Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 50-1 (2017) 12203–12209

Comparing Methods for Parameter Estimation of the Gompertz Tumor Growth Model

Spyridon Patmanidis * **Alexandros C. Charalampidis** ** **Ioannis Kordonis** *** **Georgios D. Mitsis** **** **George P. Papavassilopoulos** *

* *School of Electrical and Computer Engineering, National Technical University of Athens, Iroon Polytechniou 9, Zografou 15780, Athens, Greece (email: {spatmanid@gmail.com, spatmani@central.ntua.gr}, yorgos@netmode.ece.ntua.gr)*
** *CentraleSupélec, Avenue de la Boulaie, 35576 Cesson-Sévigné, France (email: alexandros.charalampidis@centralesupelec.fr)*
*** *University of Southern California, Viterbi School of Electrical and Computer Engineering, Ming Hsieh Department of Electrical Engineering 3740 McClintock Avenue, Los Angeles, CA 90089, USA (email: jkordonis1920@yahoo.com)*
**** *Department of Bioengineering, McGill University, 817 Sherbrooke Ave W, MacDonald Engineering Building 270, Montréal QC H3A 0C3, Canada (email: georgios.mitsis@mcgill.ca)*

Abstract: Cancer, also known as malignant tumor or malignant neoplasm, is the name given to a collection of related diseases. In all types of cancer, some of the body's cells begin to divide abnormally without stopping and have the potential to invade surrounding tissues. In this work, we focus on estimating the parameters of a model which tries to describe the growth of a cancer tumor based on the available measurements of the tumor volume and on comparing the effectiveness with respect to the accuracy of the estimation of the various methods we have tested. The Gompertz function is used as the model basis, and our analysis aims to compute the growth rate and the plateau size of the tumor. The methods used to estimate these parameters are based on least squares, maximum likelihood and the Extended Kalman Filter (EKF). In this work, we present five different methods. The results show that, when the process and measurement noise characteristics are known, maximizing the joint probability density function of the observations using numerical integration to compute the probability density functions yields most times the best results. The methods based on the EKF also yield satisfactory results.

Keywords: Tumor Growth Modeling, Biomedical Systems, Nonlinear Systems, Parameter Estimation, Maximum Likelihood, Extended Kalman Filter, Least Squares

1. INTRODUCTION

Cancer is the leading cause of death in the developed world and the second leading cause of death in the developing world (WHO, 2014). The great majority of cancers, some around 90-95% of cases, are due to environmental factors and the remaining 5-10% are due to inherited genetics (Anand et al., 2008). Treatment options − some of the primary are chemotherapy, surgery and radiation therapy − depend on the type, location and grade of the cancer (WHO, 2014, Wikipedia, 2016).

Chemotherapy is the treatment of cancer with one or more cytotoxic anti-neoplastic drugs (chemotherapeutic agents) as part of a standardized regiment. Traditional chemotherapeutic agents act by killing cells that divide rapidly, one of the main properties of most cancer cells. Mathematical modeling and optimal control techniques could help to deliver a better outcome of cancer chemotherapy. Mathematic models are used to describe the evolution of a tumor, the mechanism of drug effects and the constraints of drug use due to subsequent toxicities, while optimal control uses the developed models to design optimal chemotherapy strategies (Barbolosi and Iliadis, 2001, Dua and Pistikopoulos, 2008, Hadjiandreou and Mitsis, 2014).

In the literature, a lot of works deal with the evaluation of the ability the existing models have to describe the tumor dynamics (Benzekry et al., 2014, Nguimkeu, 2014, Ribba et al., 2014) or with the problem of predicting the future volumes of a tumor (Hadjiandreou and Mitsis, 2014, Achilleos et al., 2012, Achilleos et al., 2014). To the best of our knowledge, this is the first work comparing the performance of various methods for estimating the parameters of the Gompertz function.

Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2017.08.2289

In this work, we focus on estimating the parameters of a model that can describe the evolution of a tumor growth in an individual subject. As tumor growth depends on various parameters according to the individual patient, it is important to compute personalized models. This may lead to better chemotherapy strategies. The results depend on the method used. In our analysis, we use the Gompertz curve, a widely used model which takes into account the reduced growth rate of the tumor that is observed as its size increases. In that model, we consider that the initial tumor size is known and we need to identify two unknown parameters, the proliferative ability of the cells and the carrying capacity (the maximum size that can be reached with the available nutrients, also referred to as plateau). Both parameters are considered constant. In our approach, we have a set of measurements representing the volumes of a tumor at various time instants and we need to find the parameters that create the curve fitting best to the set of measurements.

This paper is organized as follows: Section 2 describes the tumor growth model used in this work, Section 3 describes the techniques used in order to estimate the parameters of the Gompertz function, Section 4 presents the results from the experiments we conducted, Section 5 contains the discussion about the experimental results and Section 6 contains the conclusions and suggestions for further research.

2. TUMOR GROWTH MODEL

During the early stages, cancer tumors proliferate in an exponential fashion. Later on, as the tumor size increases, the growth rate decreases and the tumor reaches a plateau size. Unlike the simple exponential model, the Gompertz function can predict this behavior successfully. This model has been widely used because of its simplicity and its ability to describe experimental data reasonably well. It is given by:

$$x_{\mathrm{k+1}} = \theta_2 exp\left(log(\frac{x_{\mathrm{k}}}{\theta_2})exp(-\frac{1}{\theta_1}T)\right) \tag{1}$$

where x_{k} (mm^3) is the tumor size, θ_1 (days) is a constant related to the proliferative ability of the cells, θ_2 (mm^3) is the carrying capacity ($\lim_{k\to\infty} x_{\mathrm{k}} = \theta_2$), T (days) is the time interval between k and $k+1$, and $k \in \mathbb{N}$ - see Dennis and Ponciano (2014).

Assuming random additive process and measurement noise the model can be formulated as follows:

$$x_{\mathrm{k+1}} = f(x_{\mathrm{k}}, \theta) + w_{\mathrm{k}} \tag{2a}$$

$$y_{\mathrm{k}} = x_{\mathrm{k}} + \upsilon_{\mathrm{k}} \tag{2b}$$

where

$$f(x_{\mathrm{k}}, \theta) = \theta_2 exp\left(log(\frac{x_{\mathrm{k}}}{\theta_2})exp(-\frac{1}{\theta_1}T)\right) \tag{3}$$

and

$$\theta = \begin{bmatrix}\theta_1\\ \theta_2\end{bmatrix}. \tag{4}$$

The random variables w_{k} and υ_{k}, $k \in \mathbb{N}$ are mutually independent and normally distributed with known parameters: $w_{\mathrm{k}} \sim \mathcal{N}(0, \sigma_1 {x_{\mathrm{k}}}^{e_1})$ and $\upsilon_{\mathrm{k}} \sim \mathcal{N}(0, \sigma_2 {x_{\mathrm{k}}}^{e_2})$. Measurements are available from time k=1 onwards and $x_0 = y_0 = 1$ is assumed. Because of the random components w_{k} and υ_{k}, the model (2) can also be represented via the description:

$$x_{\mathrm{k+1}} \sim p(x_{\mathrm{k+1}}|x_{\mathrm{k}}) \tag{5a}$$

$$y_{\mathrm{k}} \sim p(y_{\mathrm{k}}|x_{\mathrm{k}}) \tag{5b}$$

where $p(x_{\mathrm{k+1}}|x_{\mathrm{k}})$ is the probability density function describing the dynamics for given values of x_{k}, and $p(y_{\mathrm{k}}|x_{\mathrm{k}})$ is the probability density function describing the measurements.

3. PARAMETER ESTIMATION TECHNIQUES

In this paper, the following problem is considered: let $Y_{\mathrm{N}} = \{y_1, y_2, ..., y_{\mathrm{N}}\}$ be the available measurements of a cancer tumor volume and (2a), (2b) describe the system dynamics; compute an estimate $\hat{\theta}$ of the parameter θ based on the N available measurements, considering that the process and measurement noise parameters (σ_1, e_1, σ_2, e_2) are known. In the rest of this section, we describe the techniques used to estimate the parameters of the Gompertz function.

3.1 Naive Least Squares

The Least Squares approach used in this work is a very simple and easily applicable implementation. In this text, the method will be referred to as Naive Least Squares (NLS). In this method, the measurement noise is not taken into consideration, and as a result it is assumed that $y_k = x_k$. The goal is to find an estimate for $\theta \in \Theta$ that minimizes the error:

$$\varepsilon_\theta = \sum_{k=1}^{N}(y_k - \hat{x}_k)^2 \tag{6}$$

where

$$\hat{x}_k = f(y_{\mathrm{k\text{-}1}}, \theta) = \theta_1 exp\left(log(\frac{y_{\mathrm{k\text{-}1}}}{\theta_1})exp(-\frac{1}{\theta_2}T)\right) \tag{7}$$

with $\Theta \subseteq \mathbb{R}^2$ denoting a compact set of permissible values of the unknown parameter θ, and y_0 considered known.

3.2 Maximum Likelihood

The Maximum Likelihood approach involves maximizing the joint density (likelihood) $p_\theta(Y_{\mathrm{N}})$ of the observation:

$$\hat{\theta} = \arg\max_{\theta\in\Theta} p_\theta(y_1, ..., y_{\mathrm{N}}) \tag{8}$$

with $\Theta \subseteq \mathbb{R}^2$ denoting a compact set of permissible values of the unknown parameter θ (Schön et al., 2011). To compute this, Bayes' rule can be used in order to decompose the joint density according to

$$p_\theta(y_1, ..., y_{\mathrm{N}}) = p_\theta(y_1)\prod_{k=2}^{N} p_\theta(y_{\mathrm{k}}|Y_{\mathrm{k\text{-}1}}) \tag{9}$$

where

$$p(y_{\mathrm{k+1}}|Y_{\mathrm{k}}) = \int p(y_{\mathrm{k+1}}|x_{\mathrm{k+1}})p(x_{\mathrm{k+1}}|Y_{\mathrm{k}})dx_{\mathrm{k+1}} \tag{10}$$

$$p(x_{\mathrm{k+1}}|Y_{\mathrm{k}}) = \int p(x_{\mathrm{k+1}}|x_{\mathrm{k}})p(x_{\mathrm{k}}|Y_{\mathrm{k}})dx_{\mathrm{k}} \tag{11}$$

and

$$p(x_{k+1}|Y_{k+1}) = \frac{p(y_{k+1}|x_{k+1})p(x_{k+1}|Y_k)}{p(y_{k+1}|Y_k)}. \quad (12)$$

In this work, the permissible values Θ consist of a $n_\Theta \times n_\Theta$ grid. Since the dimension of x_k is 1, numerical integration can be used to approximate the integrals of (10) and (11).

The most crucial part in this method is to perform an accurate integration. The interval of integration is $[0, X_{max}]$ and in order to compute this definite integral the trapezoidal rule

$$\int_a^b f(x)dx \approx (b-a)\left[\frac{f(a)+f(b)}{2}\right] \quad (13)$$

is used. The smaller the interval $[a, b]$ is, the smaller the approximation error will be. For this reason, we can divide the interval $[0, X_{max}]$ to smaller intervals and compute the sum for all the integrals. However, the number of intervals required for a very small error is very large in our case and computing all these integrals at every iteration is prohibitive. Nevertheless, because of the nature of the problem, there is a way to find an interval significantly smaller than the interval $[0, X_{max}]$ where the error is also very small. Based on the Gaussian nature of the noise and using the measurement y_k as the center of the distribution for x_k with standard deviation $\sigma = \sigma_0 {y_k}^{e_0}$, we compute the interval $[y_k - 4\sigma, y_k + 4\sigma]$. This interval is divided into smaller intervals (in this work the experiments were conducted using 50 intervals), and the sum of all these intervals gives the values for the distributions of (10) and (11).

As regards the grid, there are two important parameters. The size of the grid (how many values it contains) and the range between these values. The size of the grid is the main parameter that affects the execution time, while the range between the values the grid contains affects the accuracy of the parameter estimation. In order to reduce the size of the grid, the ML method is executed three times, each time decreasing the range between the minimum and the maximum value. The first time the search interval is big enough to contain the real value of the unknown parameters. During the subsequent steps, the search interval becomes smaller, and the grid is centered at the value estimated in the previous step. The default values for the aforementioned variables will be mentioned in Section 4.

3.3 Naive Maximum Likelihood

Inspired by the NLS method in 3.1 a simplification concerning the noise can be made, which leads to equations that can be dealt with easier. If (2a) and (2b) are modified to

$$x_{k+1} = f(y_k, \theta) + d_k \quad (14a)$$
$$y_k = x_k \quad (14b)$$

where

$$f(y_k, \theta) = \theta_1 exp\left(log(\frac{y_k}{\theta_1})exp(-\frac{1}{\theta_2}T)\right) \quad (15)$$

$$\theta = \begin{bmatrix}\theta_1 \\ \theta_2\end{bmatrix} \quad (16)$$

and $d_k \sim \mathcal{N}(0, \sigma_k)$, where $\sigma_k = \sigma_1 {y_{k-1}}^{e_1} + \sigma_2 {y_{k-1}}^{e_2}$, then maximum likelihood estimation can be used to find an estimate $\hat{\theta}$ for the unknown parameter θ that maximizes the likelihood $p_{like\_\theta}(Y_N)$:

$$\hat{\theta} = \arg\max_{\theta \in \Theta} p_{like\_\theta}(y_1, ..., y_N) \quad (17)$$

where

$$p_{like\_\theta}(y_1, ..., y_N) = \prod_{k=1}^{N} p_\theta(y_k|y_{k-1}) \quad (18)$$

with $\Theta \subseteq \mathbb{R}^2$ denoting a compact set of permissible values of the unknown parameter θ, y_0 being considered known and $p_\theta(y_k|y_{k-1})$ coming from (2b). This method will be referred to in this text as Naive Maximum Likelihood (NML).

3.4 Extended Kalman Filter

In this subsection, we use the Extended Kalman Filter (EKF) to estimate the parameters of the stochastic system using a state augmentation. The EKF gives an approximation of the optimal estimate. In order to approximate the non-linearities of the system dynamics, a linearized version of the nonlinear system model around the last state estimate is created (Maybeck, 1982, Charalampidis and Papavassilopoulos, 2011, Charalampidis et al., 2016).

If, in (3), we consider that the parameters θ_1 and θ_2 are also states of the system, and we name θ_1 as x_k^2 and θ_2 as x_k^3 (these are not exponents), the equation becomes:

$$f(x_k^1, x_k^2, x_k^3) = x_k^3 exp\left(log(\frac{x_k^1}{x_k^3})exp(-\frac{1}{x_k^2}T)\right) \quad (19)$$

where $x_i^j \in (0, +\infty)$ for every $i = 1, 2, ...N$ and $j = 1, 2, 3$. Then:

$$x_{k+1}^1 = f(x_k^1, x_k^2, x_k^3) + w_k \quad (20a)$$
$$x_{k+1}^2 = x_k^2 \quad (20b)$$
$$x_{k+1}^3 = x_k^3 \quad (20c)$$
$$y_k = x_k^1 + v_k \quad (21)$$

where $w_k \sim \mathcal{N}(0, q_k)$ and $v_k \sim \mathcal{N}(0, r_k)$. If we define

$$X_{k+1} = \begin{bmatrix}x_{k+1}^1 \\ x_{k+1}^2 \\ x_{k+1}^3\end{bmatrix} = \begin{bmatrix}f(x_k^1, x_k^2, x_k^3) \\ x_k^2 \\ x_k^3\end{bmatrix} + \begin{bmatrix}w_k \\ 0 \\ 0\end{bmatrix} = F(X_k) + W_k \quad (22)$$

$$Y_k = [\, X_k \;\, 0 \;\, 0\,] + V_k = H(X_k) + V_k \quad (23)$$

where $W_k \sim \mathcal{N}(\mathbf{0}, Q_k)$, $\mathbf{0} = \begin{bmatrix}0 \\ 0 \\ 0\end{bmatrix}$, $Q_k = \begin{bmatrix}q_k & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0\end{bmatrix}$ and $V_k \sim \mathcal{N}(0, R_k = r_k)$, then we have

$$X_{k+1} = F(\hat{X}_k) + \frac{\partial F}{\partial x_k}|_{\hat{x}_k}(X_k - \hat{X}_k) + W_k \quad (24)$$

and

$$Y_{k+1} = H(\hat{X}_k) + \frac{\partial H}{\partial x_k}|_{\hat{x}_k}(X_k - \hat{X}_k) + V_k. \quad (25)$$

Therefore, the prediction step is:

$$\hat{X}^-_{k+1} = F(\hat{X}_k) \quad (26)$$

$$P^-_{X_{k+1}} = \frac{\partial F}{\partial x_k}|_{\hat{x}_k} P_{X_k} \frac{\partial F}{\partial x_k}^T |_{\hat{x}_k} + Q_k \quad (27)$$

while the correction step is:

$$\hat{Y}^-_{k+1} = H(\hat{X}^-_{k+1}) \quad (28)$$

$$P^-_{Y_{k+1}} = \frac{\partial H}{\partial x_k}|_{\hat{X}^-_{k+1}} P^-_{X_{k+1}} \frac{\partial H}{\partial x_k}^T |_{\hat{X}^-_{k+1}} + R_k \quad (29)$$

$$P^-_{X_{k+1}Y_{k+1}} = P^-_{X_{k+1}} \frac{\partial H}{\partial x_k}^T |_{\hat{X}^-_{k+1}} \quad (30)$$

$$\hat{X}_{k+1} = \hat{X}^-_{k+1} + K_{k+1}(Y_{k+1} - \hat{Y}^-_{k+1}) \quad (31)$$

$$P_{X_{k+1}} = P^-_{X_{k+1}} - K_{k+1} P^-_{Y_{k+1}} K^T_{k+1} \quad (32)$$

$$K_{k+1} = P^-_{X_{k+1}Y_{k+1}} + (P^-_{Y_{k+1}})^{-1}. \quad (33)$$

The initial conditions for the EKF are $x^1_0 = 1$, $x^2_0 = \theta^{(1)}_1$ and $x^3_0 = \theta^{(1)}_2$, where $\theta^{(1)}_1$ and $\theta^{(1)}_2$ are the estimates of the unknown parameters computed using the NML method, and $P_{X_0} = I$. After the execution of the EKF algorthm, a new estimate for the unknown parameters is available. Using that estimation as initial condition for x^2_0 and x^3_0 and repeating the same procedure, a new estimation is derived. This is repeated until the estimated parameters converge to a $\hat{\theta}$.

3.5 ML - EKF Combination

Another approach is the combination of the EKF and the ML method. This method uses the ML technique and the EKF in turn. At the first iteration, NML can be used to find an estimate $\hat{\theta}^{(1)}$ for the unknown parameter θ. At later iterations, $\hat{\theta}^{(i)}$, where $i = 2j + 1$ and $j \in \mathbb{N}$, is the estimate for the unknown parameter θ that maximizes the likelihood $p_{like\_\theta}(Y_{\mathrm{N}})$:

$$\hat{\theta}^{(i)} = \arg\max_{\theta \in \Theta} p_{like\_\theta}(y_1, ..., y_{\mathrm{N}}) \quad (34)$$

where

$$p_{like\_\theta}(y_1, ..., y_{\mathrm{N}}) = \prod_{k=1}^{N} p_\theta(y_k | \hat{x}^{i-1}_k) \quad (35)$$

and $p_\theta(y_k | \hat{x}^{i-1}_k)$ is given by (2b).

Having the estimate $\hat{\theta}^{(i)}$ available and using it as initial condition in the EKF, we can get an estimate $\hat{x}^{i+1}_k$ for the system states x_k. Using the new estimate $\hat{x}^{i+1}_k$ in the maximum likelihood estimator, we can get another estimate $\hat{\theta}^{(i+2)}$, and so on. This procedure can be repeated several times until $\hat{\theta}^{(k)}$ converges to a vector $\hat{\theta}$.

4. SIMULATIONS

In this section, we present the results of the simulations performed using Matlab. In the first experiment, we simulate the growth of 100 cancer tumors. (2a) and (2b) describe the tumor dynamics. For each tumor, the tumor's carrying capacity θ_2 (parameter θ^j_2) and the doubling time θ_1 (parameter θ^j_1), as long as the process and measurement noise parameters $(\sigma^j_1, e^j_1, \sigma^j_2, e^j_2)$, where $j = 1, ..., 100$ is the identification number of the test subject, were chosen randomly from uniform distributions. Table 1 shows the minimum and maximum possible values for each parameter. The sampling time between two consecutive measurements is two days and the number of available measurements for each tumor is 30. The initial volume for x_0 for every tumor is 1 (mm^3).

As regards the ML method described in 4.3, $n_\Theta = 11$, $\sigma_0 = \sigma_1 + 0.1$ and $e_0 = e_1$. Furthermore, in the first execution of the method, the grid's intervals are $[x_{c1} - 10, x_{c1} + 10]$ for the x-axis and $[x_{c2} - 10, x_{c2} + 10]$ for the y-axis, where $x_{c1} = 12$ and $x_{c2} = 15$. In the second and third execution, the variables x_{c1} and x_{c2} take the values $\hat{\theta}_1$ and $\hat{\theta}_2$ estimated by the previous execution of the algorithm, and the grid's intervals are $[x_{c1} - 2, x_{c1} + 2]$ for the x-axis and $[x_{c2} - 2, x_{c2} + 2]$ for the y-axis in the second iteration and $[x_{c1} - 1, x_{c1} + 1]$ for the x-axis and $[x_{c2} - 1, x_{c2} + 1]$ for the y-axis in the third iteration.

The function `fmincon` provided in The Mathworks (2016) is used to minimize the error ϵ_θ and the likelihood $p_{like\_\theta}(Y_N)$ in NLS and NML, respectively. In order to test if `fmincon` can provide the set of parameters which minimizes the error ϵ_θ and the likelihood $p_{like\_\theta}(Y_N)$, we used parameter grids and performed extensive simulations to check if there is any set from the grids which is better than the set provided by `fmincon`. In every case, `fmincon` provided the best set of parameters.

Table 1. Minimum and maximum parameter values

Parameters	Minimum Value	Maximum Value
θ_1	5	15
θ_2	8	20
σ_1, σ_2	0.1	0.3
e_1, e_2	0.3	0.5

Table 2. Simulation results

Method	Mean Div. (%)		RMS Div.	
	θ_1	θ_2	θ_1	θ_2
NLS	16.0612	7.1645	13.3365	7.3556
NML	14.5844	6.6445	10.8416	6.7939
ML	12.6019	6.8500	10.7606	7.3053
EKF/EKF − ML	14.0094	7.0537	11.9133	7.4391

Table 2 shows the absolute mean percentage and the RMS value of the divergence of the estimated value for parameters θ_1 and θ_2 from the real value for every method described in Section 3. We also present the minimum and maximum values of the absolute percentage of divergence for each of the two unknown parameters, as well as the real values of θ_1 and θ_2 for the test subjects that the parameters showed the minimum and maximum divergence from the real ones, and the estimated values $\hat{\theta}_1$ or $\hat{\theta}_2$ (based on which parameter shows the minimum or maximum). Tables 3 and 4 contain the aforementioned information.

Table 3. Minimum divergence between real and estimated parameters achieved by each method

Method	Div. (%)	θ_1	θ_2	$\hat{\theta}_1$
NLS	0.0413	13.6740	8.1862	13.6796
NML	0.5213	8.6419	18.9599	8.5969
ML	0.0722	8.7937	14.5772	8.8000
EKF/EKF − ML	0.6290	13.9511	19.1620	13.9423
				$\hat{\theta}_2$
NLS	0.0241	10.5643	8.9633	8.9654
NML	0.0368	9.0437	18.6802	18.6733
ML	0.1569	6.9781	9.2145	9.2000
EKF/EKF − ML	0.0386	5.9121	9.0397	9.0432

Table 4. Maximum divergence between real and estimated parameters achieved by each method

Method	Div. (%)	θ_1	θ_2	$\hat{\theta}_1$
NLS	71.3236	19.2754	17.0948	33.0234
NML	53.4808	8.6876	8.1531	4.0414
ML	53.2046	18.2762	15.2335	28.0000
EKF/EKF − ML	52.1723	19.0060	9.0625	9.0901
				$\hat{\theta}_2$
NLS	37.2337	19.2754	17.0948	23.4599
NML	32.3063	18.2762	15.2335	20.1549
ML	53.2046	18.2762	15.2335	22.0000
EKF/EKF − ML	37.1583	18.0459	14.0771	8.8463

By observing Table 2, we can see that estimating the doubling time of the tumor is not as accurate as estimating the carrying capacity. In order to test if this behavior is a result of the system dynamics or it occurs due to the available measurements, we conducted two different experiments. Instead of waiting to obtain 30 measurements, the tumor size measurement procedure stops when the last measured volume reaches the 75% of the carrying capacity in the first experiment and the 50% of the carrying capacity in the second one. The results from these experiments are presented in Tables 5 and 6 and are discussed in the next section.

Table 5. Simulation results (75%) of carrying capacity

Method	Mean Divergence (%)		RMS Divergence	
	θ_1	θ_2	θ_1	θ_2
NLS	35.1654	38.3545	19.3628	30.4817
NML	32.7685	38.8622	21.3486	29.9026
ML	34.4015	33.2359	31.3080	22.9401
EKF	34.3691	43.8032	19.5450	31.0698
EKF − ML	34.3689	43.4494	19.5449	29.8533

Table 6. Simulation results 50% of carrying capacity

Method	Mean Divergence (%)		RMS Divergence	
	θ_1	θ_2	θ_1	θ_2
NLS	62.4214	53.3370	58.1404	35.1616
NML	64.5582	53.1736	42.1386	35.1578
ML	73.3614	67.1030	62.8518	55.3252
EKF	66.7913	55.7251	41.6237	34.9852
EKF − ML	66.5279	55.4277	41.5188	34.5241

Lastly, the final experiment is conducted to check the effect of process and measurement noise to the estimation results. Three noise categories have been created, see Table 7, based on the values the parameters σ_1^j and σ_2^j take. For every category, the experiment described at the start of this section was repeated. The results are presented in tables 8 to 10.

Table 7. Noise categories

Parameters	Minimum σ_1^j and σ_2^j	Maximum σ_1^j and σ_2^j
Low	0.01	0.1
Medium	0.1	0.25
High	0.25	0.4

Table 8. Low system and measurement noise variance

Method	Mean Divergence (%)		RMS Divergence	
	θ_1	θ_2	θ_1	θ_2
NLS	3.6265	1.6377	3.0243	1.4587
NML	3.3309	1.5478	2.5050	1.3714
ML	3.3001	1.5336	2.7131	1.3701
EKF/EKF − ML	3.9977	2.1865	3.6205	3.2832

Table 9. Medium system and measurement noise variance

Method	Mean Divergence (%)		RMS Divergence	
	θ_1	θ_2	θ_1	θ_2
NLS	19.1271	6.9694	14.1750	7.9602
NML	17.1271	6.6480	10.8947	6.7706
ML	13.8611	6.8200	11.3370	7.4151
EKF/EKF − ML	18.6537	9.7560	10.9018	11.1491

Table 10. High system and measurement noise variance

Method	Mean Divergence (%)		RMS Divergence	
	θ_1	θ_2	θ_1	θ_2
NLS	33.3905	11.5012	19.2734	11.3903
NML	32.8010	11.4564	17.6611	12.0277
ML	22.3425	11.3136	18.3328	14.5604
EKF	30.3944	14.2893	15.5761	16.2113
EKF − ML	30.9344	14.1240	15.5761	15.4199

5. DISCUSSION

In the discussion section, we start by commenting how the implemented techniques performed and continue with presenting and explaining the observations which can be made by studying the results.

It is obvious that the implementation of the NLS method used is outperformed by the other methods (except in case of low process and measurement noise where the EKF methods have a slightly worse result). Nevertheless, it is a very simple and fast approach and its results will probably be better than random estimates. The estimates computed using this method could be used as initial conditions for the EKF methods.

A second approach used due to its simplicity is the NML method. Simulation results show that in a few cases this method performs better than all the other methods used, but this behavior is not ensured. Also, in most cases, the ML method described in 3.3 performs better than NML. NML is preferable to NLS regarding the estimation of the initial conditions.

When the noise parameters are known, the EKF method and the combination of the EKF and ML yield the same results. This does not happen if the carrying capacity has not been reached, if noise parameters are high, or if the noise parameters are unknown (this is a more complex problem and it is outside the scope of this paper). Judging from the experimental results, it is not clear if the methods using the EKF are better compared to NLS and NML. Sometimes they perform better and sometimes they perform worse. Generally, they are better at estimating the doubling time and worse at estimating the carrying capacity. However, it is clear that the ML method performs better than the EKF methods.

The last method used in order to estimate the unknown parameters of the Gompertz function is the ML method described in 3.3. As mentioned before, compared to the other methods, the estimated values of the unknown parameters have the least divergence from the real parameters. However, in order to achieve this improvement, the amount of time needed to compute the integrals numerically is many times bigger than the time needed to compute parameter estimates using the other methods, but it is still acceptable.

Table 11 shows the mean time (in seconds) that each method needs. The simulations were performed using an Intel Core i7-6700K @ 4.00GHz and 16GB of DDR4 @ 3200 MHz.

Table 11. Mean execution time for each method

Method	Mean Time (seconds)
NLS	0.5
NML	0.1
ML	224
EKF	1.8
EKF − ML	1.1

Starting from Table 2, as regards the doubling time parameter θ_1, the ML method shows the best general performance, EKF methods (EKF and EKF-ML) come second, NML performs a little worse than the EKF methods and last comes the NLS method. Regarding the carrying capacity θ_2, the NML method performs sightly better than the ML and then follows the EKF methods and the NLS method. However, the results of carrying capacity estimation do not diverge as much as the results of doubling time estimation.

In Table 3, we can see the minimum absolute divergence between the real tumor's doubling time (θ_1) and the estimated doubling time ($\hat{\theta}_1$) achieved by each method. We can also see the minimum absolute divergence between the real tumor's carrying capacity (θ_2) and the estimated carrying capacity ($\hat{\theta}_2$). These minimum values were achieved when the noise category for the system was low or medium.

In order to estimate the carrying capacity θ_2 accurately, it is necessary that the tumor volume has reached the plateau. This also explains why the real value of doubling time θ_1 is low in Table 3. Because of the way the experiment has been set up - N is chosen to be 30 (measurements) and t is 2 (days) - the doubling time has to have a low value so that the tumor is able to reach the plateau during the measurement time. Another important factor in estimating the carrying capacity is the number of measurements near the plateau size, more measurements near the plateau resulting to higher accuracy. We also need to add that the reason the value of real θ_2 in Table 3 is most times low, is that the measurement noise variance is considered to be $\sigma_2 y_k^{e_2}$. Indeed, when θ_2 has small values, the variance is low and as a result the measurement noise will probably be smaller.

As regards the cases when the methods fail to make a good estimation of the parameters, for parameter θ_1 this happens mostly when the doubling time is high and the plateau is low, and so the measurements do not provide a good description of the tumor growth procedure. However, there are also cases where the tumor grows too fast and in combination with the high divergence among measurements at the plateau, the proposed methods fail to give a good estimation. For parameter θ_2, the methods fail to give good estimates if the doubling time is high and as a result there are only a few or no measurements of the plateau size during the monitoring time. This observation led us to the next experiment, where we investigate how well can the proposed methods estimate the parameter θ_2, when the measurements have reached only the 75% or the 50% of the carrying capacity.

Judging by the first experiment only, could lead to the conclusion that it is easier to estimate θ_2 compared to estimating θ_1. This happens because there are a lot of measurements describing the plateau of the tumor. When the tumor has not reached the plateau, estimating the carrying capacity is as difficult as estimating the doubling time. Additionally, the less measurements there are, the less accurate the parameter estimation is.

The last experiment was conducted to check the effect of the noise to the parameter estimation. Three noise categories were created (low, medium and high - see Table 7), depending on parameters σ_1 and σ_2. The simulation results confirm the intuitive expectation that higher noise variance results in a less accurate estimation. Furthermore, this experiment also provides us with another important observation, that the ML method proposed in 3.3 performs better than all the other methods tested and in addition when the process and measurement noise are medium or high, ML provides significantly better results regarding the estimation of the doubling time. The estimation of the carrying capacity is also better when ML is used but there is no significant difference from the other methods. This last conclusion can be also reached for the case of low noise variance.

Regarding the execution time, the simpler a method is, the faster it performs. NLS and NML are the fastest and then follows the EKF and the EKF-ML. ML needs a lot more time than all the other, but using this method we can compute even better estimates.

6. CONCLUSIONS

In this work, we developed methods that can estimate the unknown parameters of the Gompertz function, in order to use the function to describe the evolution of tumor volumes. Furthermore, we created synthetic data representing measurements of tumor volumes by using the same function and we applied the methods we developed. The parameters we considered unknown and tried to estimate are the tumor's doubling time and the carrying capacity of the tumor. Both process and measurement noise characteristics are considered known. We used the Least Squares method, the Maximum Likelihood and the Extended Kalman Filter. We also combined the aforementioned methods to check if we could achieve better results.

The simulation results show that ML yields the best estimates in case the process and measurement noise characteristics are known. However, it requires a lot of time compared to the other methods. Using the EKF to estimate the unknown parameters, or using the EKF to estimate the states of the system and ML to estimate the unknown parameters yields the same estimates when the process and measurement noise characteristics are known. Finally, making the assumption that there is no measurement noise and thus using NLS or NML can provide a good estimate that can be used as initial condition when using the EKF.

For further research, we propose investigating the cases where the characteristics of one or both of the noises are unknown. In practice, both process and measurement noise characteristics are unknown. However, investigating the case where process or measurement noise characteristics is known may provide useful hindsights. Furthermore, we propose the implementation of Particle Filtering, a method widely used in nonlinear system filtering. Particle Filtering can be used instead of the EKF, in order to estimate the unknown states x_k. Finally, the application of these methods to real data is proposed for future work in order to test if the results match with the results from the synthetic data.

REFERENCES

Achilleos A., Loizides C., Stylianos T. and Mitsis G. D. (2012) Linear Dynamic Modelling and Bayesian Forecasting of Tumor Evolution *Proceedings of the 2012 IEEE 12th International Conference on Bioinformatics & Bioengineering*, pp. 671–676.

Achilleos A., Loizides C., Hadjiandreou M. M., Stylianos T. and Mitsis G. D. (2014) Multiprocess Dynamic Modeling of Tumor Evolution with Bayesian Tumor-Specific Predictions *Ann Biomed Eng. Kluwer Academic Publishers*, vol. 42, no. 5, pp. 1095–1111.

Anand P., Kunnumakara A. B., Sundaram C., Harikumar K., Tharakan S., Lai O. S., Sung B., and Aggarwal B. B. (2008) Cancer is a preventable disease that requires major lifestyle changes *Pharmaceutical Research*, vol. 25, no. 9, pp. 2097–2016.

Barbolosi D. and Iliadis A. (2001). Optimizing drug regimens in cancer chemotherapy: a simulation study using a PK-PD model. *Computers in Biology and Medicine*, vol. 31, pp. 157–172.

Benzekry S., Lamont C., Beheshti A., Tracz A., Ebos J. M., Hlatky L., and Hahnfeldt P. (2014) Classical Mathematical Models for Description and Prediction of Experimental Tumor Growth *PLOS Computational Biology*, vol. 10, no. 8, e1003800.

Charalampidis A. C. and Papavassilopoulos G. P. (2011) Computationally efficient Kalman filtering for a class of nonlinear systems *IEEE Trans. Autom. Control*, vol. 56, no. 3, pp. 483–491.

Charalampidis A. C., Pontikis K., Mitsis G. D., Dimitriadis G., Lampadiari V., Marmarelis V. Z., Armaganidis A., and Papavassilopoulos G. P. (2016) Calibration of a Microdialysis Sensor and Recursive Glucose Level Estimation in ICU Patients Using Kalman and Particle Filtering *Biomedical Signal Processing and Control*, vol. 27, pp. 155–163.

Dennis B. and Ponciano J. M. (2014). Density dependent state space model for population abundance data with unequal time intervals data with unequal time intervals *Ecology*, vol. 95, no.8, pp. 2069–2076.

Dua P. D. D. and Pistikopoulos E. (2008). Optimal Delivery of chemotherapeutic agents in cancer *Computers and Chemical Engineering*, vol. 32, pp. 99–107.

Hadjiandreou M. M. and Mitsis G. D. (2014). Mathematical Model of Tumor Growth, Drug-Resistance, Toxicity, and Optimal Therapy Design *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 2, pp. 415–425.

Maybeck P. S. (1982) *Stochastic Models, Estimation and Control Volume 2*. Academic Press.

Nguimkeu P. (2014) A Simple Selection Test Between the Gompertz and Logistics *Technological Forecasting and Social Change*, vol. 88, pp. 98-105.

Ribba B., Holford N. H., Magni P., Trocniz I., Gueorguieva I., Girard P., Sarr C., Elishmereni M., Kloft C., and Friberg L. E. (2014) A Review of Mixed-Effects Models of Tumor Growth and Effects of Anticancer Drug Treatment Used in Population Analysis *CPT: Pharmacometrics Systems Pharmacology*, vol.3, no.5, e113.

Schön T. B., Wills A., and Ninness B. (2011) System Identification of nonlinear state-space models *Automatica*, vol. 47, no. 1, pp. 39–49.

The Mathworks, Inc., 1994-2016. [Online] `fmincon` Retrieved 27 September 2016. Available: `http://www.mathworks.com/help/optim/ug/fmincon.html?requestedDomain=www.mathworks.com`.

World Health Organization, 2014. [Online] Cancer Fact sheet N297 Retrieved 27 September 2016. Available: `http://www.who.int/mediacentre/factsheets/fs297/en/`.

Wikipedia, the free encyclopedia, 2016 [Online] Treatment of Cancer Retrieved 27 September 2016. Available: `https://en.wikipedia.org/wiki/Treatment_of_cancer`.