MCMC and SMC algorithms for estimating parameters in Thermogravemetric Analysis.

Matthew Berry

Supervisors: Matt Moores, Mark Nelson and Brian Monaghan.

Steel Research Hub, University of Wollongong

June 27, 2022

Motivation

- What problem is motivating this work?

Motivation

- What problem is motivating this work?
- What are we trying to achieve?

Motivation

- What problem is motivating this work?
- What are we trying to achieve?
- Why are we interested in Bayesian methods?

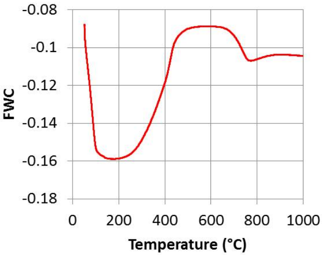
Thermogravemetric Analysis

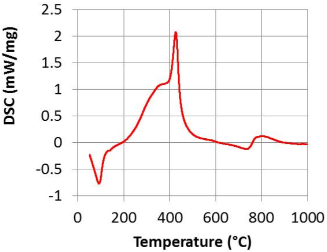
- How does the experiment work?
- What are we trying to achieve with the experiment?
- What data do we obtain from it?

Thermogravemetric Analysis

- How does the experiment work?
- What are we trying to achieve with the experiment?
- What data do we obtain from it?

We denote the mass data by **m** and the DSC by **d**.

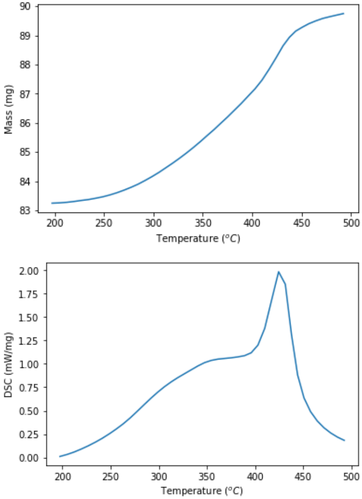




Reaction Equations

Reactions:

$$\mathrm{FeO}_{(s)} + \frac{1}{6}\mathrm{O}_{2(g)} = \frac{1}{3}\mathrm{Fe}_3\mathrm{O}_{4(s)},$$
$$\mathrm{Fe}_{(s)} + \frac{3}{4}\mathrm{O}_{2(g)} = \frac{1}{2}\mathrm{Fe}_2\mathrm{O}_{3(s)}.$$



Mathematical Model

Model Assumptions:

- The sample is homogeneous.
- The temperature does not vary within the sample and the sample is heated at a constant rate.
- The reactions are independent of each other.

Model Equation

$$\frac{dM_i}{dt} = -M_i A_i \exp\left(\frac{-E_i}{RT}\right) \quad (1)$$

$$\frac{dT}{dt} = \alpha \quad (2)$$

UOW AUSTRALIA

Combining the Reaction Data

Calculate sample weight,

$$\frac{dM(t)}{dt} = \sum w_{c,i} M_i A_i \exp\left(\frac{-E_i}{RT}\right) \tag{3}$$

Calculating the DSC data

$$D(t) = \sum Q_i M_i A_i \exp\left(\frac{-E_i}{RT}\right) \tag{4}$$

These are evaluated at some vector of times to obtain the solution vectors, $\mathbf{M}_\tau$ for mass and $\mathbf{D}_\tau$ for DSC.
We then have the residuals, $\mathbf{e}_M = \mathbf{M}_\tau - \mathbf{m}$, and $\mathbf{e}_D = \mathbf{D}_\tau - \mathbf{d}$

Inferred Parameters

Within our model we have some parameters that need to be inferred $\theta_m = [A_i, E_i, Q_i]$. Whilst the parameters $w_{c,i}$ are fixed by the choice of reactions and R is the ideal gas constant.

Inferred Parameters

Within our model we have some parameters that need to be inferred $\theta_m = [A_i, E_i, Q_i]$. Whilst the parameters $w_{c,i}$ are fixed by the choice of reactions and R is the ideal gas constant.

These model parameters are not the ideal choice to infer.

Instead we use the parameters,

$$\theta_m = [\widetilde{E}_i, T_{m,i}, \widetilde{Q}_i], \tag{5}$$

where $\widetilde{E}_i = \log_{10} E$, $\widetilde{Q}_i = \log_{10} Q$ and,

$$A_i = \frac{E_i}{RT_{m,i}^2} \exp\left(\frac{E_i}{RT_{m,i}}\right). \tag{6}$$

Bayesian Framework

We assume the residuals are normally distributed,

$$\begin{aligned} \mathbf{e}_m &\sim \mathcal{N}\left(0, \sigma_M^2 \mathbf{I}\right), \\ \mathbf{e}_d &\sim \mathcal{N}\left(0, \sigma_D^2 \mathbf{I}\right), \end{aligned}$$

Bayesian Framework

We assume the residuals are normally distributed,

$$\mathbf{e}_m \sim \mathcal{N}\left(0, \sigma_M^2 \mathbf{I}\right),$$
$$\mathbf{e}_d \sim \mathcal{N}\left(0, \sigma_D^2 \mathbf{I}\right),$$

Prior Distributions

$$\widetilde{E}_i \sim \mathcal{U}(2, 6),$$
$$\widetilde{Q}_i \sim \mathcal{U}(2, 6),$$
$$\sigma_M^2 \sim \mathcal{IG}(1, 0.001),$$
$$\sigma_D^2 \sim \mathcal{IG}(1, 0.001),$$

Prior Distribution for T_m

The Prior Distribution for T_m is informed by the experimental data available.

Prior Distribution for T_m

The Prior Distribution for T_m is informed by the experimental data available.

- Independent and uniformly distributed.

Prior Distribution for T_m

The Prior Distribution for T_m is informed by the experimental data available.

- Independent and uniformly distributed.
- Uniformly distributed over the region, $a < T_{m,1} < T_{m,2} < \ldots < T_{m,n} < b$.

Prior Distribution for T_m

The Prior Distribution for T_m is informed by the experimental data available.

- Independent and uniformly distributed.
- Uniformly distributed over the region, $a < T_{m,1} < T_{m,2} < \ldots < T_{m,n} < b$.
- Independent and normally distributed around some informed mean.

Prior Distribution for T_m

The Prior Distribution for T_m is informed by the experimental data available.

- Independent and uniformly distributed.
- Uniformly distributed over the region, $a < T_{m,1} < T_{m,2} < ... < T_{m,n} < b$.
- Independent and normally distributed around some informed mean.
- Normally Distributed conditioned on being larger than the previous temperature.

Prior Distribution for T_m

The Prior Distribution for T_m is informed by the experimental data available.

- Independent and uniformly distributed.
- Uniformly distributed over the region, $a < T_{m,1} < T_{m,2} < \ldots < T_{m,n} < b$.
- Independent and normally distributed around some informed mean.
- Normally Distributed conditioned on being larger than the previous temperature.

We use the independent and normally distributed prior.

The Intractible Likelihood

Our posterior distribution is then given by,

$$p(\boldsymbol{\theta} \mid \mathbf{y}) = \frac{p(\mathbf{y} \mid \boldsymbol{\theta})\; p(\boldsymbol{\theta})}{p(\mathbf{y})}, \tag{7}$$

with,

$$p(\mathbf{y}) = \int p(\mathbf{y} \mid \boldsymbol{\theta})\; p(\boldsymbol{\theta})$$

The Intractible Likelihood

Our posterior distribution is then given by,

$$p(\boldsymbol{\theta} \mid \mathbf{y}) = \frac{p(\mathbf{y} \mid \boldsymbol{\theta})\ p(\boldsymbol{\theta})}{p(\mathbf{y})}, \tag{7}$$

with,

$$p(\mathbf{y}) = \int p(\mathbf{y} \mid \boldsymbol{\theta})\ p(\boldsymbol{\theta})$$

Since $p(\mathbf{y} \mid \boldsymbol{\theta})$ involves solving a differential Equation, this cannot be computed directly.

Proposal Distributions

Random Walk proposal for θ_m.

$$\theta_m^* \sim \mathcal{N}(\theta_m, \Sigma),$$

Where Σ is diagonal.

Proposal Distributions

Random Walk proposal for θ_m.

$$\theta_m^* \sim \mathcal{N}(\theta_m, \Sigma),$$

Where Σ is diagonal.
Propose σ using the distribution,

$$\sigma^{2*} \sim \mathcal{IG}\left(1 + n/2, 0.001 + \frac{\mathbf{e}^T\mathbf{e}}{2}\right),$$

where n is the number of data points.

The Metropolis-Hastings Algorithm

Initialise $\boldsymbol{\theta}_0$ by sampling from the priors;

for $j = 1, \ldots, J$ **do**

  Propose new parameters $\boldsymbol{\theta}^* \sim q(\cdot \mid \boldsymbol{\theta}_{j-1})$;

  Solve $\mathbf{M}_\tau(\boldsymbol{\theta}^*)$ Using Runge–Kutta;

  Calculate $\mathbf{D}_\tau(\boldsymbol{\theta}^*)$;

  Set $\pi = \min\left\{\frac{p(\mathbf{y}|\boldsymbol{\theta}^*)\ q(\boldsymbol{\theta}_{j-1}|\boldsymbol{\theta}^*)\ p(\boldsymbol{\theta}^*)}{p(\mathbf{y}|\boldsymbol{\theta}_{j-1})\ q(\boldsymbol{\theta}^*|\boldsymbol{\theta}_{j-1})\ p(\boldsymbol{\theta}_{j-1})}, 1\right\}$;

  Sample $b \sim \mathcal{U}(0, 1)$;

  if $b \leq \pi$ **then**

    Set $\boldsymbol{\theta}_j = \boldsymbol{\theta}^*$

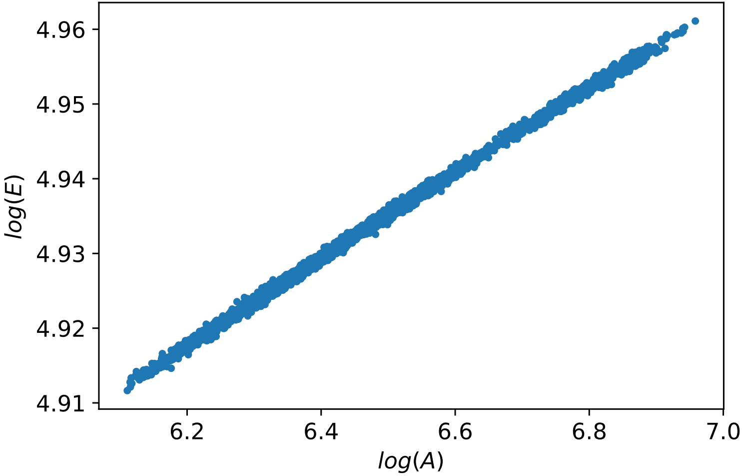
  else

    Set $\boldsymbol{\theta}_j = \boldsymbol{\theta}_{t-1}$

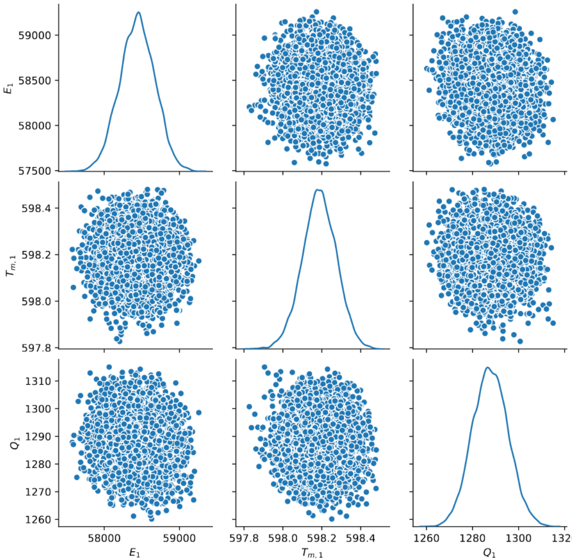
  end

end

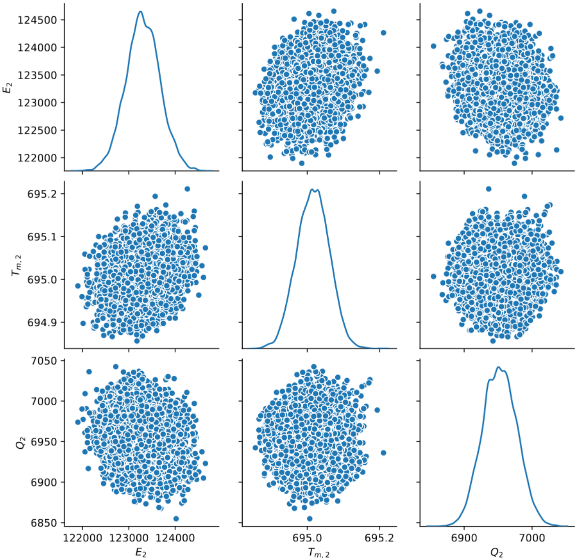
A vs E



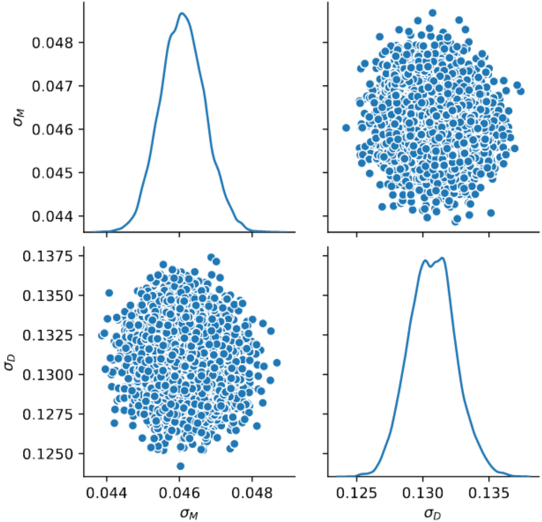
First Reaction



Second Reaction



Noise Parameters



Convergence

We assess convergence using the Split-$\widehat{R}$ and the Effective Sample Size (ESS)

	Split-$\widehat{R}$	ESS
E_1	1.00064	2006
$T_{m,1}$	1.00190	2985
Q_1	1.00118	2011
E_2	1.00166	2210
$T_{m,2}$	1.00083	2730
Q_2	1.00143	1610
σ_M	1.00050	8650
σ_D	1.00081	8342

Sequential Monte Carlo

Where does SMC fit into our problem?

Sequential Monte Carlo

Where does SMC fit into our problem?

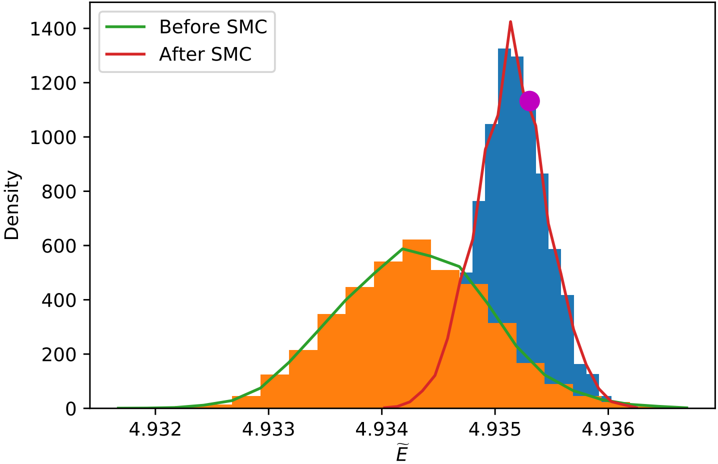
- We have multiple TGA data sets.

Sequential Monte Carlo

Where does SMC fit into our problem?

- We have multiple TGA data sets.
- We can compare different proposed reaction schemes.

Results



Before SMC
After SMC
1400
1200
1000
800
600
400
200
0
Density
4.932
4.933
4.934
4.935
4.936
$\widetilde{E}$

Future and Ongoing Work

We can extend this into multiple directions:

Future and Ongoing Work

We can extend this into multiple directions:

- Include reactions that involve multiple reactants within the sample.

Future and Ongoing Work

We can extend this into multiple directions:

- Include reactions that involve multiple reactants within the sample.
- Extend to dependant reactions.

Future and Ongoing Work

We can extend this into multiple directions:

- Include reactions that involve multiple reactants within the sample.
- Extend to dependant reactions.
- Apply this to data with more reactions.

Future and Ongoing Work

We can extend this into multiple directions:

- Include reactions that involve multiple reactants within the sample.
- Extend to dependant reactions.
- Apply this to data with more reactions.
- Apply the SMC algorithm onto our experimental data.

Future and Ongoing Work

We can extend this into multiple directions:

- Include reactions that involve multiple reactants within the sample.
- Extend to dependant reactions.
- Apply this to data with more reactions.
- Apply the SMC algorithm onto our experimental data.
- Use these posterior samples in the forward stockpile problem.

Future and Ongoing Work

We can extend this into multiple directions:

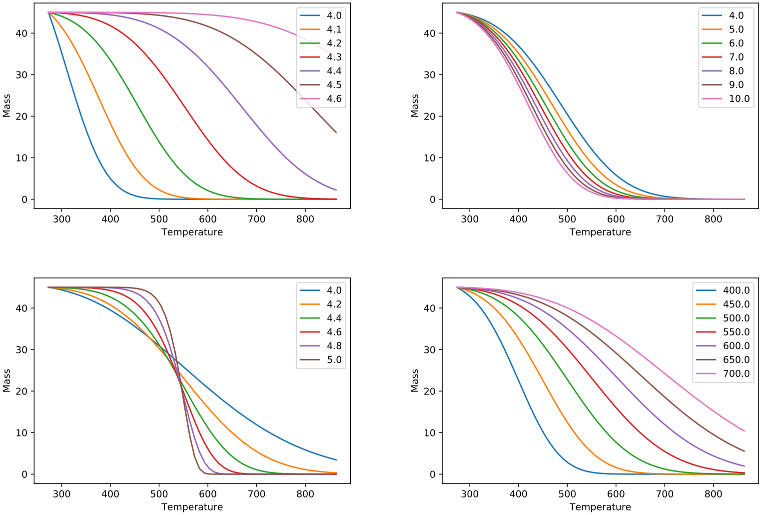
- Include reactions that involve multiple reactants within the sample.
- Extend to dependant reactions.
- Apply this to data with more reactions.
- Apply the SMC algorithm onto our experimental data.
- Use these posterior samples in the forward stockpile problem.
- Examine and compare different reaction schemes that have been proposed.

Conclusion

The Key points from our analysis.

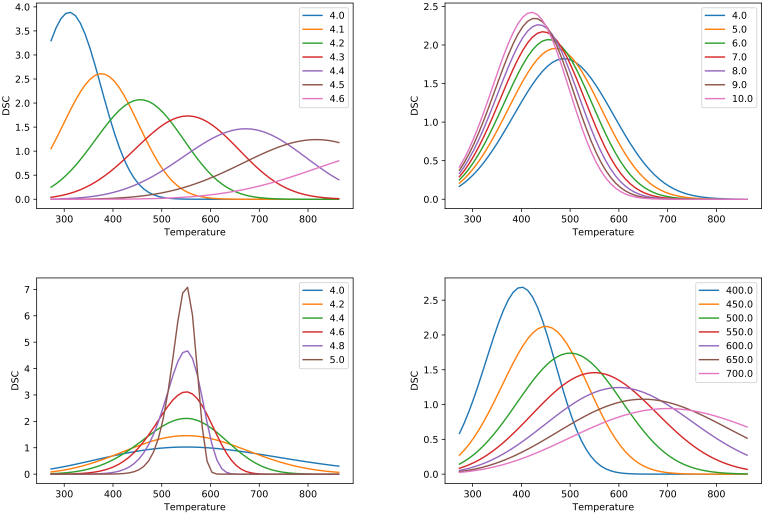
- MCMC algortihms are a valuable tool in addressing the inverse problem and quantifying the uncertainty.
- Reparameterisation of the model can be extremely valuable to improving the MCMC algorithm.
- Sequential Monte Carlo can be used to improve the uncertainty estimates of our model parameters.

$[A, E] or [T_m, E]$



Mass
Temperature
4.0
4.1
4.2
4.3
4.4
4.5
4.6
Mass
Temperature
4.0
5.0
6.0
7.0
8.0
9.0
10.0
Mass
Temperature
4.0
4.2
4.4
4.6
4.8
5.0
Mass
Temperature
400.0
450.0
500.0
550.0
600.0
650.0
700.0

$[A, E] or [T_m, E]$



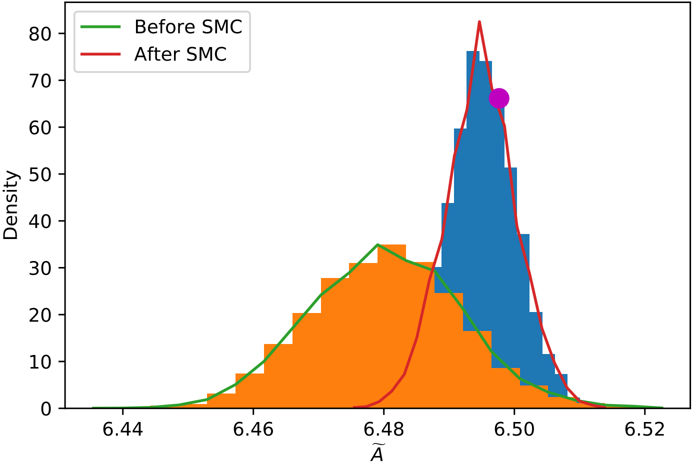
DSC
Temperature
4.0
4.1
4.2
4.3
4.4
4.5
4.6
5.0
6.0
7.0
8.0
9.0
10.0
4.8
400.0
450.0
500.0
550.0
600.0
650.0
700.0

Sequential Monte Carlo Algorithm

We utilise importance sampling to add additional information.

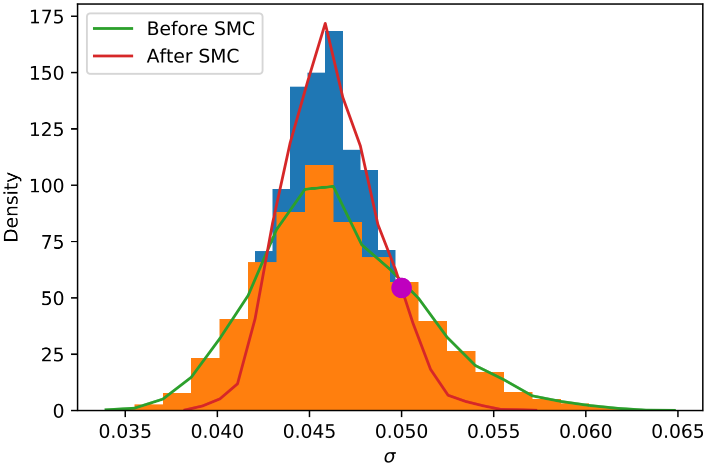
1. Obtain a sample of the posterior distribution from one experiment.
2. Weight particles using the new experimental data.
3. If ESS is below a threshold, then resample the particles.
4. Resample using a Metropolis Hastings Algorithm.
5. Repeat the reweighting process for as many additional experiments exist.

SMC A



Before SMC
After SMC
Density
0
10
20
30
40
50
60
70
80
6.44
6.46
6.48
6.50
6.52
$\widetilde{A}$

SMC σ



Before SMC
After SMC
Density
175
150
125
100
75
50
25
0
0.035
0.040
0.045
0.050
0.055
0.060
0.065
σ