From *SIAM News*, Volume 44, Number 5, June 2011

***CSE 2011***

# Bayesian Calibration, Validation, and Uncertainty Quantification for Models of Tumor Growth

*By Andrea Hawkins-Daarud*

The successful use of computational models to predict physical events depends on several fundamental concepts and processes: the mathematical model itself, the particular quantities of interest (QoI's) of the physical event, and, for predictability, experimental observations made for calibration and validation. Each of these concepts is described in detail in a two-part article, "Computer Predictions with Quantified Uncertainty," that appeared in *SIAM News* in November and December, 2010.

## Tumor Growth Example

Using models of tumor growth, the present article provides a concrete example of the abstract concepts presented in the earlier articles. Of the many types of tumor growth models, only those derived through the continuum theory of mixtures are considered here [1–3]. This multiphase theory represents 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. Details about the theory can be found in [1,4].

For illustration, a two-phase isothermal mixture consisting of tumor $u$ and non-tumor $n$ (healthy tissue, extracellular fluid, etc.), along with a representative nutrient $c$, say oxygen, is considered. Under a saturation assumption, only $u$ and $c$ are taken as unknown. $\mathcal{M}_1$ and $\mathcal{M}_2$ are two such models; they have the same boundary and initial conditions but differ in that $\mathcal{M}_1$ has constant parameters and $\mathcal{M}_2$ has time-dependent parameters:

$$\begin{aligned} u_t &= \nabla\cdot(Mu^2\nabla\mu) + Pcu - Au && \text{in } (0,T)\times\Omega,\\ \mu &= f'(u) - \varepsilon^2\Delta u - \varepsilon\chi c && \text{in } (0,T)\times\Omega,\\ 0 &= \nabla\cdot(D\nabla c) - cu && \text{in } (0,T)\times\Omega \qquad (\mathcal{M}_1)\end{aligned}$$

$$\begin{aligned} u_t &= \nabla\cdot(Mu^2\nabla\mu) + Pcu - Au && \text{in } (0,T)\times\Omega,\\ \mu &= f'(u) - \varepsilon^2\Delta u - \varepsilon\chi(t) c && \text{in } (0,T)\times\Omega,\\ 0 &= \nabla\cdot(D(t)\nabla c) - cu && \text{in } (0,T)\times\Omega \qquad (\mathcal{M}_2)\end{aligned}$$

$\mathcal{M}_2$ is used to generate (virtual) data (see Figure 1), against which the validity of $\mathcal{M}_1$ is assessed for the QoI of the final tumor volume. Snapshot images are taken at $t_1 = 3$ and at $t_2 = 6$ for calibration and validation data, respectively. The question addressed is thus: Based on the data observed in Figure 1 at $t_1$ and $t_2$, is model $\mathcal{M}_1$ invalid for predicting the QoI, $Q(u)$ = (tumor volume at $t = 9$)?

**Figure 1.** *Two-dimensional images of the progressive growth of a tumor at times t = 3 and 6 for use as calibration and validation data. The images were generated with model $\mathcal{M}_2$ with two-plane symmetry assumed.*

We begin the process by defining the model parameters of interest for calibration, the form of the prior, the data that will be used for the calibration, and the form of the likelihood function. For $\mathcal{M}_1$, the key model parameters chosen for calibration are $\mathbf{m} = (\chi, D)$. Due to "prior knowledge," the prior pdf is taken as the uniform distribution $\rho(\chi, D) = U(3,19.5) \times U(0.3,1.95)$. The simulation mesh provides an analogue of a pixel array typical of an MRI image. For both the calibration and the validation steps, therefore, we consider as observational data the $L^2$ norm of the observed pixels "recovered" from the image and the position of the pixels on which the interface lies.

For simplicity, the pdf $\theta_{\text{noise}}(\mathbf{e})$ is taken to be independent of $\mathbf{m}$ and is a bivariate, un-correlated, half-normal distribution, i.e., pdf $\theta_{\text{noise}}(\mathbf{e}) = (\theta_1\theta_2)/\pi^2 \exp(-M_{d,1}^2\theta_1^2/\pi)\exp(-M_{d,2}^2\theta_2^2/\pi)$, where $\theta_1$ and $\theta_2$ are the distributional parameters and $M_{d,i} = M_{d,i}(\mathbf{G}(\chi, D), \mathbf{d}^{\text{obs}})$ denotes the distance between the simulation and the observed data. The parameters $\theta_1$ and $\theta_2$ are chosen to correspond to 10% relative error. The given prior and likelihood functions are used to calculate the calibrated posterior pdf (Figure 2, left). The setup for the validation Bayesian update is the same as for the calibration update: Determine the prior, the data to be used, and the likelihood. Taking them to be of the same form as the calibration, we obtain the corresponding validation posterior pdf (Figure 2, right).

Finally, we check the validity of the prediction of the model $\mathcal{M}_1$ by calculating the prediction QoI, $q(u(\chi, D))$, chosen to be the tumor volume. To determine $q$, the values $(\chi, D)$ for which $\sigma_M$ and $\sigma_V$ were originally calculated are used to extend the calculation to compute the tumor volume from $u(\chi, D)$, and associate $q(u(\chi, D))$ with the value $\sigma_M(\chi, D|\mathbf{d}_c^{\text{obs}})$ (or $\sigma_V$). Both of these QoI cumulative distribution functions are shown in Figure 3. Both $\sigma_M$ and $\sigma_V$ have the same most likely estimator (MLE); the prediction from $\mathcal{M}_1$ with these values is also shown in Figure 3.

A final question remains: Is $\mathcal{M}_1$ a valid (a non-invalid) model for this QoI? The metric used to compare these two cdf's is the largest difference between the inverse of the cdf's with a tolerance $\gamma_{tol} = 10\% \times Q(\text{MLE}) = 10\% \times 4.06 = 0.406$. It is calculated that $M(q_p^C, q_p^V) = 0.1478 \leq \gamma_{tol}$, meaning that the model has not been found to be invalid.

The model having been declared "not invalid," we need to find a way to answer the question, What will the volume of the tumor be at $t = 9$? There are, in fact, various ways to answer: the most likely estimator volume ± the standard deviation, the mean of the QoI pdf ± the standard deviation, or an interval, say the 90% confidence interval associated with the QoI pdf. However the question is answered, the proposed framework offers an avenue for giving an answer with a meaningful level of uncertainty (confidence).

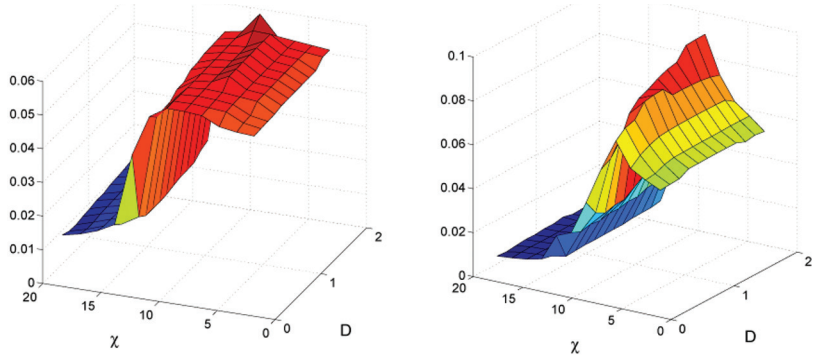
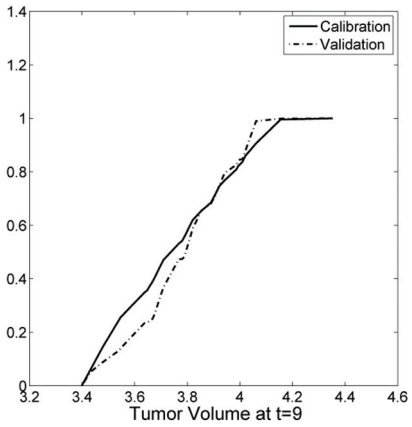



**Figure 2.** *Left: Calibrated posterior pdf* $\sigma_M(\chi, D|\mathbf{d}_c^{obs})$. *Right: Validation posterior pdf* $\sigma_V(\chi, D|\mathbf{d}_v^{obs})$.

## Conclusions

As new methodologies emerge for acquiring data on the evolution of tumors in specific subjects, the issue arises as to how these data ultimately inform the processes critical to predicting the behavior of cancer and the effects of treatments. The fidelity of such computer predictions depends on the models used, the knowledge of the key model parameters, and the quality of the data. This article presents a unified approach for statistical calibration and validation of models and prediction of quantities of interest based on Bayesian inference. Importantly, the approach can take into account uncertainties in parameters, observations, and the model itself and lead to predictions with quantifiable uncertainty. While the validation process exercised here uses models from mixture theory, the process itself is quite general and is applicable to virtually any modeling scenario.



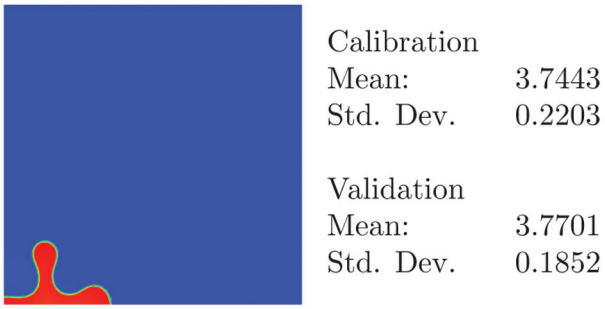




**Figure 3.** *Top: Cumulative distribution functions for the tumor volume at t = 9, as determined with both the calibration (solid line) and the validation (dashed line) posterior pdf's. Bottom: Prediction with the most likely estimate.*

## Acknowledgments

The author would like to thank J. Tinsley Oden, Serge Prudhomme, and Kristoffer van der Zee for many helpful conversations and guidance on the presented work.

## References

[1] J.T. Oden, A. Hawkins, and S. Prudhomme, *General diffuse-interface theories and an approach to predictive tumor growth modeling,* Math. Mod. and Meth. in Appl. Sci, 20:3 (2010), 447–517.

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

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

[4] R.M. Bowen, *Theory of mixtures*, in *Continuum Physics*, Volume 3, A.D. Eringen, ed., Academic Press, New York, 1976.

*Andrea Hawkins-Daarud completed a PhD in computational sciences, engineering, and mathematics at the University of Texas at Austin in May 2011.*