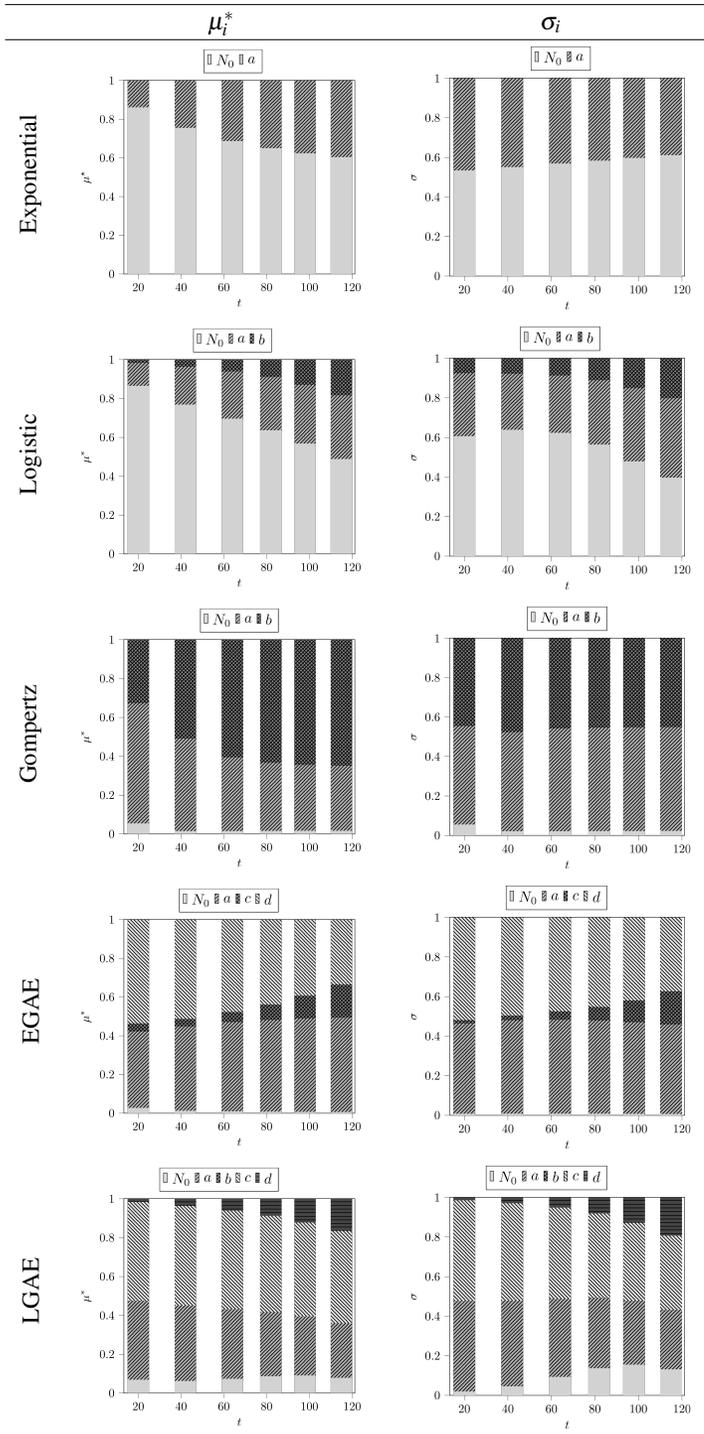
Table 4: Samples of the posterior distributions of the parameters for each model, including the initial condition N_0 and the hyperparameter σ.
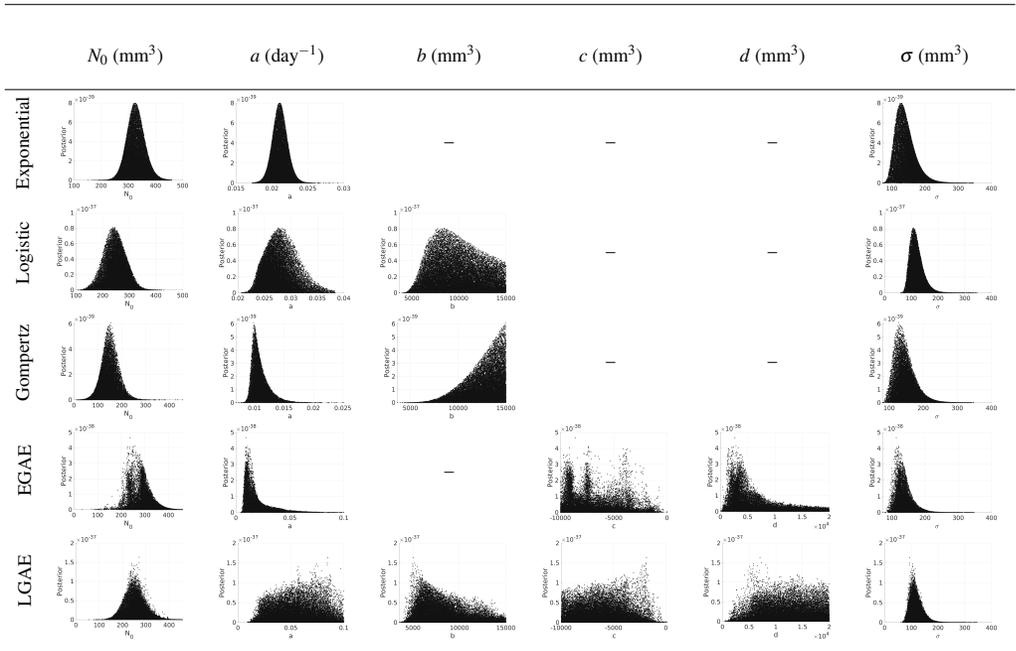

Table 5: The maximum *a posteriori* probability (MAP) estimates of calibrated parameters for each tumor growth model.

Model	N_0 (mm^3)	a (day^{-1})	b (mm^3)	c (mm^3)	d (mm^3)	σ (mm^3)
Exponential	326.339	0.021	–	–	–	124.725
Logistic	245.185	0.028	8472.914	–	–	109.408
Gompertz	155.186	0.010	14944.920	–	–	128.207
EGAE	237.783	0.009	–	-8622.679	2642.535	113.245
LGAE	271.309	0.085	5026.492	-1856.806	7023.970	94.289

parameter σ, closely followed by the logistic model. Besides, their hyperparameter uncertainties are quite similar, being the smallest among all models.

Analyzing Table 5, the c and d MAP estimates of the LGAE model indicate the presence of a weak Allee effect since $c < 0$ and $d > |c|$. On the other hand, the EGAE model does not display any Allee effect mechanism and therefore lacks a biological interpretation in the observed scenario. Notice that the initial experimental data point refers to the tumor volume with a significantly larger number of cells (5×10^6 cells were inoculated in each mouse to obtain xenografts [35]) compared to the amount analyzed by [17] (less than 200 cells per mm^3), what may explain why the Allee effect is more significant in the latter. Besides, it is noteworthy that
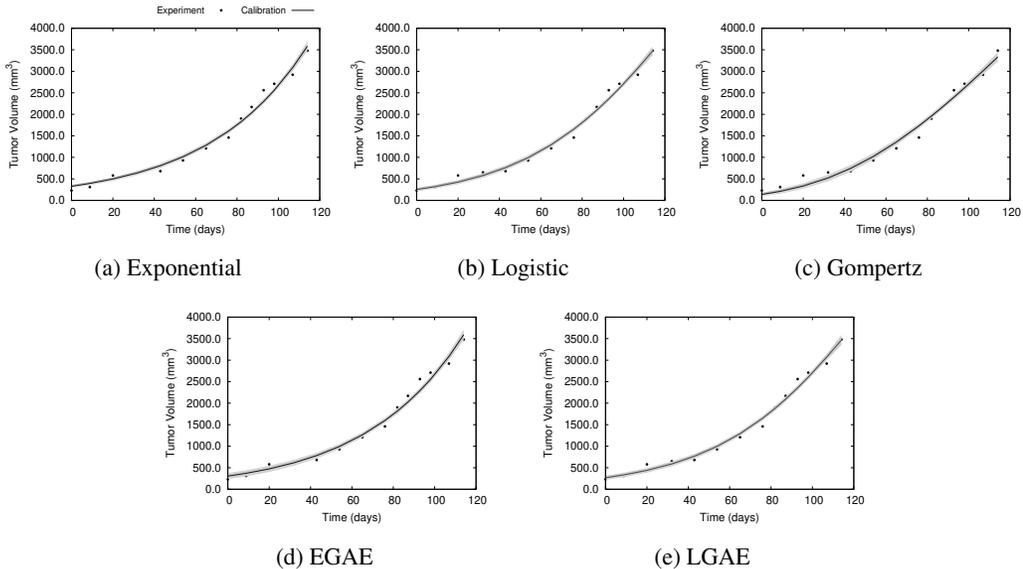
Figure 2: Simulation of the models considered in gray, with a credible interval at 95%. The black dots represent the experimental data.

the experiment of [17] was carried out on an *in vitro* culture which allows evaluating the tumor growth dynamics at low cell densities.

Figure 2 shows the tumor growth simulations for each of the calibrated mathematical models. Overall, all models represented the experimental data adequately. We notice that the credible interval is very narrow, despite the uncertainties mentioned above. Note that inspection of the figure alone does not allow to identify the best model that fits the data. This choice will be made systematically in the model selection step, but it is firstly necessary to assess the quality of the calibration process.

The sensitivity matrix condition numbers for the calibrated models are presented in Table 6. Notice that the condition number is over 1.0×10^5 in all cases, with the Extended Allee models being the most sensitive to small variations around the MAP estimates. Thus, one must be careful with the conclusions obtained through the analysis of the EGAE and LGAE models regarding the evidence of the existence of the Allee effect. Further studies are needed for more detailed identification of the influence of the Allee effect on the investigated tumor scenario. In contrast, the exponential model is the least sensitive to these changes, while logistic and Gompertz models have a little bit higher cond(J). Given the posterior distributions presented in Table 4, this result reinforces the observation made earlier that incorporating additional experimental data would improve the estimation of the carrying capacity of the logistic, Gompertz, and LGAE models, which would ultimately reduce the sensitivity matrix condition number of these models.

Table 6: Sensitivity matrix condition number (cond(J)) calculated for each tumor growth model. cond(J) $= \sigma_{max}/\sigma_{min}$, where σ_{max} and σ_{min} are the maximum and minimum singular values of J, respectively.

Model	σ_{min}	σ_{max}	cond(J)
Exponential	4.188×10^0	6.821×10^5	1.628×10^5
Logistic	3.600×10^{-2}	4.545×10^5	1.265×10^7
Gompertz	9.310×10^{-3}	8.238×10^5	8.850×10^7
EGAE	1.350×10^{-7}	1.572×10^8	1.160×10^{15}
LGAE	7.325×10^{-8}	1.659×10^6	2.264×10^{13}

Table 7: Estimated values for the Akaike (AIC), the second-order Akaike (AIC$_c$), and Bayesian (BIC) information criteria for each tumor growth model. The log-likelihood function at its maximum point ($\log(\mathscr{L}(\hat{\boldsymbol{\theta}}|\boldsymbol{y}))$), number of calibrated parameters (n), and differences of each criterion are also presented.

Model	$\log(\mathscr{L}(\hat{\boldsymbol{\theta}}|\boldsymbol{y}))$	n	AIC	AIC$_c$	BIC	ΔAIC	ΔAIC$_c$	ΔBIC
Exponential	-87.725	3	181.449	183.849	183.366	2.637	0.593	1.998
Logistic	-85.406	4	178.812	183.257	181.368	0.000	0.000	0.000
Gompertz	-87.993	4	183.986	188.431	186.543	5.174	5.174	5.174
EGAE	-85.962	5	181.924	189.424	185.119	3.112	6.167	3.751
LGAE	-84.706	6	181.413	193.413	185.247	2.601	10.156	3.879

Finally, the model selection information criteria, their corresponding differences, the log-likelihood function at its maximum point ($\log(\mathscr{L}(\hat{\boldsymbol{\theta}}|\boldsymbol{y}))$), and the total number of calibrated parameters (n) used to calculate them are shown in Table 7. Each model selection criterion yields different ordering of the evaluated models, although all three criteria selected the logistic model as the most parsimonious based on the available data. According to AIC and BIC, all models are supported by the available data, having criterion differences less than 10. However, due to the small ratio between the number of experimental samples m and the number of estimated parameters n (that can be as low as 2.333 for the LGAE model), this information can be misleading, favoring models that have more parameters. Indeed, AIC$_c$ corrects this small ratio bias and indicates the LGAE model as the worst in the candidate set. However, it is worth noting that the LGAE model has the smallest absolute value of the $\log(\mathscr{L}(\hat{\boldsymbol{\theta}}|\boldsymbol{y}))$ and the selection as the worst model indicates that the goodness-of-fit term was not good enough to compensate the cost of having more parameters together with small amount of data.

Figure 3 depicts the weights w_i which represent the probability that the ith model is the best among the candidate model set. In fact, the logistic model is the most indicated by the three model selection criteria, with AIC, AIC$_c$, and BIC weights equal to 0.548, 0.535, and 0.575, respectively. Ranking the models from the smallest to the largest AIC$_c$, the weight w_1 of the logistic model is used to calculate the evidence ratios. Compared with the exponential model,
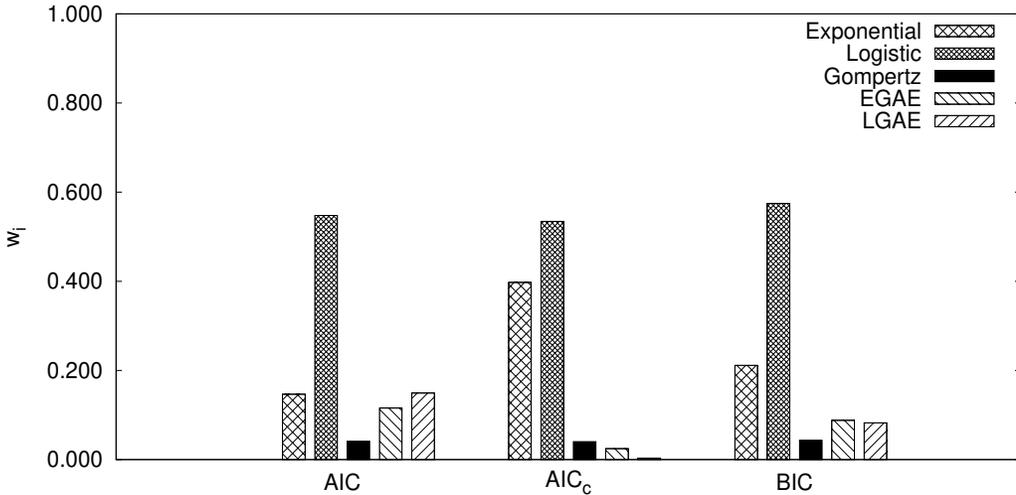
Figure 3: Probabilities (or weights) of being the best model given the data for each tumor growth model in the candidate model set.

we obtained $\left(\frac{w_1}{w_2} = 1.345\right)$, while with Gompertz, EGAE, and LGAE models, the evidences are $\left(\frac{w_1}{w_3} = 13.291\right)$, $\left(\frac{w_1}{w_4} = 21.836\right)$, and $\left(\frac{w_1}{w_5} = 160.465\right)$, respectively. These values indicate strong evidence in favor of the logistic model.

4   CONCLUDING REMARKS

In this paper, we analyzed some cancer growth models in the light of experimental data in mice. This analysis allowed the identification of the influence order among each model's parameters regarding the tumor volume evolution. The models were calibrated using the Bayesian approach, which provides not only the parameter MAP estimates but also their probability distributions, in contrast to pointwise estimates obtained by [23]. This fact adds more information to the study and may be used for more detailed analysis. In general, the available data were able to improve the prior knowledge defined for all model parameters. The sensitivity of these models in the neighborhood of the MAPs was also verified. We then applied model selection criteria to identify the best model for describing the experimental data. The strength of evidence in favor of the logistic model being the best in comparison with the exponential model is 1.345, and the evidence increases when compared to the Gompertz (13.29), EGAE (21.84), and LGAE (160.46) models.

Although the weak Allee effect was suggested by the LGAE model, this model was not selected as the best to describe the available data and its sensitivity matrix condition number indicated high sensitivity to small variations around MAP estimates. Thus, additional experiments are necessary to precisely evaluate the importance of the Allee effect for the considered experimental scenario. There is evidence in the literature on the importance of the Allee effect in breast cancer

cell culture [17], in post-resection recurrence of glioblastoma [24], in tumor necrotic nucleus formation [14], among others. Of note, other studies indicated that the Allee effect even without being considered *a priori* in the modeling process may emerge as a result of tumor dynamics. This fact occurred in the experiments performed in [5] as a result of specific regulation of the phenotypic plasticity between migratory and proliferative tumor cells. The presence of the Allee effect on tumor growth dynamics may be related to the existence of cooperative behavior between tumor cells, due to autocrine growth factors, and the production and secretion of diffuse signaling molecules by cells that increase growth and proliferation of other cells [24]. Recently, this hypothesis was further investigated in [13] using both *in vitro* and clinical trials. That reported evidence opens new avenues for tumor growth modeling in more complex scenarios which should be investigated preferably based on a larger amount of experimental data.

Overall, the proposed methodology for model analysis provides a detailed picture of calibration reliability, model assumptions, and uncertainties, and may be particularly useful in developing models to describe more complex scenarios of tumor dynamics.

Acknowledgements

REFERENCES

[1] P. Altrock, F. Michor & L. Liu. The mathematics of cancer: integrating quantitative models. *Nature Reviews*, **15** (2015), 730–745. doi:10.1038/nrc4029.

[2] D. Basanta & A.R.A. Anderson. Exploiting ecological principles to better understand cancer progression and treatment. *Interface Focus*, **3**(4) (2013), 1–11. doi:10.1098/rsfs.2013.0020.

[3] G. Box & G. Tiao. "Bayesian inference in statistical analysis". Addison-Wesley series in behavioral science: Quantitative methods. Addison-Wesley Pub. Co. (1973). doi:10.1002/9781118033197.

[4] R. Brady & H. Enderling. Mathematical Models of Cancer: when to Predict Novel Therapies, and When Not to. *Bulletin of Mathematical Biology*, **81** (2019), 3722 – 3731. doi:10.1007/s11538-019-00640-x.

[5] K. Böttger, H. Hatzikirou, A. Voss-Böhme, E.A. Cavalcanti-Adam, M.A. Herrero & A. Deutsch. An Emerging Allee Effect Is Critical for Tumor Initiation and Persistence. *PLOS Computational Biology*, **3** (2015), 1–14. doi:10.1371/journal.pcbi.1004366.

[6] K.P. Burnham & D.R. Anderson. "Model Selection and Multimodel Inference – A Practical Information-Theoretic Approach". Springer-Verlag, New York, 2nd ed. (2002).

[7]   K.P. Burnham & D.R. Anderson. Multimodel Inference: Understanding AIC and BIC in Model Selection. *Sociological Methods & Research*, **33**(2) (2019), 261 – 304. doi:10.1177/0049124104268644.

[8]   F. Campolongo, S. Tarantola & A. Saltelli. Tackling quantitatively large dimensionality problems. *Computer physics communications*, **117**(1-2) (1999), 75–85.

[9]   A. Cintrón-Arias, H.T. Banks, A. Capaldi & A.L. Lloyd. A sensitivity matrix based methodology for inverse problem formulation. *Journal of Inverse and Ill-Posed Problems*, (2009), 545–564. doi:10.1515/JIIP.2009.034.

[10] F. Courchamp, L. Berec & J. Gascoigne. "Allee effects in ecology and conservation". Oxford University Press (2008).

[11] F. Courchamp, T. Clutton-Brock & B. Grenfell. Inverse density dependence and the Allee effect. *Trends in Ecology & Evolution*, **14**(10) (1999), 405–410. doi:10.1016/s0169-5347(99)01683-3.

[12] J. da Costa, H. Orlande & W. da Silva. Model selection and parameter estimation in tumor growth models using approximate Bayesian computation – ABC. *Computational and Applied Mathematics*, **37** (2017), 2795 – 2815. doi:10.1007/s40314-017-0479-0.

[13] M. Delitala & M. Ferraro. Is the Allee effect relevant in cancer evolution and therapy? *AIMS Mathematics*, **5**(6) (2020), 7649–7660. doi:10.3934/math.2020489.

[14] P. Feng, Z. Dai & D. Wallace. On a 2D Model of Avascular Tumor with Weak Allee Effect. *Journal of Applied Mathematics*, **2019** (2019), 1–13. doi:10.1155/2019/9581072.

[15] P. Gregory. "Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with Mathematica® Support". Cambridge University Press (2005). doi:10.1017/CBO9780511791277.

[16] INCA. Instituto Nacional de Câncer. Estatísticas de câncer (2018). URL `http://www.inca.gov.br/numeros-de-cancer`.

[17] K.E. Johnson, G. Howard, W. Mo, M.K. Strasser, E.A.B.F. Lima, S. Huang & A. Brock. Cancer cell population growth kinetics at low densities deviate from the exponential growth model and suggest an Allee effect. *PLOS Biology*, **17(8)** (2019), e3000399. doi:10.1371/journal.pbio.3000399.

[18] P. Kirk, T. Thorne & M.P. Stumpf. Model selection in systems and synthetic biology. *Current Opinion in Biotechnology*, **24**(4) (2013), 767 – 774. doi:10.1016/j.copbio.2013.03.012.

[19] K.S. Korolev, J.B. Xavier & J.G. Gore. Turning ecology and evolution against cancer. *Nature Reviews Cancer*, **14** (2014), 371 – 380. doi:10.1038/nrc3712.

[20] R.J. LeVeque. "Finite Difference Methods for Ordinary and Partial Differential Equations". Society for Industrial and Applied Mathematics, Phyladelphia, PA (2007). doi:10.1137/1.9780898717839. URL `https://epubs.siam.org/doi/abs/10.1137/1.9780898717839`.

[21] J. Liepe, P. Kirk, S. Filippi, T. Toni, C.P. Barnes & M.P.H. Stumpf. A framework for parameter estimation and model selection from experimental data in systems biology using approximate Bayesian computation. *Nature Protocols*, **9** (2014), 439 – 456. doi:10.1038/nprot.2014.025.

[22] E.A.B.F. Lima, J.T. Oden, D.A. Hormuth, T.E. Yankeelov & R.C. Almeida. Selection, calibration, and validation of models of tumor growth. *Mathematical Models and Methods in Applied Sciences*, **26**(12) (2016), 2341–2368. doi:10.1142/S021820251650055X.

[23] H. Murphy, H. Jaafari & H.M. Dobrovolny. Differences in predictions of ODE models of tumor growth: a cautionary example. *BMC Cancer*, **16**(1) (2016), 163. doi:10.1186/s12885-016-2164-x.

[24] Z. Neufeld, W.v. Witt, D. Lakatos, J. Wang, B. Hegedus & A. Czirok. The role of Allee effect in modelling post resection recurrence of glioblastoma. *PLOS Computational Biology*, **17** (2017), 1–14. doi:10.1371/journal.pcbi.1005818.

[25] J.T. Oden, E.E. Prudencio & A. Hawkins-Daarud. Selection and assessment of phenomenological models of tumor growth. *Mathematical Models and Methods in Applied Sciences*, **23**(7) (2013), 1309–1338. doi:10.1142/S0218202513500103.

[26] F. Pianosi, K. Beven, J. Freer, J.W. Hall, J. Rougier, D.B. Stephenson & T. Wagener. Sensitivity analysis of environmental models: A systematic review with practical workflow. *Environmental Modelling & Software*, **79** (2016), 214 – 232. doi:https://doi.org/10.1016/j.envsoft.2016.02.008.

[27] E.E. Prudêncio & K.W. Schulz. The parallel C++ statistical library 'QUESO': Quantification of Uncertainty for Estimation, Simulation and Optimization. In "Euro-Par 2011: Parallel Processing Workshops". Springer (2012), p. 398–407. doi:10.1007/978-3-642-29737-3_44.

[28] A.C.M. Resende. "Sensitivity analysis as a tool for tumor growth modeling". Master's thesis, Laboratório Nacional de de Computação Científica (LNCC/MCTIC), Petrópolis/RJ, Brazil (2016).

[29] A. Rohatgi. WebPlotDigitizer Version: 4.1 (2018). URL https://automeris.io/WebPlotDigitizer/.

[30] A. Saltelli, K. Aleksankina, W. Becker, P. Fennell, F. Ferretti, N. Holst, S. Li & Q. Wu. Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. *Environmental Modelling & Software*, **114** (2019), 29 – 39. doi:10.1016/j.envsoft.2019.01.012.

[31] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana & S. Tarantola. "Global sensitivity analysis: the primer". John Wiley & Sons (2008). doi:10.1002/9780470725184.

[32] T. Toni, D. Welch, N. Strelkowa, A. Ipsen & M.P. Stumpf. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of The Royal Society Interface*, **6**(31) (2009), 187–202. doi:10.1098/rsif.2008.0172.

[33] D.J. Warne, R.E. Baker & M.J. Simpson. Using experimental data and information criteria to guide model selection for reaction–diffusion problems in mathematical biology. *Bulletin of Mathematical Biology*, **81**(6) (2019), 1760–1804. doi:10.1007/s11538-019-00589-x.

[34] WHO. World Health Organization. Cancer (2019). URL http://www.who.int/cancer/en/.

[35] A. Worschech, N. Chen, Y.A. Yu & et al. Systemic treatment of xenografts with vaccinia virus GLV-1h68 reveals the immunologic facet of oncolytic therapy. *BMC Genomics*, **10** (2009), 301–301. doi:10.1186/1471-2164-10-301.

APPENDIX

Table 8: Experimental data set obtained from [23] through *WebPlotDigitizer* tool. Murphy *et al.* [23] extracted these data from [35].

Time (day)	Tumor Volume (mm^3)	Time (day)	Tumor Volume (mm^3)
0.000	230.000	75.933	1460.000
8.858	310.000	81.950	1900.000
20.000	580.000	87.075	2170.000
32.033	650.000	92.869	2560.000
42.953	680.000	97.994	2710.000
53.872	930.000	106.908	2920.000
65.014	1210.000	114.039	3480.000