Annals of Biomedical Engineering, Vol. 42, No. 5, May 2014 (© 2014) pp. 1095–1111
DOI: 10.1007/s10439-014-0975-y

BMES BIOMEDICAL ENGINEERING SOCIETY™ www.bmes.org

Multiprocess Dynamic Modeling of Tumor Evolution with Bayesian Tumor-Specific Predictions

ACHILLEAS ACHILLEOS,[1] CHARALAMBOS LOIZIDES,[1] MARIOS HADJIANDREOU,[1] TRIANTAFYLLOS STYLIANOPOULOS,[2] and GEORGIOS D. MITSIS[1,3]

[1]KIOS Research Center for Intelligent Systems and Networks, Department of Electrical and Computer Engineering, University of Cyprus, P.O. Box 20537, Kallipoleos 75, Nicosia, Cyprus; [2]Department of Mechanical and Manufacturing Engineering, University of Cyprus, P.O. Box 20537, Kallipoleos 75, Nicosia, Cyprus; and [3]Department of Bioengineering, McGill University, 817 Sherbrooke St. W., Macdonald Engineering Building, Room 270, Montreal, QC H3A 0C3, Canada

(*Received 29 July 2013; accepted 9 January 2014; published online 1 February 2014*)

Associate Editor Konstantinos Konstantopoulos oversaw the review of this article.

Abstract—We propose a sequential probabilistic mixture model for individualized tumor growth forecasting. In contrast to conventional deterministic methods for estimation and prediction of tumor evolution, we utilize all available tumor-specific observations up to the present time to approximate the unknown multi-scale process of tumor growth over time, in a stochastic context. The suggested 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. Inference can be carried out using the full, possibly multimodal, posterior, and predictive distributions instead of point estimates. In our simulation study we illustrate the superiority of the suggested multi-process dynamic linear model compared to the single process alternative. The validation of our approach was performed with experimental data from mice. The methodology suggested in the present study may provide a starting point for personalized adaptive treatment strategies.

Keywords—Computational oncology, Tumor-specific statistical modeling, Bayesian forecasting, Gompertz growth law.

INTRODUCTION

Cancer arises through a very complex, multi-step process that involves a large degree of genetic instability, leading to the initiation of biological processes essential for tumor initiation, development, and maintenance. The quantitative characteristics of tumor growth, from a macroscopic perspective, have been extensively studied over the past decades. The development of mathematical models describing the evolution of a tumor over time has had a significant effect on understanding biological growth dynamics,[6] the evolution of resistance to anticancer therapy and the design of optimal control strategies through constrained optimization.[24] Even a quick look at the literature reveals a large number of growth models, either deterministic or probabilistic/stochastic, such as compartmental ordinary differential equations, cellular automata, mechanical models, and many more.[1,2,4,6,7,9,10,17,19–23,25–28] Some recent examples include Bodnar *et al.*,[4] who proposed a modified delayed Gompertz model with an extra term describing treatment effect; Stylianopoulos *et al.*[27] and MacLaurin *et al.*[19] have developed deterministic mechanical models for tumor growth. The former takes into account growth-induced, externally applied and fluid pressure and the latter models the buckling of capillaries in cancer tumours using nonlinear solid mechanics. Hadjiandreou and Mitsis[14–16] have suggested compartmental differential equation models that take into account drug toxicity and resistance and have used these to design optimal therapy strategies. Alemani *et al.*[2] used a hybrid approach consisting of cellular automata and the lattice Boltzmann method to model tumor growth, nutrient diffusion, and immune competition. While such modeling techniques[2,19,27] are undeniably very useful, they are often computationally expensive and both analytical solutions and parameter estimation are typically infeasible. Another limitation of such approaches is that individualized tumor growth modeling is very difficult to implement when only macroscopic measurements are available, as is often the case in practice. The more

Address correspondence to Georgios D. Mitsis, KIOS Research Center for Intelligent Systems and Networks, Department of Electrical and Computer Engineering, University of Cyprus, P.O. Box 20537, Kallipoleos 75, Nicosia, Cyprus. Electronic mail: mitsis.georgios@ucy.ac.cy

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 may arise. 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 behavior.[24] However, the performance of such models depends crucially on the reliable estimation of their unknown parameters. For animal models, these parameters are typically estimated by carrying out experiments where the tumor size is observed and recorded across time for each subject. Based on the observations, parameter estimates for each subject may be obtained, since each subject has a unique response to carcinogenesis that dictates the assumption of individualized models. Unlike other diseases,[18] this approach is rarely feasible for models describing tumor growth for human patients. Thus, because of the small sample size, any kind of model estimates are difficult to obtain and bound to have small precision (large variability), yielding results that are characterized by considerable uncertainty. Hence, even though theoretically attractive, the aforementioned models for tumor growth dynamics should be improved in order for them to be directly applicable for either forecasting or personalized optimal therapy planning.

In the above context, motivated by the urgent need for tumor-specific treatment strategies, we propose an approach that addresses some of the important shortcomings of existing growth models. Our approach is applicable in cases where only a handful of macroscopic measurements (tumor volume) may be available, as is often the case in practical situations, in contrast to more detailed mechanistic models such as, e.g., cellular automata models. Specifically we hypothesize that, 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. In the next section we discuss the growth model and the suggested statistical framework. In the "Results" section, we present the results from both extensive simulations and the application of the model to experimental data. A discussion on the main findings of our study and an outline of the conclusions and prospects of the suggested methodology follows.

MATERIALS AND METHODS

The Growth Law

We are interested in modeling and forecasting a univariate time series, i.e., tumor volume across time. In the following, we consider a model that obeys the Gompertz-law,[20,24] 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

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

Different values of the positive real valued parameters c_1 and c_2 can be used to determine the evolution of different tumor types,[22] 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 biochemical factors, as well as the spatial and temporal heterogeneity of tumors, it is quite common to observe discrepancies within and between subjects and/or tumors as well as 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.

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, e.g., an expert opinion that may alter our beliefs on how the series Y_{t+k}, for $k > 0$, will evolve.

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

$$p(\boldsymbol{\theta}|\mathcal{D}) = \frac{p(\mathcal{D}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{\int_\Theta p(\mathcal{D}|\boldsymbol{\theta}')p(\boldsymbol{\theta}')\mathrm{d}\boldsymbol{\theta}'},$$

$$\text{i.e., posterior} = \frac{\text{Likelihood} \times \text{Prior}}{\text{Normalizing Constant}},$$

where $\boldsymbol{\theta}$ is some unknown parameter vector of interest and $\mathcal{D}$ is the available data. Thus, in order to have an explicit expression of the posterior, the integral in the denominator must be analytically calculated. One way to avoid the analytic calculation of the integral and still obtain the exact posterior distribution is to use conjugate priors. This can be done by choosing the prior, $p(\boldsymbol{\theta})$, such that it belongs to the same family of distributions as the likelihood.[5] Note, however, that the analytical calculation is not always feasible, even in the case of a conjugate prior.

An appropriate dynamic linear model[11,29] (DLM) describing the progression of Eq. (1) over time is defined by $Y_t = \alpha_{1,t} + \alpha_{2,t} X_t + v_t$, with $X_t = \exp(-c_2 t)$ and c_2 fixed, $v_t \sim N(0, V_t)$, $\alpha_{1,t} = \alpha_{1,t-1} + w_{1,t}$, $\alpha_{2,t} \sim (0, V_t)$, $\alpha_{1,t} = \alpha_{1,t-1} + w_{1,t}$, $\alpha_{2,t} = \alpha_{2,t-1} + w_{2,t}$, and $\mathbf{w}_t \sim N(0, \mathbf{W}_t)$. 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 the superscript t denotes the set of all observed data up to and including time point t, e.g., $\mathbf{Y^{t-1}} = \{Y_0, Y_1, \ldots, Y_{t-1}\}$.

The updating recurrence relationships for the parameters of our model and the one-step ahead forecast error $e_t = Y_t - E(Y_t|\mathbf{Y}^{t-1})$ are derived using the concept of conditional independence (the future is independent of the past, given the present) and the Bayes theorem. 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 $\boldsymbol{\theta}$ of size a with respect to $p(\boldsymbol{\theta})$ if the conditions: (i) $P(\boldsymbol{\theta} \in \boldsymbol{R_a}) = \boldsymbol{a}$ and (ii) $p(\boldsymbol{\theta_1}) \geq p(\boldsymbol{\theta_2})$ for all $\boldsymbol{\theta_1} \in \boldsymbol{R_a}$ and $\boldsymbol{\theta_2} \notin \boldsymbol{R_a}$, hold.[8,29]

Estimating the Variances

Firstly, let us assume that the observational variance is constant, i.e., $V_t = V$. Then we can address the uncertainty on V by using standard Bayesian conjugate analysis.[7,11] 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 $1/S_0$, where $S_0 = d_0/n_0$, is a prior point estimate of V. When V is unknown, it can be shown[29] 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. We will use the notation $T_n[m, C]$ to denote the non-standardized t-distribution with n degrees of freedom, mean m, and scale parameter C.

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 $E(\phi_{t-1}|I_{t-1}) = E(\phi_t|I_{t-1})$. Now, we set $\phi_t = \gamma_t \, \phi_{t-1} \, / \, 0.95$, where $\gamma_t \sim B(0.95\, n_{t-1}/2, 0.05\, n_{t-1}/2)$ 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 covariance matrix $\mathbf{W}_t$ using discounted variance learning, i.e., by defining $\mathbf{W}_t$ to be a fixed proportion of $\mathbf{C}_{t-1} = Var(\boldsymbol{\theta_{t-1}}|I_{t-1})$. Particularly, we use $\mathbf{W}_t = \delta \times \mathbf{C}_{t-1}$, for all $t > 0$, with $\delta \geq 0$. This results in an increase in the variance of the prior distribution $p(\boldsymbol{\theta_t}|I_{t-1})$ at time point t, compared to the posterior $p(\boldsymbol{\theta_{t-1}}|I_{t-1})$ at time $t-1$, by a factor $1 + \delta$.

Model Specification

Let $\mathbf{F}_t = (1, X_t)'$. After discussing the choice of distributions for the variance components and the ramifications on the other distributions,[11,29] we can now write our model and the one step-ahead forecast distributions as follows:

Observation: $Y_t = \mathbf{F}_t' \boldsymbol{\theta_t} + v_t, \quad v_t \sim N(0, 1/\phi_t)$
State: $\boldsymbol{\theta_t} = \mathbf{G}\boldsymbol{\theta_{t-1}} + \mathbf{w}_t, \quad \mathbf{w}_t \sim T_{n_{t-1}}(0, \mathbf{W}_t)$
Precision: $\phi_t = \gamma_t \, \phi_{t-1}/0.95, \quad \gamma_t \sim B(0.95_{n_{t-1}}/2, 0.05_{n_{t-1}}/2)$
Forecast: $Y_t|I_{t-1} T_{0.95 n_{t-1}}(f_t, Q_t)$

with $f_t = \mathbf{F}_t' \mathbf{a}_t$ and $Q_t = \mathbf{F}_t' \mathbf{R}_t \mathbf{F}_t + S_{t-1}$. We also have that

$$(\boldsymbol{\theta_{t-1}}|I_{t-1}) \sim T_{n_{t-1}}(\mathbf{m}_{t-1}, \mathbf{C}_{t-1})$$

$$(\boldsymbol{\theta_t}|I_{t-1}) \sim T_{n_{t-1}}(\mathbf{a}_t, \mathbf{R}_t)$$

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

$$(\phi_t|I_{t-1}) \sim G(0.95 n_{t-1}/2, 0.95 d_{t-1}/2)$$

with $\mathbf{a}_t = \mathbf{G}_t \mathbf{m}_{t-1}, \mathbf{R}_t = \mathbf{G}_t \mathbf{C}_{t-1} \mathbf{G}_t' + \mathbf{W}_t$ and $S_{t-1} = d_{t-1}/n_{t-1}$. The sequential updating of our model is based on the following recurrence relationships:

$$(\boldsymbol{\theta_t}|I_t) \sim T_{n_t}(\mathbf{m}_t, \mathbf{C}_t)$$

$$(\phi_t|I_t) \sim G(n_t/2, d_t/2)$$

using $\mathbf{m}_t = \mathbf{a}_t + \mathbf{A}_t e_t, \mathbf{C}_t = (S_t/S_{t-1})(\mathbf{R}_t - \mathbf{A}_t \mathbf{A}_t' Q_t), n_t = 0.95 n_{t-1} + 1, d_t = 0.95 d_{t-1} + S_{t-1} e_t^2/Q_t, S_t = d_t/n_t, e_t = Y_t - f_t$ and $\mathbf{A}_t = \mathbf{R}_t \mathbf{F}_t / Q_t$.

For the above to be consistent with the model we describe in "The Gaussian Dynamic Linear Model" section, we need to set $\mathbf{G} = I_2$, the 2×2 identity matrix, for all t. It is worth noting that, when $\mathbf{G}$ is the identity matrix, the mean, $\mathbf{m}_t$, of the posterior distribution of $\boldsymbol{\theta}_t$ is equal to the mean, $\mathbf{m}_{t-1}$ of its prior distribution plus a correction term which is proportional to the forecast error $e_t = Y_t - f_t$. The first obvious conclusion is that the greater the error, the greater the difference between the two posterior means. The matrix $\mathbf{A}_t = \mathbf{R}_t\mathbf{F}_t/Q_t$ controls the magnitude of the correction term at each time point. It can be seen as a measure of the posterior precision of $\boldsymbol{\theta}_t$ relative to the posterior prediction of the forecast, with the adaptive coefficient $\mathbf{A}_t$ having less impact when the posterior of $\boldsymbol{\theta}_t$ is more precise. Also of note is that the posterior precision $\mathbf{C}_t^{-1}$ is always larger than the corresponding prior $\mathbf{R}_t^{-1}$, 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.

Multi-Process Modeling

Whereas any single DLM defines a process model, a combination of several DLMs defines a multi-process model (MPDLM), also known as mixture model.[29] In many occasions, in order to satisfy the local linearity assumptions of the DLM, we conventionally assume that some quantities, denoted by $\mathbf{c}$, which would be otherwise considered random, are fixed and known. For these, we usually use a plug-in estimate in the model that is obtained from the training data. Multi-process modeling provides us the means to improve on the plug-in approach when we deal with this kind of situation.

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.[29]

Let Ξ_t be a vector of random quantities of interest; for example, the state vector $\boldsymbol{\theta}_t$, the future observation Y_{t+l}, for $l > 0$, *etc*. Now consider that we fit more than one such DLMs, where for each we choose a different value for $\mathbf{c}$. Inference about Ξ_t in each DLM is based on $p(\Xi_t|\mathbf{c}, I_t)$; thus many such densities may exist, one for each $\mathbf{c} \in \mathbb{A}$. 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:

$$p(\Xi_t|I_t) = \int_{\mathbb{A}} p(\Xi_t|\mathbf{c}, I_t)p(\mathbf{c}|I_t)d\mathbf{c}. \tag{2}$$

In practice, we can approximate the integral in Eq. (2) with a finite sum. Specifically, we choose a fixed and finite grid of points for the parameter vector $\mathbf{c}$, that is $\{\mathbf{c}_1, \ldots, \mathbf{c}_k\}$. 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. In our case the dimension of $\mathbf{c}$ is one, thus k can be chosen large enough, so that the points cover the parameter space well, more easily.

Thereafter, c is considered a discrete random variable, the posterior distribution of which has a mass function p_t at time t:

$$p_t(j) = p(c_j|I_t) \quad (j = 1, \ldots, k)$$

with the initial prior distribution denoted by $p_0(j)$. Hence, the integral in Eq. (2) is approximated and replaced by the sum:

$$p(\Xi_t|I_t) = \sum_{j=1}^{k} p(\Xi_t|\mathbf{c}, I_t)p_t(j). \tag{3}$$

Considering the one-step ahead forecast problem (i.e., $\Xi_t = Y_t$), the posteriors are calculated according to:

$$p_t(j) = \kappa_t p_{t-1}(j)p(Y_t|c_j, I_{t-1}),$$

where $p(Y_t \mid c_j\ ,\ I_{t-1})$ is the one-step ahead forecast posterior density at time $t-1$. The normalizing constant κ_t can be obtained by:

$$\kappa_t^{-1} = \sum_{j=1}^{k} p_{t-1}(j)p(Y_t|c_j, I_{t-1}).$$

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 $\beta_t = \lambda\,\beta_{t-1} + w_{\beta,t}$, where $w_{\alpha,t}$ an $w_{\beta,t}$ are the evolution variances, $\lambda = \exp(-c_2)$ and $0 < \lambda < 1$. The differences compared to the model we present in "Model Specification" section is that here, $\mathbf{F_t} = (1,1)'$ while $\mathbf{G}$ is no longer the identity matrix but $\mathbf{G} = \mathrm{diag}(1,\lambda)$. 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.[29]

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.

RESULTS

Simulations

We generate data according to Eq. (1) subsequently adding a zero-mean Gaussian noise with a constant (with respect to time) signal-to-noise ratio (SNR). Here, SNR is defined as:

$$SNR(t) = \frac{\mu(t)}{\sigma(t)}, \tag{4}$$

where $\mu(t)$ is the output value of the deterministic equation (Eq. (1)) at time t after we input the values of the simulation parameters c_1, c_2, and N_0 (For all simulated datasets we use $N_0 = 110$ mm^3). Thus adding noise with a constant $SNR = C$, means that at each time point the standard deviation of the noise is $\sigma(t) = \mu(t)/C$.

We consider a $3 \times 3 \times 2 \times 2 \times 2$ grid of parameter values. Specifically, the parameters and their respective values used in our simulation study are: $c_1 = \{1.5, 1.8, 2.1\}$, $c_2 = \{0.2, 0.24, 0.28\}$, $\hat{c}_1 = \{(1+q_{11})c_1, (1+q_{12})c_1\}$, $\hat{c}_2 = \{(1+q_{21})\ c_2, (1+q_{22})c_2\}$ and $SNR = \{10, 4\}$. The values of c_1 and c_2 were chosen based on the range of parameter values obtained from the experimental data ("Experimental Data" section); thus, they are physiologically plausible but they also span a wider range of possible behaviors. Their combinations yield a wide range of curves, both in terms of growth rate and in terms of maximum carrying capacity. The quantities $\hat{c}_1$ and $\hat{c}_2$ denote the initial estimates of c_1 and c_2 respectively (in practice these could be obtained from a set of available training data). The quantities q_{ij}, $i, j = 1, 2$ are realizations from Beta distributed random variables and are included to accommodate the possibly large deviations of our initial estimates from the true underlying parameters (for a particular tumor). We use $q_{11} \sim B(10, 90)$, $q_{12} \sim B(50, 50)$, $q_{21} \sim B(30, 70)$, and $q_{22} \sim B(80,20)$. We remind that the expected value of a Beta distributed r.v. $q \sim B(a, b)$, is $E(q) = a/b$. Therefore the values of q_{ij} determine how much our initial estimates deviate from the true parameters of the underlying curve. In addition, for the MPDLM we choose the set of values for λ to be $\mathbb{L} = \{0.2, 0.4, 0.6, 0.8, 0.95\}$.

For each of the 72 parameter combinations we perform 100 realizations, where in each of these realizations we generate a different noise sample from the underlying Gaussian distribution according to the SNR and a new pair of samples (q_{1i}, q_{2j}) from their respective Beta distributions, in order to generate a new set of initial estimates This results in a total of 7200 simulated growth curves. We compare the performance of the MPDLM and the DLM as well as a naive deterministic method, where the fitted model is based on the initial estimates. For each simulation in our study we compute the mean absolute deviation (MAD) of the forecasts from the simulated data, for all three methods. For a single growth curve $y_1, y_2 \ldots y_T$ and forecasts $\hat{y}_2, \ldots, \hat{y}_T$, the MAD is defined as:

$$MAD = \frac{1}{T-1}\sum_{t=2}^{T} |y_t - \hat{y}_t| = \frac{1}{T-1}\sum_{t=2}^{T} |e_t|. \tag{5}$$

In our simulation study we use $T = 26$. The pooled MAD values from all 36 parameter sets for each SNR are presented in Fig. 1.

In terms of raw average numbers, the MPDLM outperforms (smaller MAD) both the DLM and the naive method for all 72 parameter combinations. However in some cases, the differences between the MPDLM and DLM are relatively small, while the naive method always fares much worse. In order to obtain more formal results, we compare the MAD values obtained by the three methods using Friedman's two-way analysis of variance by ranks, implementing both the omnibus test that compares all three methods and the *post-hoc* tests for pairwise comparisons.[12] The use of Friedman's test is necessitated by the fact that we calculate the quantitative response (MAD) for each method using the same sample. Therefore each sample defines a block in the test design. The three groups are defined by the independent variable/factor which in this case is the method used to obtain the MAD and has three levels, Naive, Single DLM and MPDLM. We

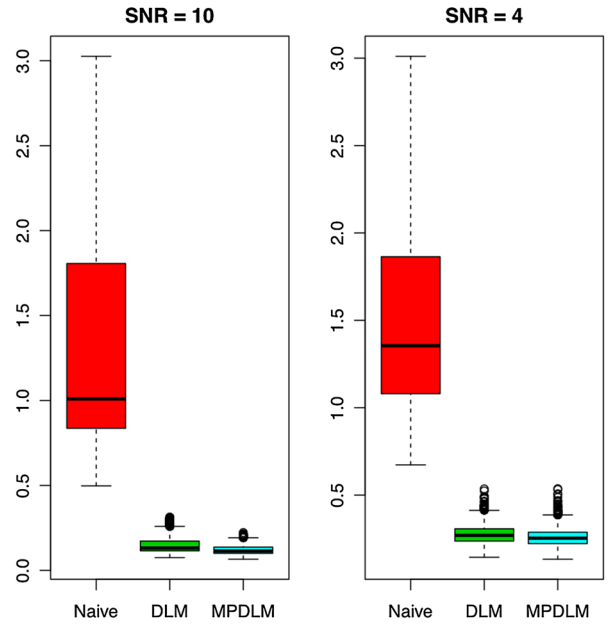

FIGURE 1. Boxplots from the pooled MADs for the two SNR levels considered. Each boxplot is obtained from 3600 realizations (MAD samples). The improvement in performance yielded by the MPDLM was found to be statistically significant ("Simulations" section).

choose the non-parametric Friedman's test over the parametric repeated measures ANOVA since the within-level distributions of the MADs appeared skewed and deviated from the Normal distribution. Additionally the assumption of sphericity is violated since differences involving MADs from the Naive method are more variable than the differences between MADs obtained from the DLM and MPDLM methods.

The DLM and the naive method were found to be statistically better than the MPDLM in none of the 72 tests we considered. The naive approach is always the worst; on the other hand the difference between the performance of the DLM and the MPDLM is non-significant in a few occasions according to the *post-hoc* tests. Specifically, those occasions occur only when we use $SNR = 4$, i.e., when we generate more noisy data. The choice of SNR appears to be the most important factor that determines whether the MPDLM is statistically better than the DLM, in terms of MAD. In all 36 parameter combinations we used with $SNR = 10$ (less noisy data) the MPDLM was significantly better than the DLM. Among the 36 parameter combinations we used with $SNR = 4$ (more noisy data) the MPDLM was significantly better than the DLM in 28 while in the remaining 8 combinations the MAD values yielded by the two methods were not statistically different. Figure 2 shows four representative realizations of our simulations, two for each SNR value, where in the two realizations with $SNR = 10$ and one realization with $SNR = 4$ the MPDLM is found to perform better than the DLM in a statistical sense, while in the other $SNR = 4$ realization, the performance of the two tests was not deemed different statistically.

The other factor that appears to affect the relative performance of the two methods is the deviation of the ratio $\hat{c}_1/\hat{c}_2$ for the initial estimates from the true parameter ratio c_1/c_2. The deviation of the ratio $\hat{c}_1/\hat{c}_2$ from its true value c_1/c_2 is determined by the quantity R_{ij} defined as follows:

$$R_{ij} := \frac{\hat{c}_1/\hat{c}_2}{c_1/c_2} = \frac{1+q_{1i}}{1+q_{2j}}, \quad i,j = 1,2. \tag{6}$$

The closer to 1 R_{ij} is, the closer our initial estimates are to the true ratio. Because the values of the random variables q_{ij} are different in the 100 simulations for each parameter combination, we use the estimate $\hat{R}_{ij} = E(1+q_{11})/E(1+q_{22})$, where E denotes expected value, when we compare the obtained MAD values for different parameter combinations. The MPDLM performs much better compared to the DLM when the initial estimate of the ratio is not very accurate. For a fixed i and j for the variables (q_{1i},q_{2j}), the values of the remaining three simulation parameters (c_1, c_2, SNR) yield 18 combinations. Table 1 shows the relative performance of the MPDLM compared to the DLM for all pairs $(i,j) \in \{1,2\} \times \{1,2\}$. It is evident that the further $\hat{R}_{ij}$ is from 1, the better the relative performance of the MPDLM compared to the DLM becomes.

Representative results for the time evolution of the estimated model coefficients $\hat{\boldsymbol{\theta}}_t = (\hat{\alpha}_t, \hat{\beta}_t)$ are given in Fig. 3). When the MPDLM is used, the estimated coefficients are more flexible and are able to adjust and converge close to their true values faster. The top panels of Fig. 3 correspond to a case where the MPDLM outperforms the DLM, while the bottom panel corresponds to a case where the two methods are not statistically different. The real values of the parameters are calculated using the inverse transformations $\alpha_t = c1/c2$ and $\beta_t = -(c1/c2 - \log(N_0))\exp(-c_2 t)$. Therefore, using the MPDLM framework allows us to be more lenient with the initial values and still obtain reliable results; this is particularly evident in the case of the top panel. This is due to the fact that the MPDLM has the ability to adjust and choose a combination from a predetermined set of single-process DLMs, as shown in Fig. 4, where the one-step ahead forecast, the model weights as well as the convergence of the model parameters is given for the case presented in the top panel of Fig. 3. This is a case where the initial estimates are far from their true values. In the early stage, the component model with $\lambda = 0.2$ has the largest weight, then the

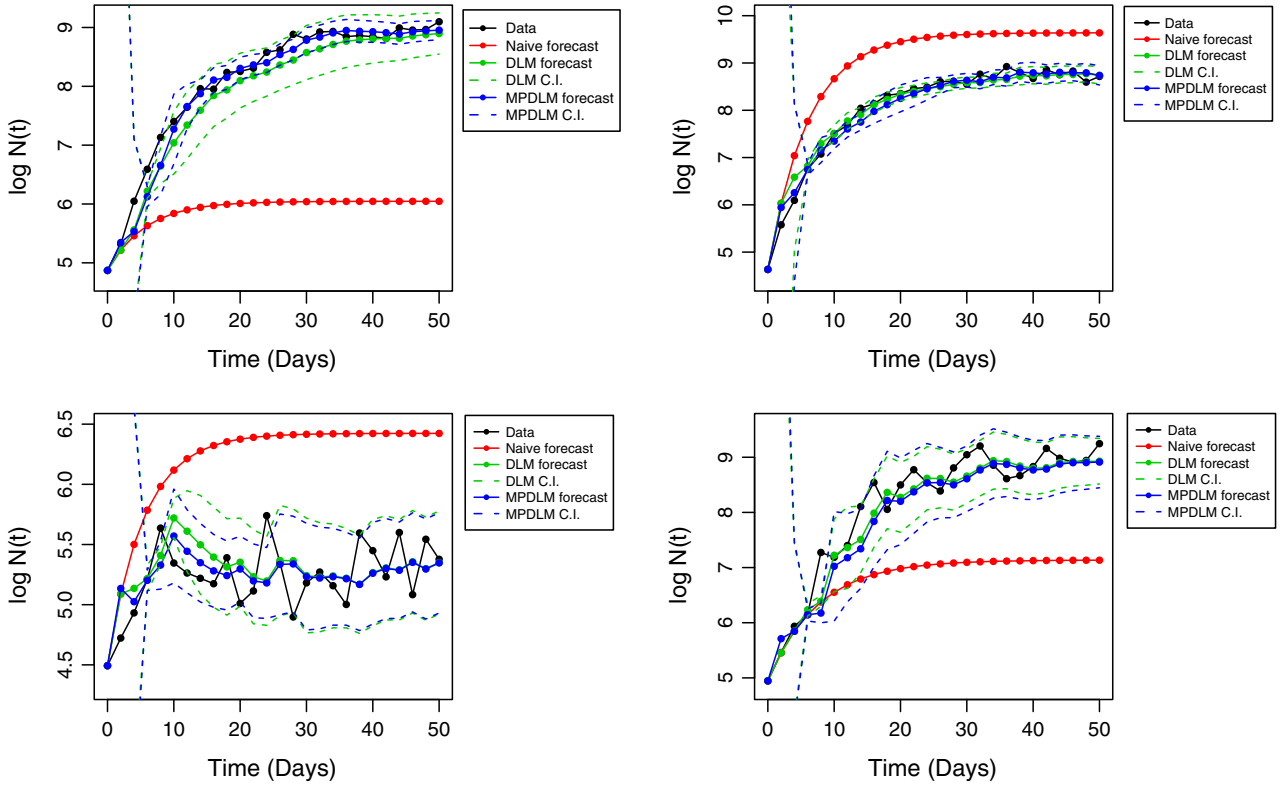

FIGURE 2. Four realizations from the simulation study with the forecasts obtained from the three methods and their corresponding HPD intervals, where applicable. Top panels: *SNR* = 10, Bottom panels: *SNR* = 4. The top and bottom-left panels correspond to parameter combinations where the MPDLM provided statistically better performance than the DLM while in the bottom-right panel the two methods yielded MAD values that were found not to be statistically different. The naive method has the worst performance in all cases. (The values of the parameters (c_1, c_2, q_{1i}, q_{2j}) are, from left to right and top to bottom: (1.8, 0.2, q_{11}, q_{22}), (2.1, 0.24, q_{12}, q_{21}), (1.5, 0.28, q_{12}, q_{21}), and (1.8, 0.2, q_{11}, q_{21})).

TABLE 1. Effect of the initial estimates ratio $\hat{c}_1/\hat{c}_2$ on the relative performance of the MPDLM compared to the DLM.

	q_{11} $(E = 0.1)$	q_{12} $(E = 0.5)$
q_{21} $(E = 0.3)$	$\hat{R}_{11} \approx 0.85$ [12/18]	$\hat{R}_{12} \approx 1.15$ [18/18]
q_{22} $(E = 0.8)$	$\hat{R}_{21} \approx 0.61$ [18/18]	$\hat{R}_{22} \approx 0.83$ [16/18]

The values inside the square brackets correspond to the occasions that the MPDLM statistically outperforms the DLM out of the 18 possible parameter combinations for (c_1, c_2, *SNR*).

model with $\lambda = 0.6$ becomes the one with the largest posterior weight and finally, the model with $\lambda = 0.8$ dominates from $t = 18$ and onwards. This alternation between different models, helps the MPDLM achieve the flexibility and fast convergence we see in Fig. 3—top panel. Overall, our simulation study shows that the MPDLM framework performs better than the DLM in a statistical sense.

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, which are selected according to two different scenarios, 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.

Experimental Methods and Materials

Tumors were prepared as described in previous work[26] 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 *U*87 (flank, number of mice subjects $n_s = 6$), human fibrosarcoma *HT*1080 (flank, $n_s = 11$), murine

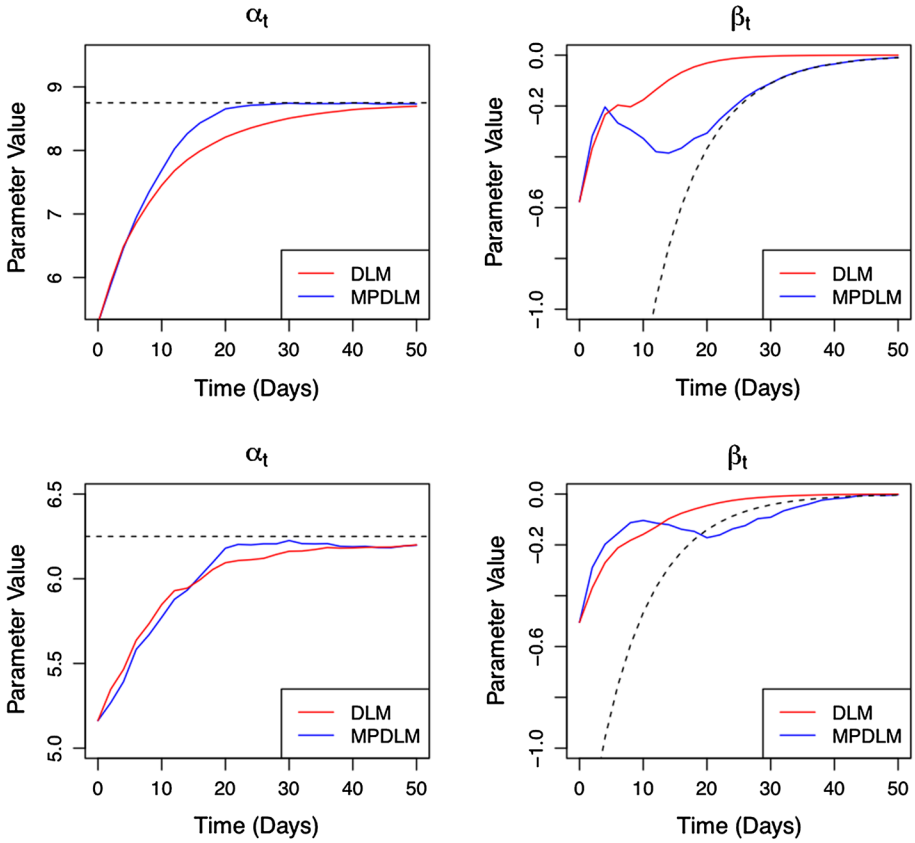

FIGURE 3. Evolution of the model parameters α_t and β_t for two representative cases. We show the average behavior of these parameters across all 100 simulated realizations from a given parameter combination. Top panels: The estimates correspond to a case where the MPDLM outperforms the DLM in a statistical sense. Bottom line panels: The estimates correspond to a case where the performance of the two methods was not found to be different in a statistical sense. (The set of simulation parameters (*SNR*, c_1, c_2, q_{1i}, q_{2j}) for the two examples are: (10, 2.1, 0.24, q_{11}, q_{22}) on the top panel and (4, 1.5, 0.24, q_{11}, q_{21}) on the bottom panel).

mammary adenocarcinoma 4*T*1 (mfp, $n_s = 12$), and murine mammary adenocarcinoma *E*0771 (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, the tumor volume V equals

$$V = \frac{4\pi}{3}\frac{xyz}{8} = \frac{\pi}{6}(xy)^{3/2}.$$

In Fig. 5 we present the course of tumor growth for all available subjects, where the tumor volume $N(t)$ is given in mm^3 and time is given in days. Each curve describes the growth of a tumor in a single mouse.

Choice of the Initial Values and Hyperparameters

Consider that we have tumor progression data that come from individuals from the population of interest, having the same type of cancer, denoted by $Y_{i,1}, \ldots, Y_{i,j_i}$, with $i = 1, \ldots, n_r$, 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 (Eq. (1)). The resulting estimates of $c_{1,i}$ and $c_{2,i}$ are denoted by $\tilde{c}_{1,i}$ and $\tilde{c}_{2,i}$ respectively for $i = 1, \ldots, n_r$. Then, we construct an "average" subject having a maximum carrying capacity equal to $\tilde{\alpha} = \frac{\tilde{c}_1^{\text{med}}}{\tilde{c}_2^{\text{med}}}$ where $\tilde{c}_2^{\text{med}} = \text{median}_{i=1,\ldots,n_r}\{\tilde{c}_{2,i}\}$ and $\tilde{c}_1^{\text{med}} = \tilde{c}_2^{\text{med}} \times \tilde{\alpha}^{\text{med}}$, with $\tilde{\alpha}^{\text{med}} = \text{median}_{i=1,\ldots,n_r}\left\{\frac{\tilde{c}_{1,i}}{\tilde{c}_{2,i}}\right\}$. 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 $\mathbf{m}_0 = (\tilde{c}_1^{\text{med}}/\tilde{c}_2^{\text{med}}, \log(N_0) - \tilde{c}_1^{\text{med}}/\tilde{c}_2^{\text{med}})'$. For a single process DLM we can choose $\lambda = \exp(-\tilde{c}_2^{\text{med}})$, while for a MPDLM with k component models we use $\mathbb{L} = \{\lambda_{(1)}, \ldots, \lambda_{(k)}\}$, 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 $Y_{i,j}, i = 1, \ldots, n_r$ and $j = 1, \ldots, j_i$, respectively.

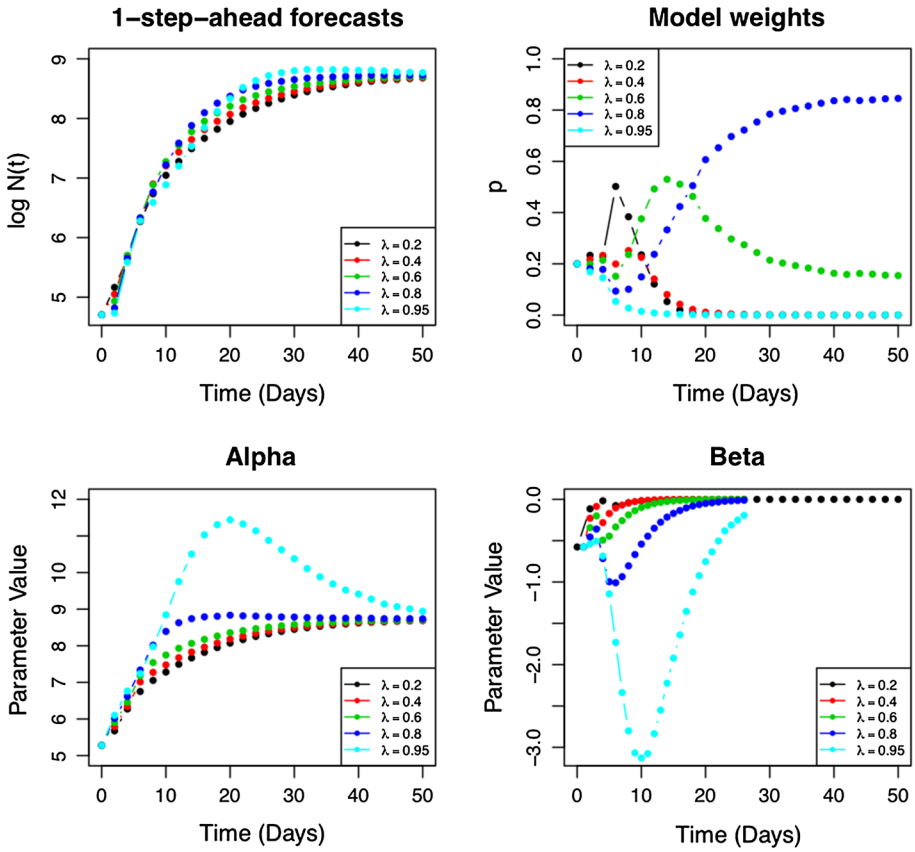

FIGURE 4. Model characteristics for the component models of the MPDLM—representative simulated case. Top left: One-step-ahead-forecasts for the five component models. Top right: Posterior weights $p_t(j) = p(\lambda_j \mid I_t)$. Bottom line panels: Evolution of the parameters α_t and β_t for the five component models. (The simulation parameters for this example are $(SNR, c_1, c_2, q_{1i}, q_{2j}) = (10, 2.1, 0.24, q_{11}, q_{22})$).

In the "Simulations" section, it was shown that the MPDLM framework yields better performance when the values of the initial estimates for the unknown model parameters deviate from their true values, allowing us to be more lenient regarding the selection of these values. In our experiments, we treated the values for the hyperparameters n_0, d_0, C_0, and δ as tuning parameters. Particularly, in order to assess the effects of these tuning parameters on the predictive performance of the model, a Monte Carlo simulation experiment was performed using a factorial design for all possible combinations of parameter values from pre-specified grids. The predictive performance of each model was assessed by the following penalty criterion:

$$L(M) = \prod_{t=1}^{n} p(Y_t | I_{t-1})$$

which is the product of the one step-ahead forecast densities evaluated at the actual observation, where n is the observed sample size. Note that the larger the value of the penalty criterion L, the better the predictive performance of the model M. The values corresponding to the best overall (for all cancer cell lines) performance were chosen. Of course, using different values for each cancer cell line would improve the performance of both the single and multiprocess DLMs even further but, at the same time, observation-specific bias should also be considered.

Data Analysis and Validation

In order to assess the predictive performance of the multi-process DLM we split the data in each of the U87, HT1080, 4T1, and E0771 cell lines into training and testing sets. Initially, 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 set. Since the testing set is randomly selected from the total number of mice used for each cancer cell line in the study, we repeat the allocation ${}_{n_s}C_{n_T}$ times to account for the effect of randomness, where ${}_uC_v$ denotes the total number of ways of picking v unordered outcomes from u possibilities. Moreover, we generated the training set by following a different procedure, i.e., by using only one mouse in this set for all possible combinations of subjects and tumor types in order to

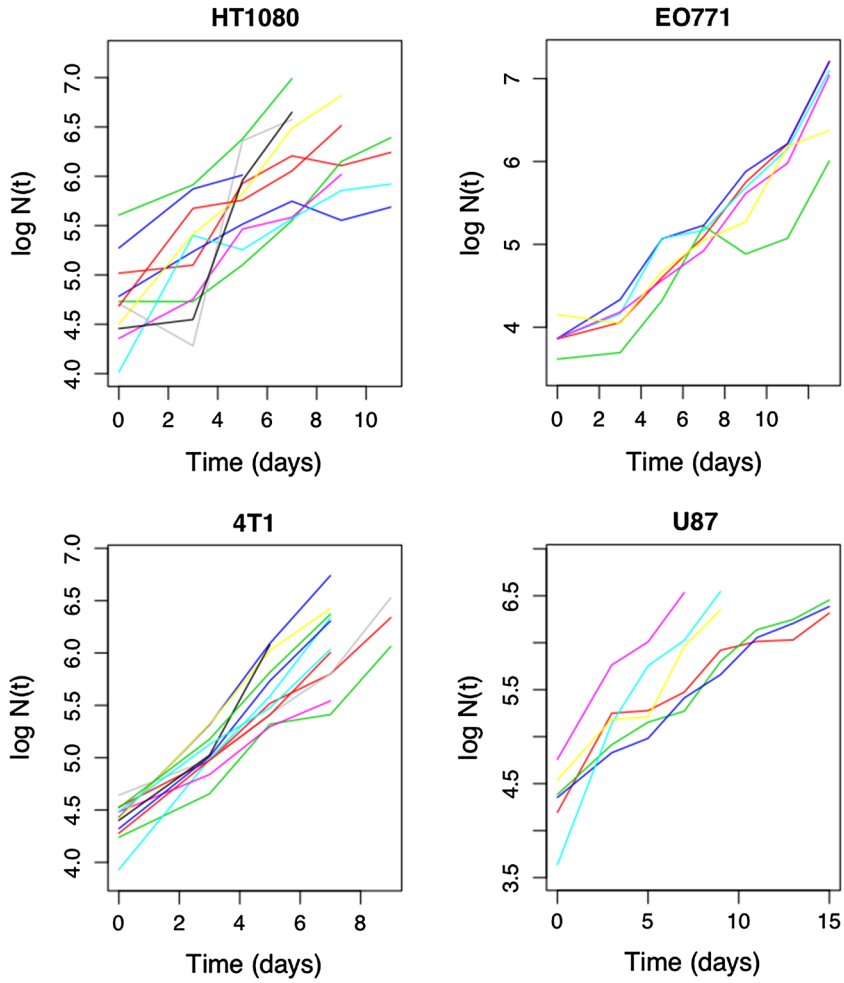

FIGURE 5. Experimental data for the course of tumor growth. Each curve describes growth of a tumor (log volume) in one mouse; from top left to bottom right: 4*T*1, *EO*771, *HT*1080, and *U*87 cell line.

show that the suggested models perform well even in cases when little prior information is available.

As mentioned above, 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, and let $\lambda_{(1)}, \ldots, \lambda_{(5)}$ denote the ordered (ascending) estimated values of λ corresponding to each component model (for each cancer cell line, given the training set) using the approach described in the "Choice of the Initial Values and Hyperparameters" section. Based on a factorial design using the training set, we set the initial parameters δ, n_0, d_0, and matrix C_0 equal to $0.25, 1, 0.001$, and $\mathrm{diag}(0.01, 0.01)$ 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.

The comparisons are based on the MAD of the forecasts obtained from each method (Figs. 6 and 7). In order to formally assess the differences in MAD, we used the non-parametric Friedman's two-way ANOVA, both for omnibus and *post-hoc* pairwise comparisons (Table 2). The results indicate that the MPDLM is statistically superior to both the DLM and Naive methods in all cases but the EO771, where the difference is not statistically significant, even though in the case where we use one training mouse the corresponding p-value is marginally non-significant ($p = 0.058$—Table 2, line 7).

In addition to the boxplots presenting the full results, we also show some representative cases, visualizing the one-step ahead predictive performance of the MPDLM for the various cancer types we considered in our study. We overlay the performance of the naive and single DLM methods as well (Figs. 8–11). The red dots (red dashed lines) denote the one-step ahead predictions (90% HPD intervals) of the MPDLM. Analogously, the blue dots (blue dashed lines) denote the one-step ahead predictions (90% HPD intervals) of the single process DLM. The deterministic predictions of the naive

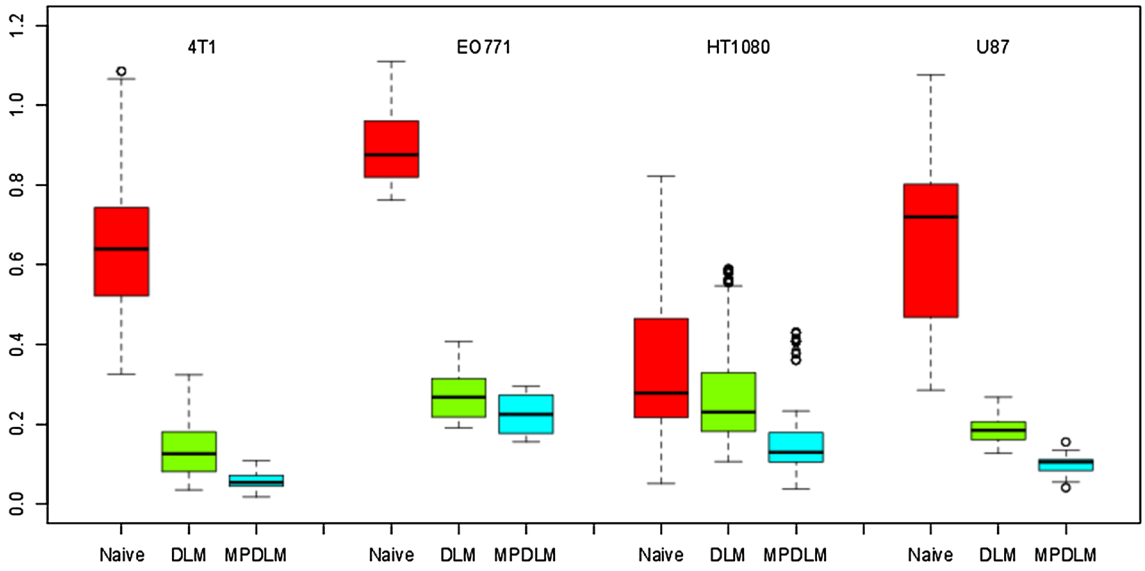

FIGURE 6. Boxplots of mean absolute deviation (MAD) values for each cancer cell line using the Naive, single DLM, and MPDLM methods. Training set consists of 2/3 of mice.

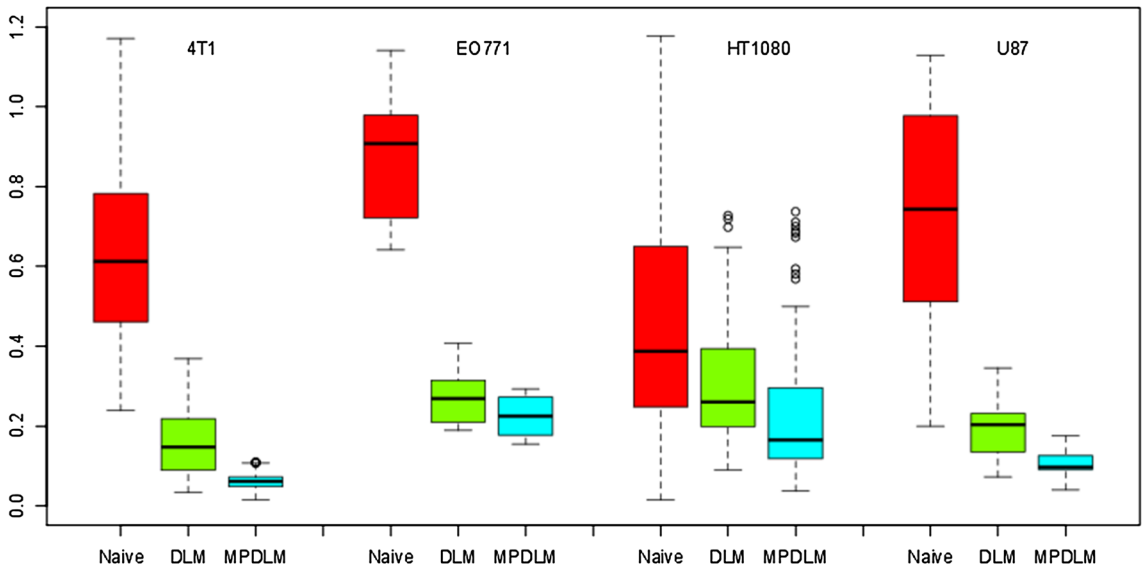

FIGURE 7. Boxplots of mean absolute deviation (MAD) values for each cancer cell line using the Naive, single DLM, and MPDLM methods. Training set consists of only one mouse per line.

approach are denoted by the light-blue dots. The black dots represent the measured tumor volume at each time point. The right top panels show the posterior probabilities (weights) for the component models of the MPDLM as time evolves. The bottom panels show the posterior forecast distributions for the MPDLM (left) and single-process DLM (right) as time evolves. Clearly, the MPDLM is initially multimodal (it can have up to five modes) and then as new information about the specific tumor becomes available, the component model or the combination of component models that better describe the evolution of the observed data gain more weight (posterior probability). On the other hand, the forecast distributions of the DLM are unimodal t-distributions that become more diffused when uncertainty increases and steeper (concentrated around the mode) when the model prediction error is small.

The single process DLM performs better than the naive method in terms of one-step ahead predictions, but not as well as the MPDLM. For example, for the subject belonging to the testing set of the HT1080 cancer cell line (Fig. 8), the tumor growth rate is faster than the average rate. The naive approach is thus underestimating the tumor volume. The performance

TABLE 2. *P*-values for the omnibus and *post-hoc* tests (Friedman's two-way ANOVA) for the MAD values obtained from the naive, DLM, and MPDLM methods.

		Cell lines			
Training set	Test	4T1	EO771	HT1080	U87
2/3 Data	Omnibus	$<10^{-6}$	$<10^{-6}$	$<10^{-6}$	$<10^{-6}$
	Naive vs. MPDLM	$<10^{-6}$	$<10^{-6}$	$<10^{-6}$	$<10^{-6}$
	DLM vs. MPDLM	$<10^{-6}$	0.097	$<10^{-6}$	3.6×10^{-4}
	DLM vs. Naive	$<10^{-6}$	4.5×10^{-6}	8.4×10^{-4}	3.2×10^{-4}
1 Observation	Omnibus	$<10^{-6}$	$<10^{-6}$	$<10^{-6}$	$<10^{-6}$
	Naive vs. MPDLM	$<10^{-6}$	$<10^{-6}$	$<10^{-6}$	$<10^{-6}$
	DLM vs. MPDLM	$<10^{-6}$	0.052	1.8×10^{-4}	0.002
	DLM vs. Naive	$<10^{-6}$	9.6×10^{-6}	0.016	1.0×10^{-4}

The MPDLM is statistically superior to both the DLM and Naive methods in all but the EO771 cell line. Note that even in this case the difference is marginally non-significant when one subject was used in the training set.

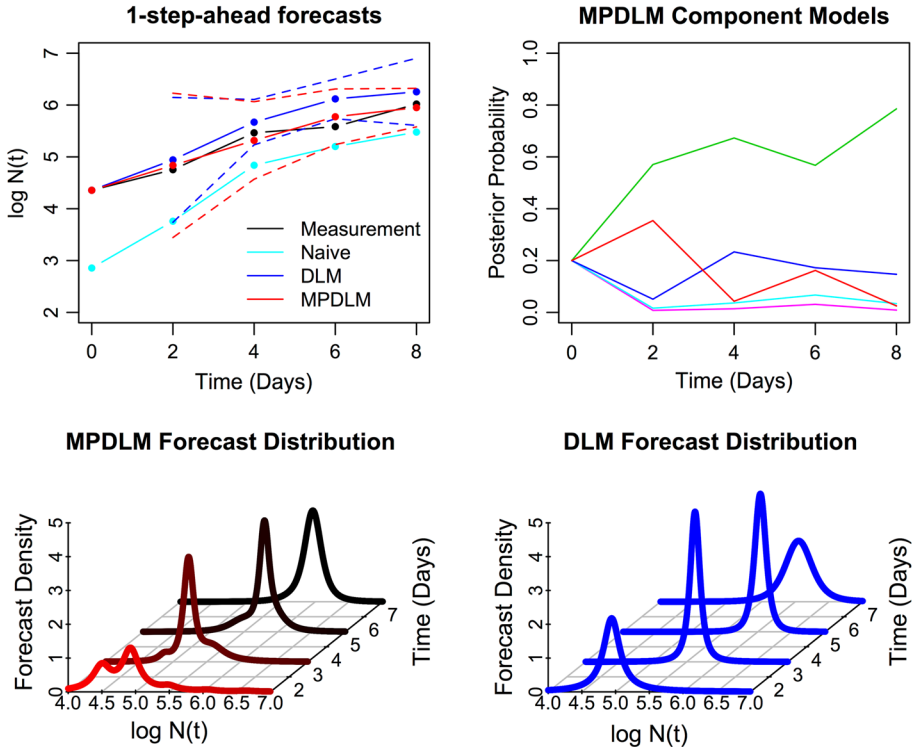

FIGURE 8. Top left panel: One-step ahead predictive performance of MPDLM, single process DLM and naive methods for a representative mouse from the test set (cell line HT1080). The dashed lines represent the respective 90% HPD intervals of DLM and MPDLM. Right top panel: Posterior weights for the component models of the MPDLM with respect to time. Red: $\lambda = 0.55$, green: $\lambda = 0.65$, blue: $\lambda = 0.75$, light-blue: $\lambda = 0.85$, pink: $\lambda = 0.95$. Bottom line panels: Forecast distributions of the MPDLM (bottom left) and single DLM (bottom right) with respect to time. Overall, the MPDLM has better performance, yielding a narrower HPD interval at the end while is also able to include all observations within its HPD interval.

of the other two methods is better but the MPDLM is more accurate—note importantly that the observation of the tumor size at the sixth day lies outside the 90% HPD intervals of the DLM. The component model corresponding to the second smallest value of λ, i.e., $\lambda_{(2)}$, receives most of the support from the data, with a posterior weight lying above 0.58 for all time points. In the examples from the cancer cell lines *E*0771 and *U*87, the component model corresponding to the largest value of λ eventually dominates its rivals (Figs. 9 and 11—top right panels), while for cancer cell line 4*T*1 the two models corresponding to the largest λ values result in the largest posterior weights (Fig. 10). This is also apparent from the bottom left panel, where the forecast distribution of the MPDLM is unimodal at all time points, apart from the first two points in time, where it is bimodal (Fig. 11). From the latter figure, it can be also seen that the MPDLM adapts faster than

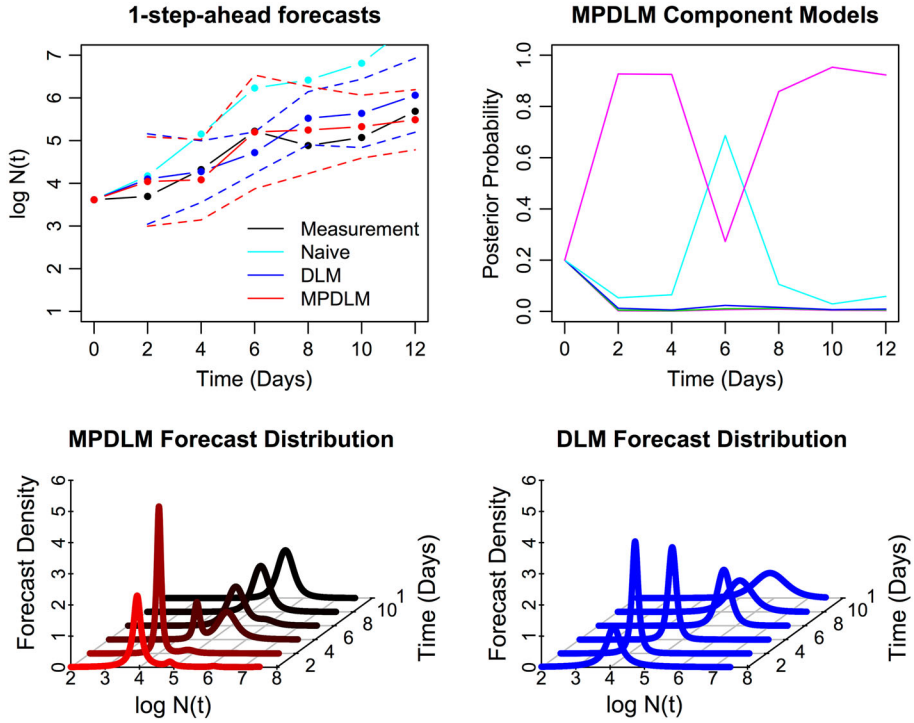

FIGURE 9. Top left panel: One-step ahead predictive performance of MPDLM, single process DLM and naive methods for a representative mouse from the test set (cell line EO771). The dashed lines represent the respective 90% HPD intervals of DLM and MPDLM. Right top panel: Posterior weights for the component models of the MPDLM with respect to time. Red: $\lambda = 0.9$, green: $\lambda = 0.9225$, blue: $\lambda = 0.945$, light-blue: $\lambda = 0.9675$, pink: $\lambda = 0.99$. Bottom line panels: Forecast distributions of the MPDLM (bottom left) and single DLM (bottom right) with respect to time. The MPDLM manages to capture all observations within its HPD interval by alternating between two of its component models.

the single process DLM to the slower growth rate of this tumor, compared to the average growth rate of the training set. The HPD intervals of the MPDLM become considerably tighter (forecast distributions are less diffused) than their DLM counterparts after a few measurements from this specific subject have been incorporated.

Finally, Fig. 12 shows the predictive performance of the three methods for another representative set of mice from all four cell lines, in the case where one mouse was used to construct the training set. The naive model, as expected, performs even worse than in the case where the training set consists of more subjects. The MPDLM performs better overall, with the MPDLM yielding both smaller MAD values (Fig. 7) but also narrower HPD intervals and/or better forecast accuracy. Specifically, in the examples we present, the MPDLM outperforms the DLM in terms of forecast precision (narrower HPD interval) in the HT1080 and U87 cell lines. In the EO771 cell line, even though the DLM yields tight HPD intervals, its forecast is not accurate, with the last three observations being marginally included or outside the HPD interval. On the other hand, the MPDLM is able to include all observations within its HPD interval. Something similar is observed in the 4T1 cell line, where a rapid reduction of the growth rate in the next to last observation, causes the DLM to make an erroneous forecast by leaving the observation outside the forecast's HPD interval. This example also demonstrates the flexibility of the MPDLM which is able to adjust quickly to changes in growth rate.

DISCUSSION

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 points 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 dynamic models has promising potential. Both the DLM and MPDLM are shown to yield superior performance to the naive approach. For both the experimental and simulation data, the MPDLM yielded better performance than the DLM in

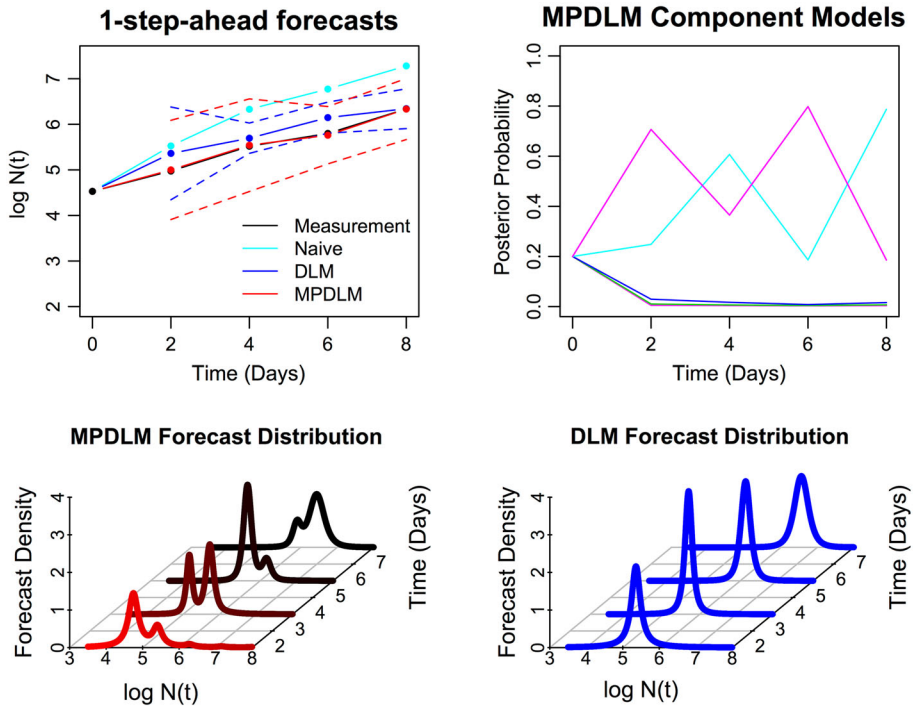

FIGURE 10. Top left panel: One-step ahead predictive performance of MPDLM, single process DLM and naive methods for a representative mouse from the test set (cell line 4T1). The dashed lines represent the respective 90% HPD intervals of DLM and MPDLM. Right top panel: Posterior weights for the component models of the MPDLM with respect to time. Red: $\lambda = 0.8$, green: $\lambda = 0.8475$, blue: $\lambda = 0.895$, light-blue: $\lambda = 0.9425$, pink: $\lambda = 0.99$; Bottom line panels: Forecast distributions of the MPDLM (bottom left) and single DLM (bottom right) with respect to time. The MPDLM demonstrates impressive accuracy *via* a bi-modal posterior forecast distribution, achieved by using two of its component models.

a statistical sense in almost all cases (Figs. 1, 6, and 7; Table 2).

The uncertainty of our predictions is quantified 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 will be similar to the baseline behavior given the available data. However, the degree of uncertainty is high, thus at time point $t = 1$ the HPD interval is often quite large. The HPD intervals are generally a decreasing function of time, because, as time evolves, our confidence on the one step-ahead predictions increases due to the new information that becomes available from the specific tumor. Nonetheless, since estimation is solely based on macroscopic measurements, as well as because of the small sample size and/or measurement errors, the HPD intervals often remain relatively large by the end of the prediction period to accommodate the underlying uncertainty regarding what will happen in the near future.

The reasons for this uncertainty are multiple. 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. Therefore, we use a macroscopic dynamic model to describe a complex multiscale phenomenon. Moreover, we have considerable measurement errors (tumor size is measured using calipers), small sample size, estimation of tumor volume with only two dimensions, as well as 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. This could have considerable potential for more reliable prognosis of the state of a subject and/or towards the development of model-based online/adaptive tumor-specific therapies. For instance, Gatenby *et al.*[13] and Ariosto *et al.*[3] suggested an empirical sequential treatment plan, where the main principle is that cancer treatment should be as dynamic as the tumor under treatment. Thus, tumor-specific treatment strategies should adapt to treatment outcome at each point in time by changing the combination of drugs, dose, and length of time intervals. To this end, they suggest an empirical/heuristic method

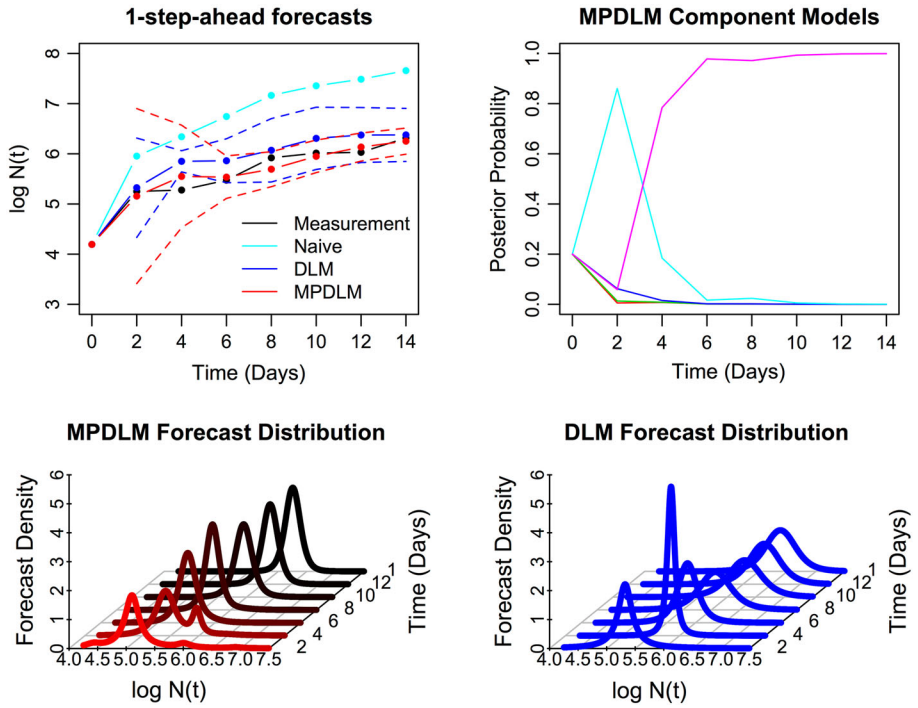

FIGURE 11. Top left panel: One-step ahead predictive performance of MPDLM, single process DLM and naive methods for a representative mouse from the test set (cell line U87). The dashed lines represent the respective 90% HPD intervals of DLM and MPDLM. Right top panel: Posterior weights for the component models of the MPDLM with respect to time. Red: $\lambda = 0.8$, green: $\lambda = 0.8475$, blue: $\lambda = 0.895$, light-blue: $\lambda = 0.9425$, pink: $\lambda = 0.99$; Bottom line panels: Forecast distributions of the MPDLM (bottom left) and single DLM (bottom right) with respect to time. The MPDLM yields more precision in its forecasts (narrower HPD intervals).

for individualised treatment strategies. Improved strategies could be designed by incorporating stochastic models that are updated at each time point, as in the present paper. Another important feature of the MPDLM is the observed faster convergence of its parameters to the ones that optimally describe the underlying mechanism. In real life applications this feature could be crucial since we typically have a very limited amount of data available and the condition of an individual may change (deteriorate) rapidly. Compared to the other methods we study, the MPDLM achieves a faster convergence and/or tighter HPD intervals in the majority of the cases (Figs. 8–11).

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 behavior 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 implies that the corresponding model is the most appropriate to describe the growth dynamics of the tumor, although no simple model actually generates the time-series Y_t. From Eq. (3) we can see that all posterior distributions, both for $\boldsymbol{\theta}_t$ and Y_t, are discrete probability mixtures of non-standardized t-distributions. Thus, it is possible that the posterior mixture distribution will not be unimodal. Particularly, with a mixture of k symmetric and unimodal distributions, as is the case with the $T_n[m, C]$ distributions in the model, we may have a posterior density with up to k modes.

The results presented here identify considerable room for improvement regarding short-term tumor-specific growth prediction, with prospects to develop 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 may allow inclusion of various extra covariates, e.g., genetic information, at little additional computational cost by simply including additional elements in the state vector. We aim to incorporate the suggested approach in the design of optimal therapy protocols by combining with pharmacokinetic-pharmacodynamic models. To this end, we will use additional data from experiments in transgenic mice, which are currently underway.

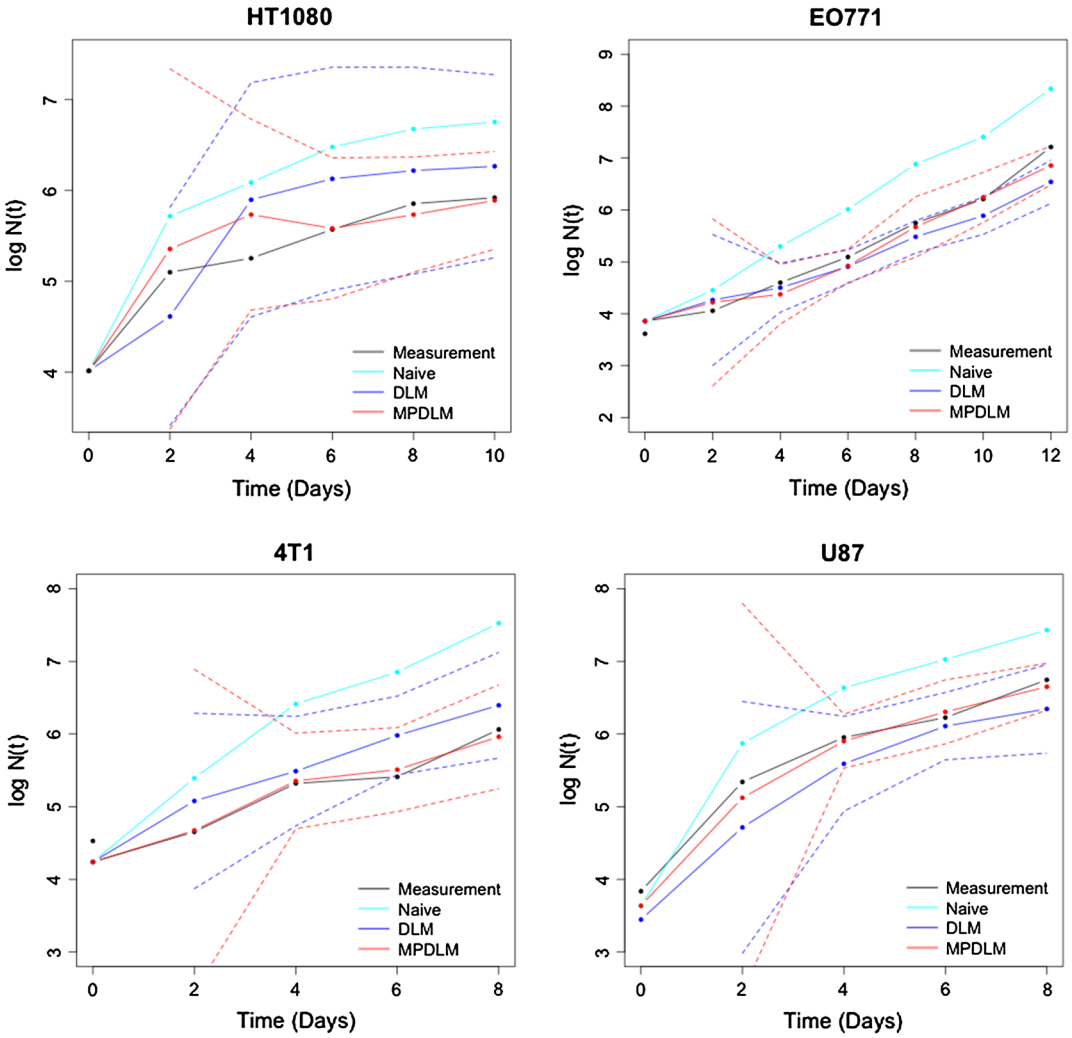

FIGURE 12. One-step ahead predictive performance of MPDLM, DLM, and naive methods for representative subjects from the test set for cancer types (from top left to bottom right) HT1080, EO771, 4T1, and U87 when only one subject was used to construct the training set. The dashed lines represent the respective 90% HPD intervals of DLM and MPDLM. The MPDLM yields narrower HPD interval for cell lines HT1080 and U87. For cell lines EO771 and 4T1 the DLM is not able to capture all observations within its HPD intervals.

ACKNOWLEDGMENTS

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%) as well as by the European Regional Development Fund and the Republic of Cyprus through the Research Promotion Foundation (Project Didaktor/0609/48).

REFERENCES

[1]Achilleos, A., C. Loizides, T. Stylianopoulos, and G. D. Mitsis. Linear dynamic modelling and Bayesian forecasting of tumor evolution. In: Proceedings of the 12th IEEE International Conference on BioInformatics and BioEngineering (BIBE), Larnaca, Cyprus, Nov 2012, pp. 671–676.

[2]Alemani, D., F. Pappalardo, M. Pennisi, S. Motta, and V. Brusic. Combining cellular automata and lattice Boltzmann method to model multiscale avascular tumor growth coupled with nutrient diffusion and immune competition. *J. Immunol. Methods* 376:55–68, 2012.

[3]Ariosto, S. S., Y. Kam, Z. P. Khin, S. E. Minton, R. J. Gillies, and R. A. Gatenby. Evolutionary approaches to prolong progression-free survival in breast cancer. *Cancer Res.* 72(24):6362–6370, 2012.

[4]Bodnar, M., M. J. Piotrowska, and U. Forys. Gompertz model with delays and treatment: mathematical analysis. *Math. Biosci. Eng.* 2013. doi:10.3934/mbe.2013.10.551.

[5]Bolstad, W. M. Introduction to Bayesian Statistics. New York: Wiley-Interscience, 2007.

[6]Byrne, H. Dissecting cancer through mathematics: from the cell to the animal model. *Nat. Rev. Cancer* 10:221–230, 2010. doi:10.1038/nrc2808.

[7]Chaplain, M., M. Ganesh, and I. Graham. Spatio-temporal pattern formation on spherical surfaces: numerical simulation and application to solid tumour growth. *J. Math. Biol.* 42:387–423, 2001.
[8]Chatfield, C. Time-Series Forecasting. Boca Raton: CRC Press, 2001.
[9]De Pillis, L., and A. Radunskaya. A mathematical tumor model with immune resistance and drug therapy: an optimal control approach. *Comput. Math. Methods Med.* 3:79–100, 2001.
[10]Deisboeck, T., and G. Stamatakos. Multiscale Cancer Modeling. Boca Raton: CRC Press, 2010.
[11]Durbin, J., S. Koopman, and A. Atkinson. Time Series Analysis by State Space Methods. New York: Oxford University Press, 2001.
[12]Friedman, M. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J. Am. Stat. Assoc.* 32:675–701, 1937.
[13]Gatenby, A. R., S. S. Ariosto, R. J. Gillies, and F. Roy. Adaptive therapy. *Cancer Res.* 69:4894–4903, 2009.
[14]Hadjiandreou, M. M., and G. D. Mitsis. Towards tumor growth control subject to reduced toxicity. In: American Control Conference (ACC2012), Montreal, QC, Canada, June 2012, pp. 5592–5597.
[15]Hadjiandreou, M. M., and G. D. Mitsis. Mathematical modeling of tumor growth, drug-resistance, toxicity, and optimal therapy design. *IEEE Trans. Biomed. Eng.* 61(2):415–425, 2014. doi:10.1109/TBME.2013.2280189.
[16]Hadjiandreou, M. M., and G. D. Mitsis. Model-based control of cancer progression subject to drug-resistance. In: Decision and Control (CDC), 51st IEEE Conference, Maui, HI, Dec 2012, pp. 350–355.
[17]Kozusko, F., and Z. Bajzer. Combining Gompertzian growth and cell population dynamics. *Math. Biosci.* 185:153–167, 2003.
[18]Liu, W., and F. Tang. Modeling a simplified regulatory system of blood glucose at molecular levels. *J. Theor. Biol.* 252:608–620, 2008.
[19]MacLaurin, J., J. Chapman, G. W. Jones, and T. Roose. The buckling of capillaries in solid tumours. *Proc. R. Soc. A* 468:4123–4145, 2012.
[20]Martin, R. Optimal control drug scheduling of cancer chemotherapy. *Automatica* 28:1113–1123, 1992.
[21]Noble, S., 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. *J. Theor. Biol.* 264:990–1002, 2010.
[22]Norton, L. A Gompertzian model of human breast cancer growth. *Cancer Res.* 48:7067–7071, 1988.
[23]Pinho, S., H. Freedman, and F. Nani. A chemotherapy model for the treatment of cancer with metastasis. *Math. Comput. Model.* 36:773–803, 2002.
[24]Preziosi, L. Cancer Modelling and Simulation. Boca Raton: CRC Press, 2003.
[25]Sherratt, J., and M. M. Chaplain. A new mathematical model for avascular tumour growth. *J. Math. Biol.* 43:291–312, 2001.
[26]Stylianopoulos, T., 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, L. L. Munn, and R. K. Jain. Causes, consequences, and remedies for growth-induced solid stress in murine and human tumors. *Proc. Natl. Acad. Sci.* 109:15101–15108, 2012.
[27]Stylianopoulos, T., J. D. Martin, M. Snuderl, F. Mpekris, S. R. Jain, and R. K. Jain. Co-evolution of solid stress and interstitial fluid pressure in tumors during progression: implications for vascular collapse. *Cancer Res.* 2013. doi: 10.1158/0008-5472.CAN-12-4521.
[28]Swierniak, A., M. Kimmel, and J. Smieja. Mathematical modeling as a tool for planning anticancer therapy. *Eur. J. Pharmacol.* 625:108–121, 2009.
[29]West, M., and J. Harrison. Bayesian Forecasting and Dynamic Models. New York: Springer, 1997.