ICES REPORT 10-44

November 2010

Bayesian Calibration, Validation, and Uncertainty Quantification of Diffuse Interface Models of Tumor Growth

by

Andrea Hawkins-Daarud, Serge Prudhomme, Kristoffer G. van der Zee, and J. Tinsley Oden

The Institute for Computational Engineering and Sciences
The University of Texas at Austin
Austin, Texas 78712

Reference: Andrea Hawkins-Daarud, Serge Prudhomme, Kristoffer G. van der Zee, and J. Tinsley Oden, "Bayesian Calibration, Validation, and Uncertainty Quantification of Diffuse Interface Models of Tumor Growth", ICES REPORT 10-44, The Institute for Computational Engineering and Sciences, The University of Texas at Austin, November 2010.

Bayesian Calibration, Validation, and Uncertainty Quantification of Diffuse Interface Models of Tumor Growth

Andrea Hawkins-Daarud, Serge Prudhomme,
Kristoffer G. van der Zee, and J. Tinsley Oden[1]

Institute for Computational Engineering and Sciences
The University of Texas at Austin
November 1, 2010

Abstract

The idea that one can possibly develop computational models that predict the emergence, growth, or decline of tumors in living tissue is enormously intriguing as such predictions could revolutionize medicine and bring a new paradigm into the treatment and prevention of a class of the deadliest maladies affecting humankind. But at the heart of this subject is the notion of predictability itself, the ambiguity involved in selecting and implementing effective models, and the acquisition of relevant data, all factors that contribute to the difficulty of predicting such complex events as tumor growth with quantifiable uncertainty. In this work, we attempt to lay out a framework, based on Bayesian probability, for systematically addressing the questions of *Validation*, the process of investigating the accuracy with which a mathematical model is able to reproduce particular physical events; and *Uncertainty Quantification*, developing measures of the degree of confidence with which a computer model predicts particular quantities of interest. For illustrative purposes, we exercise the process using virtual data for models of tumor growth based on diffuse-interface theories of mixtures utilizing virtual data.

Keywords: Bayesian probability, Calibration, Validation, Uncertainty Quantification, Tumor growth models

1 Predictive Modeling of Physical Events

The idea that one can possibly develop computational models that predict the emergence, growth, or decline of tumors in living tissue is enormously intriguing as such predictions could revolutionize medicine and bring a new paradigm into the treatment and prevention of a class of the deadliest maladies affecting humankind. In principle, there is nothing about the classical scientific method that suggests this is an impossible goal, and the growing literature on tumor modeling over the past decade attests that significant progress has been made in this direction.

But at the heart of this subject is the notion of predictability itself, the ambiguity involved in selecting and implementing effective models, and the acquisition of relevant data, all factors that contribute to the difficulty of predicting such complex events as tumor growth with quantifiable uncertainty. What do we mean by a predictive model? Can models be

[1]Corresponding author at The University of Texas at Austin, Austin, TX 78712, USA. Fax: (512)471-8694. Email address: oden@ices.utexas.edu

validated as predictive tools? And, if not, can legitimate validation processes be defined that can increase our confidence in the predictive power of the given model classes? In addition, can specific model outputs, the quantities of interest (the "QoI's", such as tumor volume at the end of a given time interval) be predicted with quantifiable uncertainty? Which model is "best"? That is, can the predictive power of members of a class of models be quantified in a meaningful way for specific quantities of interest? Finally, if it is possible to determine that a model is inadequate (invalid), how can it be improved or replaced by a better model? These issues are prevalent in many fields and have been the concern of many research efforts in computational science and engineering (see [1, 2, 3] and references therein). To the best of the authors knowledge, such questions have not been considered carefully on issues related to tumor growth models, but as these models improve these questions must be addressed for their integration into the medical community for aiding with decision making.

In this work, we attempt to lay out a framework, based on work in [4] and [5], for systematically addressing these questions. We believe that most of the issues are in the province of the emerging fields of V and V and UQ, defined as: *Validation*, the process of investigating the accuracy with which a mathematical model is able to reproduce particular physical events (quantities of interest); *Verification*, the process of considering the accuracy with which the code is capable of reproducing results delivered by the mathematical model, and *Uncertainty Quantification*, the process of developing measures of the degree of confidence with which a computer model predicts particular quantities of interest. Further, we assert that the unifying framework provided by Bayesian inference methods can bring into play powerful approaches of statistical inverse analysis that have direct bearing on the key questions underlying predictive modeling. We remark that while we believe the question of verification is extremely important, this work will only discuss at length the issues of validation and uncertainty quantification.

In the discussions following this introduction, we present the basic ideas underlying Bayesian calibration, validation, and quantifying uncertainty in predictions. We then describe a family of models of interacting media based on diffuse-interface models derived from continuum mixture theory [5], and we select from this general framework three representative models that form a model class $\mathcal{M}$ for possibly modeling tumor growth. Here, model class refers to a group of models derived via similar methods each meant to simulate the same phenomena, in this case, tumor growth. Models within a class can differ in their physics, the nature of the parameters (e.g. time dependent, spatially dependent, etc.), or the stochastic character of the parameters (e.g. normal, uniform, deterministic). Bayesian calibration methods are then applied, using virtual observational data, to demonstrate how these methodologies can address the questions posed earlier and to provide a step toward predictive modeling of tumor growth.

2 Bayesian Calibration of Mathematical Models

In this exposition, we understand a mathematical model to be a collection of mathematical constructions designed to represent our complete knowledge of a physical system. Traditionally, knowledge is acquired via the two avenues of the scientific method: theory, which involves hypotheses about realities and which provides mathematical abstractions of the

physical universe, and observations, which attempt to gain knowledge through perception of reality gained through the senses or with instruments. The application of a mathematical representation of a theory usually involves two fundamental attributes: the set $\mathbf{m}$ of parameters defining the specific environment in which the system is expected to operate (the moduli, coefficients, rates, ..., etc.), and the scenario S in which the particular instantiation of the model is considered (the domains of the solution, be it one- two- or three-dimensional, boundary and initial conditions, and data source terms, etc.). In the deterministic case, we can express the model in the abstract form,

$$\mathcal{A}\left(\mathbf{m}, S, u\left(\mathbf{m}, S\right)\right) = 0 \tag{1}$$

where $\mathcal{A}\left(\cdot\right)$ is the collection of operators, constraints, and conditions defining the model, S is a particular scenario, and $u\left(\mathbf{m}, S\right)$ is a solution for the particular parameter set $\mathbf{m}$ and the scenario S. Once $u\left(\mathbf{m}, S\right)$ is known (assuming solutions to (1) exist), we compute the quantities of interest (the QoI's) defined by functionals Q on the space of possible solutions:

$$Q\left(u\left(\mathbf{m}, S\right)\right) \in \mathbb{R}. \tag{2}$$

Before making the calculation (2), the model should be calibrated by considering calibration scenarios S_C and performing experiments that lead to observational data $\mathbf{d}_C^{obs}$. There is generally a tradeoff that must be considered in deciding which calibration data to use. If possible, it is best to use data from scenarios representing the predictive scenario as closely as possible. However, it is probable that gathering such data is very expensive or the data is not accessible. Traditionally, one wishes to use the simplest data that will provide adequate information. So, we tend to think of calibration experiments as coming from simple scenarios such as in vitro experiments, one-dimensional domains, etc. Classical inverse analysis backs out corrected parameters, $\mathbf{m}$ for which the model best agrees with the given measurements. While it may provide some useful information, this classical process is often very inadequate, and frequently meaningless for it neither takes into account the uncertainty in the parameters $\mathbf{m}$ nor in the observational data $\mathbf{d}_C^{obs}$. In Bayesian inverse analysis, this issue is addressed by representing $\mathbf{m}$ and the data each with a probability density function (pdf) characterizing the uncertainty in these values. The fundamental idea is then that, instead of a single set of model parameters, the solution to the statistical inverse problem is now a probability density function characterizing how certain one is in the value of the parameters depending on how certain one is about the observational data and prior information about the model parameters.

We introduce the notation $\mathbf{M}$ to denote the random variable characterizing the model parameters and $\mathbf{D}$ to denote the random variable characterizing the measurement with $\mathbf{m}$ and $\mathbf{d}$ representing specific realizations of these random variables respectively. It is assumed that these random variables are related through the general relation $\mathbf{D} = \mathbf{G}\left(\mathbf{M}, \mathbf{E}\right)$ where $\mathbf{E}$ is the random variable characterizing the noise in the measurement. The function $\mathbf{G}$ can be thought of as the process of solving the model (1) and deriving from the solution u predicted values of the measurement values $\mathbf{d}$. Prior to any actual observation being made, there is some prior information available on the model parameters which we assume can be represented through a probability density function, i.e. we assume $\mathbf{m} \mapsto \rho_M\left(\mathbf{m}\right)$ which is

the *prior probability density distribution*. In practice, this change in representation results in (1) becoming a stochastic forward problem, denoted

$$\mathcal{A}\left(\rho_M\left(\mathbf{m}\right), S, u\left(\rho_M\left(\mathbf{m}\right), S\right)\right) = 0, \tag{3}$$

and thus, the solution is also a random field $u\left(\rho_M\left(\mathbf{m}, S\right)\right)$ for any particular scenario S, and the QoI's are random functions of $\mathbf{m}$, $Q\left(u\left(\rho_M\left(\mathbf{m}\right), S\right)\right) = q_S\left(\mathbf{m}\right)$, where $q_S\left(\mathbf{m}\right)$ is written to indicate that the QoI can be thought of directly as a random function of $\mathbf{m}$ for scenario S.

We further assume that given a specific value of $\mathbf{m}$, we could update our information on the measurement random variable $\mathbf{D}$ to reflect what we would expect as the measurement given $\mathbf{m}$. That is we assume we know the form of the conditional probability $\theta\left(\mathbf{d}|\mathbf{m}\right)$ representing the pdf of error in the measurement given the parameter $\mathbf{m}$. Traditionally, this is reflective of the random variable $\mathbf{E}$ and ultimately characterizes the error coming from the measurement device or human error (meaning the difference between measured observables and the values delivered by the model). Additionally, we denote the pdf containing the information regarding the measurements $\mathbf{d}$ as $\rho_D\left(\mathbf{d}\right) = \int \theta\left(\mathbf{d}|\mathbf{m}\right)\rho_M\left(\mathbf{m}\right)\ d\mathbf{m}$.

With this notation in place, we use Bayes' theorem to define the solution to the statistical inverse problem (calibration) of the stochastic model (3). The pdf of the model parameters given the specific observed data $\mathbf{d}^{obs}$, or the solution, is given as the Bayesian update

$$\begin{aligned}\sigma_M\left(\mathbf{m}|\mathbf{d}^{obs}\right) &= \theta\left(\mathbf{d}^{obs}|\mathbf{m}\right)\rho_M\left(\mathbf{m}\right)/\rho_D\left(\mathbf{d}^{obs}\right) \\ &\propto \theta\left(\mathbf{d}^{obs}|\mathbf{m}\right)\rho_M\left(\mathbf{m}\right).\end{aligned} \tag{4}$$

We comment that $\rho_D\left(\mathbf{d}^{obs}\right)$ is effectively a constant for the given value $\mathbf{d}^{obs}$ and can be thought of as the reciprocal of the normalization constant. The conditional pdf $\sigma_M\left(\mathbf{m}|\mathbf{d}^{obs}\right)$ is the posterior. For simplicity, we write $\sigma_M\left(\mathbf{m}\right) := \sigma_M\left(\mathbf{m}|\mathbf{d}^{obs}\right)$. With the updated information regarding the model parameters from the posterior, we write the statistically calibrated model and corresponding random field of the QoI's as

$$\begin{aligned}\mathcal{A}\left(\sigma_M\left(\mathbf{m}\right), S, u\left(\sigma_M\left(\mathbf{m}\right), S\right)\right) = 0, \\ Q\left(u\left(\sigma_M\left(\mathbf{m}\right), S\right)\right) = q_S^C\left(\mathbf{m}\right).\end{aligned} \tag{5}$$

where $q_S^C\left(\mathbf{m}\right)$ is the random field for the QoI for scenario S based on the calibrated pdf σ_M.

While the examples presented in this paper will not address such issues, in general, it is by no means a trivial task to evaluate the posterior pdf in (4) or equivalently to solve the stochastic system (5) when the dimension of $\mathbf{m}$ is high. The visualization of (4) often involves the use of sophisticated statistical sampling algorithms, such as Latin hypercube sampling [6], a Markov Chain Monte Carlo method, i.e. Metropolis-Hastings or Gibbs sampling, or other approaches, some of which are implemented in the codes QUESO [7] and DAKOTA [8]. A number of effective methods for solving stochastic systems characterizing the forward problem represented in (5), have appeared including occasionally traditional Monte Carlo methods, but often more powerful techniques such as polynomial chaos [9], stochastic collocation [10], or stochastic Galerkin [11, 12] methods are employed. For more information on statistical calibration under a Bayesian framework, works such as those by Tarantola [13] or Kaipio and Somersalo [14] can be consulted.

3 Model Validation Process

Recalling that validation is the process of investigating whether a mathematical model is capable of reproducing with sufficient accuracy specific physical events, we recognize that there is a need for multiple validation checks. If the data for calibration and validation are not of the exact same form as the ultimate quantity of interest, it is also important to verify that the model is capable of reproducing these sets of data. If the model fails in this respect, then the information produced via the calibration and/or validation processes may be inaccurate and either the data is inadequate or the model does not capture enough of the physics to make a reliable prediction. We will first propose a method of validation against the observed data and second describe a method for validation with respect to the quantity of interest.

3.1 Data Misfit Check

In the specification of the likelihood function, and thus of the error $\mathbf{E}$, the assumption is often implicitly (and necessarily) made that the model is capable of reproducing such data and that the error is attributable to imprecise measurements only. The first step in validation is to check the adequacy of assumptions underlying the construction of the model by analyzing $\mathbf{E}$, the data misfit. Assuming for simplicity that the misfit is an additive error, we have

$$\mathbf{D} = \mathbf{G}(\mathbf{M}) + \mathbf{E} \;\Rightarrow\; \mathbf{E} = \mathbf{D} - \mathbf{G}(\mathbf{M}). \tag{6}$$

Since the models we will consider are deterministic, if $\mathbf{m}$ is specified $\mathbf{G}(\mathbf{m})$ will take on exactly one value and $\mathbf{E}$ may be regarded as the error between $\mathbf{G}(\mathbf{m})$ and the data that should be observed. From a practical point of view, we must check that there exists at least one set of model parameters $\mathbf{m}$ such that the observed noise $\mathbf{e}^{obs}(\mathbf{m}) = \mathbf{d}^{obs} - \mathbf{G}(\mathbf{m})$ is likely in some meaningful sense. That is, we must check that there exists $\mathbf{m}$ such that the specific realization of $\mathbf{e}^{obs}$ corresponding to the realization $\mathbf{d}^{obs}$ is in some sense probable. For this reason, we write the likelihood in terms of the error $\mathbf{e}$ rather than $\mathbf{d}$. To do this, we write

$$\theta(\mathbf{d}|\mathbf{m}) = \int \theta(\mathbf{d}|\mathbf{m}, \mathbf{e})\, \theta_{noise}(\mathbf{e}|\mathbf{m})\; d\mathbf{e}, \tag{7}$$

where $\theta_{noise}(\mathbf{e}|\mathbf{m})$ denotes the probability of noise level $\mathbf{e}$ given parameter $\mathbf{m}$. Noting that when both $\mathbf{m}$ and $\mathbf{e}$ are specified, $\mathbf{d}$ is completely specified from (6), we can write $\theta(\mathbf{d}|\mathbf{m}, \mathbf{e}) = \delta(\mathbf{d} - \mathbf{G}(\mathbf{m}) - \mathbf{e})$. Substituting this into the previous formula (7) we see that $\theta(\mathbf{d}|\mathbf{m}) = \theta_{noise}(\mathbf{d} - \mathbf{G}(\mathbf{m})|\mathbf{m}) = \theta_{noise}(\mathbf{e}(\mathbf{d}, \mathbf{m})|\mathbf{m})$. Thus, we may rewrite Bayes formula as

$$\sigma_M\left(\mathbf{m}|\mathbf{e}^{obs}\right) \propto \theta_{noise}\left(\mathbf{e}^{obs}|\mathbf{m}\right) \rho_M(\mathbf{m}). \tag{8}$$

Here $\mathbf{e}^{obs} = \mathbf{e}^{obs}\left(\mathbf{d}^{obs}, \mathbf{m}\right) = \mathbf{d}^{obs} - \mathbf{G}(\mathbf{m})$. With these considerations in mind, we define $T_{noise}(\mathbf{e}|\mathbf{m})$ as the likelihood cumulative distribution function (cdf), allowing the investigation of the level of error by finding the value of $T_{noise}\left(\mathbf{e}^{obs}(\mathbf{m})|\mathbf{m}\right)$. In this framework, the concept of highly probable would mean that the value of $T\left(\mathbf{e}^{obs}(\mathbf{m})|\mathbf{m}\right)$ is not out in the tails of the distribution. Thus, setting some tolerances γ_{cdf}^{up} and γ_{cdf}^{down} indicative of the

upper and lower tail cutoffs, we check that there is at least some $\mathbf{m}$ such that the following is satisfied

$$\gamma_{cdf}^{down} \leq T\left(\mathbf{e}^{obs}|\mathbf{m}\right) \leq \gamma_{cdf}^{up}. \tag{9}$$

A pictorial example is shown in Figure 1. We refer to this process as the data misfit check.

3.2 Validation of Prediction

If a model passes both the data misfit checks for the calibration data and the validation data, one must check the validity with respect to the quantity of interest. The validation experiments involve generating a new set of validation observables $\mathbf{d}_V^{obs}$ for a validation scenario S_V that is believed to depict as closely as possible the environment in which the final predicted QoI exists. Ideally, the validation experiments/observations are designed to check the acceptability of various assumptions made in developing the model for scenarios as close as feasible to that for which the prediction is to be made. As we also want to see if the information gained from the new experiments is different from the calibration experiments, the prior can be chosen to be the same as in the calibration process. It follows that the validation pdf $\sigma_V(\mathbf{m})$ is the posterior computed using the Bayesian relation,

$$\sigma_V(\mathbf{m}) := \sigma_V\left(\mathbf{m}|\mathbf{d}_V^{obs}\right) \propto \rho_M(\mathbf{m})\,\theta_{noise}\left(\mathbf{e}_V^{obs}|\mathbf{m}\right) \tag{10}$$

We may thus envision a prediction pyramid of escalating complexity, beginning with possibly several different component-level calibration experiments at the lowest level for one or more calibration scenarios, S_C, generating the posterior pdf $\sigma_M(\mathbf{m})$, and leading to one or more full system level validation experiments at the validation scenario level, S_V (see Figure 2). The target output, the goal of the simulation, is on the prediction scenario S_P on which the ultimate prediction of the principal QoI is made. Along the way, the various observations $\left(\mathbf{d}_C^{obs}, \mathbf{d}_V^{obs}\right)$ may be regarded as intermediate QoI's at various levels of the pyramid.

We can now consider the calibration predicted QoI and the validation predicted QoI for the full prediction scenario S_P,

$$Q\left(u\left(\sigma_M(\mathbf{m}), S_P\right)\right) = q_P^C(\mathbf{m}) \quad \text{and} \quad Q\left(u\left(\sigma_V(\mathbf{m}), S_P\right)\right) = q_P^V(\mathbf{m}). \tag{11}$$

These equations have employed everything we know about the model and the QoI. This includes prior information on the parameters, the theory and measurement error as embodied in likelihood pdfs, and the predictions of the so-called forward problems, and their predictions of the random fields q_P^C and q_P^V representing the target output. Thus, we arrive at the fundamental question: is the original calibrated model (11a) valid for this particular QoI?

Alas, it is not possible to validate a model; the best that can be said is that the model is not invalid if q_P^C and q_P^V are close enough together with respect to some appropriate metric. Of many possible metrics, one is

$$M\left(q_P^C, q_P^V\right) = \sup_{y\in[\alpha_1,\alpha_2]} \left|F_C^{-1}(y) - F_V^{-1}(y)\right| \tag{12}$$

where F_C and F_V are the cumulative probability distributions associated with the pdf's of the random fields q_P^C and q_P^V, respectively, and α_1 and α_2 are chosen so as to exclude

comparison of the tails of the distributions. We can then specify a tolerance γ_{tol} and declare the model *not invalid* if $M\left(q_P^C, q_P^V\right) < \gamma_{tol}$.

If this criterion is passed, it still remains to quantify the uncertainty in the QoI, namely the random field q_P^C. But this is now an exercise in elementary probability: if $q_P^C(\mathbf{m})$ is known, we may compute the mean, variance, and other moments of its corresponding pdf and, thereby, quantify the uncertainty in the prediction. A schematic diagram describing the entire validation process is given in Figure 2.

4 Some Model Classes Derived from Continuum Mixture Theory

The discussion to this point has been primarily aimed at introducing concepts of model calibration and validation. We now change focus to illustrating how such methods may be applied to phenomenological models of tumor growth. Many traditional models of tumor growth treat the tumor mass as a single phase with a sharp interface between the tumor and other constituents (e.g. [15, 16, 17, 18, 19, 20]). However, these paradigms ignore the general nature of tumors as a complex biomass comprised of different cell types varying in their degree of aggressiveness, extracellular proteins and fluid, and blood vessels. Further, the interface between the tumor and the healthy tissue may not be well defined. Thus, we restrict our attention to those models coming from what we consider to be a more natural approach to characterizing tumors and their surrounding environment: models derived through the continuum theory of mixtures (e.g. [21, 22, 5]). This multiphase theory represents the presence of each constituent as a volume fraction, allowing multiple constituents to be present at the same point at the same time and handles the interface between phases as a feature of the solution. This is done by representing surface energies of interfaces between constituents through the appearance of gradients of volume fractions, leading to so-called diffuse-interface models of the Cahn-Hilliard type. A brief summary of this theory is provided here. We assume the constituents are non-polar and that only isothermal behavior occurs in all constituents. For the general case, see [5].

4.1 A Continuum Theory of Mixtures

The fundamental idea underlying mixture theory is that a material body $\mathcal{B}$ can be composed of N constituent species $\mathcal{B}_1, \ldots, \mathcal{B}_N$ that occupy a common portion of space at the same time. Each spatial position is then allowed to be occupied by N constituents and each constituent is assigned a mass density $\hat{\rho}_\alpha$, regarded as a function of position and time, $(\mathbf{x}, t)$, which represents the mass of the α^{th} constituent per unit volume of the mixture. Then the mass density of the mixture is

$$\rho(\mathbf{x}, t) = \sum_{\alpha=1}^{N} \hat{\rho}_\alpha(\mathbf{x}, t) = \sum_{\alpha=1}^{N} \rho_\alpha(\mathbf{x}, t)\varphi_\alpha(\mathbf{x}, t) \tag{13}$$

where ρ_α is the mass per unit volume and φ_α the volume fraction of the αth constituent.

Each of the N constituents must satisfy its own balance laws which differ from those of classical continuum mechanics due to the presence of interaction terms representing the exchange of mass, momentum, and energy between constituents. For a general mixture occupying an open region Ω in $\mathbb{R}^3$ over a time interval $(0,T)$, the volume fractions, or mass concentrations, and other independent field variables must satisfy the following local forms of the balance laws for all α, $1 \leq \alpha \leq N$, all $\mathbf{x} \in \Omega$, and $t \in (0,T)$.

Balance of mass:
$$\frac{\partial \rho_\alpha \varphi_\alpha}{\partial t} + \nabla \cdot (\rho_\alpha \varphi_\alpha \mathbf{v}_\alpha) = \gamma_\alpha - \nabla \cdot \mathbf{j}_\alpha \tag{14a}$$

Balance of linear momentum:
$$\rho_\alpha \varphi_\alpha \frac{d^\alpha \mathbf{v}_\alpha}{dt} = \nabla \cdot \mathbf{T}_\alpha + \rho_\alpha \varphi_\alpha \mathbf{b}_\alpha + \hat{\mathbf{p}}_\alpha \tag{14b}$$

Balance of angular momentum:
$$\mathbf{M}_\alpha = \mathbf{T}_\alpha - \mathbf{T}_\alpha^T \tag{14c}$$

Balance of energy:
$$\rho_\alpha \varphi_\alpha \frac{d^\alpha e_\alpha}{dt} = \operatorname{tr} \mathbf{T}_\alpha^T \mathbf{L}_\alpha - \nabla \cdot \mathbf{q}_\alpha + \rho_\alpha \varphi_\alpha r_\alpha + \hat{\varepsilon}_\alpha + \Upsilon_\alpha \tag{14d}$$

In (14a), γ_α is the mass supplied to constituent α by other constituents and $\mathbf{j}_\alpha$ is the mass flux due to changes in the chemical potential defined in terms of gradients in concentrations and changes in nutrient concentrations. In the momentum equations, (14b) and (14c), $\mathbf{v}_\alpha$ is the species velocity field and $\mathbf{T}_\alpha$ and $\mathbf{M}_\alpha$ are the partial Cauchy stress tensor and the intrinsic moment of momentum vectors, respectively, with $\mathbf{b}_\alpha$ the body force density and $\hat{\mathbf{p}}_\alpha$ the momentum supplied by other constituents. In (14d), e_α is the internal energy per unit mass, $\mathbf{L}_\alpha = \nabla \mathbf{v}_\alpha$ is the velocity gradient, $\mathbf{q}_\alpha$ is the heat flux, r_α is the heat supplied per unit mass and unit time, and $\hat{\varepsilon}_\alpha$ is the energy supplied to constituent α by other constituents. The remaining term is defined by

$$\Upsilon_\alpha = \sum_{\beta=1}^{N} \nabla \cdot \left(\boldsymbol{\sigma}_{\alpha\beta} \frac{d^\alpha \varphi_\beta}{dt} \right) + \sum_{\beta=1}^{L} \zeta_{\alpha\beta} \frac{d^\alpha m_\beta}{dt}, \tag{15}$$

and is critical in making the connection between mixture theory and diffuse-interface models. In (15), the quantity $\boldsymbol{\sigma}_{\alpha\beta}$ is a generalized surface traction that is conjugate to time-changes in species volume fractions on constituent interfaces, and the second term characterizes the diffusion of chemical or biological constituents due to chemo- or bio-taxis in terms of time rates of m_β, the multiple of φ_β giving the effect of nutrient β, and $\zeta_{\alpha\beta}$ the corresponding conjugate fluxes. These effects were introduced in the models of Cristini *et al.* [22], Frieboes *et. al.* [23] and Wise *et al.* [24]. See also [5].

The equations in (14) are closed by the appropriate choice of constitutive equations. Such choices must be consistent with both the second law of thermodynamics for the mixture and the classical balance laws for the entire mixture. By introducing the Helmholtz free energy per unit mass ψ_α for each constituent, or the free energy Ψ_α per unit volume, the classical

Coleman-Noll method makes use of the constraints from classical balance laws and the second law to supply sufficient conditions on forms of key constitutive equations in terms of appropriate derivatives of the free energy (see [25, 22, 5]). In this work, we will confine our attention to materials for which the Helmholtz free energy is of the form

$$\Psi_\alpha = \Psi_\alpha \left(\varphi_1, \ldots, \varphi_N, \nabla\varphi_1, \ldots, \nabla\varphi_N, m_1, \ldots, m_N\right). \tag{16}$$

As the results shown later are primarily concerned with (14a), we note that it can be argued that a resulting consistent constitutive equation for $\mathbf{j}_\alpha$ is of the form

$$\mathbf{j}_\alpha = -\sum_{\beta=1}^{N} M_{\alpha\beta}(\boldsymbol{\varphi}, \mathbf{m}_\alpha)\nabla(\mu_{\beta\alpha}/\rho_\alpha). \tag{17}$$

Here $M_{\alpha\beta}$ is a positive semi-definite, second-order mobility tensor, $\boldsymbol{\varphi} = [\varphi_1, \ldots, \varphi_N]^T$, $\mathbf{m}_\alpha = [\varphi_\alpha m_1, \ldots, \varphi_\alpha m_L]^T$ and $\mu_{\alpha\beta}$ is interpreted as the chemical potential, defined as

$$\mu_{\alpha\beta} = \frac{\partial\Psi_\alpha}{\partial\varphi_\beta} + \delta_{\alpha\beta}\sum_{\gamma=1}^{L} m_\gamma\zeta_{\alpha\beta}/\varphi_\alpha - \nabla\cdot\frac{\partial\Psi_\alpha}{\partial\nabla\varphi_\beta}, \tag{18}$$

$\delta_{\alpha\beta}$ being the Kronecker delta. Other details on this theory are given in [5].

4.2 A Class of Models of Tumor Growth

To investigate how such models are applied to tumor growth, we consider an organ or gland that occupies a region Ω_0 in $\mathbb{R}^3$ at time $t = 0$, such as the prostate. The gland is in contact with other tissue and vascular structure that supply traction along the boundary $\partial\Omega_0$. The system is not closed: oxygenated blood and other nutrients may flow into and out of Ω over a time interval of interest $[0, T]$, during which Ω traces out a family of configurations. Within this domain, we assume there is a subdomain ω, possibly not connected, which is comprised of tumorous tissue with material properties different than its surrounding tissue. Both the tumorous domain and the surrounding tissue are comprised of various types of cells, such as epithelial, endothelial, muscle, stroma, and blood cells. Each cell type may have unique material properties and thus could be modeled as separate constituents. We begin, for simplicity, with the case of a two-phase isothermal mixture consisting of tumor u and non-tumor n (i.e. healthy tissue, extracellular proteins, and fluid) and focus on methods of assessing model predictability. Additionally, we include in our model a representative nutrient, c, say oxygen. It is assumed that the mixture is saturated, i.e. $u + n = 1$. Thus, only the variables u and c need to be taken as unknown. Note that as oxygen will occupy a volume fraction significantly smaller than that of the tumor or non-tumor, it is not considered in the mixture balance laws. Rather, it will be simulated by a separate, coupled reaction-diffusion equation and represented in the free energy via the factors m_α in (18).

We first postulate a form of the free energy functional for the tumor phase. As we are particularly interested in a diffuse-interface methodology, we choose a form similar to that used in the general Cahn-Hilliard theory but augment it with a term corresponding to a lower energy when the tumor is able to interact with the chemical species, oxygen:

$$\Psi_u\left(u, \nabla u, c\right) = \gamma u^2\left(u-1\right)^2 + \epsilon^2/2\left|\nabla u\right|^2 - \chi c u, \tag{19}$$

where γ, χ, and ϵ are positive constants. The first term is a double-well potential with minima at $u = 0$ and $u = 1$, the second term can be thought of as energy due to surface tension, and the final term is energy due to reactions with the nutrient and accounts for cell chemotaxis. With this free energy, the form of the chemical potential follows from (18):

$$\mu_{uu} = \mu = \gamma\left(4u^3 - 6u^2 + 2u\right) - \epsilon^2\Delta u - \chi c. \tag{20}$$

Note that $\mu_{un} = 0$. As we expect the mobility to be positive semi-definite and zero where there is no tumor, a reasonable form of the mobility tensor M_{uu} is as follows,

$$M_{uu} = Mu^2, \tag{21}$$

where M is a positive constant. An example tumor mass exchange term might have the phenomenological form

$$\gamma_u = Pcu - Au. \tag{22}$$

Here P and A are both positive constants. A positive value of γ_u signals growth in u, whereas negative γ_u corresponds to a decline in the concentration of u. This form is consistent with the assumption that tumor cells proliferate where there is enough oxygen and that they may die at some constant rate. We further assume that the oxygen diffuses throughout the domain and is uptaken by the cells proliferating most quickly, i.e. the tumor cells. Mathematically, this can be written as follows:

$$0 = \nabla \cdot (D\nabla c) - cu \tag{23}$$

where D is a positive constant. We note the oxygen is not explicitly considered as part of the mixture, that is, it is not included in saturation assumption, $\sum_\alpha \varphi_\alpha = 1$. This is done due to the extremely small volume of the oxygen concentrations compared to cells.

Within the framework of the assumptions made thus far, particularly the assumptions of isothermal behavior of a non-polar two phase mixture governed by (14) and (19), infinitely many models of tumor growth and decline exist. Of course, these classes constitute a subclass of a much larger set of mathematical models falling under the general theory governed by (14), some of which are discussed in [5]. We wish to describe methods for assessing the predictability of representative models in this subclass, i.e. the validity. The level of complexity is not necessarily a factor in the acceptability of a model; all depends on the quantities of interest (the target outputs of the simulation), and the data in hand.

We consider a class $\mathcal{M}$ of three models, $\mathcal{M} = \{\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3\}$ differing in the choice for (21), (22), dependence on c, and time dependency of parameters. These are defined by the following initial-boundary value problems for concentrations u and c :

$\mathcal{M}_1$: Simple Proliferation Model

$$\begin{array}{rll} u_t &= \nabla \cdot (M\nabla\mu) + Pu & \text{in } (0,T)\times\Omega, \\ \mu &= f'(u) - \epsilon^2\Delta u & \text{in } (0,T)\times\Omega, \\ \nabla u \cdot n &= \nabla\mu \cdot n = 0 & \text{on } (0,T)\times\partial\Omega, \\ u(0,\mathbf{x}) &= u_0 & \text{in } \{0\}\times\Omega \\ f'(u) &= \gamma\left(4u^3 - 6u^2 + 2u\right) & \end{array} \tag{24}$$

$\mathcal{M}_2$: Proliferation/Apoptosis Model with Degenerate Mobility and Oxygen Dependence

$$\boxed{\begin{array}{rll} u_t &= \nabla\cdot\left(Mu^2\nabla\mu\right)+Pcu-Au & \text{in } (0,T)\times\Omega,\\ \mu &= f'(u)-\epsilon^2\Delta u-\epsilon\chi c & \text{in } (0,T)\times\Omega,\\ 0 &= \nabla\cdot\left(D\nabla c\right)-cu & \text{in } (0,T)\times\Omega,\\ \nabla u\cdot n &= \nabla\mu\cdot n=0 \quad c=1 & \text{on } (0,T)\times\partial\Omega,\\ u(0,\mathbf{x}) &= u_0 & \text{in } \{0\}\times\Omega\\ f'(u) &= \gamma\left(4u^3-6u^2+2u\right) & \end{array}} \tag{25}$$

$\mathcal{M}_3$: Proliferation/Apoptosis Model with Degenerate Mobility and Oxygen Dependence and Time Dependent Parameters

$$\boxed{\begin{array}{rll} u_t &= \nabla\cdot\left(Mu^2\nabla\mu\right)+Pcu-Au & \text{in } (0,T)\times\Omega,\\ \mu &= f'(u)-\epsilon^2\Delta u-\epsilon\chi(t)\,c & \text{in } (0,T)\times\Omega,\\ 0 &= \nabla\cdot\left(D(t)\nabla c\right)-cu & \text{in } (0,T)\times\Omega,\\ \nabla u\cdot n &= \nabla\mu\cdot n=0 \quad c=1 & \text{on } (0,T)\times\partial\Omega,\\ u(0,\mathbf{x}) &= u_0 & \text{in } \{0\}\times\Omega\\ f'(u) &= \gamma\left(4u^3-6u^2+2u\right) & \end{array}} \tag{26}$$

Both processes of model calibration and model validation are inherently tied to the chosen quantity of interest. For certain quantities of interest, simpler models may perform acceptably, while for more complicated quantities of interest they may not be sufficient. For the purposes of this work, we investigate the situation for the global quantity of interest of the final tumor volume calculated as

$$Q(u) := \int_\Omega u(\mathbf{x},T)\ d\mathbf{x} \tag{27}$$

5 Illustrative Cases of Bayesian Calibration and Model Selection

We now describe two examples of the processes of model calibration and validation for representative scenarios portrayed in the sequences of (virtual) images given in Figures 3 and 8 generated through models $\mathcal{M}_2$ and $\mathcal{M}_3$ respectively. In both figures, we observe a two-phase characterization of the growth of a well-defined solid tumor mass captured by (for instance) MRI imaging, that evolves from an initial shape at $t = 0$, which we take, for simplicity, to be elliptical (see Figure 3).

Snapshot images are taken at $t = t_1 = 3$ and later at $t = t_2 = 6$, that represent the calibration and validation scenarios. The physical "reality" of the tumor at $t = t_3 = 9$ for both scenarios is shown in the bottom right image in Figures 3 and 8. We wish to test the prediction of the tumor volume at this time, based on the calibration data taken at $t = t_1$. For illustration purposes, we will carry out the processes of calibration and validation twice,

Parameter	Value	Parameter	Value
M	200	A	0
P	0.1	ϵ	0.005
χ	10	γ	0.045
D	1		

Table 1 *Values of parameters used for the simulation generating the virtual data.*

once for model $\mathcal{M}_1$ using virtual data generated from model $\mathcal{M}_2$ and once for model $\mathcal{M}_2$ based on data from model $\mathcal{M}_3$. Thus, we address the following types of questions: based on the data observed in Figure 3 at times $t = 3$ and $t = 6$, is model $\mathcal{M}_1$ invalid for predicting a particular quantity of interest such as $Q(u) =$(tumor volume at $t = 9$)?

Remark 1. *The scenarios described here represent a set of virtual experiments and images designed only to demonstrate the validation process. The observed structures shown are generated using models $\mathcal{M}_2$ and $\mathcal{M}_3$ for a deterministic calculation using the parameter values given in Table 1a and in Equation (35), and geometric symmetry is assumed to further simplify the presentation. But the process itself is general and is applicable to much more complex situations. For instance, the actual sequence of MRI images of the evolution of a tumor seeded in a laboratory mouse model, as shown in Figure 4, could conceivably be used as a basis for similar calibration and validation processes.*

5.1 Assessment of $\mathcal{M}_1$ against data from $\mathcal{M}_2$

To begin the process of model calibration and validation, one must begin by defining the model parameters of interest for calibration, the form of the prior, the data which will be used for the calibration, and the form of the likelihood function (which will depend on the data). For model $\mathcal{M}_1$, we choose as the key model parameter to calibrate $\mathbf{m} = P$ (recall (24)). We assume prior knowledge of this quantity suggests it belongs to the bounded interval $0 \leq P \leq 0.3$. The values of M, ϵ and γ will be taken from Table 1. Due to our "prior knowledge", we choose the prior probability density function to be the uniform distribution

$$\rho_M(P) = U(0, 0.3). \tag{28}$$

The nonlinear system of partial differential equations characterizing each model are solved numerically using mixed finite element methods. Details on the discretized models and their solutions are given in Appendix A. The finite element mesh provides an analogue of a pixel array typical of an MRI image, but generally somewhat coarser. Thus, for data, we assume we have available an entire MRI "image" at $t = 3$ and $t = 6$ (for calibration and validation respectively). So for both the calibration and validation steps we consider as observational data the array of pixels recovered from the image with each pixel output defined by the L^2 norm of the observed solution per pixel/element. We are interested in the shape of the interface, so we also select as data the position of the elements on which the interface lies. We choose to work with $\theta_{noise}(\mathbf{e}|\mathbf{m})$ rather than the equivalent $\theta(\mathbf{d}|\mathbf{m})$ (recall (7)), and thus need to define the relation $\mathbf{e} = M_d(\mathbf{G}(P), \mathbf{d})$, i.e. the method of comparing model

output against observed data. Due to the nature of our data, the comparison between the data and simulation will not be a linear operation, and thus simple subtraction will not be an appropriate operation for M_d. Rather, we think of M_d as a metric returning a two-dimensional vector. We make the simplifying assumption that the distribution of $\mathbf{e}$ is independent of P, giving us $\theta_{noise}(\mathbf{e}) = \theta_{noise}(M_d(\mathbf{G}(P), \mathbf{d}))$. Since $M_d(\mathbf{G}(P), \mathbf{d}) > \mathbf{0}$ for all $\mathbf{d}$, in this example, we anticipate that the pdf $\theta_{noise}(\mathbf{e})$ will have the shape of a bivariate, uncorrelated, half-normal distribution, i.e.

$$\theta_{noise}(\mathbf{e}) = \frac{\theta_1\theta_2}{\pi^2} \exp\left(-M_{d,1}^2\theta_1^2/\pi\right) \exp\left(-M_{d,2}^2\theta_2^2/\pi\right), \tag{29}$$

where θ_1 and θ_2 are the parameters of the half normal distributions related to the standard deviation of the distribution and $M_{d,i} = M_{d,i}\left(\mathbf{G}(P), \mathbf{d}^{obs}\right)$ denotes the i^{th} component of M_d. Before defining θ_1 and θ_2, let us first consider how M_d is to be evaluated.

Without loss of generality, we let the L^2 norm of the solution per element be the first component of the data and the location of the interface the second. We evaluate the first component of $M_d(\mathbf{G}(P), \mathbf{d})$ as follows:

$$M_{d,1} = \left(\sum_{\mathrm{k}} \left(\int_{\mathrm{k}} u(x,3)^2 \, d\mathbf{x} - \int_{\mathrm{k}} u^{data}(x)^2 \, d\mathbf{x}\right)^2\right)^{1/2}. \tag{30}$$

Here the symbol k indicates the subdomains corresponding to the pixels in the data images. For the evaluation of the second component, we define the set of interface elements

$$\mathbb{I}(u) = \{\mathrm{k} : \exists\, \mathbf{x} \in \mathrm{k} \text{ where } u(\mathbf{x}) = 0.5\}. \tag{31}$$

Further, let $C(k)$ denote the centroid of a pixel k. Then, the difference in the second component is evaluated like a Hausdorff measure,

$$\begin{aligned} M_{d,2} = \frac{1}{2}\Bigg\{ & \max_{\hat{k}\in\mathbb{I}\left(u^{data}\right)} \left(\min_{k\in\mathbb{I}(u(\mathbf{x},3))} d\left(C(k), C\left(\hat{k}\right)\right)\right) \\ & + \max_{\hat{k}\in\mathbb{I}(u(\mathbf{x},3))} \left(\min_{k\in\mathbb{I}\left(u^{data}\right)} d\left(C(k), C\left(\hat{k}\right)\right)\right)\Bigg\}, \end{aligned} \tag{32}$$

where $d(\cdot,\cdot)$ is the traditional Euclidean distance.

To determine θ_1 and θ_2, we note that these parameters are related to the standard deviation, σ_i, of the corresponding full normal distribution $N(0, \sigma_i)$, i.e.

$$\theta_i = \frac{\sqrt{\pi/2}}{\sigma_i}. \tag{33}$$

The parameters θ_1 and θ_2 are chosen to correspond to a σ equivalent to 10% relative error. For the first data misfit value corresponding to the L^2 pixel data, we choose to define the error relative to the maximum possible error, i.e. if the simulation solution is equal to the unit constant function and the data is zero everywhere. For the case considered here, this value is 2.304. From this, we determine $\theta_1 = \sqrt{\pi/2}/.2304 = 5.439$. For the interface data,

we use as an error cutoff the distance equivalent to a four element distance, i.e. for a 71x71 mesh, $4 \times h = 0.7211$. Thus, $\theta_2 = \sqrt{\pi/2}/0.7211 = 1.738$.

Given the data from the image in Figure 3b and having defined the form of the likelihood, our next step is to perform the data misfit check. We take a uniform sample from the prior distribution for P and calculate $u(x, 3)$ using discretization techniques as described in the appendix. The marginal of the likelihood function for each data variable is calculated and the cumulative distribution, $T(\mathbf{e})$ is determined so as to define $T\left(\mathbf{e}^{obs}(P) = \mathbf{d}^{obs} - \mathbf{G}(P)\right)$ for each P sampled. The tolerances γ_{cdf}^{down} is taken to be 0 (due to the half-normal distribution) and γ_{cdf}^{up} as 0.7. The plots of the marginal cdfs with the cutoff level indicated in red are seen in Figure 5. Since the likelihood function is assumed to be independent of P, we need only to check that the minimum value of $T_{noise}^{L^2}\left(\mathbf{e}^{obs}\right)$ and $T_{noise}^{inter}\left(\mathbf{e}^{obs}\right)$ over all P is below the cutoff $\gamma_{cdf}^{up} = 0.7$. We find $\min T_{noise}^{L^2}\left(\mathbf{e}^{obs}\right) = 0.402$ and $\min T_{noise}^{inter}\left(\mathbf{e}^{obs}\right) = 0.683$. As these both meet the requirements, we continue with the calibration.

Introducing (28) and (29) into (4) we arrive at the calibrated posterior pdf, to within a constant,

$$\sigma_M\left(P|\mathbf{d}_c^{obs}\right) = k\frac{5.439 \times 1.738}{\pi^2}\exp\left(-D_1^2\left(5.439\right)^2/\pi\right)\exp\left(-D_2^2\left(1.738\right)^2/\pi\right) \tag{34}$$

where D_1 and D_2 are defined in equations (30) and (32) and evaluated with $\mathbf{d}_c^{obs}$. Note that the constant from the prior uniform distribution is included in the normalization constant k. Figure 6 shows a visualization of the posterior pdf. We note that, at least from the points tested, the most likely estimator (MLE) is at the point $P = 0.03375$. The $t = 3$ snapshot of this simulation is also shown in Figure 6. Examining this image, it is observed that the model $\mathcal{M}_1$ is unable to capture the initiating bump in the calibration data.

The depiction is performed by calculating σ_M at many different values of P and linearly interpolating between the points, where the function $\mathbf{G}(P)$ is evaluated for specific values of P through calculations based on the finite element models and time integration methods discussed in Sections A.2 and A.3. This method is chosen over traditional Monte Carlo methods due to the length of time required to calculate $\mathbf{G}(P)$ for any given P.

The chosen validation scenario encompasses the image of the tumor at time $t = 6$. The set up for the validation Bayesian update is the same as for the calibration update, i.e. we determine the prior, the data to be used, and the likelihood. The prior used is the same as that used in the calibration stage. Further, since we intend to use the same types of data from the tumor image at $t = 6$ as at $t = 3$, θ_{noise} takes the same form as (29) but with D_1 and D_2 calculated using $\mathbf{d}_v^{obs}$ and $u(x, 6)$ instead of $u(x, 3)$. To perform the data misfit check, we again take a uniform sample of P, calculate $\mathbf{G}(P)$, and look at the CDFs of the marginals with $\gamma_{cdf}^{down} = 0$ and $\gamma_{cdf}^{up} = 0.7$. For this set of data, we find $\min T_{noise}^{L^2}\left(\mathbf{e}^{obs}\right) = 0.726$ and $\min T_{noise}^{inter}\left(\mathbf{e}^{obs}\right) = 0.998$. Thus, with respect to the given tolerances, we see that the model fails with respect to both distances and we must reject the model. This data misfit check is visualized in Figure 7.

5.2 Assessment of $\mathcal{M}_2$ against data from $\mathcal{M}_3$

As a second example, we calibrate model $\mathcal{M}_2$ with respect to the parameters χ and D. The data used for this case will be analogous to that of the previous example, i.e. the L^2 norm

per element and the location of the interface, but the data are now generated with $\mathcal{M}_3$, shown in Figure 8, using the following time dependent values of χ and D:

$$\chi = \begin{cases} 7.27 & t \in [0, 1.705) \\ 10.91 & t \in [1.705, 3.038) \\ 8.30 & t \in [3.038, 4.965) \\ 9.33 & t \in [4.965, 6.96) \\ 11.11 & t \in [6.96, 7.1) \\ 12.08 & t \in [7.1, 7.648) \\ 7.76 & t \in [7.6848, 8.246) \\ 12.52 & t \in [8.246, 9] \end{cases} \qquad D = \begin{cases} 1.18 & t \in [0, 1.705) \\ 1.98 & t \in [1.705, 3.038) \\ 1.87 & t \in [3.038, 4.965) \\ 1.43 & t \in [4.965, 6.96) \\ 1.36 & t \in [6.96, 7.1) \\ 1.45 & t \in [7.1, 7.648) \\ 1.05 & t \in [7.6848, 8.246) \\ 1.58 & t \in [8.246, 9] \end{cases} . \quad (35)$$

The idea that data comes from a scenario in which parameters change over time is analogous to thinking of data coming from a tumor which is continually evolving and is thus phenotypically changing over time. The parameters which are not time dependent in model $\mathcal{M}_3$ are taken from Table 1.

Once again, it is necessary to define what the form of the prior probability density function is for the parameters of interest, χ and D, and what form the likelihood, $\theta_{noise}(\mathbf{e}|\chi, D)$, takes. For these steps, we will proceed in a similar manner to the previous example. Due to our prior knowledge, we begin by considering a uniform prior probability density function:

$$\rho(\chi, D) = U(3, 19.5) \times U(0.3, 1.95) \quad (36)$$

with U once again indicating a uniform distribution. Since our data is of the same type as that used earlier we also use the same likelihood function (29) with D_1 and D_2 evaluated as in equations (30) and (32), but with data from the image in Figure 8b.

We first check that the model passes the data misfit check. Again taking a uniform sample of the prior distribution for χ and D, and computing $\mathbf{G}(\chi, D)$, we see the model passes this check with $\min T^{l2}_{noise}\left(\mathbf{e}^{obs}\right) = 0.001$ and $\min T^{inter}_{noise}\left(\mathbf{e}^{obs}\right) = 0.0$ against the tolerances given in the previous section. The marginal CDFs and evaluated points are shown in Figure 9. Thus, there certainly exists (χ, D) such that our data misfit check criterion is satisfied, and we can continue with the calibration process.

Since we have assumed a uniform prior and the likelihood function as in (29), the posterior pdf is given by

$$\sigma_M\left(\chi, D|\mathbf{d}^{obs}_c\right) = k\frac{5.439 \times 1.738}{\pi^2}\exp\left(-D_1^2\left(5.439\right)^2/\pi\right)\exp\left(-D_2^2\left(1.738\right)^2/\pi\right), \quad (37)$$

the shape of which is shown in Figure 10. Of the points tested, the most likely estimator is at the point $(\chi, D) = (9.0, 1.95)$, the snapshot at $t = 3$ of this simulation is also shown in Figure 10.

This surface is generated by evaluating equation (37) at specific values of χ and D, using the discretization methods as described in the appendix, and then linearly interpolating between the points.

We now proceed to the validation stage, which we define to be the stage of the tumor at $t = 6$ (shown in Figure 8c). Proceeding as before, we again perform the data misfit check

and find that the model passes with $\min T_{noise}^{L^2}\left(\mathbf{e}^{obs}\right) = 0.003$ and $\min T_{noise}^{inter}\left(\mathbf{e}^{obs}\right) = 0.083$. These plots are shown in Figure 11.

To create the validation posterior pdf, we employ the same likelihood function and note that the posterior has the same analytic form, but is evaluated with the validation data. This pdf is shown in Figure 12. From the points tested, we again find the point $(\chi, D) = (9, 1.95)$ as the most likely estimator. The snapshot at $t = 6$ of this simulation is also shown in Figure 12. We see that while there are slight differences, there is in fact high agreement with the data.

Finally, to check the validity of model $\mathcal{M}_2$, we must calculate the prediction: the quantity of interest $Q(u)$ at $t = 9$, chosen to be the tumor volume as calculated through (27). Since our model is deterministic, we can view the problem of finding the pdf for the QoI as a change of variables in both σ_M and σ_V from (χ, D) to $q =$ (tumor volume at $t = 9$), i.e. $q = Q(\chi, D)$, (see [26] pg. 25) i.e.

$$p(q) = k\, \sigma_M\left(Q^{-1}(q)\,|\mathbf{d}_c^{obs}\right). \tag{38}$$

To determine q, we take each pair of values of (χ, D) for which σ_M and σ_V were originally calculated, extend the calculation to compute the QoI, $Q(\chi, D)$, and associate the value $\sigma_M\left(\chi, D|\mathbf{d}_c^{obs}\right)$ (or σ_V) with $p(q) = Q(\chi, D)$, and renormalize. We show both of these QoI cdfs ($P(q_c)$ and $P(q_v)$) in Figure 13. Both σ_M and σ_V had the same MLE; the prediction of the simulation using these values is also shown in Figure 13. The targeted outputs delivered by this model appear to be in good agreement with the data.

We now come to the central question: "Is $\mathcal{M}_2$ a valid (a non-invalid) model for this QoI?" We use the tolerance $\gamma_{tol} = 10\% \times Q(9.0, 2.1) = 10\% \times 4.06 = 0.406$. Using (12), we find

$$M\left(q_p^C, q_p^V\right) = 0.1478 \leq \gamma_{tol} \tag{39}$$

meaning that the model has not been found to be invalid.

While the model has now been declared not invalid, we still must determine how to answer the question "What will the volume of the tumor be at $t = 9$?" There are, in fact, various ways to answer this. One answer may be given in terms of the volume associated with the most likely estimator $\pm$ the standard deviation, in this case, 4.06 ± 0.2203. Another answer may be stated with regard to the mean of the QoI pdf $\pm$ the standard deviation, namely, 3.7443 ± 0.2203. (Note that we must use the calibrated pdf, since the one used for the validation of the calibration pdf has not yet been validated itself.) We may also choose to offer an interval, say the 90% confidence interval associated with our QoI pdf. For the present case this corresponds approximately to the interval $[3.43, 4.1]$. However answered, the proposed framework offers an avenue to answer the question with a level of uncertainty (confidence) associated with it. The issue of making decisions based on these types of answers is addressed by the field of decision theory, a subject of research outside the current objectives of this paper.

6 Conclusion

As new methodologies emerge for acquiring data on the evolution of tumors in specific subjects, ranging from various minimally invasive imaging modalities to biopsies, blood

tests, and other more invasive observational procedures, the issue arises as to how these data ultimately inform the processes critical to predicting the growth or decline of cancer and the effects of various treatments. A fundamental premise of computational science is that properly derived computational models can be used to map relevant data into predictions of critical features of physical events that, in time, can be used as a basis for making important decisions. The fidelity of such computer predictions depends upon the models used, the knowledge of the key model parameters, and the quality of the data.

In the present exposition, we describe a unified approach for statistical calibration and validation of models and prediction of quantities of interest based on Bayesian inference. The models we study here are based on continuum mixture theory and diffused-interface characterizations of the free energy of a volume of interacting constituents, the behavior of which is constrained by conservation laws of mechanics. Importantly, the approach can take into account uncertainties in parameters, observations, and the model itself and lead to predictions with quantifiable uncertainty.

While we demonstrated the validation process using phenomenological models from mixture theory, the process itself is quite general, and is applicable to virtually any modeling scenario, including those that attempt to model events at the cellular level. We believe that complementing these validation processes in conjunction with a rich source of relevant data can lead to a truly predictive approach to tumor growth modeling.

Acknowledgments: The authors gratefully recognize the support by the Department of Energy under Award Number DE-FC52-08NA28615. The authors also acknowledge useful discussions with Ivo Babuška, Hector Gomez, Luca Dede, and Jesse Windle.

References

[1] W. L. Oberkampf and C. J. Roy. *Verification and Validation in Scientific Computing*. Cambridge University Press, 2010.

[2] H. W. Coleman and W. G. Steele. *Experimentation, Validation, and Uncertainty Analysis for Engineers*. John Wiley & Sons, 2009.

[3] P. J. Roache. *Fundamentals of Verification and Validation*. Hermosa Publishers, New Mexico, 2nd edition, 2009.

[4] I. Babuška, F. Nobile, and R. Tempone. A systematic approach to model validation based on bayesian updates and prediction related rejection criteria. *CMAME*, 197:2517–2539, 2008.

[5] J. T. Oden, A. Hawkins, and S. Prudhomme. General diffuse-interface theories and an approach to predictive tumor growth modeling. *Mathematical Models and Methods in Applied Sciences*, 0(1):1–41, 2010.

[6] W. L. Loh. On latin hypercube sampling. *Ann. Stat.*, 25(5):2058–2080, 1996.

[7] E. Prudencio and K. Schulz. The Parallel C++ Statistical Library 'QUESO': Quantification of Uncertainty for Estimation, Simulation and Optimization. *IEEE IPDPS*, 2010. Submitted.

[8] B. M. Adams, W. J. Bohnhoff, K. R. Dalbey, J. P Eddy, M. S. Eldred, D. M. Gay, K. Haskell, P. D. Hough, and L. P Swiler. DAKOTA, A multilevel parallel object-oriented framework for design optimization, parameter estimation, uncertainty quantification, and sensitivity analysis: Version 5.0 user's manual. Technical Report SAND2010-2183, Sandia National Laboratory, December 2009.

[9] R. G. Ghanem and P. D. Spanos. *Stochastic Finite Elements: A Spectral Approach*. Springer-Verlag, 1991.

[10] I. Babuška, F. Nobile, and R. Tempone. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM J. Num. Anal.*, 45(3):1005–1034, 2007.

[11] M. Kleiber and T. D. Hien. *The Stochastic Finite Element Method*. John Wiley & Sons Ltd., 1992.

[12] W. K. Liu, T. Belytschko, and A. Mani. Random field finite elements. *Int. J. Num. Meth. Engng.*, 23:1831–1845, 1986.

[13] A. Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM Publishers, Philadelphia, 2005.

[14] J. Kaipio and E. Somersalo. *Statistical and Computational Inverse Problems*. Springer, 2005.

[15] D. Ambrosi and F. Mollica. On the mechanics of a growing tumor. *Int. J. Eng. Sci.*, 40(12):1297–1316, 2002.

[16] A. R. A. Anderson and M. A. J. Chaplain. Continuous and discrete models of tumor-induced angiogenesis. *Bull. Math. Biol.*, 60:857–899, 1998.

[17] H. M. Byrne and S. A. Gourley. The role of growth factors in avascular tumour growth. *Math. Comput. Modelling*, 4:35–55, 1997.

[18] M. A. J. Chaplain and B. D. Sleeman. Modelling the growth of solid tumours and incorporating a method for their classification using nonlinear elasticity theory. *J. Math Biol.*, 31:431–473, 1993.

[19] P. A. DiMilla, K. Barbee, and D. A. Lauffenburger. Mathematical model for the effects of adhesion and mechanics on cell migration speed. *J. Biophys.*, 60:15–37, 1991.

[20] J. P. Ward and J. R. King. Mathematical modelling of avascular tumour growth. *IMA J. Math Appl. Med. Biol.*, 14:36–69, 1997.

[21] R. P. Araujo and D. L. S. McElwain. A mixture theory for the genesis of residual stresses in growing tissues i: A general forumuation. *SIAM J. Appl. Math.*, 65:1261–1284, 2005.

[22] V. Cristini, X. Li, J. S. Lowengrub, and S. M. Wise. Nonlinear simulations of solid tumor growth using a mixture model: invasion and branching. *J. Math. Biol.*, 58:723–763, 2009.

[23] H. B. Frieboes, F. Jim, Y. L. Chuang, S. M. Wise, J. S. Lowengrub, and V. Cristini. Three-dimensional multipspecies nonlinear tumor growth–ii: Tumor invasion and angiogenesis. *J. Theor. Biol.*, 264:1254–1278, 2010.

[24] S. M. Wise, J. S. Lowengrub, H. B. Frieboes, and V. Cristini. Three-dimensional diffuse-interface simulation of multispecies tumor growth-I. Model and numerical method. *J. Theor. Biol.*, 253(3):523–543, 2008.

[25] B. D. Coleman and W. Noll. The thermodynamics of elastic materials with heat conduction and viscosity. *Arch. Ration. Mech. Anal.*, 13:167–178, 1963.

[26] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin. *Bayesian Data Analysis*. Chapman & Hall/CRC, second edition, 2004.

[27] C. M. Elliott. The Cahn–Hilliard model for the kinetics of phase separation. In J.F. Rodrigues, editor, *Mathematical Models for Phase Change Problems: Proceedings of the European Workshop held at Óbidos, Portugal, October 1–3, 1988*, volume 88 of *International Series of Numerical Mathematics*, pages 35–73. Birkhäuser, 1989.

[28] C. M. Elliott and H. Garcke. On the Cahn–Hilliard equation with degenerate mobility. *SIAM J. Math. Anal*, 27(2):404–423, 1996.

[29] L. A. Caffarelli and N. E. Muler. An L^∞ bound for solutions of the Cahn–Hilliard equation. *Arc. Rat. Mech. Anal.*, 133:129–144, 1995.

[30] D. J. Eyre. Unconditionally gradient stable time marching the Cahn–Hilliard equation. In J. W. Bullard, L.-Q. Chen, R. K. Kalia, and A. M. Stoneham, editors, *Computational and Mathematical Models of Microstructural Evolution*, volume 529 of *Mater. Res. Soc. Symp. Proc.*, pages 39–46, Warrendale, PA, 1998. Materials Research Society.

[31] S. M. Wise, C. Wang, and J. S. Lowengrub. An energy-stable and convergent finite-difference scheme for the phase field crystal equation. *SIAM J. Num. Anal.*, 47(3):2269–2288, 2009.

[32] S. M. Wise. Unconditionally stable finite difference, nonlinear multigrid simulation of the Cahn–Hilliard–Hele-Shaw system of equations. *J. Sci. Comp.*, 44:38–68, 2010.

A Discretization of the models

In this section, we consider the discretization and numerical solution of the three models described in section 4.2. We consider a mixed finite element approximation in space combined with a semi-implicit time-stepping algorithm. We focus on model $\mathcal{M}_3$ given in (26). The discretization of $\mathcal{M}_1$ can be derived from $\mathcal{M}_3$ as a special case. Of course, $\mathcal{M}_2$ is equivalent to $\mathcal{M}_3$ up to time-independent parameters.

A.1 Weak formulation

Anticipating possible low regularity in solutions and in preparation for general mixed finite element approximations, we first consider a weak-mixed formulation of (26). Let V, V_g and W_{u_0} denote the spaces

$$\mathsf{V} := L^2\left(0,T;H^1\left(\Omega\right)\right)\ , \qquad \mathsf{V}_g := \left\{v \in \mathsf{V} : v = g \text{ on } \partial\Omega\ ,\ \text{a.e. } t\right\},$$
$$\mathsf{W}_{u_0} := \left\{v \in \mathsf{V} \cap L^\infty(0,T;L^\infty(\Omega))\,,\ v_t \in \mathsf{V}' = L^2\big(0,T;(H^1(\Omega))'\big)\,,\ v(0) = u_0\right\}.$$

In particular, note that $\mathsf{V}_0 = L^2\left(0,T;H_0^1(\Omega)\right)$ and $\mathsf{V}_1 = \{1\} + \mathsf{V}_0$. We shall now define a weak solution of (25) to be a triple $\mathbf{u} := (u,\mu,c) \in \mathsf{W}_{u_0} \times \mathsf{V} \times \mathsf{V}_1$ such that

$$\int_0^T \Big(\left\langle u_t,\varphi\right\rangle + B(\mathbf{u};\mathbf{w})\Big)\,\mathrm{d}t = 0 \qquad \forall \mathbf{w} := (\varphi,\eta,\xi) \in \mathsf{V}\times\mathsf{V}\times\mathsf{V}_0\ , \tag{40}$$

where $B\left(\cdot;\cdot\right)$ is the semi-linear form,

$$\begin{aligned} B\left(\mathbf{u};\mathbf{w}\right) = &\left(Mu^2\nabla\mu,\nabla\varphi\right) - \left(Pcu - Au,\varphi\right) \\ &+ \left(\mu - f'\left(u\right),\eta\right) + \epsilon\left(\chi\, c,\eta\right) - \epsilon^2\left(\nabla u,\nabla\eta\right) \\ &+ \left(D\,\nabla c,\nabla\xi\right) + \left(cu,\xi\right)\ , \end{aligned} \tag{41}$$

$\left\langle\cdot,\cdot\right\rangle$ denotes the duality pairing between $\left(H^1\left(\Omega\right)\right)'$ and $H^1\left(\Omega\right)$, $\left(\cdot,\cdot\right)$ denotes the L^2 inner product over Ω, and $\left(\cdot,\cdot\right)_{\partial\Omega}$ is the L^2 inner product over $\partial\Omega$.

The well-posedness of the above weak formulation (40) is an open question not considered here. The major difficulties are the nonlinear mobility Mu^2 and the nonlinear coupling between c and u, i.e. the reaction terms cu. On the other hand, the simpler model $\mathcal{M}_1$ is a compact perturbation of the Cahn–Hilliard equation, the well-posedness of which can be found in, for instance [27] and [28]. The semilinear form (41) is bounded above for u bounded (in $\mathsf{W}_{u_0} \subset L^\infty(0,T;L^\infty(\Omega))$), but this may not be true for certain cases (see [29] for a related result for the Cahn-Hilliard equation). Thus, existence and well-posedness of solutions to (40) is an open question.

A.2 Semi-discrete finite element approximations

We now address the construction of semi-discrete conforming finite element approximations of the system (40). Such a formulation involves a discretization of the spatial variations of (u,μ,c) keeping, for the moment, the time variations continuous in time.

Let $\mathcal{T}^h$ denote a member of a family of partitions of domain Ω into meshes of non-overlapping convex finite elements Ω_k such that $\overline{\Omega} = \bigcup_{k=1}^{N(h)} \overline{\Omega}_k$ and $\Omega_k \cap \Omega_j = \emptyset\,,\ k \neq j$.

Each Ω_k is the image of a master element $\tilde{\Omega}$ under an invertible (generally affine) map F_k. We construct finite-dimensional subspaces of $H^1(\Omega)$ and $H_0^1(\Omega)$ defined by

$$\mathcal{S}^h := \left\{ v^h \in H^1(\Omega) : v^h|_{\Omega_k} = \hat{v} \circ F_k^{-1}, 1 \le k \le N(h), \hat{v} \in \mathbb{P}^K\left(\overline{\tilde{\Omega}}\right) \right\} , \tag{42}$$

and $\mathcal{S}_0^h := \mathcal{S}^h \cap H_0^1(\Omega)$, respectively. We further set $\mathcal{S}_1^h := \{1\} + \mathcal{S}_0^h$. Here $\mathbb{P}^K(\tilde{\Omega})$ is either the space of polynomials of degree $\le K$ defined on the closure of $\tilde{\Omega}$ or the space of tensor products of polynomials of degree K on the closure of $\tilde{\Omega}$. With these notations and conventions in place, the semi-discrete approximation $\mathbf{u}^h := (u^h, \mu^h, c^h) : [0,T] \to \mathcal{S}^h \times \mathcal{S}^h \times \mathcal{S}_1^h$ is defined as follows:

$$\left(u_t^h(t), \varphi^h\right) + B(\mathbf{u}^h(t); \mathbf{w}^h) = 0 \qquad \forall \mathbf{w}^h := (\varphi^h, \eta^h, \xi^h) \in \mathcal{S}^h \times \mathcal{S}^h \times \mathcal{S}_0^h , \tag{43}$$

for a.e. $t \in (0,T)$. The initial condition for u^h is set by $u^h(0) = \mathcal{I}^h u_0$, where $\mathcal{I}^h : H^1(\Omega) \to \mathcal{S}^h$ is a suitable interpolation or projection.

Let us denote by $\{\varphi_i(\mathbf{x})\}_{i=1}^N$ a set of basis functions of $\mathcal{S}^h$ generated on a partition $\mathcal{T}^h$ by the finite element approximations: $\mathcal{S}^h = \operatorname{span}\{\varphi_i\}$. Then each of the component approximations is of the form

$$u^h(t,\mathbf{x}) = \sum_{i=1}^N u_i(t)\,\varphi_i(\mathbf{x}) , \quad \mu^h(t,\mathbf{x}) = \sum_{i=1}^N \mu_i(t)\,\varphi_i(\mathbf{x}) , \quad c^h(t,\mathbf{x}) = \sum_{i=1}^N c_i(t)\,\varphi_i(\mathbf{x}) . \tag{44}$$

Here $u_i(t)$, $\mu_i(t)$ and $c_i(t)$ are continuous functions in time. Upon introducing these into (43), we obtain a system of nonlinear ordinary differential equations (ODE's) in the unknown discrete variables, $u_i(t)$, $\mu_i(t)$ and $c_i(t)$. We proceed by developing a time discretization of these ODE's.

A.3 Time discretization

An attractive class of robust schemes for time-integration of Cahn–Hilliard-type problems are semi-implicit schemes [30, 31, 32]. Semi-implicit schemes are preferred for their superior stability properties allowing large time-steps to be taken. These schemes split the free energy f in a contractive and expansive part, $f(u) = f_c(u) - f_e(u)$, and accordingly treat f_c implicitly and f_e explicitly. The splitting is such that f_c and f_e are both convex functions, at least in the domain of interest (for $u \in [0,1]$). We shall employ the splitting $f_c(u) = \frac{3}{2}u^2$ and $f_e(u) := f_c(u) - f(u)$.

Let Δt denote the time step and $t_n = n\Delta t$, for $n = 0, \ldots, N := T/\Delta t$ corresponding discrete times. We denote the approximation to $\mathbf{u}^h(t_n)$ by $\mathbf{u}_n$. Then, the fully discrete approximation $\mathbf{u}_n := (u_n, \mu_n, c_n) \in \mathcal{S}^h \times \mathcal{S}^h \times \mathcal{S}_1^h$ is defined recursively by

$$\left(\frac{u_n - u_{n-1}}{\Delta t}, \varphi^h\right) + \left(M u_{n-1}^2 \nabla\mu_n, \nabla\varphi^h\right) - \left(P c_{n-1} u_n - A u_n, \varphi\right) = 0 \quad \forall \varphi^h \in S^h , \tag{45a}$$

$$\left(\mu_n - f_c'(u_n) + f_e'(u_{n-1}), \eta^h\right) + \epsilon\left(\chi(t_n)\, c_n, \eta^h\right) - \epsilon^2\left(\nabla u_n, \nabla\eta^h\right) = 0 \quad \forall \eta^h \in S^h , \tag{45b}$$

$$\left(D(t_n)\,\nabla c_n, \nabla\xi^h\right) + \left(c_n u_{n-1}, \xi^h\right) = 0 \quad \forall \xi^h \in S_0^h , \tag{45c}$$

for $n = 1, \ldots, N$. The scheme is initialized using the initial condition $u_0 = u^h(0) = \mathcal{I}^h u_0$.

Note that analogous to f, we have treated the other nonlinear terms ($Mu^2\nabla\mu$, cu) in a semi-implicit fashion. Since our splitting of $f(u)$ resulted in a linear contractive part $f_c'(u)$, the above scheme requires the solution of a linear system of equations at each time step.

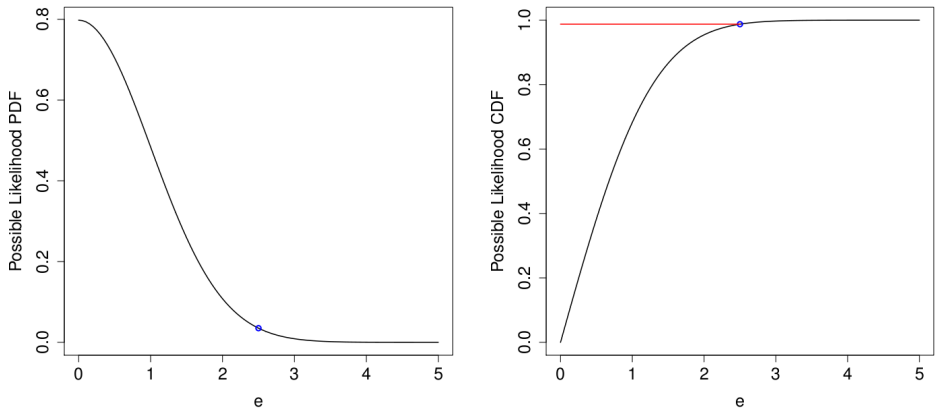


Figure 1 *Example data misfit check for a likelihood pdf distributed as a half-normal distribution* $HN(\sqrt{\pi/2})$. *Left: Likelihood pdf (–); example* $T\left(\mathbf{e}^{obs} = \mathbf{d}^{obs} - \mathbf{G}\left(\mathbf{m}\right) = 2.5\right)$ *indicated with a circle (○). Right: Likelihood cdf with example* $\mathbf{e}^{obs} = 2.5$ *indicated with a circle. In this example,* $T\left(\mathbf{e}^{obs}|\mathbf{m}\right) = 0.9875$, *thus if* $\gamma_{cdf}^{up} \leq 0.9875$, *the model would fail this check for this* $\mathbf{m}$.

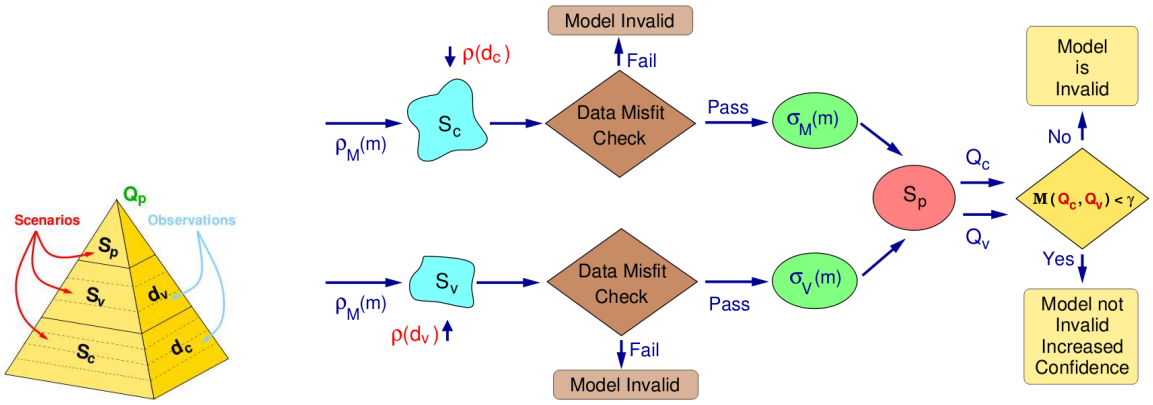


Figure 2 *The prediction pyramid, and a flow chart describing a possible statistical validation process (inspired by [4]).*

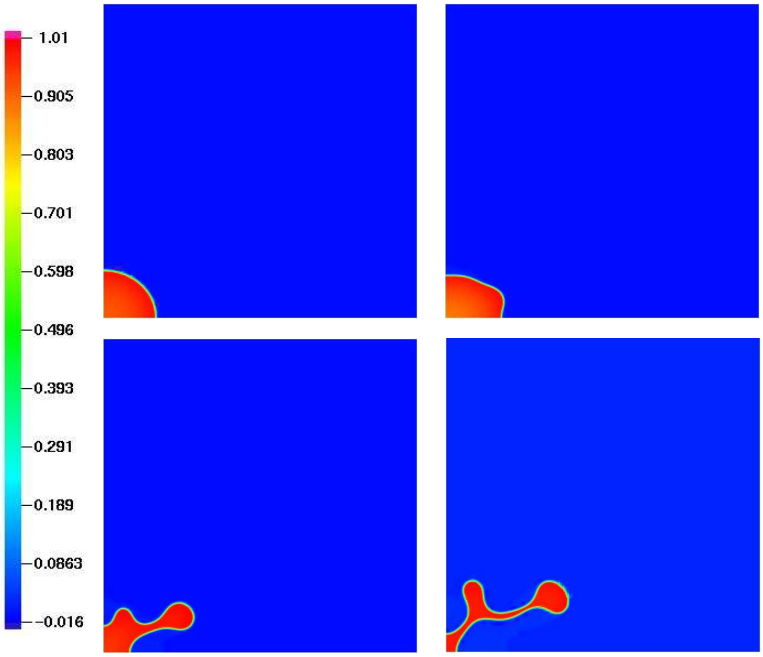


Figure 3 *Sequential two-dimensional images of the progressive growth of a tumor at time* $t = 0, 3, 6, 9$*, originating from an elliptical shape,* $(x/2.1)^2 + (y/1.9)^2 = 1$ *at* $t = 0$*, with a quarter of the domain* $\Omega = (-12.8, 12.8)^2$ *shown and two-plane symmetry assumed. Images generated using* $\mathcal{M}_2$ *with parameters given in Table 1. These (virtual) images are to be used for calibration* $(t = 3)$ *and validation* $(t = 6)$ *for* $\mathcal{M}_1$*.*

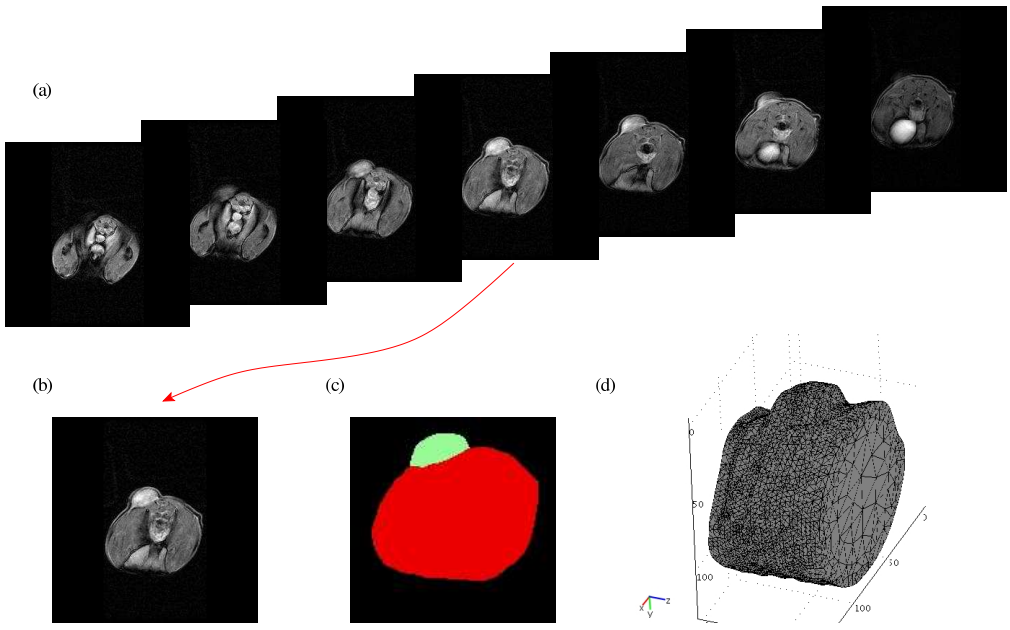


Figure 4 *(a) Spatial series of MRI images of a tumor implanted on a mouse's back. (b) Two-dimensional cross-slice containing the largest area of tumor. (c) Segmentation of (b) into tumor and non-tumor regions. (d) Mesh generated by images in (a).*

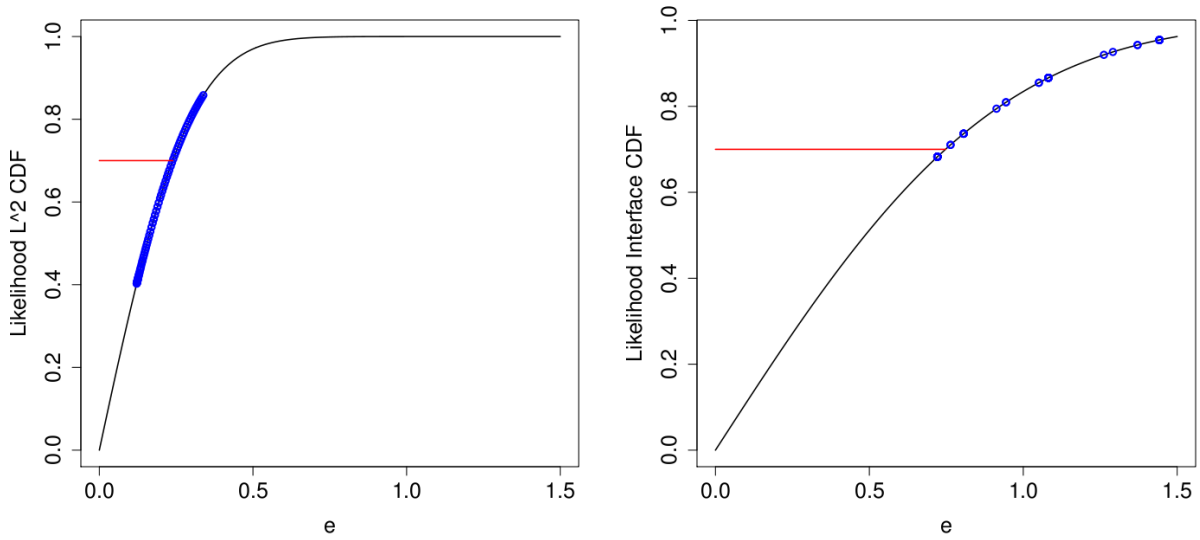


Figure 5 *Plots of the calibration likelihood function marginal corresponding to the L^2 pixel data (left) and interface data (right) CDF. Circles marks $\mathbf{e}^{obs}$ for all $\mathbf{m}$ tested. Line corresponding to $\gamma_{cdf}^{up} = 0.7$ is shown.*

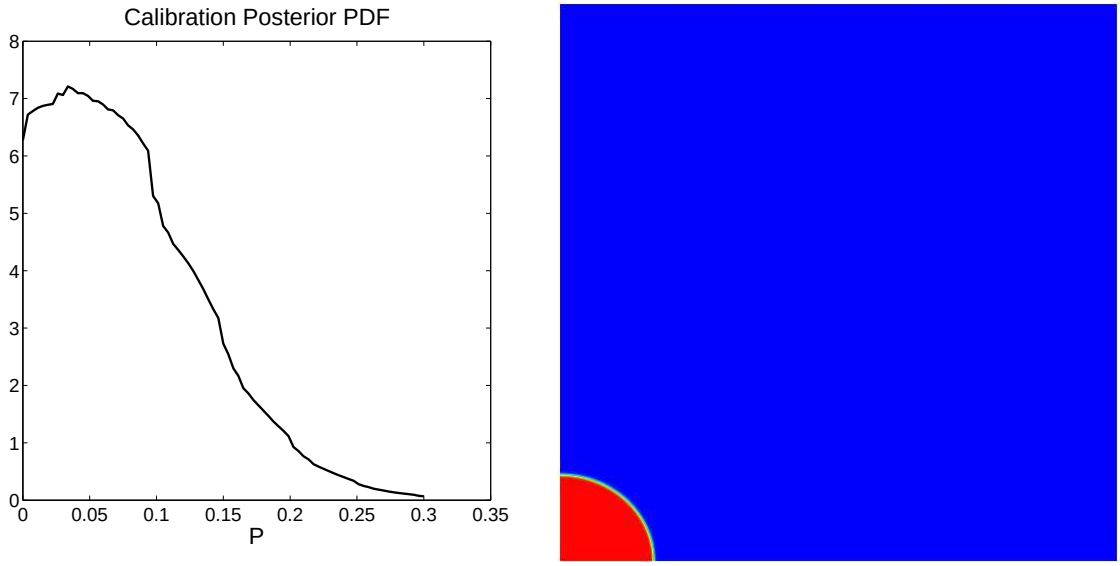


Figure 6 *Left:Calibrated posterior pdf $\sigma_M\left(P|\mathbf{d}^{obs}\right)$. Right: Result of simulation using $\mathcal{M}_1$ with the most likely estimator for $P = 0.03375$ at $t = 3$ with the data values as mapped through $\mathbf{G}_c\left(P\right)$.*

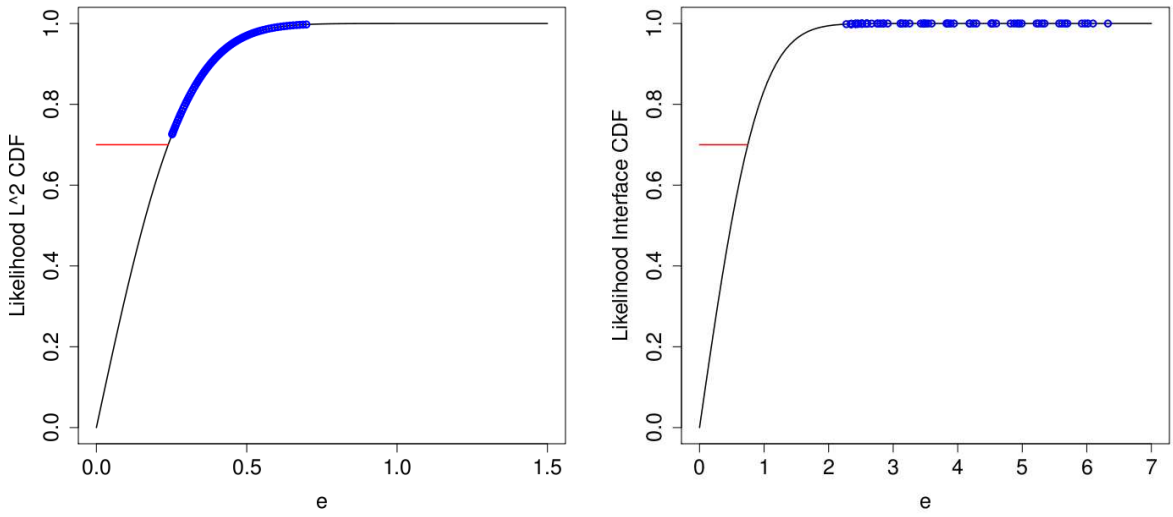


Figure 7 *Plots of the marginals of the likelihood function under the validation scenario. Circle marks* $\mathbf{e}$ *for all* $\mathbf{m}$ *tested. Line corresponding to* $\gamma_{cdf}^{up} = 0.7$ *is shown.*

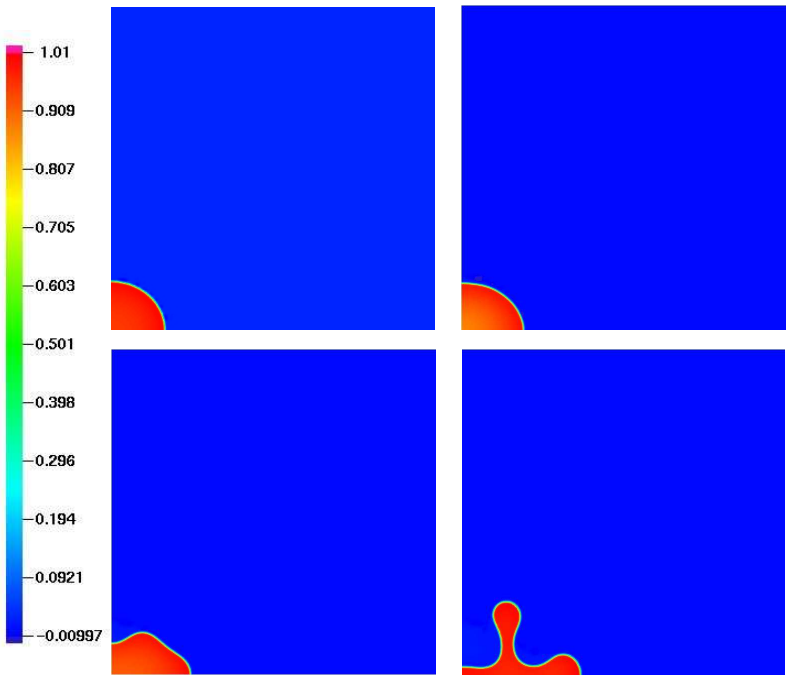


Figure 8 *A sequence of two-dimensional images of the progressive growth of a tumor at time* $t = 0, 3, 6, 9$*, originating from an elliptical shape,* $(x/2.1)^2 + (y/1.9)^2 = 1$ *at* $t = 0$*, with a quarter of the domain* $\Omega = (-12.8, 12.8)^2$ *shown and two-plane symmetry assumed. Images generated with model* $\mathcal{M}_3$ *with parameters in (35). These (virtual) images are to be used for calibration (image at* $t = 3$*) and validation (image at* $t = 6$*) for* $\mathcal{M}_2$*.*

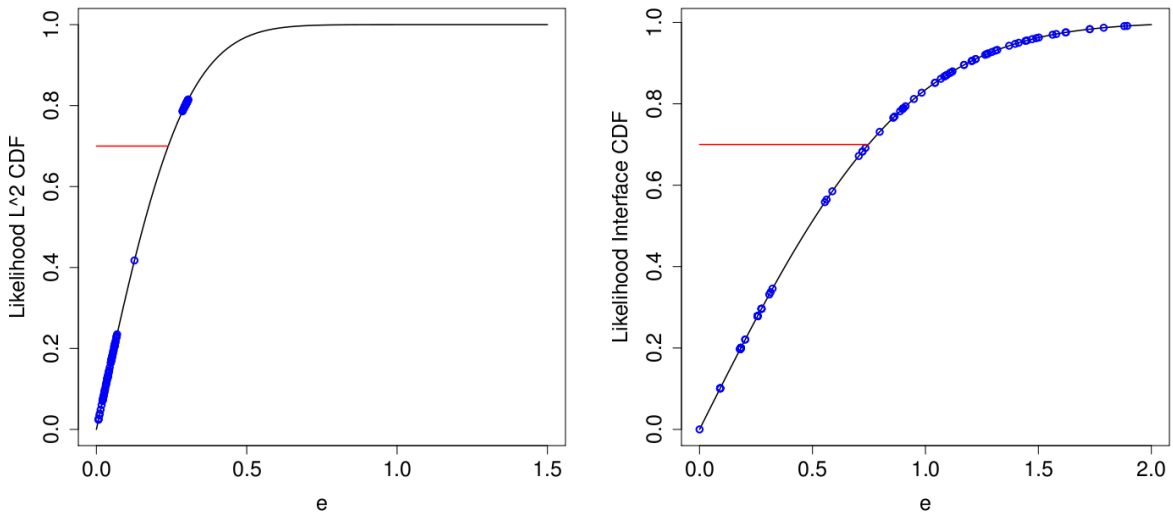


Figure 9 *Plots of the marginals of the likelihood function with calibration data from* $\mathcal{M}_3$ *for* $\mathcal{M}_2$ *(29). Circle marks values of* $\mathbf{e}^{obs}$ *for all tested values of* (χ, D)*. The cutoff level is marked with a red line.*

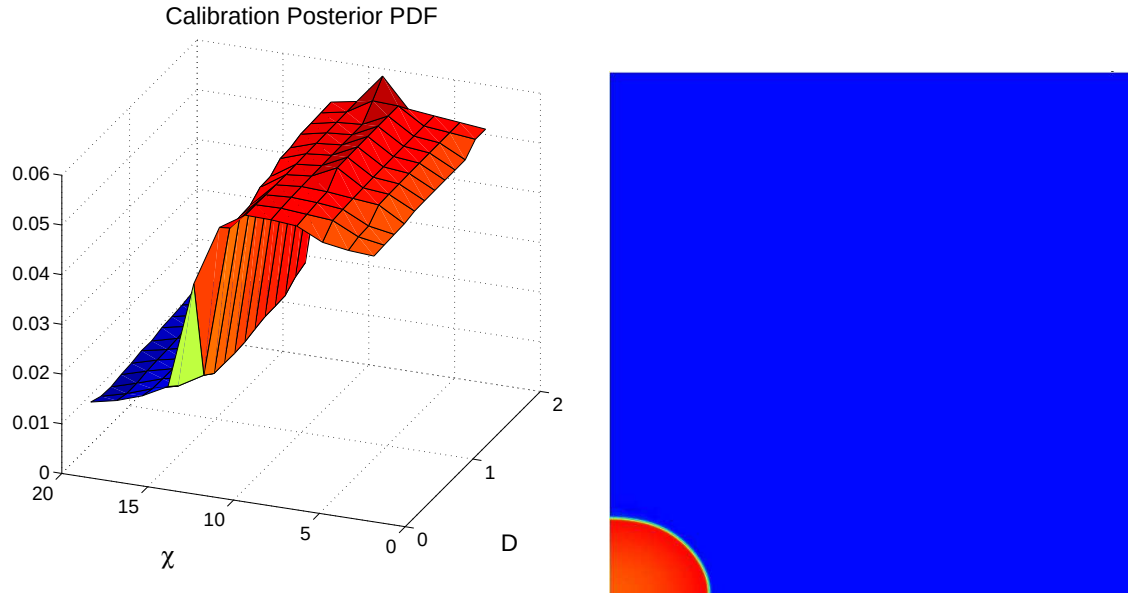


Figure 10 *Left: Calibrated posterior pdf* $\sigma_M\left(\chi, D|\mathbf{d}_c^{obs}\right)$*. Right: Result of simulation using* $\mathcal{M}_2$ *with the most likely estimator of* $(\chi, D) = (9.0, 1.95)$ *at* $t = 3$*.*

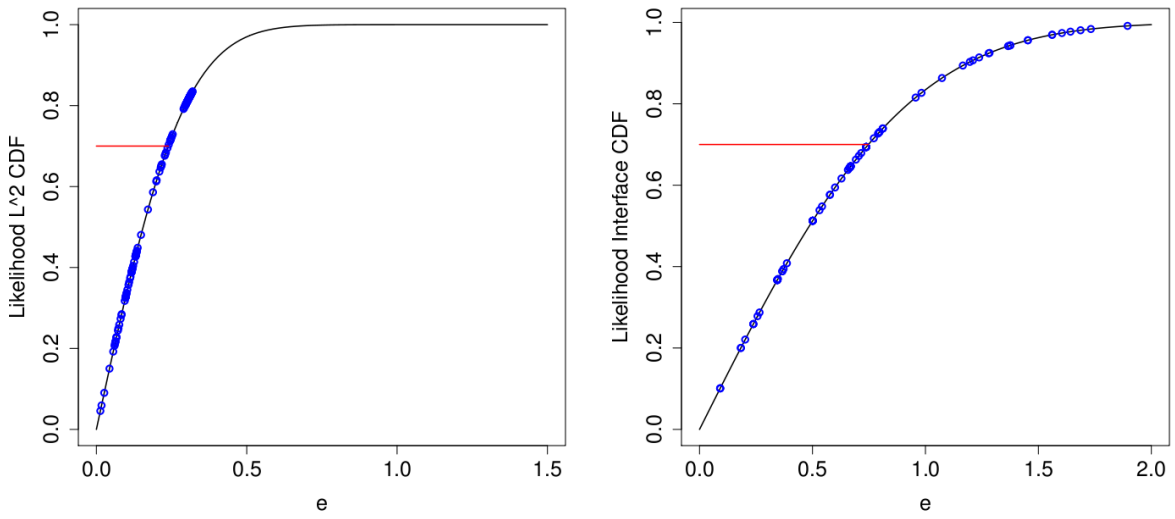


Figure 11 *Plots of the marginals of the likelihood function with validation data from* $\mathcal{M}_3$ *for* $\mathcal{M}_2$ *(29). Circle marks values of* $\mathbf{e}^{obs}$ *for all tested values of* (χ, D)*. The cutoff level is marked with a red line.*

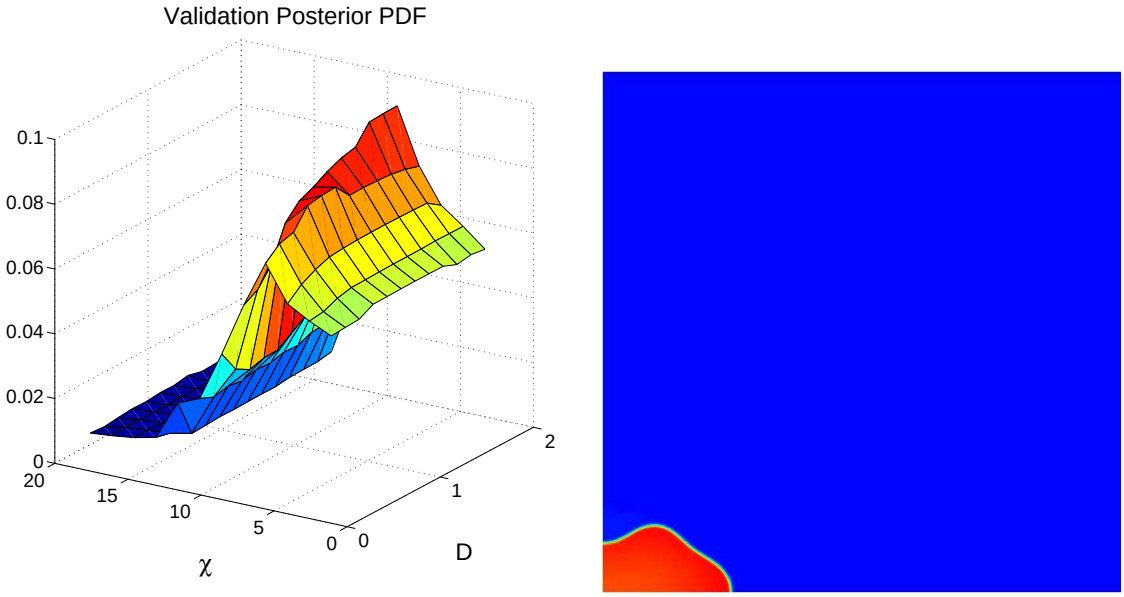
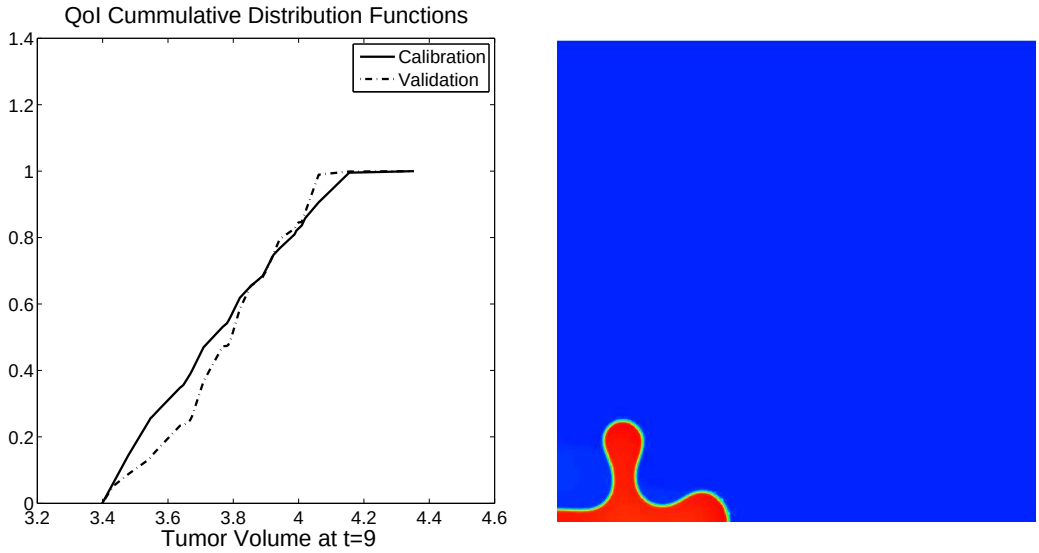


Figure 12 *Left: Validation posterior pdf* $\sigma_V\left(\chi, D|\mathbf{d}_v^{obs}\right)$*. Right: Result of simulation using* $\mathcal{M}_2$ *with the most likely estimator of* $(\chi, D) = (9.0, 1.8)$ *at* $t = 6$*.*



Calibration		Validation	
Mean:	3.7443	Mean:	3.7701
Std. Dev.	0.2203	Std. Dev.	0.1852

Figure 13 *Left: Cumulative distribution functions for the tumor volume at* $t = 9$ *as determined with both the calibration (solid line) and the validation (dashed line) posterior pdfs. Right: Prediction using the MLE.*