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

Andrea Hawkins-Daarud* **Serge Prudhomme*** **Kristoffer G. van der Zee †** **J. Tinsley Oden***

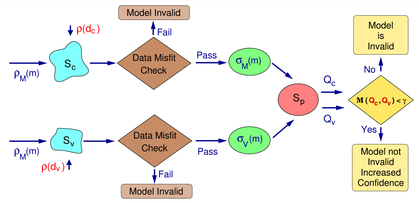
* Institute for Computational Engineering and Sciences (ICES), The University of Texas at Austin
† Department of Mechanical Engineering, Eindhoven University of Technology

■ **Goal:** Lay out a framework based on Bayesian probability, for systematically addressing the questions of *Validation* (the process of investigating the capability of a mathematical model to reproduce particular physical events) and *Uncertainty Quantification* (developing measures of the degree of confidence in computer model predictions of quantities of interest) for Tumor Growth Models.

Predictive Modeling

The successful use of computational models to predict physical events depends upon several fundamental concepts and processes:

1. **The mathematical model itself**: the manifestation of a scientific theory cast in mathematical structures intended to provide a meaningful abstraction of reality.
2. **The particular quantities of interest (QoI's)** of the physical event of interest that are targets of the prediction must be clearly specified in advance.
3. **Experimental observations must be made for** two purposes: 1) **calibration**, to reduce uncertainty in the model parameters for the environment of interest and 2) **validation** to subjectively determine if the model is capable of predicting the quantities of interest with sufficient accuracy.



The prediction pyramid, and a flow chart describing a possible statistical validation process.

The Bayesian framework

Every step in both calibration and validation encounters uncertainties in the model parameters $\mathbf{m}$, the observational data $\mathbf{d}$, the choice of a theoretical model, and the design of the validation process itself.

■ One must characterize all of these uncertainties, trace their propagation through the solution processes, and determine the uncertainty in the target QoI's, such as tumor volume or tumor shape.

Calibration

As described in Kaipio and Somersalo, *Statistical and Computational Inverse Problems*, 2005

- Given:
  - a *forward model* $\mathbf{D} = G(\mathbf{M}, \mathbf{E})$ relating *model parameters* $\mathbf{M}$ and noise $\mathbf{E}$ with *observables* $\mathbf{D}$ (described by the likelihood function $\theta(\mathbf{d}|\mathbf{m})$)
  - actual *observations* $\mathbf{d}^{\text{obs}}$ and the distribution of $\mathbf{E}$, $\theta^{\text{noise}}(\mathbf{e})$ (the prior pdf of the noise, including observational and modeling error)
  - a *"prior"* estimate, of model parameters and their uncertainty (described by $\rho_{\text{prior}}(\mathbf{m})$, the prior pdf of model parameters)
- Seek a *statistical characterization* of model parameters consistent with observations, forward model, and prior model: the *posterior pdf of model parameters* is given by

$$\sigma(\mathbf{m}|\mathbf{d}) = \sigma_M(\mathbf{m}) = \frac{\rho_{\text{prior}}(\mathbf{m})\theta(\mathbf{d}_{\text{obs}}|\mathbf{m})}{\rho(\mathbf{d}_{\text{obs}})}$$

■ For additive noise independent of $\mathbf{M}$, G representing a forward model solve, the error model is $\mathbf{D} = G(\mathbf{M}) + \mathbf{E}$ and the posterior pdf is

$$\sigma_M(\mathbf{m}) = \frac{\rho_{\text{prior}}(\mathbf{m})\theta_{\text{noise}}(\mathbf{d}_{\text{obs}} - G(\mathbf{m}))}{\rho(\mathbf{d}_{\text{obs}})}$$

Validation

There are two steps to validation:

1. **Data Misfit Check:** Verify that the model is capable of reproducing the data used for calibration and validation.
2. **Validation of Prediction**: Verify the accuracy of the QoI.

Data Misfit Check

■ 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."

- Define ranges of $\mathbf{e}^{obs}$ that are "likely" based off of the likelihood function.
- Check that there exists allowable $\mathbf{m}$ such that $\mathbf{e}^{obs}(\mathbf{m})$ is in these ranges.

Validation of Prediction

■ Check that different data does not drastically alter prediction of QoI

- Use different observational data, d_V^{obs}, and define the validation posterior pdf $\sigma_V(\mathbf{m})$ is $\sigma_V(\mathbf{m}) := \sigma_V(\mathbf{m}|\mathbf{d}_V^{obs}) \propto \rho_M(\mathbf{m})\theta_{noise}(\mathbf{e}_V^{obs}|\mathbf{m})$.
- Calculate the calibration predicted QoI and the validation predicted QoI $q_P^C(\mathbf{m})$ and $q_P^V(\mathbf{m})$.
- Define meaningful metric between two pdf's $M(\cdot,\cdot)$.
- Specify a tolerance γ_{tol} and declare the model *not invalid* if $M(q_P^C, q_P^V) < \gamma_{tol}$.

Uncertainty Quantification

■ If validation criterion is passed, quantify the uncertainty in the QoI $q_P^C(\mathbf{m})$ by computing the mean, variance, etc. of its corresponding pdf.

Phase-Field Models of Tumor Growth

- We begin with the case of a two-phase isothermal mixture consisting of tumor u and non-tumor n (i.e. healthy tissue and extracellular fluid).
- Include a representative nutrient, c, say oxygen.

We use the following two models each utilizing the same boundary and initial conditions, but differing in that one has time dependent parameters:

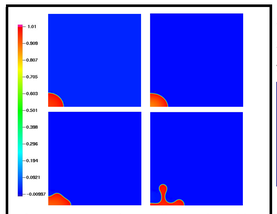
$\mathcal{M}_1$: Proliferation/Apoptosis Model

$$\begin{aligned} u_t &= \nabla\cdot\left(Mu^2\nabla\mu\right) + Pcu - Au \\ \mu &= f'(u) - \epsilon^2\Delta u - \epsilon\chi c \\ 0 &= \nabla\cdot(D\nabla c) - cu \\ f'(u) &= \gamma\left(4u^3 - 6u^2 + 2u\right) \end{aligned}$$

$\mathcal{M}_2$: Proliferation/Apoptosis Model with Time Dependent Parameters

$$\begin{aligned} u_t &= \nabla\cdot\left(Mu^2\nabla\mu\right) + Pcu - Au \\ \mu &= f'(u) - \epsilon^2\Delta u - \epsilon\chi(t)c \\ 0 &= \nabla\cdot(D(t)\nabla c) - cu \\ f'(u) &= \gamma\left(4u^3 - 6u^2 + 2u\right) \end{aligned}$$

Illustrative Example



A sequence of two-dimensional images of the progressive growth of a tumor at time $t = 0, 3, 6, 9$. Images generated with model $\mathcal{M}_2$. These (virtual) images are to be used for calibration (image at $t = 3$) and validation (image at $t = 6$) for $\mathcal{M}_1$.

Use $\mathcal{M}_2$ to generate (virtual) data against which the validity of $\mathcal{M}_1$ is assessed for the QoI of the final tumor volume.

Address the question: based on the data observed at times t_1 and t_2, is model $\mathcal{M}_1$ invalid for predicting the QoI $Q(u)$ =(tumor volume at $t = 9$)?

Setup

- Model parameters to be calibrated:

$$\mathbf{m} = (\chi, D).$$

- The prior pdf is a uniform distribution:

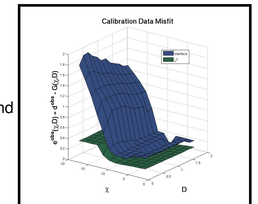
$$\rho(\chi, D) = U(3, 19.5) \times U(0.3, 1.95).$$

- The simulation mesh is an analogue of a typical MRI image. For data use:
  - L^2 norm of the observed pixels from the image
  - Position of the interface pixels.
- Assume the pdf $\theta_{noise}(\mathbf{e})$ is a bivariate, uncorrelated, half-normal distribution.

Calibration

Calibration Data Misfit Check

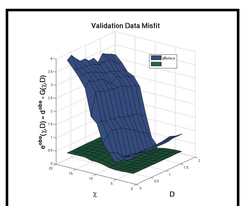
- Want $\mathbf{e}^{obs}(\mathbf{m}) \in [0, 0.25]$ for L^2 data and $\mathbf{e}^{obs}(\mathbf{m}) \in [0, 0.9]$ for interface data.
- Model passes the data misfit check.



Plot of absolute observed error for both data types. Interface error indicated in blue and L^2 indicated in green.

Calibration Posterior PDF

- Calculated calibration posterior pdf using the given prior and likelihood functions.



Plot of absolute observed error for both data types. Interface error indicated in blue and L^2 indicated in green.

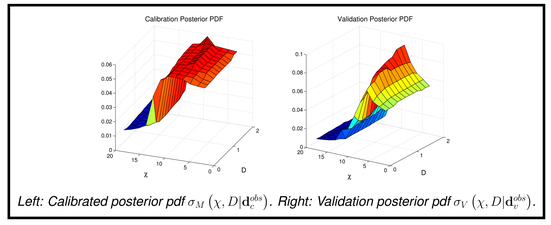
Validation

Validation Data Misfit Check

- Want $\mathbf{e}^{obs}(\mathbf{m}) \in [0, 0.25]$ for L^2 data and $\mathbf{e}^{obs}(\mathbf{m}) \in [0, 0.9]$ for interface data.
- Model passes the data misfit check.

Validation Posterior PDF

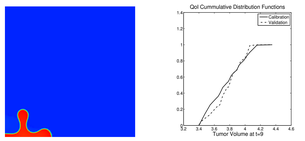
- Calculate the validation posterior pdf using the given prior and likelihood functions.



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

Validation of Prediction

- Calculate the QoI chosen to be the tumor volume.
  - Extend calculation for each pair of values of (χ, D) for which σ_M and σ_V were originally calculated to compute the tumor volume at $t = 9$.
  - Associate $q(u(\chi, D))$ with the value $\sigma_M(\chi, D|\mathbf{d}_c^{obs})$ (or σ_V).



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

Left: $\mathcal{M}_1$ tumor simulation at $t = 9$ with $(\chi, D) = (9, 1.8)$, the most likely a posterior value. Center: 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.

- Using a metric of the largest difference between the inverse of the cdfs and a tolerance $\gamma_{tol} = 10\% \times q(MLE)$, we find the model to be not invalid.

Uncertainty Quantification

It still remains to determine how to answer the question "What will the volume of the tumor be at $t = 9$?"

Many ways:

- Volume associated with the most likely estimator $\pm$ the standard deviation,
- Mean of the QoI pdf $\pm$ the standard deviation,
- With an interval, say the associated 90% confidence interval.

However answered, the proposed framework offers an avenue to answer the question with a level of uncertainty (confidence) associated with it.

■ **Conclusions:** 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. 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 models from mixture theory, the process itself is quite general, and is applicable to virtually any modeling scenario. We believe that complementing these validation processes in conjunction with a rich source of relevant data can lead to a predictive approach to tumor growth modeling.

Acknowledgements

The authors gratefully recognize the support by the Department of Energy under Award Number DE-FC52-08NA28615.