Inference of H_2O_2 thermal decomposition rate parameters from experimental statistics

SAND2017-4535C

Tiernan Casey
Mohammad Khalil, Habib Najm

Combustion Research Facility, Sandia National Laboratories
Livermore, CA.

US National Combustion Meeting, April 24, 2017

Motivation

- Kinetic rate parameters (e.g. A, n, E_a) are uncertain, arising from the uncertainty in the underlying experimental data $\Rightarrow$ motivates need for uncertainty quantification (UQ)

Motivation

- Kinetic rate parameters (e.g. A, n, E_a) are uncertain, arising from the uncertainty in the underlying experimental data $\Rightarrow$ motivates need for uncertainty quantification (UQ)
- Experimental data are typically not reported, results expressed in terms of rate constants at various temperatures, or an Arrhenius expression

Motivation

- Kinetic rate parameters (e.g. A, n, E_a) are uncertain, arising from the uncertainty in the underlying experimental data $\Rightarrow$ motivates need for uncertainty quantification (UQ)
- Experimental data are typically not reported, results expressed in terms of rate constants at various temperatures, or an Arrhenius expression
- Experimental uncertainty often expressed as uncertainty factors on rate constants, or attached to the Arrhenius pre-exponential, or error bars on each parameter

Motivation

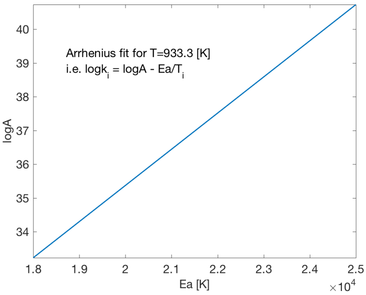
- Kinetic rate parameters (e.g. A, n, E_a) are uncertain, arising from the uncertainty in the underlying experimental data $\Rightarrow$ motivates need for uncertainty quantification (UQ)
- Experimental data are typically not reported, results expressed in terms of rate constants at various temperatures, or an Arrhenius expression
- Experimental uncertainty often expressed as uncertainty factors on rate constants, or attached to the Arrhenius pre-exponential, or error bars on each parameter
- UQ perspective: ideally infer the uncertainty in model parameters given said model and data to arrive at a (posterior) parameter PDF

Motivation

- Kinetic rate parameters (e.g. A, n, E_a) are uncertain, arising from the uncertainty in the underlying experimental data $\Rightarrow$ motivates need for uncertainty quantification (UQ)
- Experimental data are typically not reported, results expressed in terms of rate constants at various temperatures, or an Arrhenius expression
- Experimental uncertainty often expressed as uncertainty factors on rate constants, or attached to the Arrhenius pre-exponential, or error bars on each parameter
- UQ perspective: ideally infer the uncertainty in model parameters given said model and data to arrive at a (posterior) parameter PDF
- Is it still possible to conduct this inference in the absence of data?

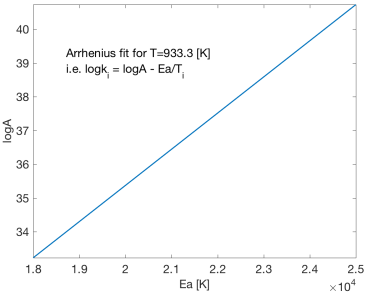
Rate parameters - correlation

- How uncertainty is reported is crucial. Arrhenius parameter uncertainty is often highly correlated and this is typically not reported



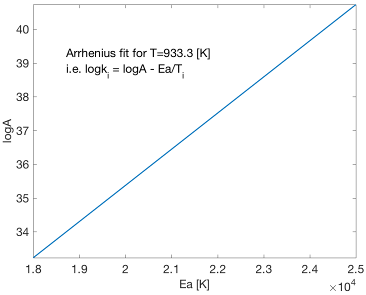
Rate parameters - correlation

- How uncertainty is reported is crucial. Arrhenius parameter uncertainty is often highly correlated and this is typically not reported
- Take measurements at some temperature, fit the data with a reaction rate



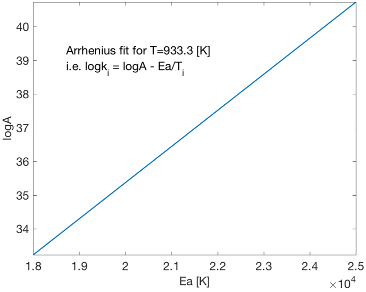
Rate parameters - correlation

- How uncertainty is reported is crucial. Arrhenius parameter uncertainty is often highly correlated and this is typically not reported
- Take measurements at some temperature, fit the data with a reaction rate
- Express the rate constant in Arrhenius form, defines a line in logA-E_a space for that temperature



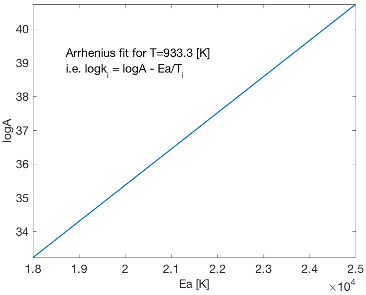
Rate parameters - correlation

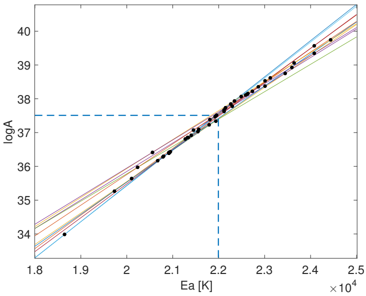
- How uncertainty is reported is crucial. Arrhenius parameter uncertainty is often highly correlated and this is typically not reported
- Take measurements at some temperature, fit the data with a reaction rate
- Express the rate constant in Arrhenius form, defines a line in logA-E_a space for that temperature
- Ideally, the consensus rate is the intersection of the lines for each temperature



Rate parameters - correlation

- How uncertainty is reported is crucial. Arrhenius parameter uncertainty is often highly correlated and this is typically not reported
- Take measurements at some temperature, fit the data with a reaction rate
- Express the rate constant in Arrhenius form, defines a line in logA-E_a space for that temperature
- Ideally, the consensus rate is the intersection of the lines for each temperature

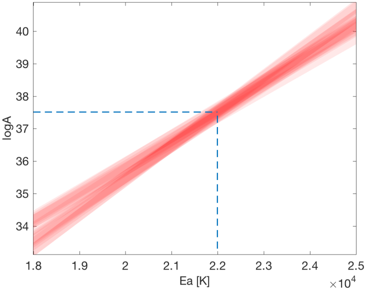




In reality no consensus due to model inadequacy and/or data error, require regression to find 'best-fit' Arrhenius parameters

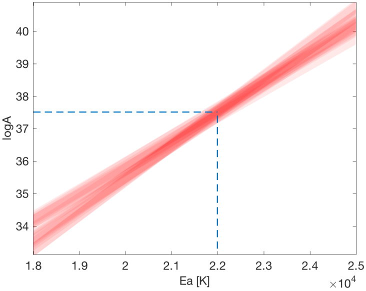
Rate parameters - correlation

- Same picture when including uncertainty



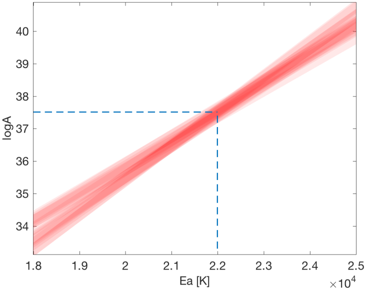
Rate parameters - correlation

- Same picture when including uncertainty
- The rate at each temperature defines an area in logA-E_a space



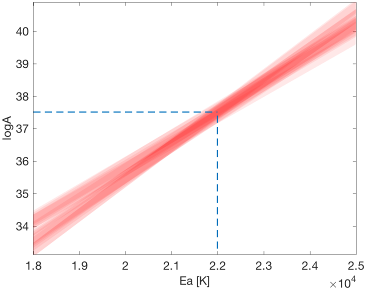
Rate parameters - correlation

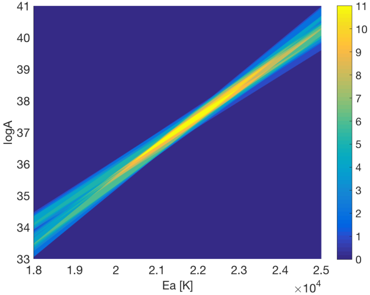
- Same picture when including uncertainty
- The rate at each temperature defines an area in logA-E_a space
- Overlapping regions imply model consensus across temperatures and represent more likely parameter values



Rate parameters - correlation

- Same picture when including uncertainty
- The rate at each temperature defines an area in logA-E_a space
- Overlapping regions imply model consensus across temperatures and represent more likely parameter values





Summing the overlaps gives some notion of an approximate parameter density. In the limit of infinite experimental temperature stations could arrive at a PDF. Regardless, parameter correlation is strongly expressed.

Data and parameter inference

- Data and parameter inference procedure proposed by Berry et al. [1] based on maximum entropy arguments and utilizing Approximate Bayesian Computation (ABC) methods [2]

[1] Berry et al. J. of Comp. Phys. 231 (2012)
[2] Beaumont et al. Genetics 162 (2002)

Data and parameter inference

- Data and parameter inference procedure proposed by Berry et al. [1] based on maximum entropy arguments and utilizing Approximate Bayesian Computation (ABC) methods [2]
- Propose a measurement data set, accept if it is consistent with the available reported statistics on the fitting parameters of the missing data (e.g. reported error bars on the Arrhenius parameters deduced from the missing data by the experimentalist)

[1] Berry et al. J. of Comp. Phys. 231 (2012)
[2] Beaumont et al. Genetics 162 (2002)

Data and parameter inference

- Data and parameter inference procedure proposed by Berry et al. [1] based on maximum entropy arguments and utilizing Approximate Bayesian Computation (ABC) methods [2]
- Propose a measurement data set, accept if it is consistent with the available reported statistics on the fitting parameters of the missing data (e.g. reported error bars on the Arrhenius parameters deduced from the missing data by the experimentalist)
- Pose problem in a Bayesian framework to explore data space to discover consistent data sets

[1] Berry et al. J. of Comp. Phys. 231 (2012)
[2] Beaumont et al. Genetics 162 (2002)

Data and parameter inference

- Data and parameter inference procedure proposed by Berry et al. [1] based on maximum entropy arguments and utilizing Approximate Bayesian Computation (ABC) methods [2]
- Propose a measurement data set, accept if it is consistent with the available reported statistics on the fitting parameters of the missing data (e.g. reported error bars on the Arrhenius parameters deduced from the missing data by the experimentalist)
- Pose problem in a Bayesian framework to explore data space to discover consistent data sets
- Algorithm composed of an outer inference on data subject to constraints, with a nested inner inference on parameters (rate coefficients and data error variance) with an error model

[1] Berry et al. J. of Comp. Phys. 231 (2012)
[2] Beaumont et al. Genetics 162 (2002)

Bayesian framework

Bayes' Law

$$p(A,B) = p(A|B)\,p(B) = p(B|A)\,p(A)$$

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

Bayesian framework

Bayes' Law

$$p\left(A,B\right)=p\left(A|B\right)p\left(B\right)=p\left(B|A\right)p\left(A\right)$$

$$\Rightarrow p\left(A|B\right)=\frac{p(B|A)p(A)}{p(B)}$$

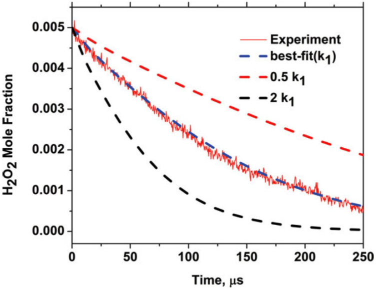
Bayesian Inference

$$p\left(A|B\right)\propto p\left(B|A\right)\pi\left(A\right)$$

$p\left(A|B\right)$: posterior distribution of A given information B
$p\left(B|A\right)$: likelihood function, probability that A is consistent with B
$\pi\left(A\right)$: prior distribution of A

H_2O_2 decomposition - target experiment

- $H_2O_2 + M \longrightarrow 2\,OH + M$
  consider low pressure results



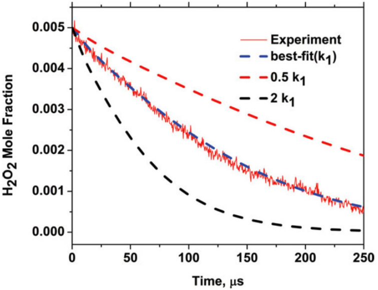
P [atm]	T [K]	$H_2O_{2\text{initial}}$ [%]
1.15	1012.5	0.42
1.08	1186.8	0.45
2.33	997.54	0.47
2.35	1166.6	0.50

Table: Experimental conditions

[3] Sajid et al. Int. J. Chem. Kinet. (2014)
[4] Hong et al. Combustion and Flame 158 (2011)

H_2O_2 decomposition - target experiment

- $H_2O_2 + M \longrightarrow 2\,OH + M$
  consider low pressure results
- Shock tube study, laser absorption by H_2O_2 [3]



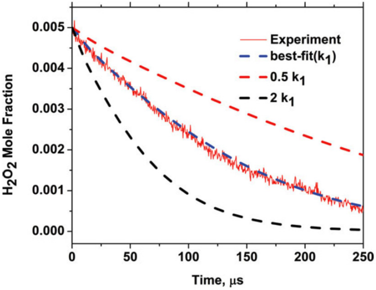
P [atm]	T [K]	$H_2O_{2\,initial}$ [%]
1.15	1012.5	0.42
1.08	1186.8	0.45
2.33	997.54	0.47
2.35	1166.6	0.50

Table: Experimental conditions

[3] Sajid et al. Int. J. Chem. Kinet. (2014)
[4] Hong et al. Combustion and Flame 158 (2011)

H_2O_2 decomposition - target experiment

- $H_2O_2 + M \longrightarrow 2\,OH + M$
  consider low pressure results
- Shock tube study, laser absorption by H_2O_2 [3]
- Rate constant determined by solving ODE system with a calibration chemical mechanism [4]



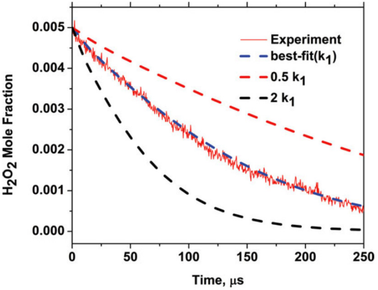
P [atm]	T [K]	$H_2O_{2\text{initial}}$ [%]
1.15	1012.5	0.42
1.08	1186.8	0.45
2.33	997.54	0.47
2.35	1166.6	0.50

Table: Experimental conditions

[3] Sajid et al. Int. J. Chem. Kinet. (2014)
[4] Hong et al. Combustion and Flame 158 (2011)

H_2O_2 decomposition - target experiment

- $H_2O_2 + M \longrightarrow 2\,OH + M$
  consider low pressure results
- Shock tube study, laser absorption by H_2O_2 [3]
- Rate constant determined by solving ODE system with a calibration chemical mechanism [4]
- Data are concentration vs. time decay profiles of H_2O_2 derived from the device signal



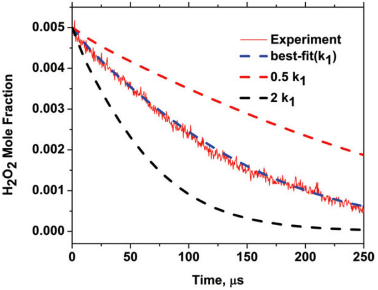
P [atm]	T [K]	$H_2O_{2initial}$ [%]
1.15	1012.5	0.42
1.08	1186.8	0.45
2.33	997.54	0.47
2.35	1166.6	0.50

Table: Experimental conditions

[3] Sajid et al. Int. J. Chem. Kinet. (2014)
[4] Hong et al. Combustion and Flame 158 (2011)

H_2O_2 decomposition - target experiment

- $H_2O_2 + M \longrightarrow 2\,OH + M$
  consider low pressure results
- Shock tube study, laser absorption by H_2O_2 [3]
- Rate constant determined by solving ODE system with a calibration chemical mechanism [4]
- Data are concentration vs. time decay profiles of H_2O_2 derived from the device signal
- Reported fit:
  $k(T)=10^{16.29\pm0.12}\exp(-21993\pm301/T)$



P [atm]	T [K]	$H_2O_{2\,\mathrm{initial}}$ [%]
1.15	1012.5	0.42
1.08	1186.8	0.45
2.33	997.54	0.47
2.35	1166.6	0.50

Table: Experimental conditions

[3] Sajid et al. Int. J. Chem. Kinet. (2014)
[4] Hong et al. Combustion and Flame 158 (2011)

Data and parameter inference procedure

Missing data model (Gaussian error/noise, $\varepsilon \sim \mathcal{N}(0,1)$):

$$[\mathrm{H_2O_{2data}}] = [\mathrm{H_2O_{2model}}] + \sigma\varepsilon$$

Data and parameter inference procedure

Missing data model (Gaussian error/noise, $\varepsilon \sim \mathcal{N}(0,1)$):

$$[\mathrm{H_2O_{2data}}] = [\mathrm{H_2O_{2model}}] + \sigma\varepsilon$$

Data inference, data likelihood (ABC):

$$p(\mathbf{D}|S) = \frac{p(S|\mathbf{D})\pi(\mathbf{D})}{p(S)} \qquad p(S|\mathbf{D}) = \prod_{i=1}^{4} \exp\left[-\delta_i \left(\frac{S_{i,\mathbf{D}} - S_{i,\mathrm{experiment}}}{S_{i,\mathrm{experiment}}}\right)\right]$$

Data and parameter inference procedure

Missing data model (Gaussian error/noise, $\varepsilon \sim \mathcal{N}(0,1)$):

$$[\mathrm{H_2O_{2data}}] = [\mathrm{H_2O_{2model}}] + \sigma\varepsilon$$

Data inference, data likelihood (ABC):

$$p(\mathbf{D}|S) = \frac{p(S|\mathbf{D})\pi(\mathbf{D})}{p(S)} \qquad p(S|\mathbf{D}) = \prod_{i=1}^{4} \exp\left[-\delta_i\left(\frac{S_{i,\mathbf{D}} - S_{i,\mathrm{experiment}}}{S_{i,\mathrm{experiment}}}\right)\right]$$

Parameter inference (to establish $S_{\mathbf{D}}$), parameter likelihood:

$$p(\lambda|\mathbf{D}) = \frac{p(\mathbf{D}|\lambda)\pi(\lambda)}{p(\mathbf{D})} \qquad p(\mathbf{D}|\lambda) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{\left([\mathrm{H_2O_{2data}}]_{i,j} - [\mathrm{H_2O_{2model}}]_{i,j}\right)^2}{2\sigma^2}\right)$$

Data and parameter inference procedure

Missing data model (Gaussian error/noise, $\varepsilon \sim \mathcal{N}(0,1)$):

$$[\mathrm{H_2O_{2data}}] = [\mathrm{H_2O_{2model}}] + \sigma\varepsilon$$

Data inference, data likelihood (ABC):

$$p(\mathbf{D}|S) = \frac{p(S|\mathbf{D})\pi(\mathbf{D})}{p(S)} \qquad p(S|\mathbf{D}) = \prod_{i=1}^{4} \exp\left[-\delta_i\left(\frac{S_{i,\mathbf{D}} - S_{i,\mathrm{experiment}}}{S_{i,\mathrm{experiment}}}\right)\right]$$

Parameter inference (to establish $S_\mathbf{D}$), parameter likelihood:

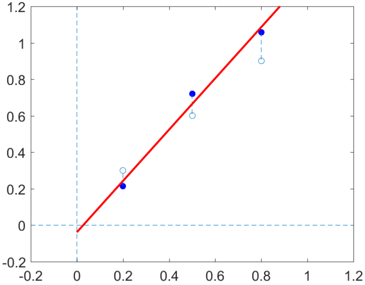
$$p(\lambda|\mathbf{D}) = \frac{p(\mathbf{D}|\lambda)\pi(\lambda)}{p(\mathbf{D})} \qquad p(\mathbf{D}|\lambda) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{\left([\mathrm{H_2O_{2data}}]_{i,j} - [\mathrm{H_2O_{2model}}]_{i,j}\right)^2}{2\sigma^2}\right)$$

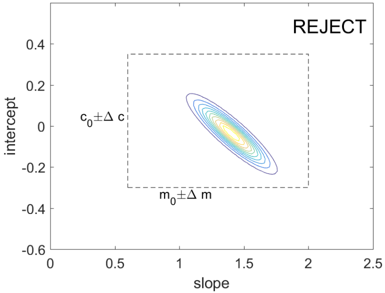
Knowledge of the posterior PDFs are constructed using Markov Chain Monte Carlo sampling for both the outer (data) and inner (parameter) inference.

Data inference - parameter line model

Data model:

$$y_{data} = y_{model} + y_{error} = mx + c + y_{error} \tag{1}$$

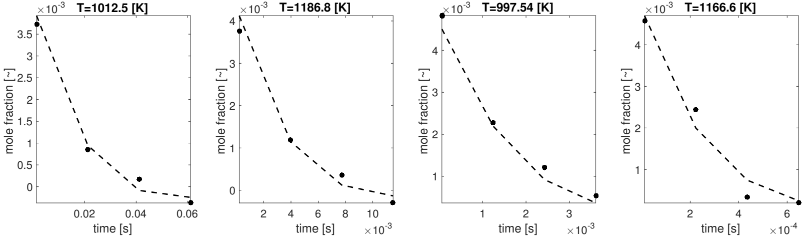




Accept data consistent with constraints, i.e. mean line slope and intercept (m_0, c_0) and marginal error bars on m,c, interpreted as standard deviations

Inferred H_2O_2 data

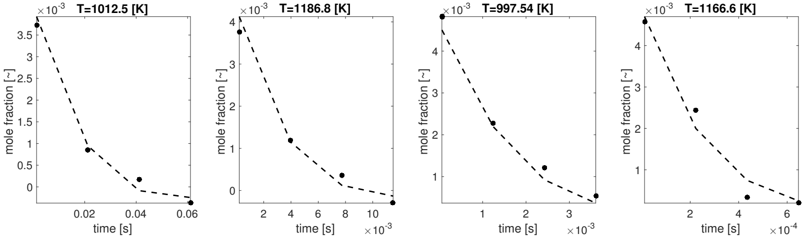
- Replicate the context of the experiment as closely as possible based on reported information



Consistent data sample arrising from the data inference procedure (dashed lines are model predictions using nominal rate parameters).

Inferred H_2O_2 data

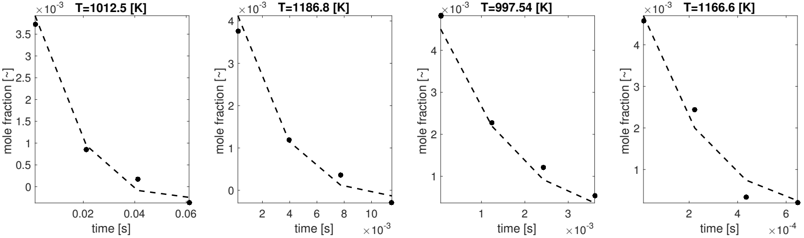
- Replicate the context of the experiment as closely as possible based on reported information
- Use identical chemical mechanism and thermochemistry database for forward model



Consistent data sample arrising from the data inference procedure (dashed lines are model predictions using nominal rate parameters).

Inferred H_2O_2 data

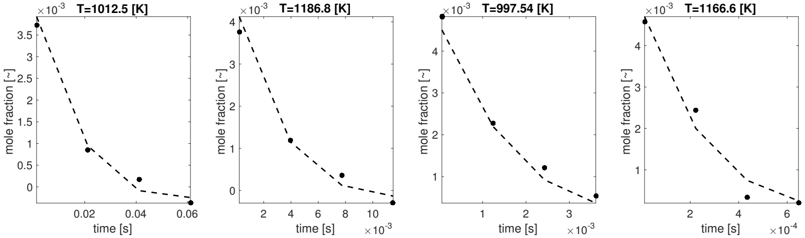
- Replicate the context of the experiment as closely as possible based on reported information
- Use identical chemical mechanism and thermochemistry database for forward model
- Use reported initial conditions to initialize model evaulations at the reported temperatures



Consistent data sample arrising from the data inference procedure (dashed lines are model predictions using nominal rate parameters).

Inferred H_2O_2 data

- Replicate the context of the experiment as closely as possible based on reported information
- Use identical chemical mechanism and thermochemistry database for forward model
- Use reported initial conditions to initialize model evaulations at the reported temperatures
- Enforce constraints on the data in the form of symmetric error bars on the reported Arrhenius parameters, i.e. $k(T)=10^{16.29\pm0.12}\exp(-21993\pm301/T)$



Consistent data sample arrising from the data inference procedure (dashed lines are model predictions using nominal rate parameters).

Pooled parameter PDF

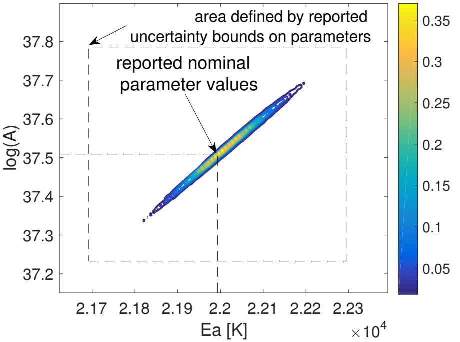


Figure: contours of probability density, joint parameter posterior PDF

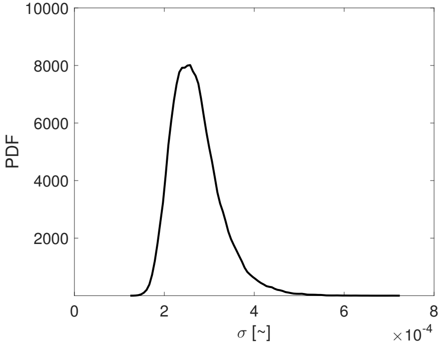


Figure: marginal PDF of data error

- marginalize over data in joint data-parameter posterior

Pooled parameter PDF

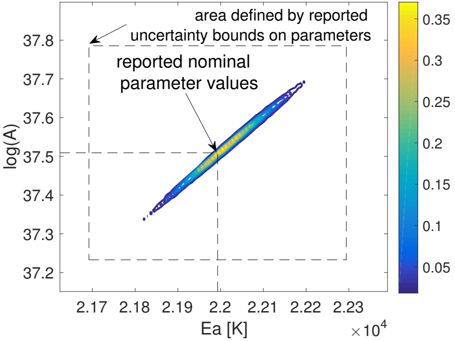


Figure: contours of probability density, joint parameter posterior PDF

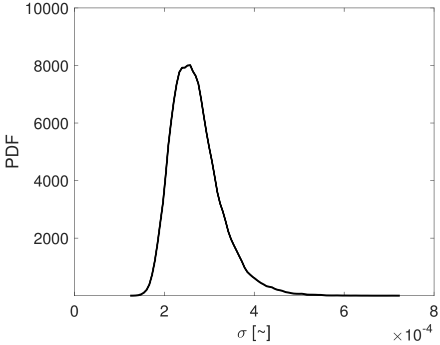


Figure: marginal PDF of data error

- marginalize over data in joint data-parameter posterior
- PDF reveals parameter correlation

Pooled parameter PDF

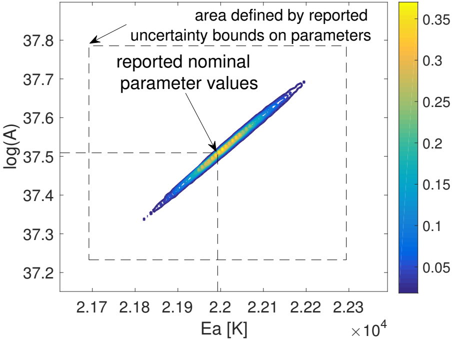


Figure: contours of probability density, joint parameter posterior PDF

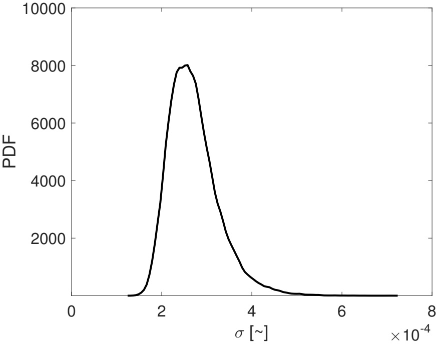


Figure: marginal PDF of data error

- marginalize over data in joint data-parameter posterior
- PDF reveals parameter correlation
- nominal values occur in regions of high probability mass

Pooled parameter PDF

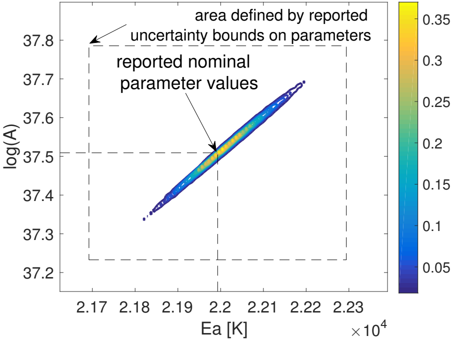


Figure: contours of probability density, joint parameter posterior PDF

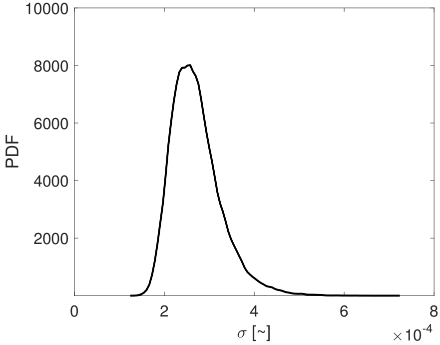


Figure: marginal PDF of data error

- marginalize over data in joint data-parameter posterior
- PDF reveals parameter correlation
- nominal values occur in regions of high probability mass
- noise amplitude is also inferred

Application: H_2-O_2 explosion limits

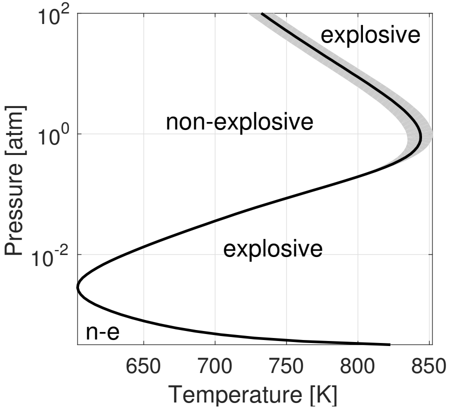


Figure: stoichiometric H_2-O_2 explosion limit curve

- H_2O_2 decomposition controls the 3rd explosion limit in the H_2-O_2 system

Application: H_2-O_2 explosion limits

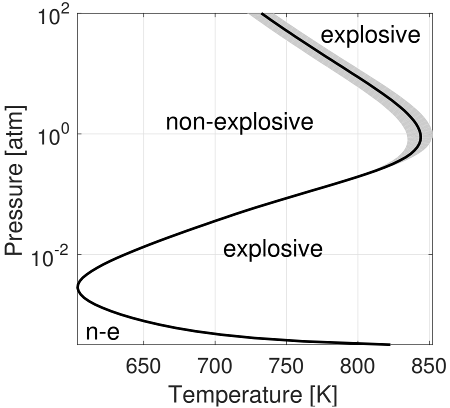


Figure: stoichiometric H_2-O_2 explosion limit curve

- H_2O_2 decomposition controls the 3rd explosion limit in the H_2-O_2 system
- forward model: homogenous ignition delay

Application: H_2-O_2 explosion limits

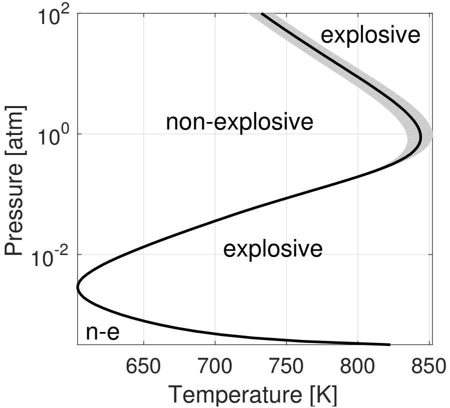
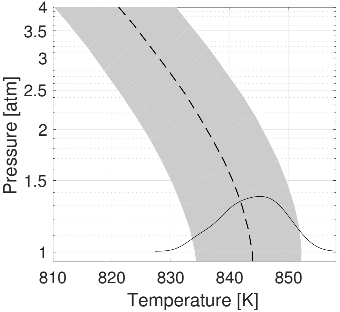


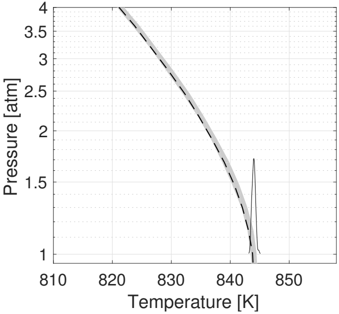
Figure: stoichiometric H_2-O_2 explosion limit curve

- H_2O_2 decomposition controls the 3rd explosion limit in the H_2-O_2 system
- forward model: homogenous ignition delay
- binary classifier establishes the limit curve within ≈ 0.1 K.

Uncertainty propagation



(a)



(b)

Sampling from uniform space defined by symmetric error bounds predicts large uncertainty in the third limit location, while sampling from the correlated posterior parameter PDF results in much smaller uncertainty

Summary

Data inference

Demonstrate capability for generating representations of missing experimental data given constaints on the data space in the form of statistics on reported Arrhenius parameters

Summary

Data inference

Demonstrate capability for generating representations of missing experimental data given constaints on the data space in the form of statistics on reported Arrhenius parameters

'Data-free' parameter inference

Marginalizing over the consistent data delivers the correlated joint parameter posterior PDF consistent with the reported experimental statistics and conditions, i.e. parameter inference in the absence of data

Summary

Data inference

Demonstrate capability for generating representations of missing experimental data given constaints on the data space in the form of statistics on reported Arrhenius parameters

'Data-free' parameter inference

Marginalizing over the consistent data delivers the correlated joint parameter posterior PDF consistent with the reported experimental statistics and conditions, i.e. parameter inference in the absence of data

Importance of experimental data

Highlight the importance of publishing raw experimental data for enabling uncertainty quantification. However many rate expressions in production mechanisms are based on legacy experiments whose data is lost.

Summary

Data inference

Demonstrate capability for generating representations of missing experimental data given constaints on the data space in the form of statistics on reported Arrhenius parameters

'Data-free' parameter inference

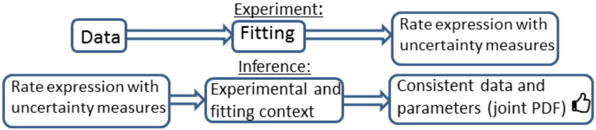
Marginalizing over the consistent data delivers the correlated joint parameter posterior PDF consistent with the reported experimental statistics and conditions, i.e. parameter inference in the absence of data

Importance of experimental data

Highlight the importance of publishing raw experimental data for enabling uncertainty quantification. However many rate expressions in production mechanisms are based on legacy experiments whose data is lost.

Data combination

Presents a framework for pooling data across experiments for determining consensus rate expresssions.



This work was supported by the US Department of Energy (DOE), Office of Basic Energy Sciences (BES) Division of Chemical Sciences, Geosciences, and Biosciences. Sandia National Laboratories is a multi-program laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under contract DE-AC04-94-AL85000.