Multi-process Dynamic Modeling of Tumor-specific Evolution

A. Achilleos†, C. Loizides†, T. Stylianopoulos‡, G. D. Mitsis†
† KIOS Research Center for Intelligent Systems and Networks - Department of Electrical and Computer Engineering
‡ Department of Mechanical and Manufacturing Engineering
The University of Cyprus

***Abstract*—We suggest a multi-process dynamic model and a sequential bayesian forecasting method of tumor-specific growth. The mixture model uses prior information obtained from the general population and becomes more individualized as more observations from the tumor are sequentially taken into account. In this study we propose utilizing all available tumor-specific information up to date to approximate the unknown multi-scale process of tumor growth over time, in a stochastic context. The validation of our approach was performed with experimental data from mice and the results show that after few observations from a tumor are obtained and included in the model, the latter becomes more individualized, in the sense that its parameters are adjusted in order to reflect the growth of each individual tumor, yielding more precise estimates of its size.**

***Index Terms*—Statistical Modeling, Mixture Dynamic Model, Bayesian forecasting, Gompertz-law of growth, xenograft model.**

I. Introduction

The development of mathematical models describing the evolution of a tumor over time has had a significant effect on understanding biological growth dynamics, the evolution of resistance to anti-cancer therapy and the design of optimal control strategies through constrained optimization. Even a quick look at the literature reveals a large number of growth models, either deterministic or probabilistic/stochastic [1], [2], [3], [4], [5]. The more complex a model is, the more physiological phenomena or processes it may potentially take into account. However at the same time, both the theoretical and practical analyses may become considerably harder, e.g. identifiability problems. Consequently, a trade-off between these two aspects must be sought. Based on experimental data, the growth rate of a malignant tumor has been characterized as exponential at the beginning of the course of the disease, followed by a linear growth towards an asymptote which is usually termed as maximum carrying capacity. Several mathematical, mostly phenomenological growth curves, have been used to model this or a similar behaviour [4]. However, the performance of such models depends crucially on the reliable estimation of their unknown parameters which is often not possible in practice. Compared to the deterministic framework, a more realistic description of complex physiological phenomena, such as the dynamic growth of a tumor in a given subject, with many possible different sources of uncertainty, can be achieved by obtaining probabilistic inferences about any quantity of interest. In this context, the objective of this study is to suggest (i) a statistical tumor-specific growth model, (ii) an individualized recurrent updating procedure for the model parameters and (iii) a sequential approach to forecast tumor dynamics under the Bayesian framework.

II. Dynamic Modeling of Tumor Growth

A. The Growth Law

In this paper we are interested in modeling and forecasting of a univariate time series, that is, the volume of the tumor across time. In the following, we consider a model that obeys the Gompertz-law [1], [4], due to its simplicity, popularity and ability to fit tumor growth experimental data relatively well. The generalization of our model in order to accommodate for any continuous function that describes the temporal tumor dynamics is straightforward.

Let Y_t denote the log tumor volume at time $t > 0$. The deterministic Gompertz-type tumor growth function is defined by

$$Y_t = \frac{c_1}{c_2} - \left\{ \frac{c_1}{c_2} - \ln(N_0) \right\} \exp(-c_2 t). \tag{1}$$

Different values of the positive real valued parameters c_1 and c_2 can be used to determine the evolution of different tumor types [6], [1], with $N_0 > 0$ being the initial tumor volume and c_1/c_2 being the maximum carrying capacity for a specific subject. Nonetheless, because of various physical, mechanical and chemical factors, as well as the spatial and temporal heterogeneity of tumors, it is quite common to observe discrepancies within and between subjects and deviations from the baseline response to a particular form of cancer. Therefore, it is reasonable, in the face of uncertainty, to treat the parameters of the model as random variables.

B. The Gaussian Dynamic Linear Model

Let I_t be the set of all available information at time t. As time passes, new observations become available for inclusion in the model thus updating our information. We can express this information as: $I_t = \{Y_t, I_t^*, I_{t-1}\}$, with $I_0 = \{Y_0, I_0^*\}$, where I_t^* represents any additional relevant information obtained at time t. Then an appropriate dynamic linear model [7], [8] (DLM) describing the progression of (1) over time is defined by

$$Y_t = \alpha_{1,t} + \alpha_{2,t} X_t + v_t,$$

978-1-4799-3163-7/13/$31.00 ©2013 IEEE

with $X_t = \exp(-c_2 t)$ and c_2 fixed, $v_t \sim N(0, V_t)$, $\alpha_{1,t} = \alpha_{1,t-1} + w_{1,t}$, $\alpha_{2,t} = \alpha_{2,t-1} + w_{2,t}$, and $\mathbf{w}_t \sim N(0, W_t)$. Here, v_t is independent of $(\boldsymbol{\theta}^t, \mathbf{Y}^{t-1})$, $\mathbf{w}_t$ is independent of $(\boldsymbol{\theta}^{t-1}, \mathbf{Y}^{t-1})$, where $\mathbf{w}_t = (w_{1,t}, w_{2,t})'$, $\boldsymbol{\theta}_t = (\alpha_{1,t}, \alpha_{2,t})'$ and $\mathbf{u}^t = \{u_1, \ldots, u_t\}$.

The updating recurrence relationships for the parameters of our model and the one-step ahead forecast error $e_t = Y_t - E\big(Y_t|\mathbf{Y}^{t-1}\big)$ are derived using the concept of conditional independence (the future is independent of the past, given the present) and the Bayes' theorem, according to which $p(\theta|x) \propto p(x|\theta)p(\theta)$. All the information regarding the future is based on inference drawn from the posterior distribution. In the sequel, whenever we quantify the uncertainty for our model predictions, we refer to the highest predictive density (HPD) regions with respect to these predictions. A region $R_a \subset \Theta$ is said to be a highest density region for θ of size a with respect to $p(\theta)$ if the conditions: (i) $P(\theta \in R_a) = a$ and (ii) $p(\theta_1) \geq p(\theta_2)$ for all $\theta_1 \in R_a$ and $\theta_2 \notin R_a$, hold [7].

Assuming that the observational variance is constant, i.e. $V_t = V$, we can address the uncertainty on V by using standard Bayesian conjugate analysis [8], [9]. Let $\phi = 1/V$ be the precision variable. We assume that ϕ follows a gamma distribution (equivalently, inverse gamma for V) and define the prior distribution $(\phi|I_0) \sim G(n_0/2, d_0/2)$, where n_0, $d_0 \in \mathbb{R}^+$ and G denotes the gamma distribution. Note that the mean of this prior distribution is $1/S_0$, where $S_0 = d_0/n_0$ is a prior point estimate of V. When V is unknown, it can be shown [7, ch. 2, 4, 17] that the posterior distributions of the model coefficients as well as the forecast distribution are no longer normal but rather follow non-standardized t-distributions.

Nevertheless, in our particular case the observational variance is not constant through time. Therefore, we assume that the precision of the observation error is subject to a random disturbance at each time point. Specifically, in order to model the transition from ϕ_{t-1} to ϕ_t we use a random walk which satisfies $E(\phi_{t-1}|I_{t-1}) = E(\phi_t|I_{t-1})$. Following the approach of [7, ch. 10], we set $\phi_t = \gamma_t \phi_{t-1}/0.95$, where $\gamma_t \sim B(0.95n_{t-1}/2, 0.05n_{t-1}/2)$ is a random variable that is independent of ϕ_{t-1}, with B representing the beta distribution, for $0 < \gamma_t < 1$ with $E(\gamma_t|I_{t-1}) = 0.95$. At time point $t-1$, the posterior distribution of ϕ_{t-1} is given by $(\phi_{t-1}|I_{t-1}) \sim G(n_{t-1}/2, d_{t-1}/2)$. We estimate the evolution variance matrix using discounted variance learning, i.e. by defining $\mathbf{W}_t$ to be a fixed proportion δ of $\mathbf{C}_{t-1} = \mathrm{Var}(\boldsymbol{\theta}_{t-1}|I_{t-1})$.

It is worth noting that, at each time $t > 0$, the mean of the posterior distribution of the state vector $\boldsymbol{\theta}_t$ is equal to the mean of its prior distribution at $t-1$ plus a correction term which is proportional to the forecast error e_t. The first obvious conclusion is that the greater the error the greater the difference between the two posterior means. Also of note is that the posterior precision $\mathbf{C}_t^{-1}$ is always larger than the corresponding prior, hence the posterior of the coefficient vector will never be more diffuse (less informative) than the prior. The recurrence relationship of n_t converges to a constant, with the limit being equal to 20. This implies that while trying to accommodate the various sources of uncertainty that influence our observations, we restrain the precision of the estimated variance at each time point by bounding the degrees of freedom of the posterior forecast distribution, T_{n_t}, and not allowing it to become approximately normal with thinner tails compared to a t-distribution. A value smaller than 0.95 can be used as discount parameter, increasing the variance of the random variable γ_t, but at the same time the limiting value of n_t will decrease.

III. Multi-process Model Specification

Whereas any single DLM defines a process model, a combination of several DLMs defines a multi-process model (MPDLM), also known as mixture model [10]. Let $M_t(\mathbf{c})$ be a DLM for a given value of $\mathbf{c} \in \mathbb{A}$, where $\mathbb{A}$ is the parameter space. We consider the class of MPDLMs under which we assume that there is a true but unknown value of $\mathbf{c}$, say $\mathbf{c}^*$. Then, our objective is to choose a mixture, from a finite collection of competing models, that is as close as possible to $M_t(\mathbf{c}^*)$, even if no single component dynamic model is appropriate to describe the evolution of the process [7].

Let Ξ_t be a vector of random quantities of interest; for example, the state vector $\boldsymbol{\theta}_t$ or the future observation Y_{t+l}, for $l > 0$. Now consider that we fit in parallel more than one such DLMs, where for each we choose a different value for $\mathbf{c}$. Starting with an initial prior $p(\mathbf{c}|I_0)$ for $\mathbf{c}$, as data become available we can sequentially update $p(\mathbf{c}|I_t)$ using Bayes' rule. Now the posterior $p(\mathbf{c}|I_t)$ quantifies, in a sense, the support that each $\mathbf{c}$ receives from the data we collected up to time t.

In order to infer on Ξ_t without reference to any particular value of $\mathbf{c}$, we need to integrate $\mathbf{c}$ out, that is $p(\Xi_t|I_t) = \int_{\mathbb{A}} p(\Xi_t|\mathbf{c}, I_t)p(\mathbf{c}|I_t)d\mathbf{c}$. In practice, we approximate this integral with a finite sum using a fixed and finite grid of points $\{\mathbf{c}_1, \ldots, \mathbf{c}_k\}$.

The analysis is thus based on the use of a finite collection of DLMs, updated in parallel, each corresponding to a different choice of the parameter value. Thereafter, c is considered a discrete random variable, the posterior distribution of which has a mass function p_t at time t given by $p_t(j) = p(c_j|I_t), \quad (j = 1, \ldots, k)$ with the initial prior distribution denoted by $p_0(j)$. Considering the one-step ahead forecast problem (i.e. $\Xi_t = Y_t$), the posteriors are calculated according to $p_t(j) \propto p_{t-1}(j)p(Y_t|c_j, I_{t-1})$. The normalizing constant κ_t can be obtained by $\kappa_t^{-1} = \sum_{j=1}^{k} p_{t-1}(j)p(Y_t|c_j, I_{t-1})$.

A. Reparameterization of the dynamic model

We define the system $Y_t = \alpha_t + \beta_t + v_t$, with $\alpha_t = \alpha_{t-1} + w_{\alpha,t}$ and $\beta_t = \lambda\beta_{t-1} + w_{\beta,t}$, for some correlated parameters α, β and γ, with $\gamma > 0$, where $w_{\alpha,t}$ an $w_{\beta,t}$ are the evolution variances, $\lambda = \exp(-\gamma)$ and $0 < \lambda < 1$.

It is now evident that, for all $t > 0$, under our framework, we treat α_t (and β_t) as locally constant (and decreasing by λ a few points back and ahead in time), with changes over longer periods of time to be expected but modelled as purely stochastic. Therefore, we can employ the multi-process modeling techniques for the parameter λ. The parameter space is restricted between 0 and 1, thus, lacking any prior information

on λ, a reasonable approach is to consider a set of possible values $\mathbb{L}$ for λ from an equally spaced sequence between 0 and 1. If an initial estimate of λ is available, the boundaries of the sequence can be chosen to be tighter. The more data one has, the denser this sequence may be. Note that the larger (smaller) the value of λ is, the slower (faster) the corresponding growth rate is.

IV. Data Analysis and Validation

A. Choice of the initial values

Consider that we have tumor progression data that come from individuals from the population of interest, having the same type of cancer, denoted by $Y_{i,1}, \dots Y_{i,j_i}$, with $i = 1, \dots n_r$, where n_r is the number of subjects in the training set and j_i is the total number of observations for subject i in the training set. In order to obtain the initial values, we fit a non-linear least squares regression model to each subject in our training set using the parametrization (1). The resulting estimates of $c_{1,i}$ and $c_{2,i}$ are $\tilde{c}_{1,i}$ and $\tilde{c}_{2,i}$, respectively, for $i = 1, \dots n_r$. Then, we construct an "average" individual having a maximum carrying capacity equal to $\tilde{\alpha} = \frac{\tilde{c}_1^{med}}{\tilde{c}_2^{med}}$ where $\tilde{c}_2^{med} = \text{median}_{i=1,\dots n_r}\{\tilde{c}_{2,i}\}$ and $\tilde{c}_1^{med} = \tilde{c}_2^{med} \times \tilde{\alpha}^{med}$, with $\tilde{\alpha}^{med} = \text{median}_{i=1,\dots n_r}\left\{\frac{\tilde{c}_{1,i}}{\tilde{c}_{2,i}}\right\}$. We avoid using the mean instead of the median due to effect that possible extreme outliers might have. Once we have these estimates, we set $\mathbf{m}_0 = (\tilde{c}_1^{med}/\tilde{c}_2^{med}, \log(N_0) - \tilde{c}_1^{med}/\tilde{c}_2^{med})'$. For a single process DLM we can choose $\lambda = \exp(-\tilde{c}_2^{med})$, while for a MPDLM with k component models we use $\mathbb{L} = \{\lambda_{(1)}, \dots, \lambda_{(k)}\}$, where $\mathbb{L}$ is an equally spaced grid with $\lambda_{(1)}$ and $\lambda_{(k)}$ representing the smallest and largest estimated λ values from all available $Y_{i,j}$, $i = 1, \dots n$ and $j = 1, \dots T$, respectively.

B. Experimental data

In this section we study the predictive performance of the proposed models using data from mice experiments. At time $t = 0$ we start forecasting the future tumor volume (one step-ahead forecast) initially using the available information from the training data and then as time passes and more data are collected, we update our model and adapt the tumor volume forecasts to reflect the tumor's unique characteristics.

1) Methods and Materials: Tumors were prepared as described in previous work [11] by implanting a small piece (1 mm^3) of viable tumor tissue from a source tumor animal into the flank or mammary fat pad (mfp) of a severe combined immunodeficient (SCID) mouse. Specifically, the following four cancer cell lines were used: human glioblastoma $U87$ (flank, number of mice subjects $n_s = 6$), human fibrosarcoma $HT1080$ (flank, $n_s = 11$), murine mammary adenocarcinoma $4T1$ (mfp, $n_s = 12$) and murine mammary adenocarcinoma $E0771$ (mfp, $n_s = 6$). Tumor growth was monitored on a daily basis and its planar dimensions (x, y) were measured with a digital calliper every 2 days. The volume of the tumor was estimated from its planar dimensions using the volume of an ellipsoid and assuming that the third dimension z is equal to $\sqrt{xy}$. Therefore, the tumor volume V equals $V = \frac{4\pi}{3}\frac{xyz}{8} = \frac{\pi}{6}(xy)^{3/2}$.

2) Results: In order to assess the predictive performance of the multi-process DLM we randomly split the data in each of the U87, HT1080, 4T1 and E0771 cell lines into training and testing sets. Specifically, we randomly select two thirds of the mice in each cell line to comprise the training data set and the remaining third to comprise the testing data set. As mentioned before, the initial mean parameter $\mathbf{m_0}$ is chosen based on the non-linear least square estimates $\tilde{c}_1$ and $\tilde{c}_2$. From the analysis of the training set for each cancer cell line we choose a set of five values for λ defining a five-component mixture model. Based on a factorial design using the training data set, we set the initial parameters δ, n_0, d_0 and matrix C_0 equal to 0.25, 1, 0.001 and $\text{diag}(0.01, 0.01)$ respectively. We found this choice to work well for the various types of cancer cell lines we considered, even though the latter exhibit various scales and features.

In Figure 1 we present the one-step ahead forecasts and 90% HPD intervals for some representative randomly selected mice from the testing sets corresponding to the $HT1080$, $E0771$, $4T1$ and $U87$ cancer cell lines. The x-axis corresponds to the total time (days) elapsed since $t = 0$. In both figures the bottom row shows the corresponding posterior probabilities for each of the five component models at each time step. The first prediction after $t = 0$ is solely based on the baseline (initial) values of the model parameters. Then, as time passes, the model parameters are updated based on the new observations and the resulting one step-ahead forecasts exhibit increasingly improved performance in terms of mean value prediction and reduced uncertainty. For example, in the top left panel of Figure 1, the subject's initial tumor growth rate is slower than the expected rate given the training data set. Thus at the first time point after $t = 0$ the model prediction overestimates the observed value Y_1. As soon as the new information enters our model, the model parameters are updated towards the unique characteristics of the given subject.

From the posterior probabilities shown in the right column panels in Figure 1 we can gain valuable information about the support that each of the component models receives from the data. For example, in the first subject (top row) the model corresponding to $\lambda = 0.75$ has the largest weight showing evidence in favour of an average growth, i.e. close to what we expect given the training set, with the true value of λ possibly lying in the interval $[0.65, 0.75]$.

V. Discussion and Conclusions

We propose a multi-process dynamic linear model for tumor growth and an individualized Bayesian forecasting method to predict short-term tumor evolution, given a limited number of observed data as well as some a priori information about the average response of the population in which the subject belongs, to a specific type of cancer. Overall, our results suggest that the one-step ahead prediction performance of the proposed model has promising potential. Although cancer encompasses many biological processes occurring on many levels, e.g. genomic and molecular, we are only observing tumor size, which is a situation that arises often in practice.

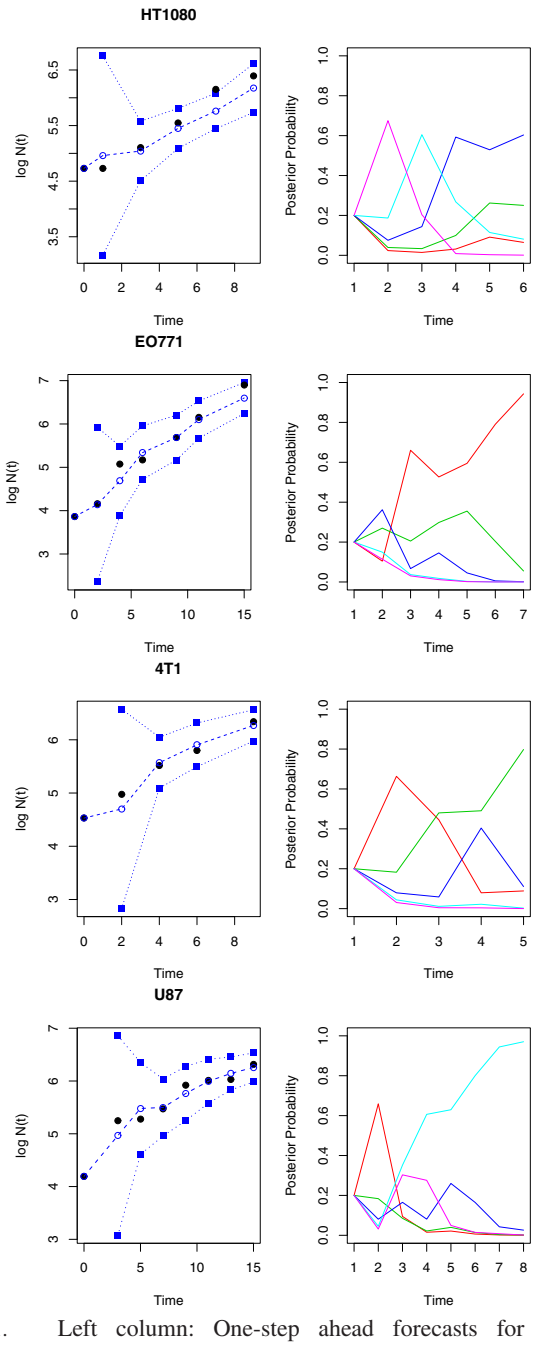

Fig. 1. Left column: One-step ahead forecasts for a randomly selected mouse from the testing set. Blue-White dots: Forecast values f_t, Black dots: Observations Y_t, Blue dotted lines and squares: 90% posterior probability intervals; Right column: Posterior probabilities of the component models; From top to bottom: Red: $\lambda = 0.55, 0.9, 0.8, 0.8$, Green: $\lambda = 0.65, 0.9225, 0.8475, 0.8475$, Blue: $\lambda = 0.75, 0.945, 0.895, 0.895$, Light-blue: $\lambda = 0.85, 0.9675, 0.9425, 0.9425$, Pink: $\lambda = 0.95, 0.99, 0.99, 0.99$.

Therefore, we use a macroscopic dynamic model to describe a complex multi-scale phenomenon. Moreover, we have considerable measurement errors (tumor size is measured using callipers), small sample size, estimation of tumor volume with only two dimensions, between and within subject variation. All these reasons render the quantification of uncertainty very important. Unlike the vast majority of cancer growth models that do not take uncertainty into account, the HPD intervals of our model quantify the aggregate effect of the heterogeneous sources of variation on the prediction uncertainty. It is therefore necessary to account for this temporal variability when designing treatment strategies in practice.

The proposed mixture approach is used in order to alleviate the fact that there is no a priori information available for the tumor under examination (we only have access to information on the population to which the subject belongs). Therefore, we use a finite collection of rival models, each corresponding to a different growth behaviour, to identify the model (or the combination of the competing models) that better describes the evolution process in the particular tumor. Convergence to a single value of λ identifies the corresponding model as the most appropriate to describe the growth dynamics of the tumor, although no simple model actually generates the time-series Y_t. The results presented here identified considerable room for improvement of short-term tumor-specific growth prediction, suggesting the development of individualized treatment strategies utilizing all available information through the deployment of posterior and predictive distributions, rather than using point estimates and predictions. Also, it is worthwhile noting that our model allows inclusion of various extra covariates, e.g. genetic information, at little additional computational cost.

ACKNOWLEDGEMENT

We thank Dr. Rakesh K. Jain, Director of the Steele Lab, Massachusetts General Hospital, where the experiments were performed. This work is co-funded by the European Territorial Cooperation Programmes (80%) and by national funds of Greece and Cyprus (20%).

REFERENCES

[1] R. Martin, "Optimal control drug scheduling of cancer chemotherapy," *Automatica*, vol. 28, no. 6, pp. 1113–1123, 1992.
[2] L. De Pillis and A. Radunskaya, "A mathematical tumor model with immune resistance and drug therapy: an optimal control approach," *Comput. Math. Meth. Med.*, vol. 3, no. 2, pp. 79–100, 2001.
[3] F. Kozusko and Ž. Bajzer, "Combining gompertzian growth and cell population dynamics," *Math. Biosci.*, vol. 185, no. 2, pp. 153–167, 2003.
[4] L. Preziosi, *Cancer modelling and simulation*. CRC Press, 2003, vol. 3.
[5] T. Deisboeck and G. Stamatakos, *Multiscale cancer modeling*. CRC Press, 2010, vol. 34.
[6] L. Norton, "A gompertzian model of human breast cancer growth," *Cancer Res.*, vol. 48, no. 24 Part 1, pp. 7067–7071, 1988.
[7] M. West and J. Harrison, *Bayesian forecasting and dynamic models*. Springer Verlag, 1997.
[8] J. Durbin, S. Koopman, and A. Atkinson, *Time series analysis by state space methods*. Oxford University Press Oxford, 2001, vol. 15.
[9] W. M. Bolstad, *Introduction to Bayesian statistics*. Wiley-Interscience, 2007.
[10] P. Damien, P. Dellaportas, N. Polson, and D. Stephens, *Bayesian Theory and Applications*. Oxford University Press, 2013.
[11] T. Stylianopoulos, J. D. Martin, V. P. Chauhan, S. R. Jain, B. Diop-Frimpong, N. Bardeesy, B. L. Smith, C. R. Ferrone, F. J. Hornicek, Y. Boucher *et al.*, "Causes, consequences, and remedies for growth-induced solid stress in murine and human tumors," *Proceedings of the National Academy of Sciences*, vol. 109, no. 38, pp. 15 101–15 108, 2012.