JOURNAL OF COMPUTATIONAL BIOLOGY
Volume 25, Number 12, 2018

Pp. 1285–1300
DOI: 10.1089/cmb.2017.0130

Research Articles

Simultaneous Model Selection and Model Calibration for the Proliferation of Tumor and Normal Cells During In Vitro Chemotherapy Experiments

JOSÉ M.J. COSTA,[1,2] HELCIO R.B. ORLANDE,[2] VIVIANE O.F. LIONE,[3] ANTONIO G.F. LIMA,[3] TAYNÁ C.S. CARDOSO,[3] and LEONARDO A.B. VARÓN[2,4]

ABSTRACT

In vitro experiments were conducted in this work to analyze the proliferation of tumor (DU-145) and normal (macrophage RAW 264.7) cells under the influence of a chemotherapeutic drug (doxorubicin). Approximate Bayesian Computation (ABC) was used to select among four competing models to represent the number of cells and to estimate the model parameters, based on the experimental data. For one case, the selected model was validated in a replicated experiment, through the solution of a state estimation problem with a particle filter algorithm, thus demonstrating the robustness of the ABC procedure used in this work.

Keywords: approximate Bayesian computation, chemotherapy, DU-145 cells, RAW 264.7 cells, state estimation.

NOMENCLATURE

A	Holling's type 2 constant that controls the drug intake by cells
$d(\mathbf{z}, \mathbf{z}^*)$	distance function
K	support capacity of cells
$\mathbf{M}$	vector that indexes the models
N	number of particles (samples of the approximate posterior distribution)
$N(t)$	number of cells
P	number of populations (iterations)
$Q(t)$	mass of chemotherapeutic drug
$\mathbf{z}$	vector of measurements
$\mathbf{z}^*$	vector of estimated state variables
Greek	
α	cell growth rate
β	saturation rate
ε	user-prescribed tolerance

[1]Department of Statistics, Federal University of Amazonas—UFAM, Manaus, Brazil.
[2]Department of Mechanical Engineering, Federal University of Rio de Janeiro—UFRJ, Rio de Janeiro, Brazil.
[3]Laboratory of Pharmaceutical Bioassays, Faculty of Pharmacy, Federal University of Rio de Janeiro—UFRJ, Rio de Janeiro, Brazil.
[4]Department of Bioengineering, School of Engineering, University of Santiago de Cali, Cali, Colombia.

λ	decay rate of the chemotherapeutic drug
μ	rate of cell reduction due to the chemotherapy
$\pi(\cdot\,\vert\,\cdot)$	conditional probability
$\boldsymbol{\theta}$	vector of model parameters
Subscripts	
1, 2, 3, 4	designate the model
p	population number
Superscript	
i	particle number

INTRODUCTION

In accordance with the world health organization, neoplastic diseases are the cause of 13% of the deaths in the world, with an estimated number of victims of 13.1 million in 2030. As a major cause of deaths worldwide, cancer has been an important topic of research, involving basic and applied sciences, mainly aiming at early diagnostics and effective treatment of the disease (Sanga et al., 2006).

Mathematical models play an important role in various aspects related to cancer, which range from the modeling of biophysical and biochemistry phenomena at the cellular scale, to larger scale and multiple scale phenomena related to imaging techniques and to treatments such as chemotherapy, radiotherapy, hyperthermia, etc. (see, e.g., Spratt et al., 1996; Gatemby, 1998; Bellomo et al., 2003; Byrne, 2003; Preziozi, 2003; Alarcón et al., 2004; Araujo and McElwain, 2004; Byrne et al., 2006; Sanga et al., 2006; Mohammadi et al., 2008; Gatenby 2009; Pinho et al., 2011, 2013; and Rodrigues et al., 2012, 2016). However, more detailed models certainly involve a larger number of input parameters and might not be more accurate than simpler mathematical models with fewer parameters, for example, to represent macroscopic variables like the number of tumor and normal cells or the mass of a chemotherapy drug in a region of the body. Indeed, model parameters represent a large source of uncertainties for the dependent variables, since, in general, they exhibit large variability from individual to individual and even for the same individual under different physiological conditions. Therefore, the choice of a mathematical model to accurately represent the phenomena under analysis, as well as the identification of the model parameters, is a subject of great importance for the individualized detection and treatment of cancer.

The main objective of this article is to apply model selection and model calibration (estimation of the model parameters) for the number of tumor and normal cells during in vitro chemotherapy experiments. Due to the scarce number of measurements, their related uncertainties were not represented in terms of classical statistical distributions and an Approximate Bayesian Computation (ABC) technique (Del Moral et al, 2007; Sisson et al., 2007; Toni et al., 2009; Toni and Stumpf, 2010a; Del Moral et al., 2012) was used for the simultaneous model selection/calibration based on the experimental data. The algorithm of Toni et al. (2009), which was successfully applied by da Costa et al. (2017) for the simultaneous model selection/calibration with synthetic measurements of the number of tumor cells in simulated experiments, is used in this work with actual experimental data involving tumor (DU-145 prostate line) and normal (RAW 264.7 macrophage) cells treated with doxorubicin. The different mathematical models examined in this work are systems of ordinary differential equations for the dependent variables, given by the number of cells and the mass of chemotherapeutic drug.

The model selected with the algorithm of Toni et al. (2009) was further validated in a replicated experiment for one specific case, by solving a state estimation problem (Kalman, 1960; Sorenson, 1970; Maybeck, 1979; Arulampalam et al., 2002; and Kaipio and Somersalo, 2004). In state estimation problems, uncertainties in the mathematical model that represents the physical phenomena and uncertainties in the measurements are appropriately accounted for within the Bayesian framework of statistics, to obtain more accurate estimates of the state variables. Sequential Monte-Carlo techniques, also referred to as particle filters, are the most general for the solution of state estimation problems and can be applied to nonlinear models with non-Gaussian uncertainties (Liu and Chen, 1998; Carpenter et al., 1999; Doucet et al., 2000; Arulampalam et al., 2002; Andrieu et al., 2004; Kaipio and Somersalo, 2004; Ristic et al., 2004; Del Moral et al., 2006; Johansen and Doucet, 2008; and Orlande et al., 2012). In this work, the particle filter algorithm of Liu and West (2001) was used for the model validation; this algorithm was successfully applied for the simultaneous estimation of state variables and parameters in the modeling of tumor growth by Costa et al. (2015).

The experiment performed in this work is presented in the next section, which is followed by a description of the mathematical models that are proposed to represent the experimental data. The algorithm of Toni et al. (2009), which is used in this study for model selection and model calibration, is then presented and applied to our experimental data. The results obtained in this work for the model selection/calibration are presented and discussed. Furthermore, the model selected for a specific case is validated with an additional set of experimental data, by using the algorithm of Liu and West (2001) of the particle filter.

EXPERIMENTS

Cell culture

DU-145 prostate cancer cells and RAW 264.7 macrophage cells (ATCC—American Type Culture Collection, Manassas, VA) were purchased from the cell bank of the Federal University of Rio de Janeiro. The cell lines were grown in 75-cm^3culture flasks with RPMI (Roswell Park Memorial Institute medium; Sigma-Aldrich) and DMEM (Dulbecco's modified Eagle's medium; Sigma-Aldrich), respectively, supplemented with 10% fetal bovine serum (FBS; Life Technologies, Grand Island, NY), 4 mM glutamine, and 100 U/mL penicillin and 100 μg/mL streptomycin. The cells were maintained at 37°C in a 5% CO_2 humidified atmosphere. The medium was changed every 2 days until 90%–95% confluence of the culture. At confluence, the cells were passaged by detaching them with 0.025% trypsin diluted in Verséne solution-EDTA (ethylenediaminetetraacetic acid; NaCl 0.14 M, Na_2HPO_4 9 mM, KCl 3 mM, phenol red 0.02%, and EDTA 0.02%) and centrifuged at 1100 RPM for 10 minutes. After centrifugation, the sediments with the DU 145 cells and the RAW 264.7 cells were diluted in RPMI and DMEM, respectively, both containing 10% FBS. The cells were then counted in a hemocytometer and seeded into 96-well plates for cell proliferation assay.

Cell proliferation assay

The cell proliferation was observed by using the colorimetric method of MTT (bromide 3-[4,5-dimethylthiazole-2-il]-2,5-diphenyltetrazolium; Sigma-Aldrich). This assay detected viable cells through the conversion of MTT into formazan by cell mitochondrial enzyme succinate dehydrogenase. In brief, DU-145 and RAW 264.7 macrophages were seeded into 96-well plates at a concentration of 10^3 and 10^4 cells per well, respectively, in culture medium supplemented with 10% FBS for cell adhesion. After 3 hours, the culture medium was removed and the cells were treated with doxorubicin (10 μM), diluted in RPMI or DMEM (depending on the cell), and supplemented with 5% of FBS for the chemotherapy experiments.

The variation of the number of cells was measured by the addition of MTT (Sigma-Aldrich) solution (1 mg/mL). At each time, the MTT was removed, the formazan crystals solubilized with isopropanol and the absorbance measured in a spectrophotometer at 490 nm (Thermo Fisher). The measurements of the number of cells were made in triplicates; the mean and the standard deviations of each triplicate are reported in the article.

Table 1 summarizes the different experiments performed in this work. DU-145 (tumor) cells were used for experiments 1 and 2, whereas RAW 264.7 (macrophage) cells were used for experiments 3–5. No chemotherapeutic drug was used for experiments 1 and 3, but the interaction of the cells with doxorubicin was analyzed in experiments 2, 4, and 5. Experiment 5 was a repetition of experiment 4.

MATHEMATICAL MODELS

There is a large number of mathematical models available in literature for the modeling of tumor growth with different degrees of refinement, for the vascular and avascular stages, which are in general based on

Table 1. Experiments

Experiment	*Cell*	*Initial number of cells*	*Chemotherapy*
1	DU-145 (tumor)	10^3	No
2	DU-145 (tumor)	10^3	Yes
3	RAW 264.7 (macrophage)	10^4	No
4	RAW 264.7 (macrophage)	10^4	Yes
5	RAW 264.7 (macrophage)	10^4	Yes

systems of ordinary and/or partial differential equations (see, e.g., Spratt et al., 1996; Gatemby, 1998; Bellomo et al., 2003; Byrne, 2003; Preziozi, 2003; Alarcón et al., 2004; Araujo and McElwain, 2004; Byrne et al., 2006; Sanga et al., 2006; Mohammadi et al., 2008; Gatenby, 2009; Pinho et al., 2011, 2013; and Rodrigues et al., 2012, 2016). We note, however, that more refined models do not necessarily predict the dependent variables of interest more accurately, in special because a large number of uncertain input parameters is required for their solutions (Beck and Arnold, 1977; Mantzaris et al., 2004; Roose et al., 2007).

Four different models, based on ordinary differential equations, were used in this work to represent the time variation of the number of cells in the wells, namely: (1) Generalized Logistic model (Spratt et al., 1996; Tsoularis and Wallace, 2002); (2) Gompertz's model (Winsor, 1932; Spratt et al., 1996); (3) Exponential model (Spratt et al., 1996); and (4) Modified Gompertz's model. They are referred hereafter as models 1, 2, 3, and 4, respectively. Ordinary differential equation models were selected because of the nature of the experiments, where the measurements do not possess a local (spatial) resolution, but represent the average number of cells in each well at a specific time. Whenever the chemotherapy drug was used in the experiments, the pharmacodynamics model was considered in the form of a Holling's type 2 function (Holling, 1959a,b; Pinho et al., 2013; Costa et al., 2015; Rodrigues et al., 2016) and a first-order pharmacokinetics model was considered to govern the mass of drug in the well. The four models are given by:

Model 1. Generalized logistic:

$$\frac{dN_1(t)}{dt} = \frac{\alpha_1}{\gamma} N_1(t)\left(1 - \frac{N_1(t)}{K_1}\right)^{\gamma} - \frac{\mu_1 N_1(t) Q(t)}{a_1 + N_1(t)} \tag{1.a}$$

Model 2. Gompertz:

$$\frac{dN_2(t)}{dt} = \alpha_2 N_2(t) \ln\left(\frac{K_2}{N_2(t)}\right) - \frac{\mu_2 N_2(t) Q(t)}{a_2 + N_2(t)} \tag{1.b}$$

Model 3. Exponential:

$$\frac{dN_3(t)}{dt} = \alpha_3 N_3(t) - \frac{\mu_3 N_3(t) Q(t)}{a_3 + N_3(t)} \tag{1.c}$$

Modelo 4. Modified Gompertz:

$$\frac{dN_4(t)}{dt} = \alpha_4 N_4(t) \ln\left(\frac{K_4}{N_4(t)}\right) e^{-\beta t} - \frac{\mu_4 N_4(t) Q(t)}{a_4 + N_4(t)} \tag{1.d}$$

where

$$\frac{dQ(t)}{dt} = -\lambda Q(t) \tag{2}$$

and $N(t)$ and $Q(t)$ are the number of cells and mass of chemotherapy drug in the well, respectively, subscripts $i = 1, 2, 3, 4$ designate the model, α is the cell growth rate, K is the cell support capacity, μ is the cell reduction rate due to the chemotherapy, a is a Holling's type 2 constant that controls the drug intake by cells, β is the saturation rate, and λ is the decay rate of the chemotherapy drug.

While models 1, 2, and 3 have been classically used to represent the growth of tumors (Spratt et al., 1996), model 4 was proposed in this work to account for reductions in the growth rate of the number of cells as time increases. Note that model 4 reduces to model 2 for small values of β.

Each model was subjected to the initial conditions given by the number of cells shown in Table 1, depending on the case analyzed. The drug was administered as a single bolus at time $t = 0$ in the chemotherapy experiments (cases 2, 4, and 5), so that the initial condition for the mass of drug was $Q_0 = 5.4352$ mg.

MODEL SELECTION AND MODEL CALIBRATION

Parameters appearing in the mathematical formulation of physical problems can be estimated by several statistical techniques (Beck and Arnold, 1977; Ozisik and Orlande, 2000; Putter et al., 2002; Brown and Sethna, 2003; Kaipio and Somersalo, 2004; Timmer et al., 2004; and Sisson et al., 2007). However, techniques within the Bayesian framework, like the Markov Chain Monte Carlo method, recently became popular because of nowadays available computer power. In the Bayesian framework, prior knowledge about the parameters, modeled in the form of a statistical distribution, $\pi(\boldsymbol{\theta})$, is used together with the measured data coded as a likelihood function, $\pi(\mathbf{z}|\boldsymbol{\theta})$, for inference on the posterior distribution function given by $\pi(\boldsymbol{\theta}|\mathbf{z}) \propto \pi(\boldsymbol{\theta})\pi(\mathbf{z}|\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is the vector of parameters and $\mathbf{z}$ is the vector of measurements (Kaipio and Somersalo, 2004). The possibility of taking into account the information about the parameters before the measurements are available is very attractive for the solution of inverse problems, because it provides regularization to the ill-posed character and allows for uncertainties in the judged ''known'' parameters to be accounted for. On the other hand, the prior distribution introduces a bias on the solution of the inverse problem.

The experimental uncertainties cannot be appropriately modeled for several practical cases and, thus, the likelihood function might not be represented in terms of analytical statistical distributions. The so-called ABC can be utilized for such cases, as well as for others, where the computation of the likelihood function becomes intractable (Beaumont et al., 2002; Marjoram et al., 2003; Toni et al., 2009; Wegmann et al., 2009; Toni and Stumpf, 2010a). The objective of ABC is to approximate the posterior distribution $\pi(\boldsymbol{\theta}|\mathbf{z})$ by $P(\boldsymbol{\theta}|d(\mathbf{z}, \mathbf{z}^*) < \varepsilon)$, where $d(\mathbf{z}, \mathbf{z}^*)$ is a distance function between the measurements $\mathbf{z}$ and the estimated observable variables $\mathbf{z}^*$ (solution of the forward problem with $\boldsymbol{\theta}$), while ε is a user-prescribed tolerance. Therefore, difficulties in the successful implementation of ABC include the specification of the distance function, as well as of the tolerance ε (Liepe et al., 2014).

ABC is used in this work for simultaneous model selection and parameter estimation because the available measured data are too scarce to be appropriately represented in terms of an analytical statistical distribution. Related works can be found in the literature, such as Toni et al. (2009), Toni and Stumpf (2010b), and Drovandi and Pettitt (2011a,b). The algorithm of Toni et al. (2009), which is an extension of the Sequential Monte Carlo algorithm of Sisson et al. (2007), is used for the selection of the four models given by Equations (1.a–d), as applied to the experimental data described previously. Therefore, the objective is to approximately obtain the combined posterior distribution (Toni et al., 2009):

$$\pi(\boldsymbol{\theta}, \mathbf{M}|\mathbf{z}) \tag{3}$$

where $\mathbf{M}$ is the vector that indexes the models that are considered in the analysis. The ABC algorithm of Toni et al. (2009) is presented in Table 2.

RESULTS AND DISCUSSIONS

For the application of the algorithm of Toni et al. (2009) to the measurements of the number of cells from the experiments presented in Table 1, the distance function was specified as

$$d(\mathbf{z}, \mathbf{z}^*) = \sum_{i=1}^{N} [\mathbf{z}^*(i) - \mathbf{z}]^T [\mathbf{z}^*(i) - \mathbf{z}] \tag{4}$$

where $\mathbf{z}$ and $\mathbf{z}^*(i)$ are the vectors of measurements and of the solution of the forward problem with particle i, respectively. A particle represents a sample of the approximate posterior distribution for the model and its parameters. Based on numerical experiments with simulated measurements, the number of particles was set to $N = 1000$ (da Costa et al., 2017).

The tolerance for $d(\mathbf{z}, \mathbf{z}^*)$ (see steps 1 and 5 in Table 2) was started with a large value to avoid too many rejections of particles in the first populations (Toni et al., 2009; Toni and Stumpf, 2010b; Liepe et al., 2014; da Costa et al., 2017), specially because priors with very large variances were used for the parameters. The tolerance was then gradually reduced to a value based on Morozov's Discrepancy Principle (Morozov, 1966) as the populations advanced. The Discrepancy Principle has been commonly used to select Tikhonov's regularization parameter, as well as the number of iterations in Alifanov's iterative

Table 2. Approximate Bayesian Computation Algorithm

1. Define the tolerances $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_P$ for each of the populations (iterations) used for selecting the model and its parameters. Also, specify the distance function $d(\mathbf{z}, \mathbf{z}^*)$ that substitutes the likelihood function. Set the population indicator $p=0$.
2. Set the particle indicator $i=1$, where each particle represents, at each iteration, a model and its parameters.
3. Sample the model M^* from the prior distribution for the models $\pi(\mathbf{M})$. If $p=0$, sample the candidate parameters $\boldsymbol{\theta}^{**}$ from the prior distribution for the parameters of model M^*, that is, $\pi(\boldsymbol{\theta}(M^*))$. Else, sample $\boldsymbol{\theta}^*$ from the previous population $\boldsymbol{\theta}(M^*)^i_{p-1}$ with weights $w(M^*)^i_{p-1}$ and perturb this particle to obtain $\boldsymbol{\theta}^{**} \approx K_p(\boldsymbol{\theta}^*, \boldsymbol{\theta}^{**})$, where K_p is a perturbation kernel.
4. If $\pi(\boldsymbol{\theta}^{**})=0$, return to step 3. Else, simulate from the forward problem (operator f) a candidate set of observable variables with model M^* and parameters $\boldsymbol{\theta}^{**}$, that is, $\mathbf{z}^* \approx f(\mathbf{z}|\boldsymbol{\theta}^{**}, M^*)$.
5. If $d(\mathbf{z}, \mathbf{z}^*) > \varepsilon_p$ return to step 3. Otherwise, set $M^i_p = M^*$, add $\boldsymbol{\theta}^{**}$ to the population of particles $\boldsymbol{\theta}(M^*)^i_p$ and calculate its weight as $w(M^*)^i_p = \begin{cases} 1, & \text{if } p=0 \\ \dfrac{\pi\left(\boldsymbol{\theta}(M^*)^i_p\right)}{\sum_{j=1}^{N} w(M^*)^j_{p-1} K_p\left(\boldsymbol{\theta}(M^*)^j_{p-1}, \boldsymbol{\theta}(M^*)^i_p\right)}, & \text{if } p>0 \end{cases}$
6. If $i < N$, where N is the number of particles, set $i=i+1$ and go to step 3.
7. Normalize the weights.
8. If $p < P$, where P is the number of populations (iterations), set $p=p+1$ and go to step 2. Otherwise, terminate the iterations.

Toni et al. (2009).

regularization method, for the solution of inverse ill-posed problems (Ozisik and Orlande, 2000; Orlande et al., 2011). The tolerance based on the Discrepancy Principle is obtained by making

$$[\mathbf{z}^*(i) - \mathbf{z}]^T [\mathbf{z}^*(i) - \mathbf{z}] = W \tag{5}$$

In Equation (4), where W is the trace of the covariance matrix of the measurements. The tolerance for the first iteration was set as $\varepsilon_1 = 3 \times 10^{13}$ cell^2. The algorithm was then stopped either if the tolerance based on the Discrepancy Principle was satisfied or if no new particles were accepted (a clear indication that the representation of the experimental data could not be improved with the competing models).

Each of the four models given by Equations (1.a–d) was initially assumed as equally probable for the analysis of the experiments described in Table 1. Similarly, uniform priors were prescribed for each parameter. Table 3 shows the minimum and maximum values specified for these uniform distributions, which differ by several orders of magnitude for most of the parameters. Therefore, from a practical point of view, these uniform priors can be considered as noninformative despite their finite variances. The perturbation kernel K_p (see step 3 in Table 2) was given in terms of a random walk process with uniform distribution $U(-\sigma, \sigma)$ for each parameter, where σ is half of the interval of the uniform prior given by Table 3.

The results obtained for experiments 1 and 2 (Table 1) are first analyzed. These experiments dealt with the DU-145 (tumor) cells without and with chemotherapy, respectively. Figure 1 presents the number of particles selected for the models, as the populations advanced in the algorithm of Toni et al. (2009), for experiment 1. We notice in this figure that the number of particles selected for model 1 steadily decreased after population 12 and then disappeared at population 23. As the number of particles selected for model 1

Table 3. Uniform Priors for the Parameters

Parameters	*Distribution*	*Unit*
$\alpha_i, \ i=1,2,3,4$	$U(10^{-3}, \ 5)$	h^{-1}
$K_i, \ i=1,2,4$	$U(10^3, \ 10^7)$	*cell*
γ	$U(1, \ 3)$	$cell^{-1}$
$\mu_i, \ i=1,2,3,4$	$U(10^{-3}, \ 10^3)$	$mg^{-1}cell^{-1}$
$a_i, \ i=1,2,3,4$	$U(10^{-3}, \ 10)$	*cell*
λ	$U(10^{-3}, \ 10^{-2})$	h^{-1}
β	$U(10^{-3}, \ 5)$	h^{-1}

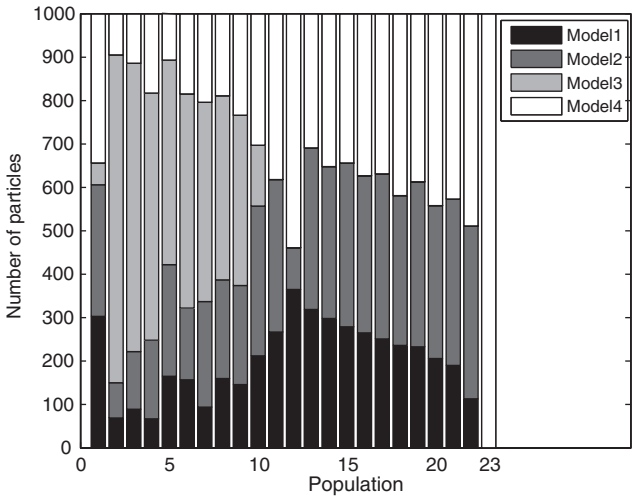

FIG. 1. Number of particles selected for the models in each population. Experiment 1 (DU-145 cells without chemotherapy).

decreased, the number of particles selected for model 4 increased after population 13. Models 2 and 4 might closely compete for the representation of the experimental data, such as observed in Figure 1, since model 4 reduces to model 2 for small values of β. On the other hand, since the tolerance for the distance function decreased as the populations advanced, model 4 provided the best representation for the measurements of experiment 1, at population 23.

A comparison of the measurements and of the number of cells obtained with model 4 and its estimated parameters are presented in Figure 2A, B, at the first and final populations, respectively, for experiment 1. These figures show that the measurements were not correctly predicted in population 1. On the other hand, as the populations advanced and the tolerances were reduced, the best model was selected and calibrated for the most accurate prediction of the measurements. The median of the distance function of the particles selected for model 4 was reduced by two orders of magnitude, from the first to the last populations, for experiment 1. Figure 2A, and B show that, without the use of the chemotherapeutic drug in experiment 1, the number of tumor cells increased very fast from the initial condition of 10^3 cells to about 3.5×10^5 cells in 60 hours.

The number of particles selected in each population is presented in Figure 3 for experiment 2, which also involved DU-145 cells, but under the effects of doxorubicin. Model 2, which is actually a special case of model 4 for small β, was selected as the one that better represents the measurements in this experiment. Figure 3 shows that both models 3 and 4 were not selected from population 26 onward. At population 26, model 2 exhibited the largest number of selected particles, except for few selections of model 1. As the tolerances were further reduced, there was a clear competition between models 1 and 2 to represent the experimental data. However, the particles of model 2 were the most selected after population 26 and, at population 36, no other model was selected.

A comparison of measurements and number of cells predicted with model 2 at the final population is shown in Figure 4. Due to the large uncertainties in the measurements for this experiment, a reduction of only 4.5-fold was obtained for the median of the distance function of particles selected for model 2, from the first to the last populations. The number of cells predicted with model 2 follows the same trend of the experimental data, as shown in Figure 4. A comparison of Figures 2 and 4 reveals that the chemotherapeutic drug was capable of limiting the number of tumor cells. On the other hand, the tumor cells have still proliferated from their initial condition, thus demonstrating that the amount of drug used in this case ($10\,\mu$M bolus at time $t=0$) was not sufficient for their eradication.

Table 4 shows the mean and 1% and 99% quantiles of the estimated parameters, for models 4 and 2 in experiments 1 and 2, respectively. It is noticed in this table that, with the exception of the parameters related to the pharmacodynamics (μ and a), all the others are estimated with quite small uncertainties.

Experiments 3 and 4, which involved RAW 264.7 (macrophage) cells, without and with chemotherapy, respectively, are now analyzed. Such as for the tumor cells (Fig. 1), model 4 was selected to represent the measurements of experiment 3 for macrophage cells without chemotherapy, as shown in Figure 5. This figure shows that the number of particles selected for model 4 gradually increased from population 14 onward. At populations 22 and 23, model 4 was the only one selected. Although the algorithm had already

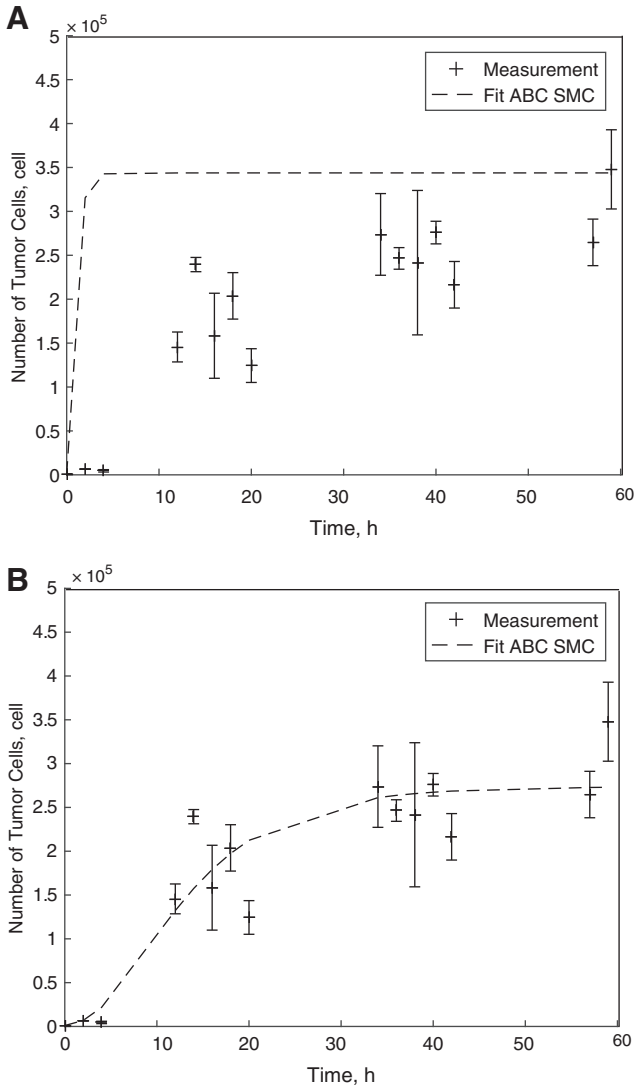

FIG. 2. Measurements and number of cells estimated with model 4 for experiment 1 (DU-145 cells without chemotherapy). **(A)** Population 1, and **(B)** Final population. ABC, Approximate Bayesian Computation; SMC, Sequential Monte Carlo.

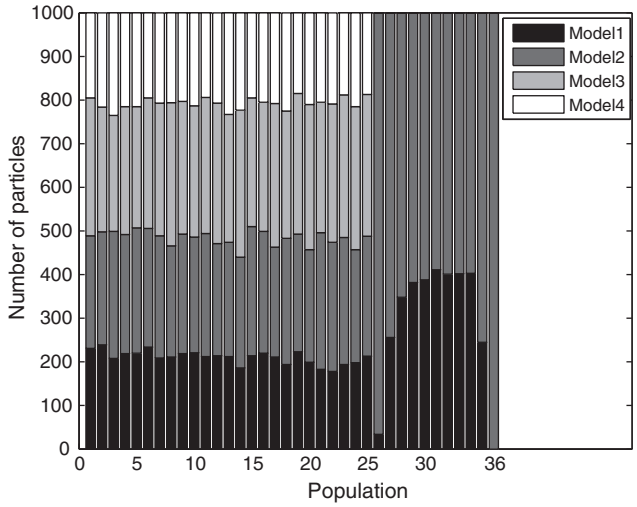

FIG. 3. Number of particles selected for the models in each population. Experiment 2 (DU-145 cells with chemotherapy).

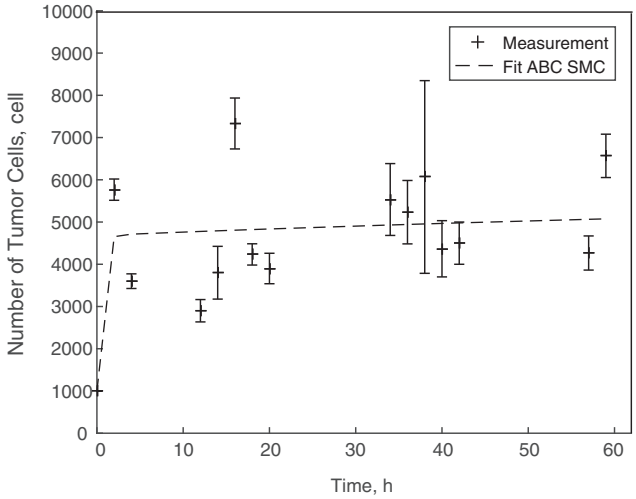

FIG. 4. Measurements and number of cells estimated with model 2 for experiment 2 (DU-145 cells with chemotherapy) at the final population.

selected model 4 among the competing models at population 22, one more population was needed to improve the parameter estimation for a more accurate representation of the experimental data. In fact, the agreement between the measurements and the estimated number of cells obtained with this model and its estimated parameters is quite good, as shown in Figure 6. Such as for experiment 1, the median of the distance function of the particles selected for model 4 was reduced by two orders of magnitude from the first to the last populations in experiment 3. Similar results were obtained for experiment 4 (macrophage cells with chemotherapy), where model 4 was also selected as the one that better represents the experimental data (Fig. 7). The comparison between measurements and estimated number of cells is presented in Figure 8 for experiment 4.

It is interesting to compare Figures 4 and 8, for experiment 2 (tumor cells with chemotherapy) and for experiment 4 (macrophage cells with chemotherapy), respectively. Although doxorubicin, with the dosage used in the experiments of a single 10 μM bolus at time $t=0$, was not capable of completely killing the tumor cells (Fig. 4), it substantially affected the macrophage cells (Fig. 8). After a sudden proliferation, the macrophage cells were practically reduced to the initial condition (10^4 cells) after 60 hours. This reduction of the number of macrophage cells was not observed in experiment 3 without chemotherapy, as shown in Fig. 6.

The estimated parameters for model 4 in experiments 3 and 4 are presented in Table 5, in terms of their mean, as well as their 1% and 99% quantiles. Model 4 was also selected for experiment 5, which is a repetition of experiment 4. The parameters estimated for experiment 5 are also presented in Table 5. Except for the support capacity, K, and the rate of cell reduction due to the chemotherapy, μ, the mean estimated for the parameters fall within the credibility intervals of the replicated experiment. Differences in the estimated support capacities and rates of cell reduction for experiments 4 and 5 can be associated to linear dependence among parameters, in special, α, K, and β.

Table 4. Estimated Parameters for Experiments 1 and 2 at the Final Population

Experiment	*Model*	*Parameter*	*Quantile 1%*	*Mean*	*Quantile 99%*
1	4	α_4, h^{-1}	0.16	0.17	0.18
		K_4, *cell*	2.78×10^5	2.80×10^5	2.84×10^5
		β_4	0.10	0.12	0.13
2	2	α_2, h^{-1}	2.63	3.47	4.25
		K_2, *cell*	5.4×10^3	5.8×10^3	6.3×10^3
		μ_2, $mg^{-1}cell^{-1}$	3.6×10^2	6.3×10^2	8.4×10^2
		a_2, *cell*	2.0	5.2	8.6
		λ, h^{-1}	0.004	0.006	0.009

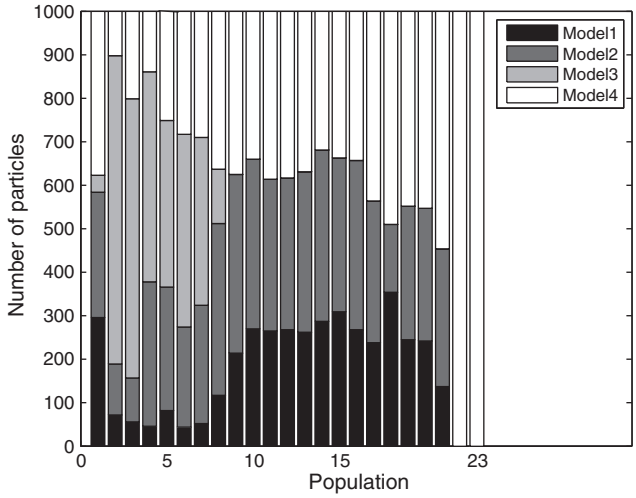

FIG. 5. Number of particles selected for the models in each population. Experiment 3 (RAW 264.7 cells without chemotherapy).

After the application of Toni et al.'s algorithm (2009) for the model selection and model calibration described previously, the particle filter algorithm of Liu and West (2001) was used for the solution of state estimation problems in the replicated experiments (experiments 4 and 5). State estimation problems are also referred to as nonstationary inverse problems (Kaipio and Somersalo, 2004). In this kind of problems, the available measured data are used together with prior knowledge about the physical phenomena and the measuring devices, to sequentially produce estimates of the desired dynamic variables. This is accomplished in such a manner that the error is minimized statistically (Maybeck, 1979; Winkler, 2003; Kaipio and Somersalo, 2004; Orlande et al., 2012). In state estimation problems, uncertainties in the mathematical model and in the measurements are taken into account for the better prediction of the state variables, which are given in the present work by the number of cells and the mass of chemotherapy drug. The algorithm of Liu and West (2001) is an extension of the Auxiliary Sampling Importance Resampling algorithm of the particle filter (Pitt and Shephard, 1999; Arulampalam et al., 2002; Doucet et al., 2001; Ristic et al., 2004) for simultaneous estimation of state variables and model parameters, where the uncertainties in the parameters are modeled in terms of a combination of Gaussian kernels (Liu and West, 2001).

Details of Liu and West's algorithm (2001) are avoided in this study for the sake of brevity, but they can be readily found in Doucet et al. (2001), Da Silva et al. (2014), Orlande et al. (2012), Costa et al. (2015),

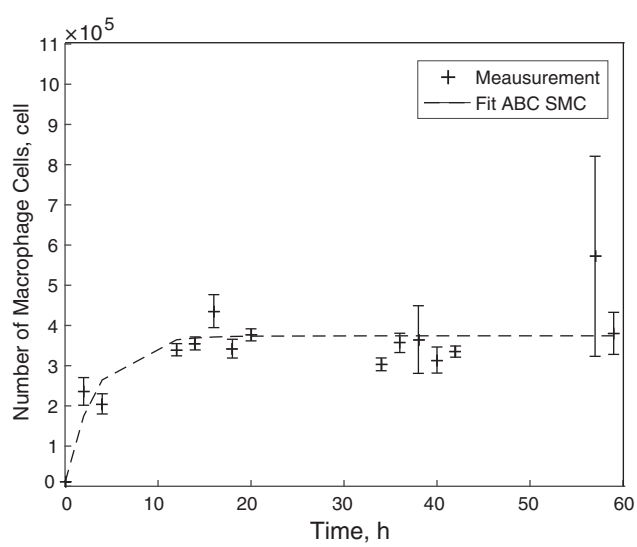

FIG. 6. Measurements and number of cells estimated with model 4 for experiment 3 (RAW 264.7 cells without chemotherapy) at the final population.

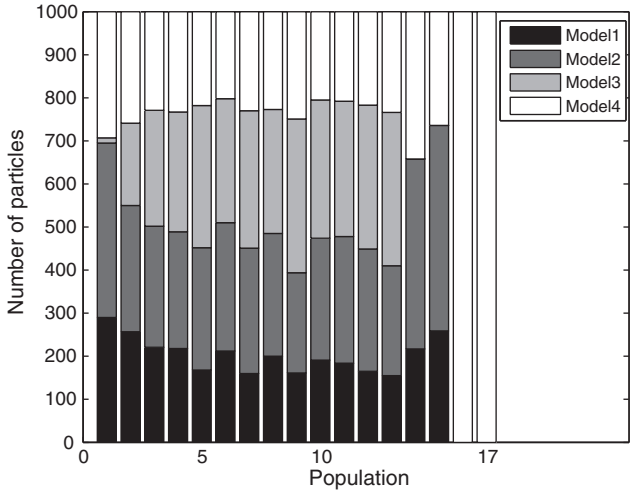

FIG. 7. Number of particles selected for the models in each population. Experiment 4 (RAW 264.7 cells with chemotherapy).

and Váron et al. (2015). Costa et al. (2015) applied Liu and West's algorithm (2001) for the simultaneous estimation of state variables and parameters in ordinary differential equation models of tumor growth, which included as state variables the number of tumor, normal and angiogenic cells, as well as the mass of chemotherapy drug in the body.

The solution of the state estimation problem is illustrated by using the measurements of experiment 4 and by initially taking the parameter values as those estimated in experiment 5, for model 4. Uncertainties in the mathematical model were assumed as additive, uncorrelated, Gaussian, with zero mean and a standard deviation of 10% of the value of the state variables. Uncertainties in the model parameters were also assumed as additive, uncorrelated, Gaussian and with zero mean, but with a standard deviation of 15% of the parameter values. The particle filter algorithm of Liu and West (2001) was applied with 2000 particles (Costa et al., 2015).

Figures 9 and 10 present the transient variations of the number of cells and of the mass of the chemotherapy drug, estimated with Liu and West's algorithm (2001). Figure 9 shows that the particle filter is capable of very accurately tracking the measured data by using model 4, which was selected above with the ABC algorithm. The mass of chemotherapeutic drug, for which there are no measurements available, can also be estimated with small uncertainties as presented in Figure 10. This figure shows the reduction of the

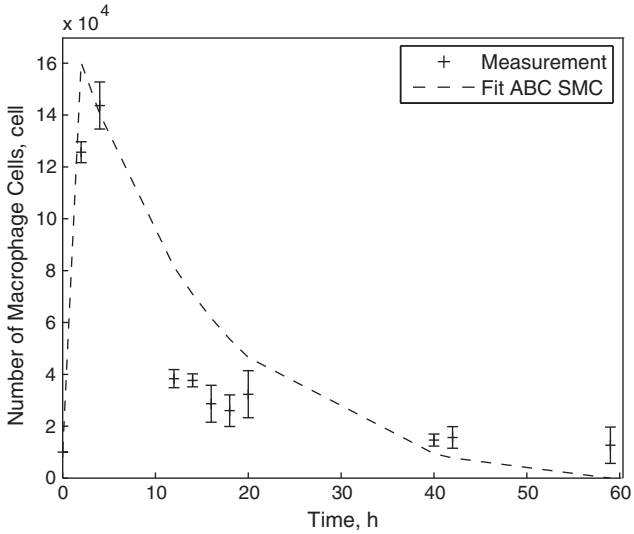

FIG. 8. Measurements and number of cells estimated with model 4 for experiment 4 (RAW 264.7 cells with chemotherapy) at the final population.

Table 5. Estimated Parameters for Experiments 3, 4, and 5 at the Final Population

Experiment	*Model*	*Parameter*	*Quantile 1%*	*Mean*	*Quantile 99%*
3	4	α_4 , h^{-1}	0.39	0.42	0.44
		K_4, *cell*	3.3×10^5	3.6×10^5	3.8×10^5
		β_4	0.45	0.50	0.53
4	4	α_4 , h^{-1}	1.5	2.4	3.4
		K_4, *cell*	1.2×10^6	4.7×10^6	8.2×10^6
		μ_4, $mg^{-1}cell^{-1}$	1.5×10^3	2.3×10^3	3.1×10^3
		a_4, *cell*	1.8	5.5	8.8
		λ, h^{-1}	0.03	0.06	0.09
		β_4	2.3	3.6	4.7
5	4	α_4 , h^{-1}	1.5	3.3	4.4
		K_4, *cell*	1.8×10^5	5.6×10^5	9.1×10^5
		μ_4, $mg^{-1}cell^{-1}$	433.4	704.8	897.5
		a_4, *cell*	1.2	5.1	9.1
		λ, h^{-1}	0.01	0.03	0.05
		β_4	1.3	3.0	4.7

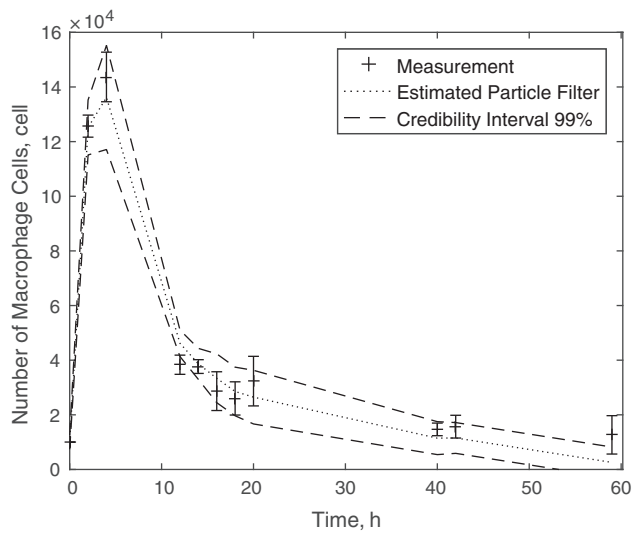

FIG. 9. Number of cells estimated with the particle filter algorithm of Liu and West (2001) for experiment 4.

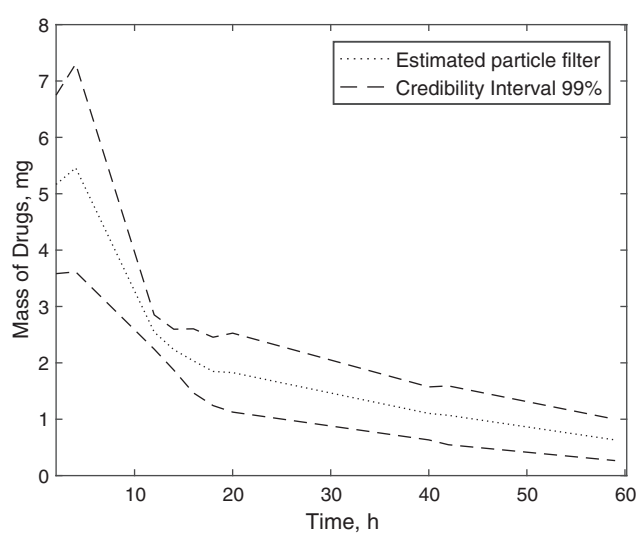

FIG. 10. Mass of chemotherapy drug estimated with the particle filter algorithm of Liu and West (2001) for experiment 4.

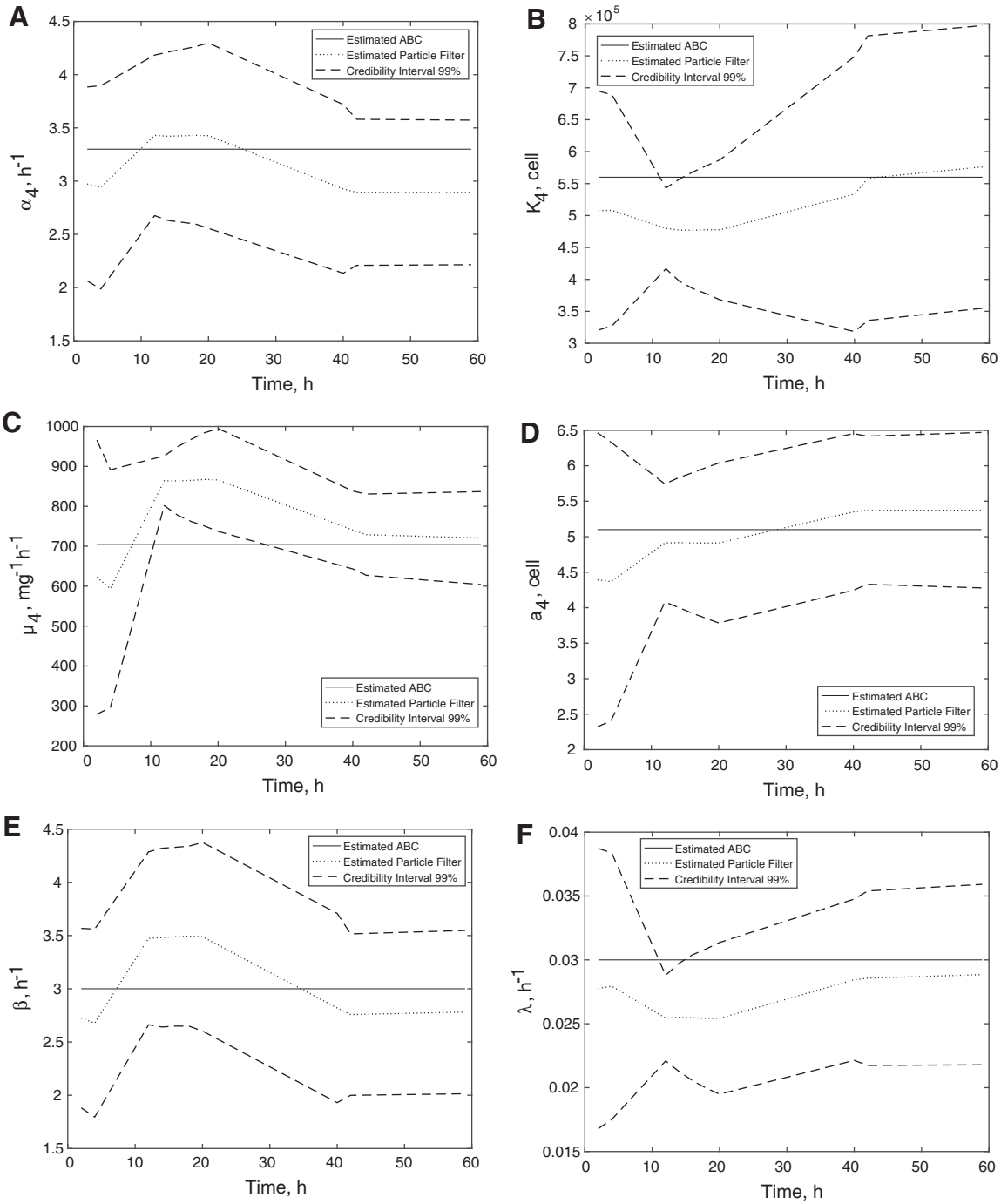

FIG. 11. Parameters estimated with the particle filter algorithm of Liu and West (2001) for experiment 4: **(A)** cell growth rate, **(B)** support capacity, **(C)** rate of cell reduction, **(D)** Holling's type 2 constant, **(E)** saturation rate, and **(F)** decay rate of the chemotherapy drug.

mass of chemotherapy drug in the well, as it is consumed by the cells during the experiment. Figure 11A–E presents the time evolution of the model parameters estimated with the particle filter algorithm of Liu and West (2001) for experiment 4. The parameter values estimated with the ABC algorithm of Toni et al. (2009; Table 5) are also presented in these figures. Figure 11A–E shows that the values estimated with the ABC algorithm of Toni et al. (2009) are all inside the credibility intervals predicted with the particle filter algorithm of Liu and West (2001).

CONCLUSIONS

In this article, we presented the application of an ABC algorithm for the simultaneous model selection and parameter estimation in chemotherapy in vitro experiments. The experiments dealt with tumor (DU-145) and normal (macrophage RAW 264.7) cells under the influence of doxorubicin. In each experiment, the time evolution of the number of cells was measured. Four models based on ordinary differential equations were proposed to represent the number of cells as a function of time, where the pharmacodynamics was considered in the form of a Holling's type 2 function and the pharmacokinetics as a first-order model. The ABC algorithm of Toni et al. (2009) used in this work was capable of selecting the model and estimating the parameter values that most accurately represent the measurements, in experiments with and without chemotherapy for the two cell lines. Such was the case, despite that all models were initially considered as equally probable and the prior information for the parameters were considered as uniform distributions with large variances.

While the Gompertz model was selected for the experiment with the tumor cells under chemotherapy, the Modified Gompertz model was selected for the other experiments. With the chemotherapy dosage used in the experiments, of one single 10 μM bolus of doxorubicin at the initial time, the tumor cells have still proliferated from their initial condition, although their support capacity was smaller than that for the case without chemotherapy. On the other hand, the chemotherapy drug substantially affected the macrophage cells, which were reduced to their initial condition at the end of the experiment. This reduction of the number of macrophage cells was not observed in the experiments without chemotherapy.

A state estimation problem was solved with the particle filter algorithm of Liu and West (2001) as a validation of the approach used in this work for model selection and parameter estimation. Indeed, in a replicated experiment for the case with macrophage cells under chemotherapy, the particle filter was capable of very accurately tracking the measured data, by using the model selected with the ABC algorithm used in this work.

ACKNOWLEDGMENTS

This work has been financed by FAPERJ, CAPES, and CNPq. The scholarship provided by FAPEAM for J.M.J. Costa is also greatly appreciated.

AUTHOR DISCLOSURE STATEMENT

The authors declare there are no relationships with any people or organizations that could inappropriately influence (bias) this work.

REFERENCES

Alarcón, T., Byrne, H.M., and Maini, P.K. 2004. Towards whole-organ modelling of tumour growth. *Prog. Biophys. Mol. Biol*. 85, 451–472.

Andrieu, C., Doucet, A., Singh, S.S. and Tadic, V.B., 2004. Particle methods for change detection, system identification, and control. Proceedings of the IEEE, 92(3), pp. 423–438.

Araujo, R.P., and McElwain, D.L.S. 2004. A history of the study of solid tumor growth: The contribution of mathematical modelling. *Bull. Math. Biol.* 66, 1039–1091.

Arulampalam, M.S., Maskell, S., Gordon, N., et al. 2002. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Trans. Signal Proc*. 50, 174–188.

Beaumont, M.A., Zhang, W., and Balding, D.J. 2002. Approximate Bayesian computation in population genetics. *Genetics*. 162, 2025–2035.

Beck, J.V. and Arnold, K.J., 1977. Parameter estimation in engineering and science. Wiley Interscience, New York.

Bellomo, N., De Angelis, E., and Preziosi, L. 2003. Multiscale modeling and mathematical problems related to tumor evolution and medical therapy. *J. Theor. Med*. 5, 111–136.

Brown, K.S. and Sethna, J.P., 2003. Statistical mechanical approaches to models with many poorly known parameters. Physical Review E, 68(2), p. 021904.

Byrne, H.M., 2003. Modelling avascular tumour growth. *In Cancer modelling and simulation*, chapter 4, (pp. 95–140). CRC Press.

Byrne, H.M., Alarcon, T., Owen, M.R., et al. 2006. Modelling aspects of cancer dynamics: A review. *Philos. Trans. A Math. Phys. Eng. Sci*. 364, 1563–1578.

Carpenter, J., Clifford, P. and Fearnhead, P., 1999. Improved particle filter for nonlinear problems. IEE Proceedings-Radar, Sonar and Navigation, 146(1), pp. 2–7.

Costa, J.M.J., Orlande, H.R., Velho, H.F., et al. 2015. Estimation of tumor size evolution using particle filters. *J. Comput. Biol*. 22, 1–17.

Costa, J.M.J., Orlande, H.R.B., and da Betencurte, W. 2017. Model selection and parameter estimation in tumor growth models using approximate Bayesian computation-ABC. *Comput. Appl. Math*. 37, 2795–2815.

Da Silva, W.B., Orlande, H.R.B. Colaço, J.M., et al. 2014. Application of particle filters to regional-scale wildfire spread. *High Temp. High Press*. 43, 415–440.

Del Moral, P., Doucet, A. and Jasra, A., 2006. Sequential Monte Carlo samplers. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 68(3), pp. 411–436.

Del Moral P, Doucet, A., Jasra A., 2007, Sequential Monte Carlo for Bayesian Computation. *Bayesian Stat,* 8, 1–34.

Del Moral, P., Doucet, A. and Jasra, A., 2012. An adaptive sequential Monte Carlo method for Approximate Bayesian Computation. Statistics and Computing, 22(5), pp. 1009–1020.

Doucet, A., Godsill, S. and Andrieu, C., 2000. On sequential Monte Carlo sampling methods for Bayesian filtering. Statistics and Computing, 10(3), pp. 197–208.

Doucet, A., Freitas, N., and Gordon, N. 2001. *Sequential Monte Carlo Methods in Practice*. Springer, New York, NY.

Drovandi, C.C., and Pettitt, A.N. 2011a. Estimation of parameters for macroparasite population evolution using approximate Bayesian computation. *Biometrics*. 67, 225–233.

Drovandi, Christopher C. and Pettitt, Anthony N. 2011. Using Approximate Bayesian Computation to Estimate Transmission Rates of Nosocomial Pathogens, Statistical Communications in Infectious Diseases: Vol. 3: Iss. 1, Article 2. DOI: 10.2202/1948-4690.1025

Gatenby, R.A., 1998. Mathematical models of tumor-host interactions. The Cancer Journal, 11(6), pp. 289–293.

Gatenby, R.A. 2009. A change of strategy in the war on cancer. *Nature*. 459, 508–509.

Holling, C.S. 1959a. The components of predation as revealed by a study of small-mammal predation of the European Pine Sawfly. *Can. Entomol*. 91, 293–320.

Holling, C.S. 1959b. Some characteristics of simple types of predation and parasitism. *Can. Entomol*. 91, 385–398.

Johansen, A.M. and Doucet, A., 2008. A note on auxiliary particle filters. Statistics & Probability Letters, 78(12), pp. 1498–1504.

Kaipio, J.P., and Somersalo, E. 2004. *Computational and Statistical Methods for Inverse Problems*. Springer, Berlin, Heidelberg.

Kalman, R.E. 1960. A new approach to linear filtering and prediction problems. *Trans. ASME Series D J. Basic Eng*. 82, 35–45.

Liepe, J., Kirk, P., Filippi, S., et al. 2014. A framework for parameter estimation and model selection from experimental data in systems biology using approximate Bayesian computation. *Nat. Protoc*. 9, 439–456.

Liu, J.S. and Chen, R., 1998. Sequential Monte Carlo methods for dynamic systems. Journal of the American Statistical Association, 93(443), pp. 1032–1044.

Liu, J. and West, M., 2001. Combined parameter and state estimation in simulation-based filtering. In Sequential Monte Carlo methods in practice (pp. 197–223). Springer, New York, NY.

Mantzaris, N.V., Webb, S. and Othmer, H.G., 2004. Mathematical modeling of tumor-induced angiogenesis. Journal of Mathematical Biology, 49(2), pp. 111–187.

Marjoram, P., Molitor, J., Plagnol, V., et al. 2003. Markov chain Monte Carlo without likelihoods. *Proc. Natl. Acad. Sci. USA*. 100, 15324–15328.

Maybeck, P.S. 1979. *Stochastic Models, Estimation, and Control*. Academic Press, New York.

Mohammadi, B., Haghpanah, V., and Larijani, B. 2008. A stochastic model of tumor angiogenesis. *Comput. Biol. Med*. 38, 1007–1011.

Morozov, V.A. 1966. On the solution of functional equations by the method of regularization. *Soviet Math. Dokl*. 1, 414–417.

Orlande, H.R., Fudym, O., Maillet, D. and Cotta, R.M. eds., 2011. Thermal measurements and inverse techniques. CRC Press., Boca Raton, Floride.

Orlande, H.R.B. and Ozisik, M.N., 2000. Inverse heat transfer: Fundamentals and applications. Taylor & Francis, New York.

Orlande, H.R.B., Colaço, J.M., Dulikravich, G.S., et al. 2012. State estimation problems in heat transfer. *Int. J. Uncertain. Quantif*. 2, 239–258.

Pinho, S.T., Rodrigues, D., and Mancera, P.F. 2011. A mathematical model of chemotherapy response to tumour growth. *Can. App. Math*. 19, 369–384.

Pinho, S.T.R., Bacelar, F.S., Andrade, R.F.S., et al. 2013. A mathematical model for the effect of anti-angiogenic therapy in the treatment of cancer tumors by chemotherapy. *Nonlinear Anal. Real World Appl*. 14, 815–828.

Pitt, M.K., and Shephard, N. 1999. Filtering via simulation: auxiliary particle filters. *J. Am. Stat. Assoc*. 94, 590–599.

Preziozi, L. 2003. *Cancer Modelling and Simulation*. Chapman and Hall, New York.

Putter, H., Heisterkamp, S.H., Lange, J.M.A. and De Wolf, F., 2002. A Bayesian approach to parameter estimation in HIV dynamical models. Statistics in Medicine, 21(15), pp. 2199–2214.

Ristic, B., Arulampalam, M.S., and Gordon, N. 2004. *Beyond the Kalman Filter. Particle Filters for Tracking Applications*. Artech House, Boston, MA.

Rodrigues, D., Pinho, S.T., and Mancera, P.F. 2012. Um modelo matemático em quimioterapia [A mathematical model for chemotherapy]. *TEMA Tend. Mat. Apl. Comput*. 13, 1–12.

Rodrigues, D.S., Mancera, P.F.A., and Pinho, S.T.R. 2016. Understanding the antiangiogenic effect of metronomic chemotherapy through a simple mathematical model. *Physica A Stat. Mech. Appl*. 464, 251–266.

Roose, T., Chapman, S.J., and Maini, P.K. 2007. Mathematical models of avascular tumor growth. *Siam Rev*. 49, 179–208.

Sanga, S., Sinek, J., and Frieboes, H. 2006. Mathematical modeling of cancer progression and response to chemotherapy. *Expert Rev. Anticancer Ther*. 6, 1361–1376.

Sisson, S.A., Fan, Y., and Tanaka, M.M. 2007. Sequential Monte Carlo without likelihoods. *Proc. Natl. Acad. Sci. USA*. 104,1760–1765.

Sorenson, H.W. 1970. Least-squares estimation: From Gauss to Kalman. *IEEE Spectrum*. 7, 63–68.

Spratt, J.S., Meyer, J.S., and Spratt, J.A. 1996. Rates of growth of human neoplasms: Part II. *J. Surg. Oncol*. 61, 68–83.

Timmer, J., Müller, T.G., Swameye, I., Sandra, O. and Klingmüller, U., 2004. Modeling the nonlinear dynamics of cellular signal transduction. International Journal of Bifurcation and Chaos, 14(06), pp. 2069–2079.

Toni, T., and Stumpf, M.P.H. 2010a. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *Bioinformatics (Oxford, England)*. 26, 104–110.

Toni, T., and Stumpf, M.P.H. 2010b. Simulation-based model selection for dynamical systems in systems and population biology. *Bioinformatics (Oxford, England)*. 26, 104–110.

Toni, T., Welch, D., Strelkowa, N., et al. 2009. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *J. R. Soc. Interface*. 6, 187–202.

Tsoularis, A., and Wallace, J. 2002. Analysis of logistic growth models. *Math. Biosci*. 179, 21–55.

Váron, L.A., Orlande, H.R.B., and Eliçabe, G. 2015. Estimation of state variables in the hyperthermia therapy of cancer with heating imposed by radiofrequency electromagnetic waves. *Int. J. Therm. Sci*. 98, 228–236.

Wegmann, D., Leuenberger, C., and Excoffier, L. 2009. Efficient approximate Bayesian computation coupled with Markov chain Monte Carlo without likelihood. *Genetics*. 182, 1207–1218.

Winkler, R.L., 2003. Introduction to Bayesian Inference and Decision. Probabilistic Publishing, Gainesville.

Winsor, C.P. 1932. The Gompertz curve as a growth curve. *Proc. Natl. Acad. Sci. USA*. 18, 1–8.

Address correspondence to:
Dr. Helcio R.B. Orlande
Department of Mechanical Engineering
Federal University of Rio de Janeiro–UFRJ
Caixa Postal 68503
Rio de Janeiro 21941972
Brazil

E-mail: helcio@mecanica.coppe.ufrj.br