*Proceedings of the 2012 IEEE 12th International Conference on Bioinformatics & Bioengineering (BIBE), Larnaca, Cyprus, 11-13 November 2012*

# Linear Dynamic Modelling and Bayesian Forecasting of Tumor Evolution

A. Achilleos†, C. Loizides†, T. Stylianopoulos‡, G. 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 consider a linear dynamic model for tumor growth evolution. A number of temporal statistical models for tumor growth exist in the literature. In the majority of these cases the employed models are formulated in a deterministic context, providing no information on their uncertainty. Some of these are theoretically well defined and very useful in practice, e.g. to define general optimal treatment protocols through nonlinear constrained optimization. Nevertheless a challenging task is the estimation of the model parameters for a specific individual since, especially in humans, it is not feasible to collect a large number of tumor size values with respect to time, as the tumor is removed immediately after diagnosis in most cases. Therefore, we suggest a probabilistic model for personalized sequential tumor growth prediction, given only a few observed data and an a priori information regarding the average response to a specific type of cancer of the population to which the subject belongs. We validated the proposed model with experimental data from mice and the results are promising.**

***Index Terms*—Personalized sequential tumor growth prediction, Linear Dynamic Modeling, Bayesian forecasting, Gompertz-law of growth, mouse xenograft model.**

## I. Introduction

Development of mathematical models describing the evolution of a tumor over time has had a significant effect on the understanding of the biological growth dynamics, the evolution of resistance to anti-cancer therapy and the design of optimal control strategies through constrained optimization. Even a shallow look at the literature reveals a large number of growth models (compartmental ordinary differential equations or partial differential equations, cellular-automata, mechanical, agent-based models and many more) describing (through various physiological phenomena) the temporal and spatial dynamics of cancer evolution before and after vascularization. The more complex a model is, the more phenomena/processes it takes into account, but at the same time, both the theoretical and practical analyses become considerably harder (e.g. identifiability problems; estimation of the unknown parameters is not always possible due to small sample size and large number of unknowns). Thus, a trade-off between these two aspects must be sought.

The main issue with the deterministic modeling techniques is that the within and between subject variability is ignored. In addition, the unique response to carcinogenesis of a specific person in the face of uncertainty is ostentatiously ignored. See, for example, [1], [2], [3] and [4] who include stochastic terms in the deterministic models, and [5] for an example of adaptive personalized optimal control.

From experimental data, the growth rate of a malignant tumor is 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 (sigmoidal shape). Therefore several statistical/phenomenological growth curves have been used to model this or similar behavior, such as exponential, logistic, Weibull, Gompertz and hyperbolastic. For example, see [3], [6] and [7]. Here we deal with a model obeying the Gompertz-law, see [8], due to its simplicity, popularity and ability to fit experimental data well. Extension of our model to any continuous function that describes the temporal tumor dynamics is straightforward.

Despite their simplicity, Gompertz-type models have proven to be appropriate to predict the average growth behavior of a tumor. Among others, see [1] and [9]. However, the performance of such models depends crucially on consistent estimation of their unknown parameters. In order to achieve this, animal experiments are usually carried out to observe a sample of the tumor size as time passes (univariate time series) for each subject on which the estimation step is based on. Note though, that each subject has a unique response to carcinogenesis, i.e. different values of the model parameters are obtained for each individual. Nonetheless, in reality, when cancer is diagnosed in a given subject, it is almost always the case (unlike other diseases, e.g. diabetes, and infections, e.g. HIV) that at most a sample of a small size can be observed, since it is very difficult and ethically inappropriate (at least for humans) to leave the tumor untreated in order to get an adequate sample of the tumor size as time evolves. In the case of a very small sample size, maximum likelihood or least squares estimation are not advisable since the resulting model's estimated parameters will have very large variance, yielding meaningless results.

Hence, even though theoretically attractive, the aforementioned models for tumor growth dynamics are not directly applicable for either forecasting or personalized optimal control. In this context, the objective of this paper is to suggest (i) a probabilistic tumor growth model, (ii) a personalized recurrent updating procedure for the model parameters and (iii) a sequential forecasting of the tumor dynamics under the Bayesian framework. The proposed model is intended to be used for adaptive optimal control towards personalized optimal

treatment strategies in mice experimentations.

In the next section we define the Gompertz growth model. The dynamic linear model (DLM) and the Bayesian forecasting technique under normality are also introduced in section 2 along with a brief discussion on the derivation details and asymptotic behavior of the proposed model and updating recurrence relationships. In section 3 we validate our model with experimental data from mice. Finally, section 4 consists of the concluding remarks and prospects of the suggested DLM.

## II. State-space Model and Growth function

Robustness is crucial for a learning system, especially in this particular case where limited data are available. In what follows we are concerned with modeling and forecasting single time series, i.e. the volume, in $mm^3$, of the tumor in time. A more complete description of complex physiological phenomena (such as the dynamic growth of a tumor in a given subject), with possibly many different sources of uncertainty, can be achieved by obtaining probabilistic inferences about any quantity of interest. In this context, we can use statistical models to predict future values of the time series of interest and the corresponding model parameters as well as the uncertainty of these estimates. As described in detail below, the Bayesian forecasting procedure can be used to this end.

### A. The Gompertz growth law

Let $N_t$ denote the volume, in $mm^3$, of the tumor at time $t > 0$. As in [3], the deterministic Gompertz-type tumor growth function is defined by the solution of the ordinary differential equation:

$$\frac{dN_t}{dt} = c_1\, N_t - c_2\, N_t \ln(N_t) \tag{1}$$

where $c_1$ represents the tumor's growth rate (depending on the proliferation rate) and $c_2$ is an anti-growth factor (e.g. due to antiangiogenic processes). The parameters $c_1$ and $c_2$ define the evolution of different tumor types and vary significantly with the type of cancer, between subjects and within subjects (as time evolves).

The solution of (1) is a sigmoidal function given by:

$$N_t = \exp\left[\frac{c_1}{c_2} - \left\{\frac{c_1}{c_2} - \ln(N_0)\right\}\exp(-c_2 t)\right] \tag{2}$$

with $N_0 > 0$ being the initial tumor's volume. The function in (2) describes an initial exponential increase followed by a linear behavior towards an asymptote. It can be shown, see [3] for example, that the largest tumor volume that a specific subject can tolerate is given by $\exp(c_1/c_2)$. After defining $Y_t = \ln N(t)$ and taking the natural logarithm on both sides of (2) we obtain:

$$Y_t = \frac{c_1}{c_2} - \left\{\frac{c_1}{c_2} - \ln(N_0)\right\}\exp(-c_2 t), \tag{3}$$

which shows, at a given $t$ and $c_2$, a linear relationship between $Y_t$ and $\exp(-c_2 t)$. Therefore, embolden from (3) we are now ready to proceed to the development of our probabilistic model for personalized sequential prediction of tumor growth dynamics.

From this point onwards, we do not represent the random quantities and their corresponding realized values separately. Thus, prior to observing the value of the tumor volume at time $t$, $Y_t$ denotes the uncertain random quantity (volume, in $mm^3$, of the tumor at time point $t$), which becomes certain when observed. The series $Y_1, Y_2, \ldots$ need not be equally spaced in time, although in the sequel we refer to equally-spaced series in time unless otherwise specified.

### B. The Dynamic Linear model

Statistical modeling of time series is commonly performed using various types of dynamic models, often termed as state-space models, that express and model the behavior of a system over time which is regarded as the driving force. When constructing a forecasting problem using dynamic models we essentially formulate a robust parametric model having parameters depending on time which allows (i) representation of all available information with probability distributions by employing conditional independence and (ii) sequential updating of the model parameters and forecasting of future observations. Among others, see [10], [11], [12] and [13] for an extensive introduction in Bayesian statistics, forecasting and time series analysis.

Let $I_t$ be the set of all available information at time $t$. As time passes, new observations become available to the model revising our information. Thus $I_t = \{Y_t, I_t^*, I_{t-1}\}$, with $I_0 = \{Y_0, I_0^*\}$, where $I_t^*$ represents any extra relevant information obtained at time $t$, e.g. an expert opinion that may alter our beliefs on how the series $Y_{t+k}$, for $k > 0$, will evolve.

**Definition 1.** *A family $F$ of probability distributions on $\Theta$ is said to be conjugate, or closed under sampling, for a likelihood function $p(\mathbf{x}|\theta)$ if, for every prior $p^* \in F$, the posterior distribution $p(\theta|\mathbf{x})$ also belongs to $F$.*

A Bayesian information updating combines information from different sources in a coherent way. This is achieved by the Bayes rule, according to which

$$p(\theta|x) = \frac{p(x|\theta)p^*(\theta|x)}{\int_\Theta p(x|\theta')p^*(\theta'|x)\,\mathrm{d}\theta'}.$$

Thus in order to have the explicit expression of the posterior we must be able to perform the integral appearing in the denominator. One way to do so is to ensure that both $p(\theta|x)$ and $p(\theta)$ belong to the same family of distributions, i.e. use conjugate priors as in definition 1. Note though that sometimes this may prove impracticable.

Now, from (3), let us propose the following representation for the growth model:

$$Y_t = \alpha_1 + \alpha_2\, X_t, \tag{4}$$

where $X_t = \exp(-\bar{c}_2 t)$, $\bar{c}_2$ is fixed and the coefficients $\alpha_1$ and $\alpha_2$ are unknown. Since it is quite common to observe

discrepancies between subjects and deviations from the baseline response to a particular form of cancer, as well as due to various physical, mechanical and chemical factors, it is reasonable (in the face of uncertainty) to treat the parameters $\alpha_1$ and $\alpha_2$ as random variables. In particular, we model the coefficients as a simple random walk that equips us with a powerful tool towards adaptation to changes in the underlying processes.

The most popular dynamic model is the Gaussian Dynamic Linear Model, referred to simply as dynamic linear model, or DLM, where normality is assumed. Such models that are associated with the normal theory are simple, flexible, analytically manageable and complete. Hence, in what follows all the conditional, marginal and joint distributions are Gaussian. The probability density function of a normally distributed random variable U, with mean $\mu$ and variance $\sigma^2$, is given by

$$p(U = u|\mu, \sigma^2) = (2\pi\sigma^2)^{-1/2} \exp\Big\{ -\frac{1}{2\sigma^2}(u-\mu)^2 \Big\}. \quad (5)$$

Under Definition 1, an appropriate DLM that describes the evolution of (4) over time is defined in the additive representation by

Observation:

$$Y_t = \alpha_{1,t} + \alpha_{2,t}\, X_t + v_t, \quad v_t \sim N(0, V_t)$$

System/State:

$$\alpha_{1,t} = \alpha_{1,t-1} + w_{1,t}$$

$$\alpha_{2,t} = \alpha_{2,t-1} + w_{2,t}$$

$$\mathbf{w_t} \sim N(0, W_t)$$

Initial Information:

$$(\alpha_{1,0}|I_0) \sim N(\bar{c}_1/\bar{c}_2, \zeta_1)$$

$$(\alpha_{2,0}|I_0) \sim N(\ln N_0 - \bar{c}_1/\bar{c}_2, \zeta_2)$$

with $X_t = \exp(-\bar{c}_2 t)$, where $\bar{c}_1$ and $\bar{c}_2$ are assumed to be fixed such that $Y_t$ has the average/baseline growth behavior for a particular population and type of cancer, with $\zeta_1, \zeta_2 \in \mathbb{R}^+$ representing our uncertainty on the initial guess of the model parameters, $v_t$ is independent of $(\theta^{\mathbf{t}}, \mathbf{Y^{t-1}})$, $\mathbf{w_t}$ is independent of $(\theta^{\mathbf{t-1}}, \mathbf{Y^{t-1}})$, where we use $\mathbf{w_t} = (w_{1,t}, w_{2,t})'$, $\theta_{\mathbf{t}} = (\alpha_{1,t}, \alpha_{2,t})'$ and the generic notation $\mathbf{u^t} = \{u_1, \ldots, u_t\}$. The choice of the observational, $V_t$, and evolution, $W_t$, variances is discussed in section II-B1.

Let $\mathbf{F_t} = (1, X_t)$. The $k$-ahead forecast function $\{f_t(k) : k = 1, 2, 3, \ldots\}$ is defined by

$$f_t(k) = E\big(Y_{t+k}|Y^t\big) = \mathbf{F}_{t+k}\, E\big(\theta_{t+k}|Y^t\big)$$

since $v_{t+k}$ is independent of $Y^t$ and $E(v_{t+k}) = 0$. Thus the distribution of the forecast $Y_t$ on the basis of the information in $\mathbf{Y^{t-1}}$ is normal with mean $f_t = E\big(Y_t|\mathbf{Y^{t-1}}\big)$ and variance $Q_t = \mathrm{Var}\,\big(Y_t|\mathbf{Y^{t-1}}\big)$. After observing $Y_t$, the likelihood for $\theta_{\mathbf{t}}$ is proportional to the observed density expressed as a function of $\theta_{\mathbf{t}}$. Consequently 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 which states that the posterior distribution of the parameter vector at any time point $t > 0$ is proportional to the prior multiplied by the observed likelihood. At any time point $t > 1$, the prior is equal to the posterior at $t - 1$. All the information regarding the future is embedded in the posterior distribution.

**Definition 2.** *A region $R_a \subset \Theta$ is said to be a highest density region for $\theta$ of size $a$ with respect to $p(\theta)$ if (i) and (ii) hold, where*

*(i) $P(\theta \in R_a) = a$*

*(ii) $p(\theta_1) \geq p(\theta_2)$ for all $\theta_1 \in R_a$ and $\theta_2 \notin R_a$.*

In the sequel, if $p(\theta)$ is either a prior, posterior or predictive density, we refer to highest prior, posterior or predictive density regions. More details and an alternative sketch of the proof of the updating recurrent equations are found, among others, in [11] and [12].

*1) Unknown Observational and Evolution Variances:* The variance $V_t$ of the observational error $v_t$ describes at any time point $t$ the uncertainty about our beliefs on the unknown random fluctuations around the average growth rate of the unique subject's tumor growth rate. The variance $W_t$ of the evolution error vector $\mathbf{w}_t$ characterizes the evolution of our system as time passes and at $t = 0$ it quantifies the uncertainty about the magnitude of the difference of the unique growth rate of the subject from the average growth rate. It is evident from experimental and clinical data that both $V_t$ and $W_t$ are not precisely known and vary with time. See, among others, [3].

We suggest explaining the uncertainty about $V$ using the standard Bayesian conjugate analysis. See [14] and [12] for example. Let $\phi = 1/V$ be the precision variable. Then, we assume that $\phi$ follows a gamma distribution (inverse gamma for V) and define the initial information

$$(\phi|I_0) \sim G(n_0/2, d_0/2),$$

where $n_0, d_0 \in \mathbb{R}^+$ and $G$ stands for the gamma prior distribution having a probability density function given by

$$p(\phi|I_0) = \frac{(d_0/2)^{n_0/2}}{\Gamma(n_0/2)} \phi^{n_0/2-1} \exp(-\phi d_0/2),$$

for $\phi \in \mathbb{R}^+$ and $\Gamma(u) = (u-1)!$. Note that the mean of this prior distribution is $1/S_0$, where $S_0 = d_0/n_0$ is the prior estimate of $V$. Thereupon, when $V$ is unknown, it can be shown (see chapters 2, 4 and 17 in [11] among others) that all distributions for the level parameters and forecasts are now based on t-distributions that replace the normal densities. In the sequel we denote the non-standardized t-distribution with $n$ degrees of freedom, mode $m$ and scale $C$ by $T_n[m, C]$, with probability density function given by:

$$p(u) \propto \Big\{ n + \frac{(u-m)^2}{C} \Big\}^{-(n+1)/2}.$$

Therefore, in order to specify our initial prior beliefs, in addition to setting the prior parameter values for the vector

$\theta_t$ we need to indicate values for $n_0$ and $d_0$ for the distribution of $V$. However, in our particular case the observational variance is not constant through time. On the contrary, it varies stochastically and unpredictably. Therefore, as in chapter 10 of [11], in our model we suppose that $V$ is subject to some steady random disturbance over the time interval $t-1$ to $t$. Hence, we generate $\phi_t$ from $\phi_{t-1}$ using some form of a random walk as follows. Let $\gamma_t \sim B\{0.95n_{t-1}/2, 0.05n_{t-1}/2\}$ be independent of $\phi_{t-1}$, where $B$ represents the beta distribution such that

$$p(\gamma_t|I_{t-1}) \propto \gamma_t^{0.95n_{t-1}/2-1}(1-\gamma_t)^{0.05n_{t-1}/2-1}$$

for $0 < \gamma_t < 1$ with $E(\gamma_t|I_{t-1}) = 0.95$. Note that $\gamma_1, \gamma_2, \ldots$ are identically and independently distributed. Then, at any time point $t > 0$ it is straightforward to deduce that under the prior

$$(\phi_{t-1}|I_{t-1}) \sim G(n_{t-1}/2, d_{t-1}/2),$$

which is the posterior at $t-1$, the resulting distribution of $\phi_t$ is still a gamma distribution. Therefore in our model the sequence $\phi_t$ is changing from $t-1$ to $t$, for all $t > 0$, by an independent random factor $\gamma_t/0.95$. For the limiting behavior of $n_t$ and consequently $S_t$ see the following section.

For the latter unknown variance a natural thought is to consider that, between observations, the addition of the error $\mathbf{w_t}$ would lead to an additive increase of the initial uncertainty. Hence, the evolution variance is usually estimated in practice using the discounted variance learning by defining $W_t$ to be a fixed proportion of $\mathbf{C_{t-1}}$, i.e. $W_t = \delta \times \mathbf{C_{t-1}}$, for all $t > 0$, with $\delta \geq 0$, where $\mathbf{C_t} = \mathrm{Var}(\theta_\mathbf{t}|Y^t)$ is a $2\times2$ diagonal matrix. Between observations, the addition of the error $\mathbf{w_t}$ leads to an additive increase of $(100 \times \delta)\%$ of the initial uncertainty $\mathbf{C}$.

### C. Model Specification

We are now in position to define our model and the one step-ahead forecasts:

Observation:

$$Y_t = \mathbf{F_t'}\theta_\mathbf{t} + v_t, \quad v_t \sim N(0, 1/\phi_t)$$

System:

$$\theta_\mathbf{t} = \theta_\mathbf{t-1} + \mathbf{w_t}, \quad \mathbf{w_t} \sim T_{n_t-1}(\mathbf{0}, \mathbf{W_t})$$

Precision:

$$\phi_t = \gamma_t\phi_{t-1}/0.95, \quad \gamma \sim B(0.95n_{t-1}/2, 0.05n_{t-1}/2)$$

Information:

$$(\theta_\mathbf{t-1}|I_{t-1}) \sim T_{n_t-1}(\mathbf{m_{t-1}}, \mathbf{C_{t-1}})$$
$$(\theta_\mathbf{t}|I_{t-1}) \sim T_{n_t-1}(\mathbf{a_t}, \mathbf{R_t})$$

with

$$\mathbf{a_t} = \mathbf{m_{t-1}}, \mathbf{R_t} = \mathbf{C_{t-1}} + \mathbf{W_t}$$
$$(\phi_{t-1}|I_{t-1}) \sim G(n_{t-1}/2, d_{t-1}/2)$$
$$(\phi_t|I_{t-1}) \sim G(0.95n_{t-1}/2, 0.95d_{t-1}/2)$$
$$S_{t-1} = d_{t-1}/n_{t-1}$$

Forecast:

$$(Y_t|I_{t-1}) \sim T_{0.95n_{t-1}}(f_t, Q_t)$$

with

$$f_t = \mathbf{F'}\alpha_\mathbf{t}, \quad Q_t = \mathbf{F_t'R_tF_t} + S_{t-1}.$$

For more details see [11] and [12]. The posterior mean $\mathbf{m_t}$ is equal to the prior mean $\mathbf{m_{t-1}}$ plus a correction term which is proportional to the forecast error $e_t$. The adaptive coefficient $\mathbf{A_t} = \mathbf{R_tF_t}/Q_t$ controls the magnitude of the correction term which is based on the relative precisions of the prior, through $\mathbf{R_t/Q_t}$, and the likelihood from the value of $\mathbf{F_t}$. An alternative representation for $\mathbf{m_t}$ as a function of the adaptive coefficient $\mathbf{A_t}$ is given by

$$\mathbf{m_t} = \mathbf{A_t}Y_t + (1 - \mathbf{A_t})\mathbf{m_{t-1}}$$

showing that $\mathbf{m_t}$ is a weighted average of the prior estimate and the observation. The posterior precision $\mathbf{C_t^{-1}}$ is always larger than the corresponding prior $\mathbf{R_t^{-1}}$, hence the posterior for the parameter vector will never be more diffuse (less informative) than the prior. It is not hard to show that $n_t$ is $O(1)$ with $S_t$ having the form of an exponentially moving average of the standardized $e_t$; this enforces adaptation to new data by discounting the old ones as time evolves.

## III. In Vivo Tumor Growth Experiments

In this section we test the predictive performance of the suggested model using experimental data. At time $t = 0$ we start forecasting the future (one step-ahead, i.e. forecasting $Y_t$ on the basis of the information in $\mathbf{Y^{t-1}}$ for all $t \geq 0$) using the initially available information (see following subsections), and then as time passes we update this behavior and adapt towards the individual's unique characteristics.

### A. Methods and Materials

Tumors were prepared as described in previous work [15] by implanting a small piece ($1mm^3$) 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 caliper 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, we have that the volume $V$ equals $(4\pi/3)(xyz/8)$, which yields

$$V = \frac{\pi}{6}(xy)^{3/2}.$$

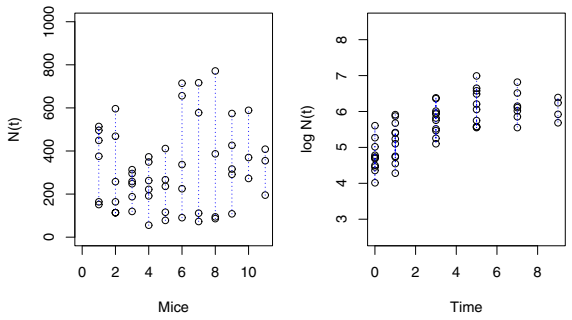


Fig. 1. Plot of $N_t$ against mice (left panel) and $Y_t$ against time (right panel) for the $HT1080$ cancer cell line.

### B. Experimental settings

Following common practice, we randomly selected one third of the observed sample as test data and the remaining two thirds served as our training data. This choice corresponds to training samples of size 4, 7, 8 and 4 for the cancer cell lines $U87$, $HT1080$, $4T1$ and $E0771$, respectively.

By reason of the large variability, we use order statistics, i.e. the sample median, rather than the sample mean to define the initial values. In particular, we construct a baseline subject having a maximum carrying capacity, i.e. $\exp(\tilde{c}_1/\tilde{c}_2)$, equal to the median estimated carrying capacity of the training set for each of the cancer cell lines, with $\tilde{c}_2$ representing the median estimated antigrowth factor of each of the training sets and $\tilde{c}_1$ being equal to the product of the log median plateau size and $\tilde{c}_2$. Consequently, we set the initial information for the model parameters as $m_0 = (\tilde{c}_1/\tilde{c}_2, \ln N_0 - \tilde{c}_1/\tilde{c}_2)$ and $C_0 = \text{diag}(0.1, 0.1)$. The choice for the a priori uncertainty is such that the log tumor growth for a given subject from $t = 0$ to $t = 1$ will have the baseline slope, estimated using the training set, with standard deviation $\sqrt{0.1}$. This value represents a large uncertainty, which reflects the large discrepancies between subjects in practice. An extensive simulation study using the training set led us to set the pilot parameters $\delta$, $n_0$ and $d_0$, used in the definition of our model, equal to $0.25$, $1$ and $0.001$ respectively. We found this choice to work well for the various types of cancer cell lines we considered, even though they had very different scales and features. Note that this choice corresponds to a small initial point estimate of the observational variance.

We assume that the only available information about each subject is the initial tumor value $N_0$ at time point $t = 0$. Then from what we have learned from the training set regarding the baseline response of mice to each cancer cell line, we predict the log volume of tumor at $t = 1$. As soon as the new observation comes in at $t = 1$, the model parameters are updated and a new prediction, that is based on the posterior distribution at $t = 1$, for the time point $t = 2$ is obtained. The sequential updating and the Bayesian one step-ahead forecasting continues for all $t \leq n - 1$. The obtained single time series for each mouse are equally spaced, i.e. the time interval from time point $t$ to $t + 1$ is equal for all $t \geq 0$.

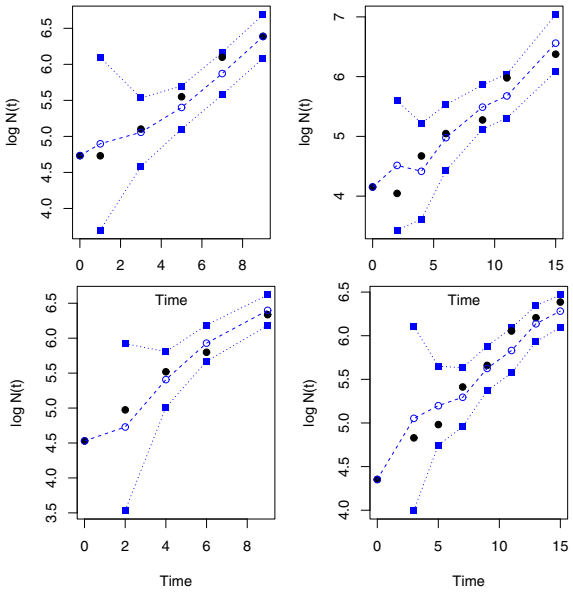


Fig. 2. From left to right: One-step ahead forecasts for a randomly selected mouse from test set corresponding to the $HT1080$, $E0771$, $4T1$ and $U87$ cancer cell line; Blue-White dots: Forecast values $f_t$, Black dots: Observations $Y_t$, Blue dotted lines and squares: $90\%$ posterior probability intervals.

### C. Results

In what follows we show only a summary of the results which were representative for the entire set of results.

The left panel of figure 1 shows the tumor volume plotted against mice for the $HT1080$ cancer cell line. There is clear evidence that the within subject standard deviation varies with subject. In the right panel we plot $Y_t$ against time for the same cancer cell line. From our data, there is no evidence against our assumptions that (i) the observational variance $V_t$ is not constant and $(ii)$ the random variable $Y_t$, at any time point $t$, is normally distributed. Figure 2 shows the one-step ahead forecasts for each subject in the test data set of each cancer cell line considered in our in vivo experiments, along with the highest posterior density intervals (HPD) as defined in Definition 2.

The overall one step-ahead predictive performance of the suggested model is satisfactory. The first prediction at time point $t = 1$ is solely based on the baseline (initial) values of the model parameters. Then, as time passes, the new observations update the model parameters and the resulting one step-ahead forecasts attain promising performance. For example, in the top right panel of figure 2, the subject's initial response to carcinogenesis is better than the baseline one of the corresponding training data set (in the sense that the tumor growth rate is smaller than the expected one). Thus at $t = 1$ the prediction overestimates the observed value $Y_1$. As soon as the new information enters our model, the model parameters

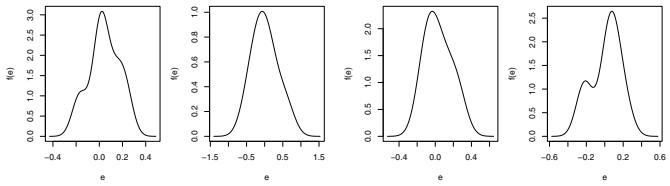

Fig. 3. Kernel density estimator for forecast error $e_t$ with plug-in bandwidth; From left to right: Randomly selected mouse - see figure 2- from test set corresponding to the $HT1080$, $E0771$, $4T1$ and $U87$ cancer cell line.

are updated towards the unique characteristics of the given subject.

The uncertainty about our predictions is explained by the HPD intervals. At $t = 0$, when no information about the individual apart from the initial tumor volume is available, we predict that the subject's response to carcinogenesis will be similar to the baseline behavior. However, we are still uncertain on our prediction, thus the HPD interval is large. The HPD interval decreases with time, because, as time evolves, our confidence on the one step-ahead predictions increases due to the new information that comes in.

Figure 3 demonstrates the kernel density estimator of the forecast error for each of the cases shown in figure 2. The smoothing parameter (bandwidth) is chosen by the plug-in approach of [16] with an adjustment parameter, due to the small sample size, equal to $3/2$. There is no evidence either against zero mean or symmetry (skewness) of the distribution of the forecast errors.

## IV. Concluding Remarks

We have proposed a dynamic linear model for tumor growth and a personalized Bayesian forecasting method to predict tumor evolution, given at most a few observed data accompanied by an a priori information about the average response, of the population in which the subject belongs, to a specific type of cancer which is under examination. From our in vivo experiments we conclude that the one-step ahead prediction performance of the suggested model is promising.

The HPD intervals quantify the aggregate effect of the heterogeneous sources of variation on the prediction uncertainty, i.e. a simple model is used to describe a complex system, small sample size, measurement error, estimation of tumor volume with only two dimensions, between and within subject variation. The HPD intervals are generally a decreasing function of time. Nonetheless, in our experiments, as commonly observed in the literature, due to the small sample size and the great uncertainty in cancer growth evolution, the HPD intervals remain relatively large by the end of the experiment to accommodate the underlying uncertainty on what will happen in the near future.

The suggested model's perspective is to incorporate the proposed methodology in adaptive optimal control using pharmacokinetic-pharmacodynamic, drug efficacy and toxicity modeling taking into account resistance.


## Acknowledgment

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


## References


[1] L. Norton, "A gompertzian model of human breast cancer growth," *Cancer research*, vol. 48, no. 24 Part 1, p. 7067, 1988.

[2] M. Chaplain, M. Ganesh, and I. Graham, "Spatio-temporal pattern formation on spherical surfaces: numerical simulation and application to solid tumour growth," *Journal of mathematical biology*, vol. 42, no. 5, pp. 387–423, 2001.

[3] L. Preziosi, *Cancer modelling and simulation*. CRC Press, 2003, vol. 3.

[4] T. Deisboeck and G. Stamatakos, *Multiscale cancer modeling*. CRC Press, 2010, vol. 34.

[5] S. Noble, E. Sherer, R. Hannemann, D. Ramkrishna, T. Vik, and A. Rundell, "Using adaptive model predictive control to customize maintenance therapy chemotherapeutic dosing for childhood acute lymphoblastic leukemia," *Journal of Theoretical Biology*, vol. 264, no. 3, pp. 990–1002, 2010.

[6] M. Tabatabai, D. Williams, and Z. Bursac, "Theoretical biology and medical modelling," *Theoretical Biology and Medical Modelling*, vol. 2, p. 14, 2005.

[7] W. Eby, M. Tabatabai, and Z. Bursac, "Hyperbolastic modeling of tumor growth with a combined treatment of iodoacetate and dimethylsulphoxide," *BMC cancer*, vol. 10, no. 1, p. 509, 2010.

[8] R. Martin, "Optimal control drug scheduling of cancer chemotherapy," *Automatica*, vol. 28, no. 6, pp. 1113–1123, 1992.

[9] I. Bassukas, "Comparative gompertzian analysis of alterations of tumor growth patterns," *Cancer research*, vol. 54, no. 16, p. 4385, 1994.

[10] A. Pole, M. West, and J. Harrison, *Applied Bayesian forecasting and times series analysis*. Chapman & Hall/CRC, 1994, vol. 1.

[11] M. West and J. Harrison, *Bayesian forecasting and dynamic models*. Springer Verlag, 1997.

[12] J. Durbin, S. Koopman, and A. Atkinson, *Time series analysis by state space methods*. Oxford University Press Oxford, 2001, vol. 15.

[13] C. Chatfield, *Time-series forecasting*. CRC Press, 2001.

[14] W. Bolstad, *Introduction to Bayesian statistics*. Wiley-Ieee, 2004.

[15] T. Stylianopoulos, J. Martin, V. Chauhan, S. Jain, B. Diop-Frimpong, B. Smith, F. Castillo, F. Hornicek, Y. Boucher, L. Munn, and R. Jain, "Growth-induced solid stress in murine and human tumors: Causes, consequences and remedies," *Under review - Nature Medicine*, 2012.

[16] S. Sheather and M. Jones, "A reliable data-based bandwidth selection method for kernel density estimation," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 683–690, 1991.