

Embedded Error Bayesian Calibration of Thermal Decomposition of Organic Materials

Ari Frankel, Ellen Wagman, Ryan Keedy, Brent Houchens, Sarah N. Scott

June 24, 2021

Abstract

Organic materials are an attractive choice for structural components due to their light weight and versatility. However, because they decompose at low temperatures relative to traditional materials they pose a safety risk due to fire and loss of structural integrity. To quantify this risk, analysts use chemical kinetics models to describe the material pyrolysis and oxidation using thermogravimetric analysis. This process requires the calibration of many model parameters to closely match experimental data. Previous efforts in this field have largely been limited to finding a single best-fit set of parameters even though the experimental data may be very noisy. Furthermore the chemical kinetics models are often simplified representations of the true decomposition process. The simplification induces model-form errors that the fitting process cannot capture. In this work we propose a methodology for calibrating decomposition models to thermogravimetric analysis data that accounts for uncertainty in the model-form and experimental data simultaneously. The methodology is applied to the decomposition of a carbon fiber epoxy composite with a three-stage reaction network and Arrhenius kinetics. The results show a good overlap between the model predictions and thermogravimetric analysis data. Uncertainty bounds capture deviations of the model from the data. The calibrated parameter distributions are also presented. The distributions may be used in forward propagation of uncertainty in models that leverage this material.

Nomenclature

β Temperature ramp rate

δ_{O_2}	Oxygen indicator variable
$\mu(x)$	Mean of x
μ_i	Mean of i -th physical parameter
ν	Mass stoichiometric coefficient
Φ	Cumulative standard normal distribution
σ_d^2	Noise in derivative observations
σ_i	Standard deviation of i -th physical parameter
σ_m^2	Noise in mass observations
θ	Model parameters
ξ	Random variable parameters
A	Arrhenius pre-factor
E	Activation energy
m	Mass fraction
m_{ij}	Mass of i -th time step in j -th experiment
n	Reaction order
T	Temperature
t	Time
T_o	Characteristic temperature
$V(x)$	Variance of x
w_j	Weight assigned to j -th experimental data

1 Introduction

Organic composites are increasing in popularity due to their low density, high strength, and versatility in applications [1–3]. For example, some commercial aircraft are now composed of up to 50% organic materials by mass [4, 5]. The use of these materials carries a safety risk due to the low temperature at which organic materials decompose relative to metallic structures. Besides the risk posed by the release of heat, smoke, fumes, and flammable gases, the decomposition of organic materials compromises the structural integrity of the full system utilizing them. It could also lead to spread of the fire should one occur, as organic composites are additional fuel that can accelerate the process.

An important part of understanding this risk is quantifying the temperature at which decomposition of a material initiates and the rate at which it does so. This task requires an understanding of the chemical kinetics and reaction pathways of the pyrolysis and oxidation processes. One common technique to explore the decomposition kinetics is thermogravimetric analysis (TGA) [6–9]. In TGA, a small sample of the material in question is placed in a heated crucible, and the temperature of the crucible is raised at a pre-determined schedule. As the temperature rises, the various decomposition mechanisms for the material are activated, and gaseous products escape from the crucible. The mass of the sample is recorded as a function of time and crucible temperature until all reactions have completed, demonstrating the characteristic temperatures and rates of each major reaction pathway.

TGA data is often used to calibrate global chemical kinetics models for use in other studies where the material may be part of a larger system exposed to high temperatures. Based on knowledge of the underlying chemical reactions or from observation of the TGA data, an analyst poses a reaction network, each reaction associated with a rate law governing the temperature dependence of the reaction. These rate laws are based on the Arrhenius equation, which carries a number of free parameters that must be determined for each reaction posed [6]. To date, the approaches considered for this problem broadly focus on optimizing the parameters to construct a single set of model parameters that most closely captures the observed experimental data [9, 10]. There are several avenues for attempting to perform this calibration. Analysts have considered changing the objective function of the optimization problem: one could minimize the mean-squared-error in the mass as a function of time [11], the derivative of the mass with respect

to time or temperature [7], linear combinations of the two, or other misfit functions besides least-squares. Other work has focused on exploring different optimization algorithms for this calibration process, comparing the performance in speed and final accuracy. Researchers have typically favored genetic algorithms for global optimization since they do not require gradient evaluations and can robustly find a local optimum [12, 13], though some quasi-Newton methods have been employed [14].

These methods are able to find a set of parameter values that can force the reaction network and rate laws to closely match the observed experiments. However a common issue in these approaches is that most do not consider the uncertainty in their models. The true chemical decomposition pathways for large molecules can consist of hundreds or thousands of individual reactions. Models calibrated to TGA data are often limited to less than 10 reactions that attempt to capture the major reactions only. The resulting models are biased. A simple Arrhenius-like rate law is unable to capture all of the trends in the data. No matter how much data is accumulated, no set of parameters will be able to exactly capture the observations because knowledge of the physical processes is limited and the models have missing components. One is left to either use simple models that are well-specified but inaccurate, or attempt to calibrate overly complex models and overfit the data [15]. Furthermore, with limited experimental data contaminated by measurement error or material variability, there is uncertainty in the parameter estimates that must be quantified, as many combinations or distributions of parameter values may be equally plausible given the available data [16]. This uncertainty must be rolled up into fire safety analyses performed with the calibrated models [17].

A common technique for addressing the quantification of uncertainty in model calibration is the use of Bayesian inference [18]. Bayesian inference makes use of Bayes’ rule to combine prior information of credible parameter values with a likelihood of observing the given data with an associated statistical model. The result is a posterior distribution of the parameter values given the data. The resulting posterior distribution may then be used in other problems for forward propagation of uncertainty to understand the range of potential results [19]. This framework is a cogent approach to addressing parametric uncertainty and has been used in a number of engineering disciplines [17, 20, 21]. Recent work has focused on extending Bayesian inference to address model-form uncertainty as well as parametric uncertainty. In (extended) model-form uncertainty, the model itself is

considered to be uncertain and has embedded error [22]. In this embedded-error formulation, the model parameters are considered to be uncertain distributions rather than uncertain variables, thus making the whole model a stochastic process. The parameters of the distributions are inferred in this framework. The resulting posterior distribution yields a set of values that provide coverage of the experimental data rather than converging to a single plausible best-fit. This work was recently applied in [23] to model embedded uncertainty in the Arrhenius parameter and activation energy of a simplified combustion kinetics mechanism.

In this work we demonstrate the application of an embedded-error formulation of uncertainty to the calibration of chemical kinetics parameters to TGA data of a multiple-component carbon fiber epoxy composite. Carbon fiber epoxy composites have been studied [5, 24–26] in TGA experiments with varying atmospheres, sample sizes, and other configuration variables. In section 2, we show the TGA data for our sample and configurations, and describe the simplified chemical reaction network and rate laws we calibrate to the data. Section 3 discusses Bayesian inference and the embedded-error formulation, as well as the numerical methods employed to perform this inference. Section 4 shows the results of the calibration process, including the comparison between the model results and experimental data as well as the parameter distributions.

2 Carbon Fiber Composite

The material we consider in this work is a carbon fiber epoxy composite similar to the materials considered in the decomposition studies of [5, 25–32]. Specifically, we use a composite consisting of AS4C carbon fiber at industrial grade HS-CP-3000 impregnated with UF3362-100 epoxy resin. Carbon fiber composites are of interest for their high durability and strength, but the binding resin (epoxy) pyrolyzes at temperatures around 300°C and thus reduces the structural integrity of the composite. Furthermore, the carbon fiber oxidizes at temperatures around 800°C in oxygen to give carbon monoxide and carbon dioxide.

2.1 TGA Results

To assess the thermal decomposition of the composite, TGA experiments were conducted at various conditions. Heating rates of $5^{\circ}\text{C}/\text{min}$ and $50^{\circ}\text{C}/\text{min}$

were applied in a nitrogen environment, and heating rates of $5^{\circ}\text{C}/\text{min}$ and $20^{\circ}\text{C}/\text{min}$ were applied in an ambient air environment. Two experimental replicates were performed for each environment and heating rate.

Figures 1 and 2 show the TGA data for the normalized mass and derivative with respect to temperature for the decomposition in nitrogen. The decomposition onset begins in both cases around 300°C , sharply increases through 400°C , and levels off slowly at 500°C . The two replicates performed under each set of experiment conditions overlap closely, although some differences are notable. The final residual mass values are somewhat different which might be attributable to differences in the sample preparation, or may even be a result of complex chemical processes in which varying residue fractions are produced depending on heating rate [33]. The work of [5, 31, 34] also showed a heat-rate dependent residual mass. One major drop in mass is observed in the mass curve, although the mass derivative curve shows two closely overlapping peaks.

Figures 3 and 4 show the TGA data in air. The nitrogen-only reactions appear to be unchanged, suggesting a purely pyrolytic process that occurs with or without oxygen, as well as two new major reactions. The other two main reactions that appear occur after the first, beginning a little before 600°C for the first and around 800°C for the second, depending on the heating rate. The composite is comprised of approximately 70% carbon fiber by mass as indicated by the manufacturer, which suggests that the second larger peak is the oxidation of carbon fiber into carbon monoxide and carbon dioxide (among other potential heavier byproducts). The smaller peak is assumed to be oxidation of remaining epoxy or epoxy residue. The data at the lower heating rate also shows a smaller reaction occurring just prior to the carbon fiber decomposition, although this is enveloped in other reactions at the higher heating rate and accounts for little mass loss. The carbon fiber decomposition was not fully completed in the higher heating rate experiments, so there is incomplete data for the onset of the carbon fiber oxidation at higher temperatures.

2.2 Decomposition Model

Given the limited test data available and the complexity of the underlying physics of the decomposition, proposing a fully encompassing decomposition model for the carbon fiber epoxy that can be accurately calibrated is pro-

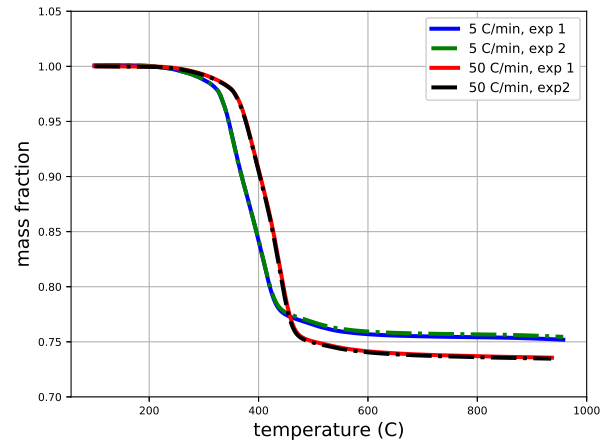


Figure 1: Mass fraction of the sample as a function of temperature for two replicates decomposing in nitrogen at two different heating rates.

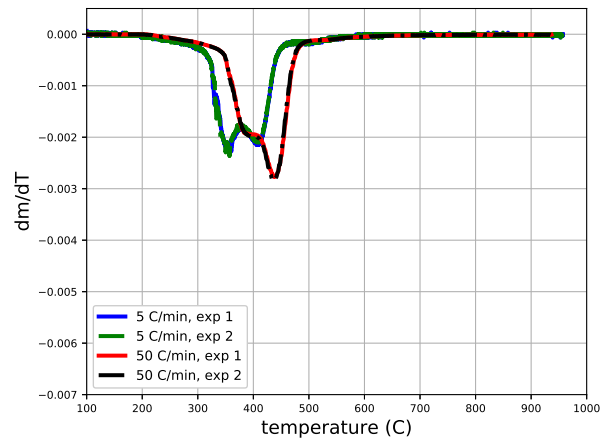


Figure 2: Rate of change of mass fraction with respect to temperature as a function of temperature for two replicates decomposing in nitrogen at two different heating rates.

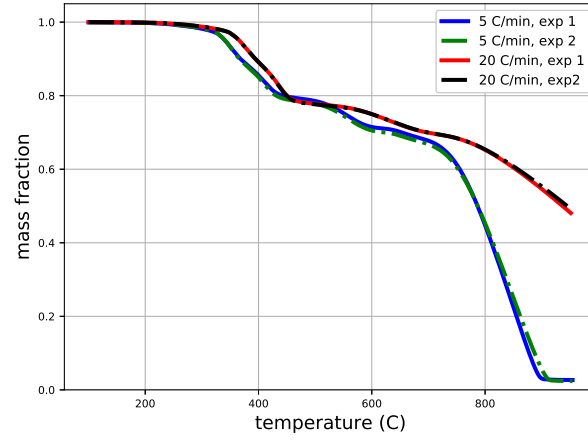


Figure 3: Mass fraction of the sample as a function of temperature for two experimental replicates of the composite decomposing in air at two different heating rates.

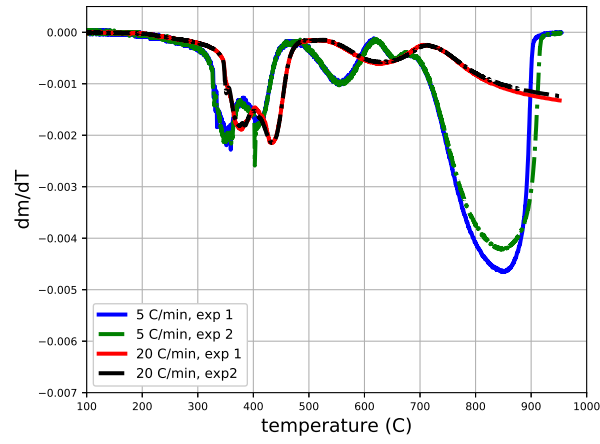
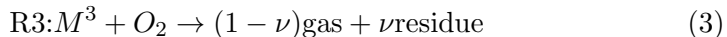


Figure 4: Rate of change of mass fraction with respect to temperature as a function of temperature for two replicates decomposing in air at two different heating rates.

hibitively difficult. The oxygen-limited decomposition of the carbon fiber is a mechanism we would be hard-pressed to calibrate appropriately given the test data only available in nitrogen-only and air-only environments. Although the epoxy component may be appropriately analyzed, the inability to characterize the full range of dependence on surface area of the porous surface also limits the model-form we can calibrate to the available data [35]. As seen in the work of [36] and [24], temperature is the main driver of the decomposition reaction between oxygen compositions of 10-21%, which covers our use case for the carbon fiber decomposition in air. To that end, we adopt a simple stylistic approach in which we calibrate individual reactions to simply capture the major trends in the TGA curves at the bounding atmospheres. We attempt to quantify the resulting model-form error as described later in this work. Though the scope of the modeling capability is limited to the laboratory settings considered, we believe the predictions with uncertainty can be informative in extrapolation to other settings.

Based on the major apparent drops in mass versus temperature and the lack of knowledge of the chemical composition of the epoxy, we propose a simplified decomposition pathway to capture the major peaks in mass loss rate. We assume that three independent moieties are reacting, two of which (M^1 and M^2) are associated with epoxy decomposition, and the third of which corresponds to the carbon fiber (M^3). Only M^1 decomposes in the nitrogen environment and may thus be associated with a pyrolysis reaction, while the other two only progress in the presence of oxygen and are thus oxidation reactions. The TGA data suggests the existence of a residual mass, which we attribute to the carbon fiber decomposition with a stoichiometric coefficient ν . No knowledge about the gas or residue composition is assumed. The reactions $R1$, $R2$, and $R3$ proposed are the following:



This model-form is deliberately simplistic in order to accommodate the lack of multiple experimental replicates and for ease of calibration. Under the assumption of an Arrhenius reaction rate law for each reaction, the associated kinetic equations for each moiety M^i with normalized mass m_i

are

$$\frac{dm_1}{dt} = -A_1 \exp\left(-\frac{E_1}{RT}\right) m_1^{n_1} \quad (4)$$

$$\frac{dm_2}{dt} = -A_2 \exp\left(-\frac{E_2}{RT}\right) m_2^{n_2} \delta_{O_2} \quad (5)$$

$$\frac{dm_3}{dt} = -A_3 \exp\left(-\frac{E_3}{RT}\right) m_3^{n_3} \delta_{O_2} \quad (6)$$

$$\frac{dm_r}{dt} = -\nu \frac{dm_3}{dt} \quad (7)$$

$$T(t) = T_0 + \beta t \quad (8)$$

where R is the ideal gas constant, A_i is the Arrhenius factor, E_i is the activation energy, T is temperature, n_i is the reaction order, δ_{O_2} is equal to 1 in the air environment and 0 in the nitrogen environment, and m_r is the mass of the residue. These differential equations are subject to the initial conditions $m_1(t=0) = m_{1,0}$, $m_2(t=0) = m_{2,0}$, $m_3(t=0) = 1 - m_{1,0} - m_{2,0}$, and $m_r(t=0) = 0$. The TGA experiments were initialized at temperature T_0 with heating rate β .

For calibration purposes, it is easier to reformulate the Arrhenius rate laws in a manner that rescales the Arrhenius pre-factor to a characteristic temperature. This is due to an ill-posedness issue in the calibration whereby the activation energy and Arrhenius pre-factor are highly correlated in the inference procedure [37]. This issue occurs when the reactions occur quickly, and the reaction rate can be captured by a linear combination of Arrhenius factors and activation energies. By reformulating the Arrhenius rate law in terms of a characteristic temperature, this correlation issue is partially alleviated. The conversion from an Arrhenius factor A_i to a characteristic temperature T_i with activation energy E_i is given by the relation

$$A_i = A_o \exp\left(\frac{E_i}{RT_i}\right) \quad (9)$$

where we define $A_o = 1s^{-1}$, and thus the reaction rate laws we calibrate are of the form

$$\frac{dm_1}{dt} = -\exp\left(-\frac{E_1}{R}\left(\frac{1}{T} - \frac{1}{T_1}\right)\right) m_1^{n_1} \quad (10)$$

$$\frac{dm_2}{dt} = -\exp\left(-\frac{E_2}{R}\left(\frac{1}{T} - \frac{1}{T_2}\right)\right) m_2^{n_2} \quad (11)$$

$$\frac{dm_3}{dt} = -\exp\left(-\frac{E_3}{R}\left(\frac{1}{T} - \frac{1}{T_3}\right)\right) m_3^{n_3} \quad (12)$$

and the residue mass rate law stays the same. In this form, an increase in the characteristic temperature tends to increase the temperature range at which the reaction progresses, while the activation energy sharpens the onset and peak of the mass-loss profile. Upon integration of these ODEs, the mass to compare against the TGA data is given by $m = m_1(t) + m_2(t) + m_3(t) + m_r(t)$, and the mass derivative with respect to temperature is also given by $\frac{dm}{dT} = \frac{1}{\beta} \frac{dm}{dt}$ which is also used for calibration data.

In total, there are 12 parameters in the reaction rate laws that must be calibrated to the observed TGA data: three activation energies E_i , three characteristic temperatures T_i , three reaction orders n_i , two initial mass fractions $m_{i,0}$, and a stoichiometric coefficient ν .

3 Bayesian Calibration

The standard approach to calibration is to identify an objective function between the model output and the experimental data to optimize, and then to apply an optimization algorithm to find the parameter values that best fit the data. However, this approach misses the presence of epistemic uncertainty; that is, there is a distribution of plausible parameter values that can fit the data, and reporting a single best set of parameters will miss this uncertainty. Furthermore, it is clear from the preceding section that our model will not be able to perfectly capture the behavior of the material no matter how many experiments we run. This is the result of model-form error, which we must also consider. In this section, we will discuss how Bayesian inference can be used to estimate model-form uncertainty, and then show we extend the inference procedure to incorporate model-form error by using a stochastic formulation of the decomposition model.

3.1 Bayesian Inference

The basis for Bayesian inference proceeds from the use of Bayes' rule of probabilities [38]. Let A and B be two events that may be related to each other. Bayes' rule states that the conditional probability of B given A , denoted by $p(B|A)$, can be related to other probabilities by

$$p(B|A) = \frac{p(A|B)p(B)}{p(A)} \quad (13)$$

Bayes' rule may be extended to parameter inference in the following way. Let θ be the vector of parameters to be inferred (the activation energies E_i , etc.), and D be the available data (the mass as a function of temperature in each experiment). We denote $p(D|\theta)$ as the likelihood function, which is the joint probability density of observing all of the experimental data given the parameter values. We also denote $p(\theta)$ as the prior probability of the parameters, which may reflect knowledge of the parameter values that analysts have about the model before having observed any experimental data. This prior probability may be the result of previous calibrations, physics-based restrictions on the model (for example that a given parameter must be positive), or subject matter expert elicitation of plausible ranges of the parameters. Then the posterior distribution for the parameters given the data $p(\theta|D)$ is given by

$$p(\theta|D) = \frac{p(D|\theta)p(\theta)}{p(D)} = \frac{p(D|\theta)p(\theta)}{\int p(D|\theta)p(\theta)d\theta} \quad (14)$$

A simple implementation may be to choose a Gaussian likelihood for the experiments with an experimental noise level σ_m^2 for the mass data and σ_d^2 for the mass derivative data. This is an approximation since the noise levels in the two datasets are expected to be related, but the noise variables are not directly of interest for the inference. We further assume N_{exp} experiments have been conducted and there are N_j data points per experiment, and approximate the observed mass and mass-derivative data as independent observations. Under this formulation, the likelihood function for calibrating against the mass and mass-derivative data jointly could be expressed as

$$p(D|\theta, \sigma_m^2, \sigma_d^2) = p(m|\theta, \sigma_m^2)p\left(\frac{dm}{dT}|\theta, \sigma_d^2\right) \quad (15)$$

$$p(m|\theta, \sigma_m^2) = \prod_{j=1}^{N_{exp}} \prod_{i=1}^{N_j} \frac{1}{\sqrt{2\pi\sigma_m^2}} \exp\left(-\frac{(m_{ij} - m_{ij,model}(\theta))^2}{2\sigma_m^2}\right) \quad (16)$$

$$p\left(\frac{dm}{dT}|\theta, \sigma_d^2\right) = \prod_{j=1}^{N_{exp}} \prod_{i=1}^{N_j} \frac{1}{\sqrt{2\pi\sigma_d^2}} \exp\left(-\frac{\left(\frac{dm}{dT}_{ij} - \frac{dm}{dT}_{ij,model}(\theta)\right)^2}{2\sigma_d^2}\right). \quad (17)$$

This model formulation assumes that the TGA data follows some underlying trend function with Gaussian distributed noise added to it. Evaluating this function amounts to the following steps:

- Choose a set of parameter values θ and noise levels σ_m^2 and σ_d^2
- Evaluate the chemical kinetics rate laws to compute the mass over time
- Plug in the pointwise comparison of experimental data to model results in the above function

The major task of Bayesian inference is the evaluation of the posterior distribution for a given dataset and model. For all but the simplest of cases, this formula can be computationally demanding to evaluate due to the integral $p(D) = \int p(D|\theta)p(\theta)d\theta$, termed the marginal likelihood or model evidence. This integral is usually not analytically tractable for computational models, and it may also be high-dimensional, making quadrature methods prohibitively expensive. The predominant method in these cases is to use Markov Chain Monte Carlo (MCMC) methods to evaluate samples from the posterior distribution directly. We will provide more details on this method in a later section.

3.2 Connection to Existing Methods

The Bayesian framework may be connected to the state-of-the-art in TGA calibration. In cases where Bayesian inference is too expensive or only a point-estimate is desired, one may resort to maximum a-posteriori (MAP) estimation. In these cases, one finds the parameter values that are at the mode of the posterior distribution. This amounts to solving the optimization problem

$$\theta^* = \arg \max_{\theta} [p(D|\theta)p(\theta)] \quad (18)$$

where the $p(D)$ term has disappeared since it does not depend on θ . Since the logarithm is a concave function, one can equivalently minimize the negative logarithm of the right-hand side of Equation 18, and using the same

Gaussian objective function as in the preceding section gives the equivalent to solving

$$\theta^* = \arg \min_{\theta} \left[\sum_{j=1}^{N_{exp}} \sum_{i=1}^{N_j} \frac{1}{\sigma_m^2} (m_{ij} - m_{ij,model}(\theta))^2 + \frac{1}{\sigma_d^2} \left(\frac{dm}{dT}_{ij} - \frac{dm}{dT}_{ij,model}(\theta) \right)^2 - \log p(\theta) \right], \quad (19)$$

which is a non-linear least-squares optimization problem for the model parameters θ with an additional penalty term $-\log p(\theta)$. Since analysts often apply uniform ranges over which to perform the optimization process, a convenient corresponding prior is a uniform distribution, and so $-\log p(\theta)$ is a constant that does not depend on θ ; thus the current approach to calibration may be seen as MAP estimation with a uniform prior on the parameters. Most research in the field has prioritized the efficient and optimal solution of this problem, as well as making arbitrary choices of noise parameters to balance between calibrating to mass versus mass derivatives, but we turn our attention instead to the larger problem of evaluating the posterior distribution.

3.3 Embedded Error Formulation

Given a large quantity of experimental data, the epistemic uncertainty associated with the parameters will shrink to zero and converge to the true “best-estimate” that minimizes the mean-square error between the model predictions and the experimental data. However, this does not guarantee that the model predictions will provide satisfactory coverage of the observations. The model is biased, and given enough data will only eventually find a solution that attempts to resolve all of the observations by compromising between all of the errors in the model predictions. Instead, we would like to find a way to embed the uncertainty of the model in the parameters as a result of these errors.

We adopt a formulation inspired by that of [22] to apply an embedded error formulation. The formulation above assumes that any misfit between the experiment and the model is the result of “noise” that can be expressed by the parameters σ_m^2 and σ_d^2 . Instead, we choose to assume that the model itself is stochastic, and that rather than treating the parameters as random variables we treat them as unknown distributions. For simplicity, we assume that the resulting likelihood function is approximately Gaussian with a variance equal to the sum of the stochastic model output variance and

experimental noise variance. That is, we now allow the parameters θ to be defined such that

$$\theta = \theta(\xi) \quad (20)$$

where ξ is a vector of hyperparameters that parameterize the distributions of θ . The work of [22, 23, 37] treated the parameters as the coefficients of polynomial chaos expansions, which is an effective strategy for small numbers of parameters and linear models. Due to the complexity of our model and the number of parameters we must infer, we choose to model each parameter as an independent Gaussian random variable. Thus ξ is a vector of the means and variances of each random variable along with the variance of the experimental noise. Under this formulation, the likelihood function takes the form of

$$p(D|\xi, \sigma_m^2, \sigma_d^2) = p(m|\xi, \sigma_m^2) p\left(\frac{dm}{dT}|\xi, \sigma_d^2\right) \quad (21)$$

$$p(m|\xi, \sigma_m^2) = \prod_{j=1}^{N_{exp}} w_j \prod_{i=1}^{N_j} \frac{1}{\sqrt{2\pi(V(m_{ij,model}(\theta(\xi))) + \sigma_m^2)}} \exp\left(-\frac{(m_{ij} - \mu(m_{ij,model}(\theta(\xi))))^2}{2(V(m_{ij,model}(\theta(\xi))) + \sigma_m^2)}\right) \quad (22)$$

$$p\left(\frac{dm}{dT}|\xi, \sigma_d^2\right) = \prod_{j=1}^{N_{exp}} w_j \prod_{i=1}^{N_j} \frac{1}{\sqrt{2\pi(V(\frac{dm}{dT}_{ij,model}(\theta(\xi))) + \sigma_d^2)}} \exp\left(-\frac{(\frac{dm}{dT}_{ij} - \mu(\frac{dm}{dT}_{ij,model}(\theta(\xi))))^2}{2(V(\frac{dm}{dT}_{ij,model}(\theta(\xi))) + \sigma_d^2)}\right) \quad (23)$$

where we have defined $\mu(m_{ij,model})$ and $V(m_{ij,model})$ as the mean and variance of the mass in the j -th experiment at the i -th time point due to the embedded uncertainty, $\mu(\frac{dm}{dT}_{ij,model})$ and $V(\frac{dm}{dT}_{ij,model})$ as the mean and variance of the mass derivative, and w_j is a weighting factor for the j -th experiment to guarantee each experiment has the same effective amount of data. This was done because some of the experiments had an order of magnitude more data than others due to the different heating rates, and we reweight the data to make the weight of each individual experiment equal. We define the weight of the j -th experiment to be

$$w_j = \frac{\frac{1}{N_j}}{\frac{1}{N_{exp}} \sum_{j=1}^{N_{exp}} \frac{1}{N_j}} \quad (24)$$

These weights have an average of 1 and inversely weight each experimental dataset by the number of points in the set. Thus we see that there is an additional contribution to the estimated variability in the model from

uncertainty inherent in the model parameters besides the additive experimental noise. This variability does not correspond to a physically observable randomness in the parameters; rather, it is a reflection of what modifications would be necessary to include in the model parameters to adequately capture the discrepancies between the model and experiments.

Evaluation of the above likelihood function is more complicated now due to the requirement of evaluating statistics of the mass versus time:

- Choose values of the mean and variance for each of the model parameters
- Perform forward propagation of uncertainty in the parameters through the model to derive the distribution of mass versus time in each experimental setting.
- Evaluate the mean and variance of the mass and mass derivative across all of the samples.
- Plug the results into the logarithm of equation 21 (for improved numerical stability) to evaluate the log-likelihood

3.4 Numerical Methods

The second step of the likelihood evaluation requires propagation of uncertainty of the parameter distributions through the computational model. This must be done numerically as the mapping from input to output is non-trivial. Although in this specific case some numerical recourse is available to approximate the output distributions, we choose to pursue a more general method that can be applied in more complex reaction scenarios. With 12 parameters to compute means and variances over, tensor-based quadrature methods are not available to us, and the model behavior with respect to the parameters is sufficiently non-linear that sparse quadrature methods may also be inaccurate. Thus we approximate the likelihood function with sampling. Specifically, we make use of the quasi-Monte Carlo method with the Sobol low discrepancy sequence [39] to generate a space-filling set of samples in the 12-dimensional uniform hypercube $[0, 1]^{12}$.

Let u be such a 12-dimensional sample with coordinates u_i . The corresponding sample in the physical parameter space may be formed from $\theta_i = \mu_i + \sigma_i \Phi^{-1}(u_i)$ where Φ^{-1} is the inverse normal cumulative distribution function, and μ_i and σ_i are the corresponding mean and standard deviation

for that parameter. This sampling scheme effectively reduces the sampling variance of the mean and variance estimators for the mass by spreading the samples far apart from each other and providing adequate coverage of the distribution tails. A large burn-in for the Sobol sequence was selected to avoid samples located on the borders of the hypercube, which would otherwise map to infinite values. 2000 such samples were drawn for each likelihood evaluation. Each sample was applied to the model evaluated at each of the 4 experimental configurations (2 heating rates and air vs. nitrogen), making for 8000 total system solutions per likelihood evaluation. Let $m_{ij,model}(\theta_k(\xi))$ be the result of the mass as a function of temperature for the physical parameters θ_k in experiment i at temperature T_j . The mean and variance of the mass are then approximated using the empirical averages over the sampled parameters θ_k given by

$$\mu(m_{ij,model}(\theta(\xi))) \approx \frac{1}{N} \sum_{k=1}^N m_{ij,model}(\theta_k(\xi)) \quad (25)$$

$$V(m_{ij,model}(\theta(\xi))) \approx \frac{1}{N} \sum_{k=1}^N (m_{ij,model}(\theta_k(\xi))^2 - \mu(m_{ij,model}(\theta(\xi)))^2) \quad (26)$$

with similar formulas for the mass derivative.

The inference procedure for the means and variances used was a Metropolis-Hastings MCMC sampling scheme [38]. The base procedure for this is as follows:

- Begin at a current set of parameters ξ_c , and evaluate $\log p(D|\xi_c)$ and $\log p(\xi_c)$
- Sample a proposal set of parameters ξ_p from a proposal distribution $p(\xi_p|\xi_c)$
- Evaluate $\log p(D|\xi_p)$ and $\log p(\xi_p)$
- Draw a uniform random number u from the range $[0, 1]$
- If $u < \exp(\log p(D|\xi_p) + \log p(\xi_p) - \log p(D|\xi_c) - \log p(\xi_c))$, accept the sample ξ_p and set that to the current sample. Otherwise, reject the proposal and maintain the current point.

For numerical stability reasons, the hyperparameters ξ were the means and log-variances of the parameters. Using the log-variance allows the parameters to get arbitrarily small without sampling a value less than zero,

whereas sampling the variances could lead to negative values. The proposal distribution $p(\xi_p|\xi_c)$ was set to be a multivariate normal distribution with an adaptive covariance matrix. The prior distribution $p(\xi)$ was assumed to be independent for each parameter. Except for the reaction orders, each parameter was given a flat prior of $p(\theta) = \text{constant}$ to minimize the informativity/subjectivity of the inference. The reaction order was selected to have an informative prior of log-normal distribution with location and scale parameters of 0.01, making a sharply peaked distribution centered at 1. This was selected in order to improve the smoothness of the fits and help guarantee physical responses, which a flat prior cannot guarantee. A flat prior may allow the reaction order to become very close to zero and lead to sharp inflection points in the mass loss curve and numerical integration difficulties in the decomposition model.

The result of the MCMC scheme is a series of samples from the posterior $p(\xi|D)$ for the means and variances of each individual parameter. What we are interested in are the aggregated distributions for the parameters, such as $p(E_1|D)$ for the activation energy of reaction 1, which must be computed by marginalizing over the chain results. This is done by evaluating the integral

$$p(E_1|D) = \int p(E_1|\xi_{E_1})p(\xi_{E_1}|D)d\xi_{E_1} \quad (27)$$

We can approximate this integral using the MCMC samples by combining each of the individual normal distributions. For each individual sample number i from the MCMC results, multiple samples were drawn from a normal distribution of mean μ_i and variance σ_i^2 . All of these samples were aggregated together to determine the final distribution of the parameters. To determine whether the MCMC results were converged, the variation in the mean and variance of the final distributions was evaluated for differing numbers of MCMC samples. The chain was considered converged when the variation in the means and variances derived from the first half of the samples and the second half of the samples was less than 0.1%, which amounted to over 200,000 samples.

4 Results

This section will show how the calibrated model with uncertainty captures the experimental data, and where model-form uncertainty helps to explain discrepancies in the data. The uncertain distribution parameters will also

be used to form an overall set of distributions for the chemical kinetics parameters.

4.1 Comparison to TGA Data

The uncertain parameter distributions were propagated through the model, drawing one random set of parameters for every other sample of means and variances in the MCMC results and evaluating the decomposition model. The statistics were aggregated and the mean, 1-percentile, and 99-percentile results were computed for the mass and mass derivative at each point in time. Figures 5, 6, 7, and 8 below show the comparison between the model predictions with uncertainty bounds and the TGA data.

In Figure 5 for the mass fraction in nitrogen, a benefit of the embedded error representation is immediately obvious. The mass fraction remaining after heating is complete is significantly different between the two heating rates, and the stochastic response of the embedded error formulation is able to represent this discrepancy. The mean model prediction is approximately halfway between the two, and the uncertainty bounds overlap with the final mass values. The differences in the inflection points of the two curves at 300 and 500°C are also captured with the uncertainty. Figure 6 shows a somewhat lower quality fit of the slope of the curve at the peaks of the decomposition around 400°C, but for the most part overlaps with the main drop in mass fraction. Although fitting additional reactions might result in a better fit, the additional parameter calibration required would increase the prediction uncertainty.

Figure 7 shows the model performance in air, where the other two reactions are activated as well. Again the model mean matches the data closely and the uncertainty bounds overlap with the mass data almost completely, despite having little data for the carbon fiber decomposition at the higher heating rate. However, Figure 8 shows that the fit for the derivative of the mass fraction is less close at the peaks around 600°C and 900°C, and the uncertainty bounds are insufficient for overlap. This is likely due to the mass fraction data comprising the majority of the likelihood function, having overall lower variance and a higher magnitude in the value of the data whereas the mass derivative is overall smaller and has a relatively higher variance. This would cause the calibration to favor fitting the mass fraction more accurately over the mass derivative. In addition, there is less data

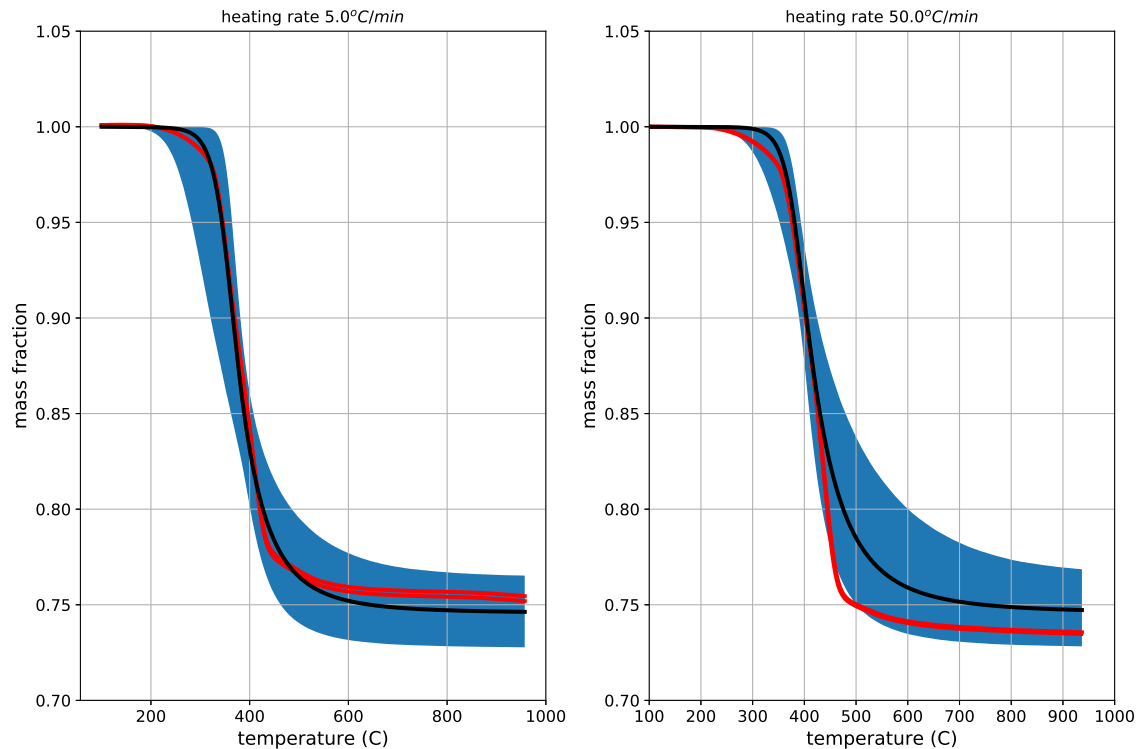


Figure 5: The average model prediction (black), the 1-percentile and 99-percentile mass (blue envelope), and experimental data (red) for the mass fraction of sample decomposing in nitrogen.

available for fitting reactions 2 and 3 as compared to reaction 1, so a poorer relative fit is to be expected. The reaction order also favors nearly straight lines in the mass curve with small reaction order (i.e. the mass derivative is nearly constant), which is why we invoked the informative prior for the reaction order to improve the smoothness of the curve.

4.2 Calibrated Distributions

To evaluate the parameter posterior distributions, 100 samples of each parameter were drawn from each set of means and variances and accumulated together. Table 1 shows the means and standard deviations of each parameter. Figure 9 shows that the distributions look very nearly Gaussian,

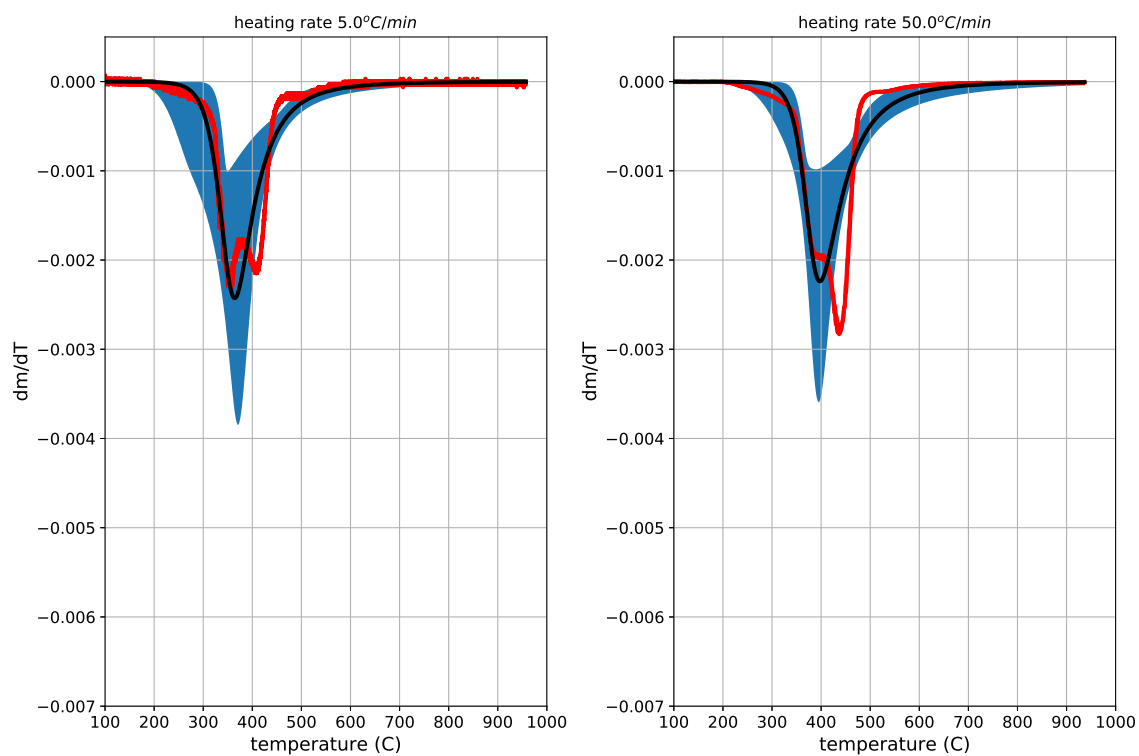


Figure 6: The average model prediction (black), the 1-percentile and 99-percentile mass (blue envelope), and experimental data (red) for the rate of mass fraction change with respect to temperature of sample decomposing in nitrogen.

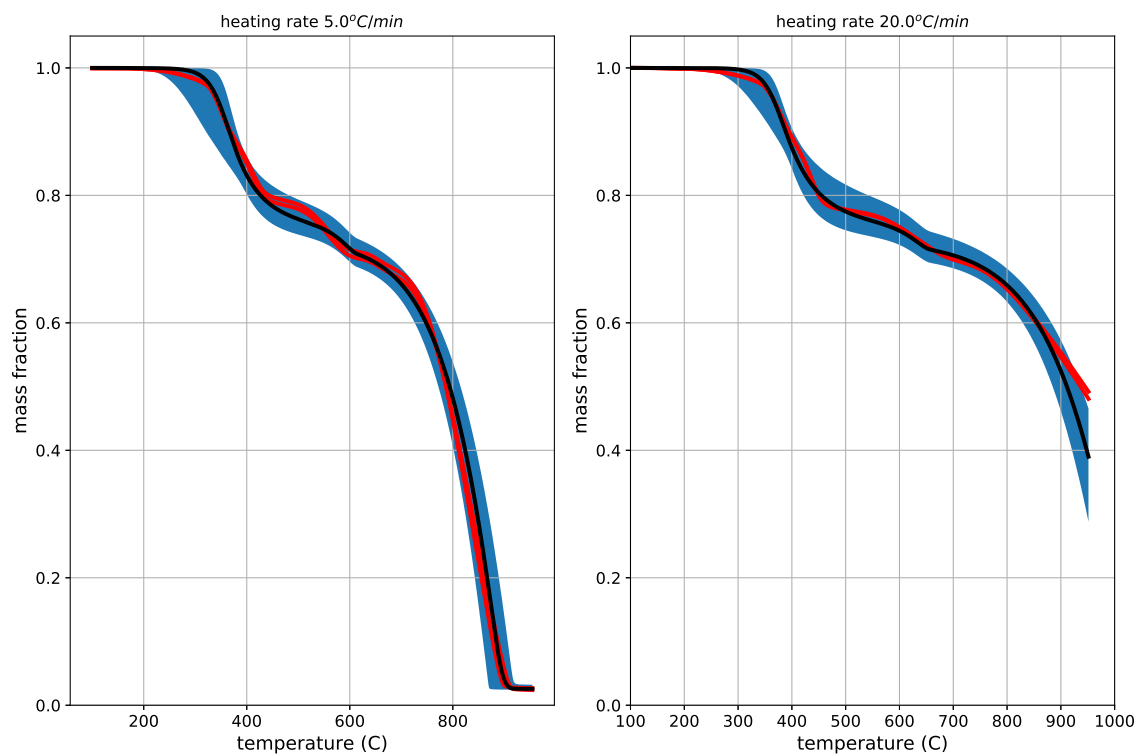


Figure 7: The average model prediction (black), the 1-percentile and 99-percentile mass (blue envelope), and experimental data (red) for the mass fraction of sample decomposing in air.

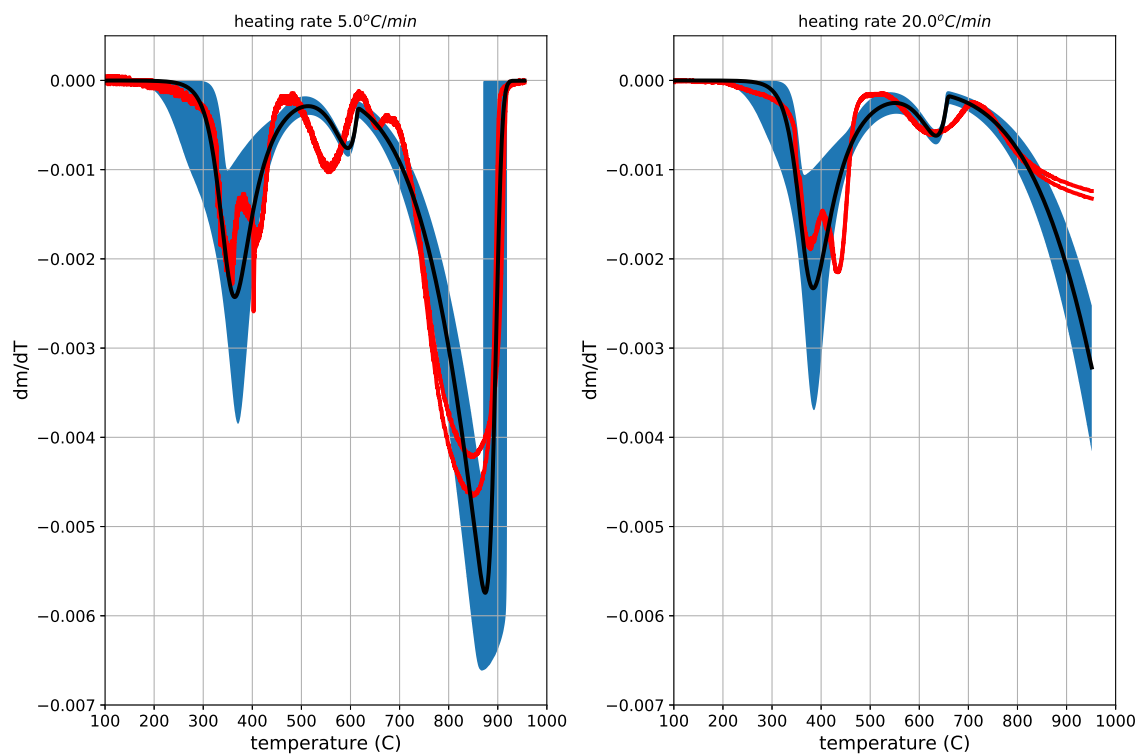


Figure 8: The average model prediction (black), the 1-percentile and 99-percentile mass (blue envelope), and experimental data (red) for the rate of mass fraction change with respect to temperature of sample decomposing in air.

Parameter	Mean (μ_i)	Standard Deviation (σ_i)	Coefficient of Variation (σ_i/μ_i)
T_{o1} (K)	651.9	0.3967	6.085×10^{-4}
$E_1/R \times 10^{-3}$ (K)	24.80	6.290	0.2536
n_1	4.153	0.1730	0.04166
$m_{1,0}$	0.2545	7.903×10^{-3}	0.03105
T_{o2} (K)	1171	16.14	0.01378
$E_2/R \times 10^{-3}$ (K)	25.02	0.5716	0.02285
n_2	0.5265	0.03908	0.07423
$m_{2,0}$	0.2956	2.801×10^{-4}	9.476×10^{-4}
T_{o3} (K)	3456	110.1	0.03186
$E_3/R \times 10^{-3}$ (K)	12.80	0.04769	3.726×10^{-3}
n_3	0.06224	3.019×10^{-3}	0.04851
ν	0.03480	1.108×10^{-4}	3.184×10^{-3}

Table 1: Reported statistics of the parameter posterior distributions: the mean, standard deviation, and coefficient of variation (standard deviation divided by mean).

such that the mean and standard deviation are sufficient to accurately represent the posterior of each parameter. The correlation coefficients of the propagated samples were also computed. Reaction 2 had parameters with substantial correlations reported below in Table 2, and reactions 1 and 3 had negligible correlations. Overall we find that the nominal values for the activation energies and Arrhenius factors are comparable to others in the literature. For example, in [40], the authors use a variety of techniques to analyze the effective activation energy and Arrhenius coefficients, and the derived values correspond closely with the mean parameter values for $R1$ in our calibration. Where our epoxy decomposition has an activation energy of approximately $206 kJ/mol$ and Arrhenius factor of $3 \times 10^{16} s^{-1}$, the paper reports a wide range of values from $180 - 280 kJ/mol$ and $10^{14} - 10^{24} s^{-1}$. One important difference in our approach versus comparable approaches in the literature is that while the kinetics are reported to take into account dependence on non-parametric formulations, ours is designed to build the uncertainty associated with using a parametric formulation into the predictions for a simpler approach.

The values of the parameters bear some consideration. The reaction orders for the first reaction and third reaction are markedly high and low respectively, at 4.1 and 0.06. In the first reaction, this would be a reflection of a particularly peaked mass loss profile with a sharp corner when the

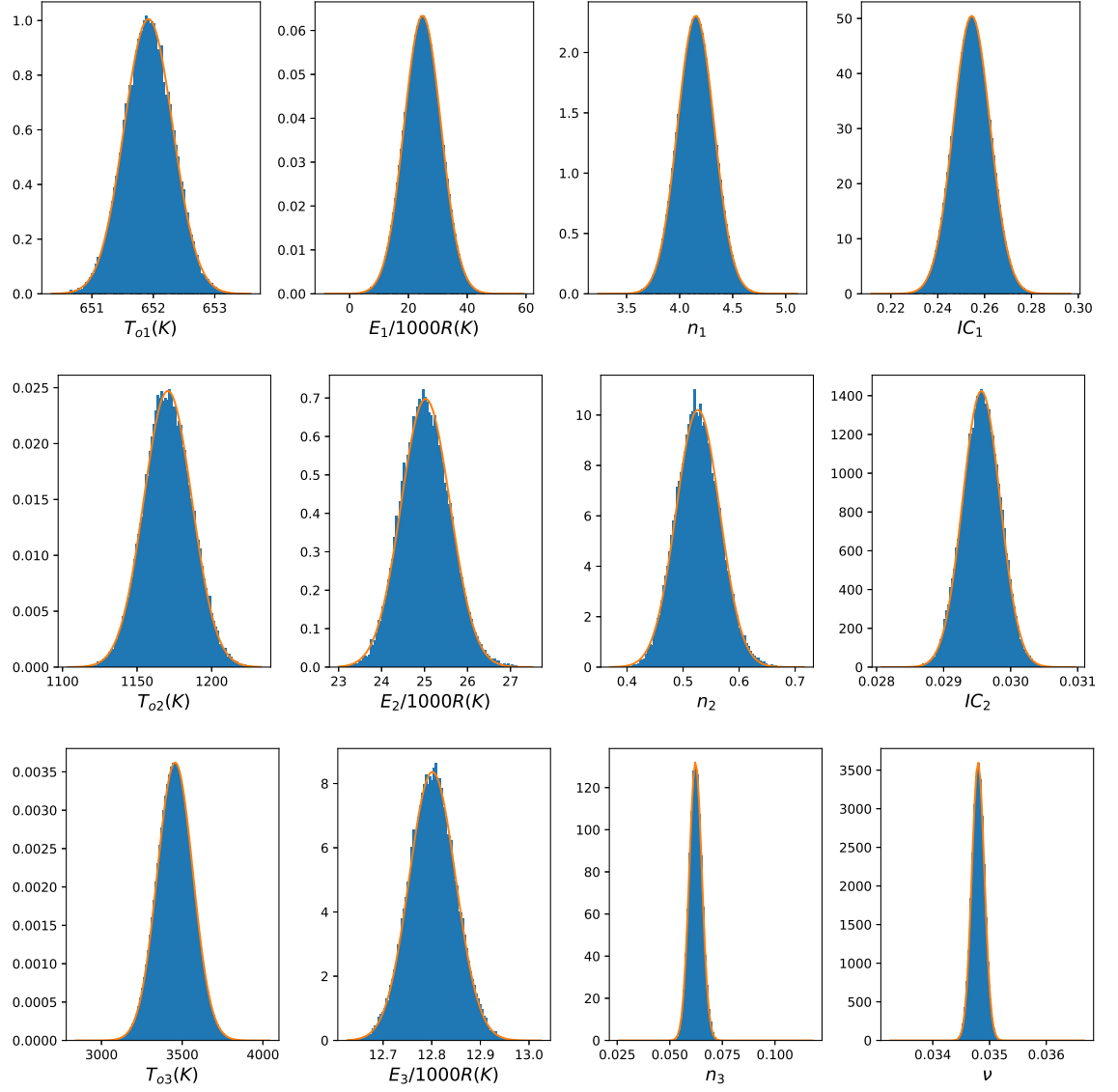


Figure 9: Histograms and Gaussian fits to the calibrated parameter distributions for each reaction.

Reaction 2	T_{o2}	E_2/R	n_2	$m_{2,0}$
T_{o2}	1	-	-	-
E_2/R	-0.8832	1	-	-
n_2	-0.8117	0.4550	1	-
$m_{2,0}$	0.1822	-0.2601	-0.01341	1

Table 2: Pearson correlation coefficients of the parameters for reaction 2. The matrix is symmetric, so dashed entries correspond to their transposed values. Values close to ± 1 indicate strong correlation, while values close to 0 indicate no correlation, and the sign indicates direct or inverse correlation.

reaction slows, for which there is data in both the nitrogen and air environments to support this. The very low reaction order for the carbon fiber decomposition reflects how the mass loss appears to be sharp and nearly linear throughout the reaction. Because the reaction did not proceed to completion in the high heating rate case, there is insufficient data to explore the curvature of the mass derivative profile and the calibration does not support a more physical reaction order. The use of an informative prior for the reaction order ameliorates this issue but does not altogether remedy it.

The characteristic temperatures correspond roughly to the relative temperatures at which the reactions initiated. This correspondence is not one-to-one due to the varying activation energies and reaction orders, but does align with the prescribed reactions. Reaction 1 corresponded to the initial epoxy pyrolysis in nitrogen which happens at a lower temperature ($350^\circ C$, close to the characteristic temperature), and also had the lowest characteristic temperature. Reactions 2 and 3 each occurred at increasing temperatures and correspondingly have higher characteristic temperatures ($900^\circ C$ and $3200^\circ C$ respectively).

The embedded uncertainties also indicate where variation was needed to capture the model-form error. The uncertainty in E_1 is very high as indicated by the coefficient-of-variation of 25%, which may be explained by the complex reaction profile to which we attempt to calibrate only a single reaction of a simple Arrhenius form. In all other cases, the reaction order coefficient-of-variation is larger than the other parameters, indicating that the mass-loss profiles are potentially the most sensitive to the reaction order and that much of the model-form error may be accommodated by varying the reaction order.

For evaluating the calibrated models, one can use the above distributions in forward propagation of uncertainty. One important point is that the distributions have been assumed to be Gaussian. For most of the parameters this is not an issue, but there is a finite probability of the parameter value taking on a negative value in uncertainty propagation. In these cases it is best to reject any sample value drawn with an unphysical value.

5 Conclusion

In this work we have outlined a technique for calibrating chemical kinetics models under uncertainty to account for model-form error. We demonstrated its application to a simplified carbon fiber epoxy composite decomposition. The results suggest a good overlap of the statistics of the model with the observed experimental data despite model-form errors. The calibration yielded uncertainty distributions for the model parameters that may be used in future simulations involving this material under comparable reaction conditions. Such studies would involve forward propagation of the calibrated uncertainties and would give a probabilistic answer for reliability of the material in fire safety engineering calculations.

Although this method is useful for calibration under uncertainty, there are improvements that can be made to the procedure and formulation. The likelihood formulation appeared to bias the mass over the mass derivative in fitting due to the scale and relative variance in the values, an observation which is consistent with other research comparing different objective functions in deterministic TGA fitting. It is possible that other likelihood formulations would yield more balanced results. The use of an informative prior for the reaction order was also an imperfect solution to regularize the calibration. Although the resulting fits are good, there remains the issue of performing this calibration without invoking ad-hoc restrictions on the parameters. And finally, the method gives uncertainty estimates at the cost of being more computationally intensive than standard calibration approaches. It requires the evaluation of many samples of the decomposition model for one single likelihood evaluation, while traditional methods would only require one evaluation. It is possible that surrogate models of reduced-order statistics of the model behavior (e.g. mass derivative peak height) would enable faster evaluations, or that approximate inference techniques would

accelerate the process. A detailed investigation of comparing the sampling approach we leveraged against quadrature-based methods would also determine whether more efficient integration techniques are available in different circumstances. All of these issues merit further investigation in future work, along with validation of the calibrated mechanisms in new laboratory settings.

Acknowledgements

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy’s National Nuclear Security Administration under contract DE-NA0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government. The authors would like to acknowledge Corinne Hagan and Brian McKay for sample preparation, Adriana Pavia Sanders and Karla Reyes for the TGA experiments, and Cosmin Safta and Andrew Kurzawski for their comments.

References

- [1] D. Rajak, D. Pagar, R. Kumar, and C. Pruncu. Recent progress of reinforcement materials: a comprehensive overview of composite materials. *Journal of Materials Research and Technology*, 8:6354–6374, 2019.
- [2] A. Mohammed, A. Manalo, W. Ferdous, Y. Zhuge, P. Vijay, and J. Pettigrew. Experimental and numerical evaluations on the behaviour of structures repaired using prefabricated frp composites jacket. *Engineering Structures*, 210:110358, 2020.
- [3] M. Al-Rubaye, A. Manalo, O. Alajarmeh, W. Ferdous, W. Lokuge, B. Benmokrane, and A. Edoo. Flexural behavior of concrete slabs reinforced with gfrp bars and hollow composite reinforcing systems. *Composite structures*, 236:111836, 2020.
- [4] V. Gurguitiu. *Structural Health Monitoring of Aerospace Composites*. Academic Press, 2016.

- [5] J. Quintiere, R. Walters, and S. Crowley. Flammability properties of aircraft carbon-fiber structural composite. Technical Report DOT/FAA/AR-07/57, Federal Aviation Administration, 2007.
- [6] J. Conesa, A. Marcilla, J. Caballero, and R. Font. Comments on the validity and utility of the different methods for kinetic analysis of thermogravimetric data. *Journal of Analytical and Applied Pyrolysis*, 58-59:617–633, 2001.
- [7] L. Abdelouahed, S. Leveneur, L. Vernieres-Hassimi, L. Balland, and B. Taouk. Comparative investigation for the determination of kinetic parameters for biomass pyrolysis by thermogravimetric analysis. *Journal of Thermal Analysis and Calorimetry*, 129:1201–1213, 2017.
- [8] K. Chetehouna, N. Grange, N. Gascoin, L. Lemee, I. Reynaud, and S. Senave. Release and flammability evaluation of pyrolysis gases from carbon-based composite materials undergoing fire conditions. *Journal of Analytical and Applied Pyrolysis*, 134:136–142, 2018.
- [9] F. Richter and G. Rein. Pyrolysis kinetics and multi-objective inverse modelling of cellulose at the microscale. *Fire Safety Journal*, 91:191–199, 2017.
- [10] F. Richter and G. Rein. Heterogeneous kinetics of timber charring at the microscale. *Journal of Analytical and Applied Pyrolysis*, 138:1–9, 2019.
- [11] D. Purnomo, F. Richter, M. Bonner, R. Vaidyanathan, and G. Rein. Role of optimisation method on kinetic inverse modelling of biomass pyrolysis at the microscale. *Fuel*, 262:116251, 2020.
- [12] L. Hasalova, J. Ira, and M. Jahoda. Practical observations on the use of shuffled complex evolution algorithm for kinetic parameters estimation in pyrolysis modeling. *Fire Safety Journal*, 80:71–82, 2016.
- [13] K. Li, X. Huang, C. Fleischmann, G. Rein, and J. Ji. Pyrolysis of medium-density fiberboard: optimized search for kinetics scheme and parameters via a genetic algorithm driven by kissinger’s method. *Energy Fuels*, 28:6130–6139, 2014.
- [14] E. Wagman, S. Scott, A. Frankel, R. Keedy, V. Brunini, M. Kury, and B. Houchens. Robust and efficient optimization of kinetic parameters in organic material thermal decomposition models. *Fire Safety Journal*, 2021.

- [15] N. Bal and G. Rein. Relevant model complexity for non-charring polymer pyrolysis. *Fire Safety Journal*, 61:36–44, 2013.
- [16] E. Kim and N. Dembsey. Parameter estimation for comprehensive pyrolysis modeling: guidance and critical observations. *Fire Technology*, 51:443–477, 2015.
- [17] H. Wang and D. Sheen. Combustion kinetic model uncertainty quantification, propagation, and minimization. *Progress in Energy and Combustion Science*, 47:1–31, 2015.
- [18] M. Kennedy and A. O’Hagan. Bayesian calibration of computer models. *J. R. Statist. Soc. B*, 63:425–464, 2001.
- [19] M. Bruns. Inferring and propagating kinetic parameter uncertainty for condensed phase burning models. *Fire Technology*, 52:93–120, 2016.
- [20] X. Huan, C. Safta, K. Sargsyan, G. Geraci, M. Eldred, Z. Vane, G. Lacaze, J. Oefelein, and H. Najm. Global sensitivity analysis and quantification of model error for large eddy simulation in scramjet design. *19th AIAA Non-Deterministic Approaches Conference*, 2017.
- [21] J. Sousa and C. Gorle. Computational urban flow predictions with bayesian inference: validation with field data. *Building and Environment*, 154:13–22, 2019.
- [22] K. Sargsyan, X. Huan, and H. Najm. Embedded model error representation for bayesian model calibration. *International Journal for Uncertainty Quantification*, 9:365–394, 2019.
- [23] R. Freitas, F. Rochinha, D. Mira, and X. Jiang. Parametric and model uncertainties induced by reduced order chemical mechanisms for biogas combustion. *Chemical Engineering Science*, 227:115949, 2020.
- [24] Jie Yang, Jie Liu, Wenbin Liu, Jun Wang, and Tao Tang. Recycling of carbon fibre reinforced epoxy resin composites under various oxygen concentrations in nitrogen–oxygen atmosphere. *Journal of Analytical and Applied Pyrolysis*, 112:253–261, 2015.
- [25] V. Biasi, G. Leplat, F. Feyel, and P. Beauchene. Heat and mass transfers within decomposing carbon fibers/epoxy resin composite materials. *11th AIAA/ASME Joint Thermophysics and Heat Transfer Conference*, 2014.

- [26] M. McKinnon, Y. Ding, S. Stoliarov, S. Crowley, and R. Lyon. Pyrolysis model for a carbon fiber/epoxy structural aerospace composite. *J. Fire Sci.*, 35:36–61, 2017.
- [27] C. Branca, C. Di Blasi, A. Galgano, and E. Milella. Thermal and kinetic characterization of a toughened epoxy resin reinforced with carbon fibers. *Thermochim. Acta*, 517:53–62, 2011.
- [28] T. Fateh, J. Zhang, M. Delichatsios, and T. Rogaume. Experimental investigation and numerical modeling of the fire performance for epoxy resin carbon fibre composites of variable thicknesses. *Fire Mater.*, 41:307–322, 2017.
- [29] N. Regnier and S. Fontaine. Determination of the thermal degradation kinetic parameters of carbon fibre reinforced epoxy using tg. *J. Therm. Anal. Calorim.*, 64:789–799, 2001.
- [30] P. Tranchard, F. Samyn, S. Duquesne, B. Estebe, and S. Bourbigot. Modelling behaviour of a carbon epoxy composite exposed to fire: part 1-characterisation of thermophysical properties. *Materials (Basel)*, 10, 2017.
- [31] P. Tranchard, S. Duquesne, F. Samyn, B. Estebe, and S. Bourbigot. Kinetic analysis of the thermal decomposition of a carbon fibre-reinforced epoxy resin laminate. *J. Anal. Appl. Pyrolysis*, 126:14–21, 2017.
- [32] D. Sikoutris, D. Vlachos, V. Kostopoulos, S. Jagger, and S. Ledin. Fire burnthrough response of cfrp aerostructures. numerical investigation and experimental verification. *Appl. Compos. Mater.*, 19:141–159, 2012.
- [33] T. Fletcher, A. Kerstein, R. J. Pugmire, M. Solum, and D. M. Grant. A chemical percolation model for devolatilization : Summary. 1998.
- [34] A. Dodd, B. Shelden, and K. Erickson. Numerical simulation of decomposition and combustion of an epoxy-carbon-fiber composite. *Interflam*, pages 407–412, 2013.
- [35] J. Su and D. Perlmutter. Effect of pore structure on char oxidation kinetics. *AIChE*, 31:973–981, 1985.
- [36] M.X. Fang, D.K. Shen, Y.X. Li, C.J. Yu, Z.Y. Luo, and K.F. Cen. Kinetic study on pyrolysis and combustion of wood under different oxygen concentrations by using tg-ftir analysis. *Journal of Analytical and Applied Pyrolysis*, 77(1):22–27, 2006.

- [37] K. Sargsyan, H. Najm, and R. Ghanem. On the statistical calibration of physical models. *Int. J.Chem. Kinet.*, 47:246–276, 2015.
- [38] E. T. Jaynes. *Probability theory: The logic of science*. Cambridge University Press, Cambridge, 2003.
- [39] I. Radovic, Sobol. I., and R. Tichy. Quasi-monte carlo methods for numerical integration: comparison of different low discrepancy sequences. *Monte Carlo Methods and Applications*, 2:1–14, 1996.
- [40] H. Li, N. Wang, H. Xuefei, H. Yuan, and J. Xie. Mechanism identification and kinetics analysis of thermal degradation for carbon fiber/epoxy resin. *Polymers*, 13, 2021.