2023

Modelling the Self-Heating of Steel Stockpiles

Matthew Berry

UNIVERSITY
OF WOLLONGONG
AUSTRALIA

Modelling the Self-Heating of Steel Stockpiles

Matthew Berry

This thesis is presented as part of the requirements for the conferral of the degree:

Doctor of Philosophy (Mathematics)

Supervisor:
Associate Prof. Mark Nelson

Co-supervisors:
Dr. M. Moores and Prof. B. Monaghan

The University of Wollongong
School of Mathematics and Applied Statistics

June 25, 2023

This research has been conducted with the support of an Australian Government Research Training Program Scholarship.

Declaration

I, *Matthew Berry*, declare that this thesis is submitted in partial fulfilment of the requirements for the conferral of the degree *Doctor of Philosophy (Mathematics)*, from the University of Wollongong, is wholly my own work unless otherwise referenced or acknowledged. This document has not been submitted for qualifications at any other academic institution.

Matthew Berry

June 25, 2023

Abstract

Understanding the uncertainty of model parameters is crucial for building predictive models. Within the field of spontaneous ignition a slight variation in the model parameters can cause a significant variation in our ability to determine if ignition occurs. We consider this problem through an application to the steel industry. A byproduct of the steelmaking process is stockpiled where oxidation can induce ignition. The resulting ignition process sinters the filter improving the durability. Understanding this process requires careful modelling and consideration of the uncertainty in the reaction kinetics. We examine some experimental data on the filter cake to determine these reaction kinetics. Due to the complex nature of the filter cake, standard estimation techniques are difficult to apply and the uncertainty in our parameters cannot be an input into the larger stockpiles. We apply a Bayesian framework for parameter estimation that considers a distribution for the parameters rather than a point estimation with an uncertainty. Using this approach we construct a Markov Chain Monte Carlo (MCMC) algorithm to sample this distribution and test this against simulated data; we capture the true values of the reaction kinetics within the target distribution. Our approach uses the experimental data, and we construct meaningful estimates for the reaction kinetics.

Once we consider multiple sets of experimental data it highlights some issues with our proposed reaction scheme and methodology. Our Sequential Monte Carlo approach identifies a discrepancy between the estimates for different experiments. We consider multiple different avenues to resolve this discrepancy. The most significant of these methods was to consider a different reaction scheme. In addition to this, the initial reactant concentrations were found to have a significant effect on the posterior distribution of parameter estimates; if we change these values then our distributions would not be overlapping. Additionally, we then consider a different approach to combining the multiple experiments; combining the parameter distributions from the MCMC method, we propose a new prior that contains this information and use this to combine the experimental data. This is effective at generating a new sample, though it does not address any of the issues that may arise from model mis-specification.

After obtaining the sample of parameter estimates we consider the scaled version of the forward stockpile model where we examine the spontaneous ignition. For generic stockpiles, the natural variation of ambient temperature had a significant effect on the

ignition behaviour of the stockpiles. We consider both Dirichlet boundary conditions and the Newtonian cooling boundary conditions. Results for both boundary conditions were consistent about the effects of the seasonal variation. We also examine the potential for using a hotspot of heated material to assist with inducing ignition. We demonstrate that such hotspots are effective at inducing ignition, with large hotspots at the centre of the stockpile proving mot effective. Inert hotspots require higher temperatures to induce ignition due to their inability to react, however large hotspots proved counter-productive for a fixed length stockpile.

Examining the ambient temperature variations and the use of hotspots provided valuable insight into the stockpile behaviour. What remained to investigate was how these results translate to the stockpiles of the steel-making byproduct. Our results are consistent with those from the scaled stockpiles. In addition to this, we examine the effect of the parameters posterior distribution, compared to point estimates have upon the predicted ignition behaviour. Using a sample from the posterior distribution we calculate the critical length the stockpile must exceed in order to induce ignition. From this distribution of critical lengths we can examine the uncertainty in the critical length and consider this when making recommendations about how to construct the stockpiles. When we compare the point estimate from the median parameter values, we find that these correspond to a critical length in the 20th percentile. Therefore if we only used this point estimate for our parameters, we can only be 20% sure that the length of the stockpiles is sufficient to cause ignition. This result highlights the importance of the uncertainty in the parameter estimates when advising on any decisions regarding inducing or preventing ignition in various applications.

Acknowledgements

In order to complete a thesis, a large amount of assistance is necessary. Firstly, I want to acknowledge the financial support I recieved from the Australian Government through the Research Training Program scholarship, and the AF Pillow Foundation for their scholarship. I would like to acknowledge all of the support that I received from my supervisors. The expertise and guidance they were able to provide me with was so valuable. In addition to my supervisors I would like to acknowledge the assistance from Raymond Longbottom who I was able to collaborate with and provided much needed insight into the experimental procedures conducted at the university on the Filter cake. The School of Mathematics and Statistics along with the National institute of Applied Statistics Research Australia, particularly Adam, Maureen and Aiden, at UOW provided a lot of support and opportunity to allow me to develop as a researcher throughout the course of this PhD.

Along with the academic support from various individuals at the university, I received a lot of mental support along the way too. My family provided some much needed grounding and a way to escape from the world of research when I was home. I would like to thank my many fellow PhD students who I was able to catch up with along the way and just talk about anything that we felt like. I would particularly like to acknowledge, Angus, Thalia, Lachlan and Jono, who I would regularly catch up with for some much needed beverages at the unibar on a Friday afternoon or Lunch. Those drinks were much needed after a long week and the banter was always enjoyed.

Finally, I want to acknowledge all the support that I received from my partner Sarah. The amount of support you provided me and the thanks I want to give you cannot be put into words. Every time I needed some assistance you were there to provide me with some. You kept me focused when I needed to be, and provided much needed distraction when it was required also. I could not have completed the assistance without all the support you have given me over many years.

Contents

List of Figures

List of Tables

Listings

Chapter 1

Introduction

In the steel industry, the on-plant utilization and recycling of by-products is crucial to achieving environmental and economic sustainability [2]. For this thesis, I investigate the self heating of dust produced in the Basic Oxygen Steelmaking (BOS) process. The dust is a by-product of the BOS process; created by blowing oxygen at supersonic velocities on liquid steel [1]. This abundant material contains valuable elements, such as iron and calcium, making this an ideal target for recycling [3, 4]. Recycling the filter cake through the blast furnace proves challenging due to the accumulation of zinc [5]. The dust is collected and compacted into a filter cake which is then stockpiled. When the BOS filter cake is stored in large stockpiles prior to recycling, the filter cake can self-sinter [5, 1, 6]. The sintered material has better handling/transport properties, such as higher strength and larger particle size, than the original filter cake. These improved properties make the sintered filter cake more desirable for recycling. Understanding the self-sintering process allows for more opportunities to increase the recycling of the BOS filter cake.

The temperature distribution within these stockpiles is of interest. The self-sintering process that occurs inside them are still being investigated. One belief is that the self-sintering process is a result of the oxidation of iron [1]. There is uncertainty in the conditions that result in the stockpiles igniting; some of the stockpiles ignite and some do not. There is currently no method for predicting which stockpiles will ignite and which ones won't. Understanding these reactions and how they affect the overall temperature will assist in developing a greater understanding of the sintering process.

When the filter cake self-sinters inside these stockpiles, exothermic oxidation reactions drive the self-heating behaviour of the stockpiles. Different possible reaction schemes have been considered [7]. Each different scheme presents its own challenge and statistical inference techniques are required to ascertain the values of the reaction kinetics. The filter cake is a very complex material and there is considerable variation within the material. It is important to consider this variation if we want to make any predictions about the behaviour of this stockpile.

1.1 Aims and Objectives

The aim of this thesis is to develop a three dimensional model to aid in the understanding of the self sintering process in large stockpiles. To do this we look to:

- Formulate a model connecting the experimental data and the stockpile problem.
- Identify the kinetics of the reactions that are occurring inside the stockpiles.
- Obtain estimates of the kinetic parameters of the reactions from the experimental data.
- Examine the effect of external factors such as temperature and moisture on stockpile ignition.
- Understand how the construction of the stockpiles can affect their propensity to ignite.
- Incorporate different forms of heat transfer to investigate stockpiles at high temperatures.
- Simulate stockpiles that are constructed by BlueScope Steel using the estimated reaction kinetics.

1.2 Significance

The self sintering of BOS filter cake improves the recyclability of the material. The sintering process improves the structure of the material and as such increases its potential use [1]. This provides a large environmental benefit. The sintered material has higher strength, better transport properties, and larger particle size. This enables it to be recycled, as a coolant, reducing the amount of material ending in landfill [1]. Furthermore, this study highlights key financial benefits for the industry. If more can be done to understand this sintering process, then the material can be recycled at a lower cost. The iron in the filter cake is of reasonable value. If there is more control over the ignition of these stockpiles then this provides a way of controlling the form of the iron so that the value can be realised.

In these stockpiles we wish to promote the oxidation of iron. This is uncommon in the spontaneous combustion literature. Much of the work on has been to prevent the combustion, as such, the models required for understanding how to promote ignition, have not been highly developed.

Additionally the current statistical techniques used to estimate the reaction kinetics are very limited. Most of this focus is on obtaining estimates for the activation energy

required for the reaction and neglect the importance of the pre-exponential factor. As part of this thesis the implemented statistical algorithms are designed to infer both of these constants. The statistical algorithms we develop have greater capacity than the previous methods to account for the uncertainty in the parameter estimates [8]. This is particularly significant in the context of using a complex sample such as the Filter Cake produced as part of the BOS process.

Chapter 2

Theory and Literature Review

The discovery of the self-sintering process within large stockpiles is comparatively with the first published observations by Longbottom et al. [5]. The filter cake produced at the BlueScope Steel facility is incredibly complex with many different compounds found within the material [1, 5]. The filter cake is a complex material and multiple experiments have been conducted to determine the compostion and the reaction kinetics [6, 7]. Given the novelty of the self-sintering filter cake, models for this process have not yet been developed. This requires us to draw on the literature from other areas of spontaneous combustion to build appropriate models for the filter cake stockpiles. Due to the abundance of experimental data we have the ability to utilise these experiments to conduct our own inference for the kinetic parameters that are necessary for the stockpile.

2.1 Experimental work

The existing research regarding the filter cake is through characterisation experiments. Longbottom et al. [1] set out to characterise this material and identify the possible reactions involved in the self-sintering process. Through understanding the experimental work, we can use models to simulate the results and use the experimental data to fit these models. These experimental models can be used to develop more effective predictions in the larger stockpiles that we are interested in.

2.1.1 Characterisation of the Material

A key aspect of this thesis is the material used to build these stockpiles. The results from Longbottom et al. [1] determined the composition of the filter cake as well as the composition after some reactions had taken place. This characterisation process was conducted using X-ray Diffraction (XRD) and X-ray fluorescence (XRF). XRD and XRF are well established means of determining composition and have been used to characterise the compostion of a range of different materials [9, 10, 11, 12, 13, 14]. The important aspects

of composition for this thesis, is the mass percentage of metallic iron and Wüstite, 10.9% and 41.8% respectively [7].

2.1.2 Experimental Procedures

Due to the availability of experimental data it is important to develop models around these. Rather than relying on arbitrary values, we can utilise the experimental data to understand the reaction kinetics. This allows the models to more accurately reflect the stockpiles.

Thermogravemetric Analysis

Thermogravemetric analysis (TGA) measures the weight change of a sample. A small sample is heated, usually at a fixed rate, and any deviation in mass is recorded. The technique is useful for assessing reactions such as oxidation and vaporisation, which have been observed from the experimental data for the filter cake [1]. Reactions such as melting and crystallisation cannot be investigated as the mass of the sample does not change [15].

In a small scale TGA experiment a 100 mg sample is placed in a crucible that is attached to a balance. Air is fed through the base of the set-up, and the sample is heated at a fixed rate. As the sample undergoes reactions the weight change is measured.

This set-up can be changed to include larger sample sizes, though this affects the quality of the data. With larger samples, significant temperature gradients within the sample can occur; as a result, diffusion needs to be considered and heating at a fixed rate becomes more difficult. Since the sample is larger, it can be representative of the stockpiled material. Having a bigger sample also improves the repeatability of the experiment.

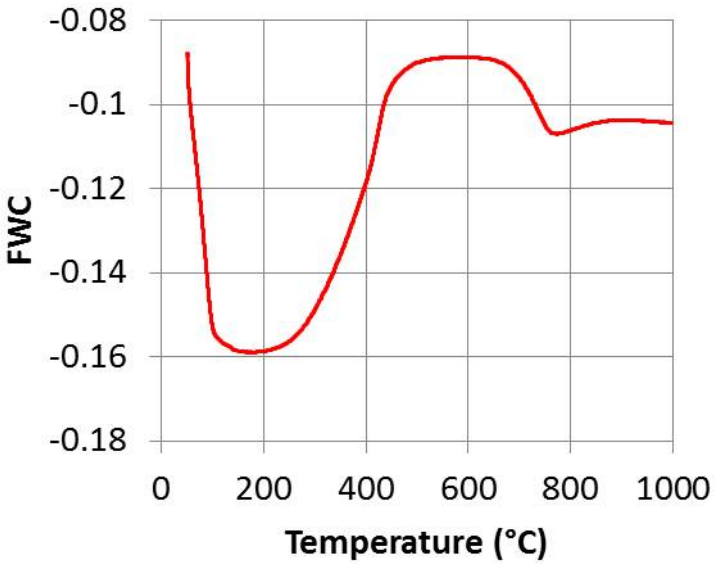


Figure 2.1: The fractional weight change (FWC) curve obtained from the TGA experiment obtained by Longbottom et al. [1].

TGA experiments have been used to characterise reaction kinetics in many different areas [16, 17, 18, 19]. One of the most significant applications is those that involve the reaction of various fuels such as coal [20, 21], biomass [22] and different fuel blends combining these two [23, 24, 25]. This demonstrates the prominence of TGA as a method to investigate reaction kinetics.

Differential Scanning Calorimetry

In addition to TGA, differential scanning calorimetry (DSC) is used to assist with the characterisation of the reactions. DSC methods have also been used to investigate fuel mixtures and are often conducted in conjuction woth TGA experiments [21, 26, 27] The experimental procedure of heating the sample is the same and as a result the two experiments can be conducted simultaneously. However, DSC measures the heat of the reactions occurring in the sample. This is done by measuring temperature differences [15].

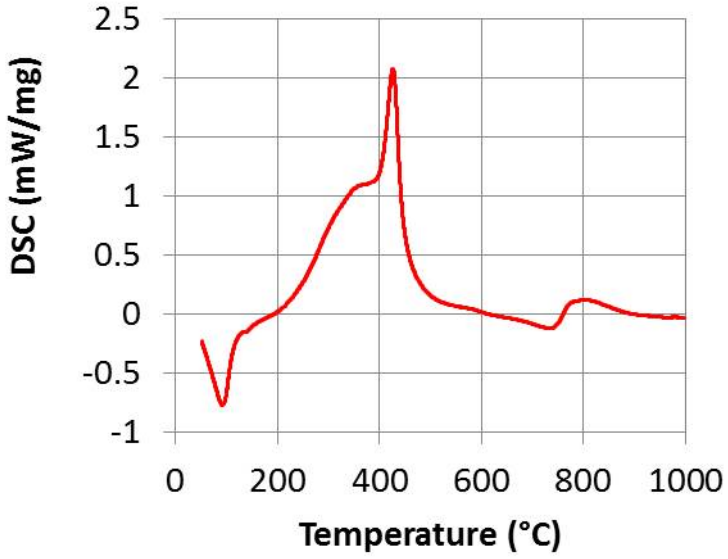


Figure 2.2: The DSC curve obtained through experimentation obtained by Longbottom et al. [1].

Through experiments Longbottom et al. [7] identified 5 stages of the self-sintering process. In the initial stage the material loses moisture resulting in a significant mass decrease. The second stage is the main stage where self-heating occurs. It has been identified [7], that there are two significant reactions occurring in this region where wüstite and iron both react with oxygen in the air and to produce hematite and magnetite. In the last three stages several reactions occur. Given that the majority of the self-heating is expected to come from this second stage we neglect these final stages of the self-sintering process.

2.2 Frank Kamentskii Theory

The fundamental theory of spontaneous combustion is known as Frank-Kamentskii (FK) theory [28]. FK theory has been used to investigate spontaneous ignition in a variety of applications, particularly in the context of safe storage [29, 30, 31, 32, 33, 34, 35]. The theory is based around temperature diffusion with an Arrhenius reaction term. The equation is,

$$\rho c \frac{\partial T}{\partial t} = k \frac{\partial^2 T}{\partial x^2} + Q \rho A \exp\left(\frac{-E}{RT}\right), \tag{2.1}$$

with domain $x \in [-L, L]$, and boundary condition $T = T_a$ It is useful to consider the steady state solution of this equation. We can rescale the variables,

$$\tau = \frac{\alpha}{L^2} t \qquad z = \frac{x}{L} \qquad u = \frac{E}{RT_a^2}(T - T_a)$$

then the steady state solution solves the differential equation,

$$\frac{d^2 u}{dz^2} = -\delta \exp\left(\frac{u}{1 + \varepsilon u}\right), \tag{2.2}$$

with boundary condition $u = 0$, and

$$\varepsilon = \frac{RT_a}{E} \qquad \delta = \frac{L^2 Q A}{k} \frac{E}{RT_a^2} \exp\left(\frac{-E}{RT_a}\right).$$

Using the Frank-Kamentskii approximation $\varepsilon = 0$, an analytic solution to the equation can be derived. This equation is the 1-dimensional analogue of Semenov's condition for ignition in a self-heating system [28]. The existence of this solution is dependant on the value for δ with two solutions existing for $\delta < \delta_{cr}$ [28]. A large amount of combustion modelling uses this critical condition as a means of developing an ignition criteria for the stockpiles [30, 31, 36, 37]

2.2.1 Theory applied to experimental work

The models used for thermogravemertic analysis have been reported in studies from Flynn and Wall [38] and Sharp [39]. In these studies the proposed model is in terms of conversion, given,

$$\frac{dC}{dt} = k f(C), \tag{2.3}$$

where k is reaction dependant. In this case the conversion, C, can be expressed as,

$$C = 1 - \left((W - W_f) / (W_0 - W_f)\right)$$

, where W is the weight, W_f is the final weight, and W_0 is the initial weight. This implies that the initial concentration, $C(0) = 0$ The expression for $f(C)$, is dependant on the reaction order. For a reaction of order n,

$$f(C) = (1-C)^n. \tag{2.4}$$

This has the effect that for higher concentrations the reaction is accelerated and slows once more of the reactant is consumed. The reaction rate k, varies with temperature per an Arrhenious expression,

$$k = A\exp\left(\frac{-E}{RT}\right), \tag{2.5}$$

where, A is a pre-exponential rate factor, E is the activation energy, and R is the ideal gas constant. Given that the Arrhenious reaction rate is temperature dependant, we use the equation,

$$\frac{dT}{dt} = \alpha, \tag{2.6}$$

for the constant heating rate.

It is not necessary to write the equations in terms of the conversion factor. Ozawa [40], uses the weight directly. The equation is,

$$-\frac{dW}{dt} = A\exp\left(\frac{-E}{RT}\right)W^n. \tag{2.7}$$

It is straightforward to show that these equations are equivalent. This requires equation (2.7), to be rescaled by the initial weight. In the two equations, the interpretation of the pre-exponential factor is different. This can be seen from dimensional analysis. For equation (2.7), $\frac{dW}{dt}$ has units MT^{-1}, where M is units of mass and T is units of time. As a result $A\exp(E/RT)W^n$ must also have units MT^{-1}. The exponential term is dimensionless, and W^n has dimension M^n. As a result the dimension of the pre-exponential factor is $M^{1-n}T^{-1}$. For equation (2.3), there is no mass dimension involved in the equation as we are dealing with the dimensionless conversion factor C. The only units involved are time and since the reaction rate k must have units T^{-1}, the pre-exponential factor, A, must have dimension T^{-1}. The units are only the same when the equation is first order, $n = 1$. In this case the reaction is invariant under scaling of W. The same equations have been used in other studies to evaluate the kinetic parameters A and E [41, 42, 43].

Kissinger [42] developed a method that can be used to determine the reaction constants, A and E. This method was to examine when the reaction rate is at its peak. The reaction rate for Equation (2.7) is at it's maximum when,

$$\frac{d}{dt}\left(\frac{dW}{dt}\right) = 0. \tag{2.8}$$

In the work carried out by Kissinger, it was assumed that the reaction order was one. This gives the equation,

$$A\exp\left(\frac{-E}{RT_m}\right) = \frac{E}{RT_m^2}\frac{dT}{dt}, \tag{2.9}$$

where T_m is the temperature at which the maximum reaction rate occurs. By taking logarithms straight line equation can be formed, for the variables, $X = 1/T_m$ and $Y = \log\left(\alpha/T_M^2\right)$ allowing a linear regression to be performed. If we do not assume the reaction order, then an additional term, the amount of converted material, is required. From this equation the reaction order and the kinetic parameters can be estimated. To estimate the parameters in this way, we must be able to determine the temperature at which the maximum reaction rate is achieved. With two reactions this is difficult to achieve and subsequently we need to implement a different method for determining these parameters. This equation is still useful in being able to relate our parameters.

2.2.2 Application to stockpile Ignition

Since the stockpiles of filter cake have not previously been modelled we draw upon other areas of stockpile modelling. Within different applications there can be slight differences into how the theory is applied but the models follow the same structure.

Coal based applications are some of the most well developed with initial models being proposed by Van Doornum in 1960 [44]. The initial model considers an ode model which has been extended into higher dimensions [29, 36, 45, 46]. The higher dimension models were based upon the FK equation (2.1). A similar approach was taken with the approach to compost models.

In the composting process the method of heating cannot simply be taken as an Arrhenious reaction. Instead a second heating function is required to model the microbial heating with Nelson et al. [32] proposing the following model,

$$\rho c_v V\frac{dT}{dt} = Q_b V F_b \frac{A_1\exp[-E_1/RT]}{1+A_2\exp[-E_2/RT]}B\left(1-\frac{B}{B_{\max}}\right) + Q_c V A_3\exp\left[\frac{-E_3}{RT}\right]C - \chi S(T-T_a). \tag{2.10}$$

This has been used as a basis for much of the following work involving compost models. This model has been extended to a two-dimensional model with temperature diffusion [47]. This latter model can be used to determine a critical ignition length for a stockpile.

One of the key components in the combustion literature is the effect of oxygen in the reactions. In both coal and compost applications oxygen plays a significant role in the reactions. Nelson et al. [48] first introduces oxygen in a spatially uniform model, before the model being later extended into two dimensions [49]. The coal literature has been more explicit with the inclusion of oxygen as this was included by Norden et al. [50]. In

the coal applications the addition of this reactant was simple whilst a new formulation for the effect of oxygen was needed within the compost literature.

When extending to multiple dimensions, an airflow is also added into the equations. In the one dimensional case, airflow can only be considered in one direction [45, 50, 51, 52]. When in two dimensions, an equation is required to express the airflow through the stockpile. This is done as a forced flow where the airflow is independent of temperature [53, 54]. For two dimensional stockpiles more natural forms of air flow can be introduced. Salinger et al [36] used Darcy's law to describe the natural flow of air through the coal stockpile and Aganetti et al. [55] used a Darcy-Brinkman equation for flow in a porous medium in a compost stockpile. The airflow inside these stockpiles have been studied further with different flows able to be described [56].

The effect of moisture on the reaction has also been added included in the model [57, 58, 59]. When modelling the moisture content, both a liquid water concentration and a water vapour concentration are needed. The concentration of liquid water affects the reaction rate; at higher water levels, there is an increased coverage of the reaction sites. The effects of ambient humidity, on the ignition criteria has been investigated [60].

Further additions to the model have been to consider different geometries [59] and also to use radiative boundary conditions [61, 62]. These developments add more complexities into the model but improves its accuracy.

The development of the models surrounding coal stockpiles have underpinned the field of spontaneous ignition and FK theory [63, 64]. Advancements in this area have been applied in the numerical simulation of Bagasse [65], and has had significant influence in understanding biomass [34, 66]. Futhermore this theory has been applied to lab-scale experiments [67, 68] where different aspects of biomass were investigated.

The adaptability of this theory into many different areas is a key reason why we use it to investigate the storage of the filter cake. One of the key differences between our approach and those that have been conducted previously is the intent behind the stockpiles. In the various different applications ignition results in loss of the reactant which is want to be preserved. However, for the filter cake, the self-heating behaviour of the stockpiles causes self-sintering which is a desired outcome.

Periodic Boundary conditions

The basic model can be extended by changing the boundary condition. One such method is to have an oscillating ambient temperature. The stockpiles are exposed to diurnal and seasonal temperature variation. Shteinberg and Khudyaev [69] proposed an ODE model,

given by,

$$\frac{d\theta}{d\tau} = a\mathrm{e}^{\theta} - \frac{1}{\delta}\left[\theta - \theta_A \sin\left(\frac{\omega\tau}{\delta}\right)\right], \tag{2.11}$$

$$\frac{da}{d\tau} = -\gamma a\mathrm{e}^{\theta}, \tag{2.12}$$

for the averaged temperature and reactant consumption across the domain. Gorel'skii et al. [70], analysed this model. These models do not consider the diffusion of temperature.

Novozhilov [71] proposed a model in one dimension with a single sinusoidal, oscillating boundary. Critical dependencies for the non-dimensionalised amplitude of the oscillations, were developed in terms of the other variables. The analysis was limited to the scaled equations rather than providing any direct application to any problems. Both convective heat transfer and fixed temperature boundary conditions were considered.

Novozhilov [72] added reactant consumption to the model, using a first order reaction scheme. The effects of having an initial condition that differs from the ambient temperature is also considered [72]. As more complexities are introduced into the model there are more parameters to investigate. The study determined critical dependencies of the scaled amplitude oscillations. Providing a critical condition for this scaled oscillation amplitude, does provide some insight into the problem. However, it does not address ways to prevent or promote ignition. This parameter does not include variables such as the stockpile length which can be controlled.

Roy [73] proposed a two dimensional model to investigate the effects of oscillating boundary conditions. The proposed model included an oscillating boundary condition along one of it's edges with fixed temperatures on the other edges. It included convection terms, though limited this to a closed system (no air flowing in or out of the domain). As such, the model is more applicable to an experimental set-up, rather than to assess the ignition in large stockpiles where the oscillations are expected on the majority of the boundary.

We see that there is comparatively limited work with oscillating boundary conditions. The current models proposed are limited in their dimensions and subsequently their application to large stockpiles. The models are restricted to one sinusoidal function, whereas a model including both seasonal and diurnal temperature variation may be more appropriate. It may also be useful to use include the weather data that has been recorded, as this will be more accurate than using a sinusoidal function.

2.2.3 Critical Ignition Criteria

The Frank-Kamanetskii theory states that for the case when $\delta < \delta_{\mathrm{cr}}$, there exists a low temperature solution and a high temperature solution [28, 74]. Various ignition criteria have been investigated. Weber et al [75] investigates the effects that different families of

initial conditions have including a constant, linear and quadratic initial temperature profiles. These initial investigations provided some useful critical ignition conditions though the model lacka any consumption of material. Brindley, Griffiths and McInstosh [76] address this by introducing consumption and also restricting their analysis to an embedded hotspot. They numerically examined some criteria required to iniate combustion waves. The existence and propogation of combustion waves have been well studied by various authors [77, 78, 79, 80, 81]. McIntosh, Brindley and Griffiths [82] build upon their work to produce an analytic approach to this problem. This extends previous work that has been done on strongly reactive material [83, 84]. The hotspots examined used a constant power source for their hotspot. For our industrial stockpiles we use a hotspot consisting of previously reacted material.

Shah et al [85] examines hotspots more consistent with the type we are interested in. This research examines the smoldering bahaviour where the reaction rate is controlled by a low oxygen concentration where the reactions are occurring. The key development is the inclusion of Oxygen and the porosity into the model. This theory has been applied to numerical and experimental studies into ignition by a heated particle [37, 86]. Similar experimental work has been carried out by Caine et al [87] who investigated a hotspot with constant power. These studies have developed the theory regarding the ignition of materials using hotspots. One area that appears to be lacking, is the use of previously reacted material.

2.2.4 Parallel Reactions

Much of the standard approach to Frank-Kamanteskii theory uses a single reaction. The work of Longbottom et al. [1] indicates that we may have several reactions occuring during the sintering process. Boddington et al. [88] developed the foundations of Frank-Kamenetskii theory with parallel reactions. They provide a method to calculate an effective activation energy and FK parameter. This then reduces the parallel reactions to a single reaction for which we can apply the existing theory. Graham-Eagle and Wake [89] extended this theory to determine the critical FK parameter in an infinite cylinder and sphere. Graham-Eagle and Wake [90] extended this further to examine the effects of introducing an endothermic reaction. All previous work had only considered exothermic reactions as these are the drivers of self-ignition. This theory has been used to simplify reaction equations in various different contexts and with differing numbers of reactions [91, 92, 93, 94].

Wake et al. [91] used this theory to examine heating in forest litter and coal. Pushpavanam and Narayanan [92] considered the effects of having an endothermic and exothermic reaction in parallel to investigate the ignition and extinction points of the steady state solution. Ajadi and Gol'dshtein [93] study the critical behaviour of a three step reaction

scheme using this technique.

2.2.5 Moisture

Moisture can be an important factor to consider. There have been instances in the literature [95, 96, 97] where wetting of dry material has been found to have caused ignition. Gray and Wake [98] model this ignition phenomenon using a series of parallel reactions. They considered the heat generated through the wetting process and also the cooling effect of evaporation. These two concepts are important to address in investigating our stockpiles.

Drying

Different drying models have been proposed [99]. This includes a reaction engineering approach that models evaporation as an Arrhenious expression with activation energy as the latent heat of vaporisation. Similarly condensation is modelled using a zero activation energy assumption. This form of drying was used by Luangwilai [100] when investigating how moisture is evaporated in compost piles.

Chen [99] proposed a model that looks at evaporation and condensation in terms of an exchange between two surfaces. The change in concentration is proportional to the difference in concentration at the inner surface and the surface area. Similarly, we have a change occurring that is proportional to the concentration difference on the outer surface. The advantage of this approach, is that it considers the concentration of water vapour within the sample to be distinct from its concentration in ambient air. The issue with this model is that it introduces additional parameters. One of the new terms introduced is the surface area of the small scale particles within the sample.

In the context of our material, this does not appear to be a feasible option, as determining the surface area is quite complicated. For simplicity, we therefore prefer to use the reaction engineering approach to drying. Due to the challenge in estimating the parameters, we were unable to include the moisture terms into our model. However, this approach remains an important consideration of future work into developing stockpile models.

Affect on the reaction

Characterising how the moisture affects the reaction is not a simple task. We can consider inhibition due to water covering the surface available for the reaction. Luangwilai [100] used this approach when modelling the combustion of organic material in compost heaps. It was assumed that if the water concentration is above a critical threshold, then the reaction stops. The function is normalised so that when there is no water then the reaction is

maximised and the effect of moisture is set to 1. Then the function decreases to 0 as the water concentration approaches some critical concentration.

More work needs to be done experimentally to characterise how the moisture levels affect the reaction. From the TGA data provided by Longbottom et al.[1], it appears as though the reactions are activated at temperatures above the evaporation point of water. This may indicate that the water would not have any effect as the reactions are not significant at those low temperatures.

2.2.6 Radiation effects on Conduction

The effect of radiation on stockpile ignition is not a common consideration. The modelled stockpiles usually are to prevent ignition [33]. Mercer and Weber [101] studied the effects of radiative transfer on combustion wave speeds. The radiation effects included was in terms of a radiative heat flux. The net radiative energy emitted and absorbed by the matter per unit time per unit volume is given by $\nabla \cdot \mathbf{q}^r$, where $\mathbf{q}^r$ is the radiative heat flux [102]. The Eddington approximation in one direction is given by the differential equation,

$$\frac{d^2 q^r(\tau)}{d\tau^2} = (1-\omega)\left[4\pi \frac{dI_b(\tau)}{d\tau} + 3q^r(\tau)\right], \tag{2.13}$$

where ω is the albedo, and $I_b(\tau)$ is the total black-body radiation intensity. We can relate the total black-body radiation intensity to the temperature,

$$I_b(x) = \frac{\sigma n^2 T(x)^4}{\pi}, \tag{2.14}$$

where σ is the Stefan-Boltzman constant and n is the refractive index. This approximation is for optically thick medium, as a result, this approximation is not as accurate near the boundaries [102]. In the literature [101, 103] the approximation has been altered to,

$$L^2 \frac{d^2 q^r(x)}{dx^2} = \left[4\pi L \frac{dI_b(x)}{dx} + 3q^r(x)\right], \tag{2.15}$$

where L is the local absorption length. When the local absorption length is small compared to the length of the medium, then a simplified expression can be used,

$$q^r(x) = -\frac{4\pi}{3} L \frac{dI_b(x)}{dx}. \tag{2.16}$$

This is the case of optically thick mediums [101].

When incorporating this into the model given by Equation (2.1) we have,

$$\rho c \frac{\partial T}{\partial t} = k\frac{\partial^2 T}{\partial x^2} + Q(T) + \frac{\partial q^r}{\partial x}. \tag{2.17}$$

This was investigated by Mercer and Weber [101]. Substituting in Equations (2.14) and 2.16 and rearranging we obtain,

$$\rho c \frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(\left(k - \frac{16}{3} L\sigma n^2 T^3\right)\frac{\partial T}{\partial x}\right) + Q(T). \tag{2.18}$$

This equation provides an effective radiative conductivity, $k_r =$ that satisfies

$$q^r(x) = -k_r \frac{dT}{dx}.$$

2.3 Parameter Estimation

TGA data is used to estimate reaction kinetics across many applications, for example wheat straw [16], rice hulls [17], and wood [18]. In particular it has been used to assess the kinetics of a modified hematite[104], which is of particular importance within the context of steel stockpiles. This highlights just how important TGA data is at obtaining information about the kinetic parameters. TGA data is not just limited to simple reactions. It has been used to understand a variety of different processes including pyrolysis and kinetics of microalgae [105], thermal stability and heat absorption for suppressant agent/combustible dust mixtures [106], and also evaporation processes [107]. This highlights the scope of kinetic analysis from TGA data. Given that obtaining information about the kinetic parameters from TGA data is so crucial to many applications of chemical engineering, it is imperative that reliable parameter estimates are made and the quantification of their uncertainty.

The traditional method that is utilised in these studies use either the Kissinger-Akahira-Sinuse (KAS) [42, 108] or the Ozawa-Flynn-Wall (OFW)[40, 109] method. These two methods are both iso-conversional methods that focus on the conversion of a particular reactant. The two methods use the model,

$$\frac{dC}{dt} = A\exp\left(\frac{-E}{RT}\right) f(C). \tag{2.19}$$

integrating this equation with heating rate $dT/dt = \alpha$,

$$g(C) = \int \frac{dC}{f(C)} = \frac{A}{\alpha}\int \exp\left(\frac{-E}{RT}\right) = \frac{A}{\alpha} I(E,T).$$

In this equation the function $g(C)$ generally cannot be computed directly nor can the integral $I(E,T)$. In both the KAS and OFW methods an approximation is made to the integral $I(E,T)$. These methods also consider the mass conversion at fixed at some C such for different heating rates and look at the temperature T_C at that value. The OFW

method uses an approximation,

$$I(E,T) = \frac{R}{E}\exp\left(-2.315\frac{0.4567E}{RT}\right).$$

This leads to the equation

$$\log(\alpha) = \log\left(\frac{AE}{g(C)R}\right) + 2.315\frac{0.4567E}{RT}. \tag{2.20}$$

describing a linear relationship between the heating rate α and the inverse of the Temperature $1/T$. Plotting the inverse of the temperature that the sample reaches some fixed conversion percentage for each reaction yields a straight line. Calculating the gradient of this line we can then obtain the activation energy.

The KAS follows a similar approach but uses the approximation,

$$I(E,T) = \frac{RT^2}{E}\exp\left(\frac{-E}{RT}\right).$$

This changes Equation (2.20) into,

$$\log\left(\frac{\alpha}{T_m}\right) = \log\left(\frac{AE}{Rg(C)}\right) - \frac{E}{R}\cdot\frac{1}{T}. \tag{2.21}$$

This equation yields a straight line relationship between the parameters $\log(\alpha/T^2)$ and $1/T$, for fixed conversion amounts. This method works well for estimating the activation energy but since $g(C)$ is unknown then we cannot determine the pre-exponential factor in this way.

Alternatively we could implement the Kissinger method using the equation,

$$A\exp\left(\frac{-E}{RT_m}\right) = \frac{E}{RT_m^2}\frac{dT}{dt}, \tag{2.22}$$

with different heating rates and corresponding maximum temperatures. The Kissinger method involves taking logarithms, we can rearrange Equation (2.22) to give

$$\log\left(\frac{\alpha}{T_m^2}\right) = \log\left(\frac{AE}{R}\right) - \frac{E}{R}\frac{1}{T_m}. \tag{2.23}$$

This is a straight line for the variables, $X = 1/T_m$ and $Y = \log\left(\alpha/T_m^2\right)$, allowing for a linear regression analysis. This is a simple method that can be used when the samples that are being experimented on are consistent and when the mass fraction is unknown. The Kissinger method also provides a means for determining the pre-exponential factor which is important for utilising these estimates in the stockpiles.

Teng et al. [17] implemented the Kissinger method for analysis of the kinetic param-

eters involved in mass loss of rice hull pyrolysis. In this paper the temperatures for the peaks of the reaction rate data were estimated and assumed to correspond to the respective reactions. This method produces estimates of the peaks that are sufficiently close when the focus is on the mass data. From our investigations into the effect this has on the DSC data, such an approach breaks down. The DSC data is more sensitive to the location of the peaks of each node and approximating these whilst ignoring the effects of the overlap in reactions causes discrepancies. Additionally in this study Teng et al. had freedom over the weight change values for the reactions whilst we consider these coefficients fixed by the choice of reaction scheme. Estimating the reaction peaks can prove particularly difficult when the reaction peaks are difficult to distinguish.

All of these methods can be quite successful when they are applied to single reactions. The iso-conversional methods rely on knowing the conversion factor of the main reactant in each reaction. In our sample we have two reactions and can only record the sample mass. These methods do not produce reliable estimates in the event of multiple reactions occurring with multiple reactants. This is highlighted by Liu et al. [110] where they investigate the combustion of agriculture briquette. In this study the TGA data indicates there are three stages where mass loss occurs. This is corroborated by the implementation of the KAS and OFW methods as the authors conclude that the activation energy changes significantly as a function of the conversion. This approach could explain a sample where two types of reactions are occurring with one reactant. We have two reactions occurring with different substances so this simple fix does not apply.

Another approach is to use the critical FK parameter, δ. The FK parameter is given by,

$$\delta = \left(\frac{L}{2}\right)^2 \frac{QCA}{k} \exp\left(\frac{-E}{RT_a}\right) \frac{E}{RT_a^2}. \tag{2.24}$$

Taking logarithms we can rearrange this to,

$$\log\left(\frac{4\delta T_a^2}{L^2}\right) = \log\left(\frac{QCA}{k} \cdot \frac{E}{R}\right) - \frac{E}{R}\frac{1}{T_a}. \tag{2.25}$$

The hot storage test heats a basket of material at a fixed ambient condition. By determining which ambient temperature causes the basket to ignite, for multiple size baskets, we can use a regression model to determine the parameters [28, 30, 111, 112]. Wang et al. used such methods to examine the critical self ignition of coal and applied these results to industrial stockpiles. Jones and Wake [113] used this method to determine the kinetic parameters for several materials. This method also has similar restrictions to KAS and OFW methods as it only considers one reaction. Jones [114] uses the work of Boddington et al. [88] on parallel reactions to modify this method for two reactions. This adjustment relies on knowing the heat of reactions. These parameters are unknown in our application. Additionally, Jones [115] highlights some of the difficulties in using these methods and

some of the limitations in respect to the safe storage for shipping.

These standard techniques are insufficient to estimate the kinetic parameters when we have multiple reactions. We cannot know what the temperature is which maximises each reaction rate. Another approach is to use an optimisation algorithm[18]. Elliott et al. [116] use a genetic algorithm as an optimisation tool for a system of parallel reactions. Zanoni et al. [117] used a Levenberg-Marquardt algorithm to determine the kinetic parameters in a reaction scheme. The objective functions in these cases are complex, exhibiting multiple local minima and maxima, as such gradient based optimisation algorithms are not as successful [116]. Reverte et al. [118] also used a Levenberg-Maraquardt algorithm. They used this to determine an optimal experimental design in order to minimise the error in the parameters estimated. They optimised the design for sample weight, temperature profile, and gas flow rate. The objective function minimised was a sum of squares of the measured data and the predicted data.

Using these techniques, quantifying the uncertainty in the model becomes more difficult. We use a Markov Chain Monte Carlo (MCMC) method to estimate our parameters. An outline of the method is provided in the Appendix A. The main advantage of using a MCMC algorithm is that we are able to draw samples from the distribution of possible parameter values. We can use this sample to estimate posterior expectations for different quantities including the critical length of the stockpiles. The optimisation based methods allow the variance around the optimum value to be determined, but does not naturally allow a sample to be obtained and parameters to be transformed.

The Bayesian approach has been used in recent studies of coal blends [119, 120]. These studies used Monte Carlo simulations in order to characterise the uncertainty in the predictor coefficients for a multiple non-linear regression analysis. This is a similar idea to our approach, however we assess the uncertainty of parameters in an ODE model.

Our Bayesian approach is designed to overcome the limitations of these previous methods. We combine a numerical solver for the mathematical model with a Markov Chain Monte Carlo (MCMC) algorithm [121, 122]. In contrast with optimization-based methods such as Levenberg–Marquardt, our MCMC algorithm explores the parameter space using a random walk. It is thus able to navigate regions of lower likelihood and is more robust to the presence of local maxima. If allowed to run for a sufficiently long time then the stationary distribution of this Markov chain produces random samples from the joint posterior distribution for the parameters. These posterior samples can then be used to calculate well-calibrated estimates and credible intervals [8].

The Frank–Kamanetskii theory has been well developed in the applications of stockpiles. This has been considered in a variety of applications from coal, compost and bagasse. The stockpile models have been extended to consider different environmental effects such as moisture. Hotspots have been established as a productive mechanism to

induce ignition in materials with exothermic reactions. An important element of the theory relates to how the kinetic parameters are estimated and the established techniques are well suited to single reaction schemes. When considering multiple reactions, then new methods are necessary in order to capture the model parameters and the uncertainty in these estimates.

Chapter 3

Definition of our Models

In this chapter we define each of the models that we use as well as some of the variations. Having a clear model is the first stage in being able to predict future behaviour. We begin by providing the base model that for the stockpiles. We then present some of the variations to the model that can be made. We do not investigate these until Chapter 6. After presenting the models for the stockpile we look at the models we use for modelling the TGA experiment.

3.1 Stockpile Model

The first model we examine is for large stockpiles. To model such stockpiles we use Frank-Kamenetskii theory (FK-theory). We consider the equation,

$$\rho c \frac{\partial T}{\partial t} = \nabla \cdot (k \nabla T) + Q(T), \tag{3.1}$$

where the parameters are defined in the nomenclature, and $Q(T)$ represents heat generation inside the stockpiles. Equation (3.1) provides the basis for which we build our models. Inside the stockpiles we have multiple reactions. Longbottem et al. [1] conducted experimental work on the material found inside the stockpiles and proposed various reaction schemes. We consider a sequence of two parallel reactions,

$$3\mathrm{Fe} + 2\mathrm{O}_2 \rightarrow \mathrm{Fe}_3\mathrm{O}_4,$$
$$6\mathrm{FeO} + \mathrm{O}_2 \rightarrow 2\mathrm{Fe}_3\mathrm{O}_4.$$

We can model these reactions through an Arrhenius reaction rate. The reaction rate is given by,

$$k(T) = A \exp\left(\frac{-E}{RT}\right). \tag{3.2}$$

Using the Arrhenius reaction rate, Equation (3.2), for two reactions, Equation (3.1) becomes,

$$\rho c \frac{\partial T}{\partial t} = \nabla \cdot (k \nabla T) + Q_1 A_1 \exp\left(\frac{-E_1}{RT}\right) + Q_2 A_2 \exp\left(\frac{-E_2}{RT}\right). \tag{3.3}$$

The parameter Q is the heat generated from the ith reaction. Equation (3.3) is a heat balance equation that describes the rate at which energy inside the stockpile changes. This formulation is valid for any geometry.

One key consideration in proposing a model is understanding the assumptions that are placed on the model. The assumptions simplify the model but we cannot make too many assumptions otherwise the information gained from the model is not useful. We have made the following assumptions:

- The filter cake is a homogeneous material.
- No material is consumed by the reactions.
- Heat is only transported via conduction.
- Oxygen is not depleted within the stockpile.
- The change in filter cake density is negligible.

These assumptions are necessary to simplify the problem into something that we can investigate. The filter cake is naturally a heterogeneous material. We are unable to quantify this heterogeneity and as such cannot factor this into the model.

Within the stockpiles oxygen, iron, and wüstite are all consumed as part of the reactions. In the combustion literature it is standard to assume that the material is not consumed [33]. The argument is that most models are interested in preventing ignition and the material depleted before ignition is often not significant. Oxygen is often added into models after the initial investigations. We eventually remove this assumption.

For outdoor stockpiles there are other mechanisms of heat transfer. Wind can transfer energy within the stockpile. Radiation is another form of energy transfer that is usually not considered. Advection is typically added into a model using simple models. These do not account for the variation in wind that occurs naturally. Radiation as a form of energy transfer is not usually considered as the effects are often small at low temperatures. Through the course of this thesis we aim to remove some of these assumptions in order to develop the model to provide more realistic predictions.

Oxygen

We add a mass balance equation for oxygen into the model. One factor to consider is the consumption of oxygen by the reaction. We modify the reaction rate equation 3.2, to

include oxygen as a factor. We assume a first order reaction for simplicity, though this may prove to be insufficient. The modified equation becomes.

$$k(T) = OA\exp\left(\frac{-E}{RT}\right). \tag{3.4}$$

This now includes oxygen as a controlling factor in the reaction rate. If the oxygen concentration is zero, then the reaction ceases. Using equation 3.4, the mass balance equation is,

$$\frac{\partial O}{\partial t} = D\nabla\cdot\nabla O - OA\exp\left(\frac{-E}{RT}\right). \tag{3.5}$$

This model is limited to include the diffusion of oxygen and does not include advection. Adding Equation 3.5 into the model removes the previous assumption we made that oxygen is not depleted within the stockpile. We still make assumptions about how oxygen moves throughout the stockpile and we also have to make assumptions as to how oxygen is exchanged at the boundary. These assumptions are more realistic than oxygen not being depleted and as such this addition is made to the model.

Material Consumption

Consumption of the material within the stockpile is usually not considered in models for spontaneous combustion. In many applications of combustion theory [33, 32, 35], the objective is to identify the conditions for self-ignition in order to prevent a loss of material. In these cases if the material has undergone a significant depletion then the objective has failed. The result is that this assumption is acceptable in these cases. Ignoring the depletion of material is a worst case scenario approach. Any conditions that are imposed on these models to limit self-heating will apply to the stockpiles with reactant depletion.

The stockpiles that we are interested in are used to recycle the filter cake. In this case it is desirable for the stockpiles to undergo self heating. It may also be useful to be able to model the stockpiles at high temperatures in order to predict where the material has sintered. In order to make the model useful at high temperatures we need to consider the consumption of reactants.

To include material consumption into our model we consider the effect it has on the reaction rate. We consider a first order Arrhenius reaction for simplicity. Equation (3.4) becomes,

$$k(T) = MOA\exp\left(\frac{-E}{RT}\right), \tag{3.6}$$

where M is the mass density of the reactant. When introducing the reactants into the model we have to make some consideration for the units. We go into detail about determining units in section 3.2.

The material within the stockpile is a solid, as such we do not consider any transport

within the stockpile. Similarly to when we introduced oxygen an assumption is removed; no depletion of reactant. New assumptions are made about how material is moved through the stockpile and the effect it has on the reaction. These assumptions are more reasonable than the material having no reactants consumed in the reaction.

3.1.1 Boundary Conditions

A partial differential equation (PDE) model requires boundary conditions. The boundary conditions can be adjusted in order to add more complexities into the model. The most basic boundary condition is the Dirichlet condition,

$$T = T_a. \tag{3.7}$$

It is often assumed that the ambient temperature is constant. This assumption does not have to be made and we remove this by investigating the effects of periodic boundary conditions. The Dirichlet boundary condition assumes that there is perfect heat transfer between the external boundary and the air. This can be overcome by introducing Newtonian cooling boundary conditions,

$$\frac{\partial T}{\partial n} = -h(T - T_a), \tag{3.8}$$

where $\frac{\partial T}{\partial n}$ is the outward unit normal derivative. Similarly to the Dirichlet condition, the ambient temperature is usually considered fixed. For the Newtonian cooling condition we have the limiting case, as $h \to \infty$, then the boundary condition approaches the Dirichlet condition, $T = T_a$. This condition indicates perfect heat transfer between the boundary and ambient air.

3.2 Thermogravimetric Analysis Model

As part of this thesis we require a model that simulates the TGA experiment. This model is used to estimate the reaction parameters used in Equation (3.2). The TGA experiment provides some data that can be used whereas the stockpiles alone cannot provide these estimates. We build this model using Equation (3.6). Equation (3.6) is the rate at which the reaction is occurring. By carefully selecting the appropriate units, we can express this as the rate at which a mass, M is consumed by the reaction. We also want to consider our choice of units in the context of the stockpiles. We do this to ensure consistency to allow for the reaction parameters to be used easily in the stockpile models. We model the TGA

experiment by the system of equations,

$$\frac{dM}{dt} = -MOA\exp\left(\frac{-E}{RT}\right), \tag{3.9}$$

$$\frac{dT}{dt} = \alpha. \tag{3.10}$$

This system represents the mass lost through a reaction as the temperature is increased at a constant rate. To use the system of Equations (3.9) and (3.10), we make the following assumptions:

- The temperature is constant throughout the sample and is increased at a fixed rate.
- Oxygen levels are constant for the duration of the experiment.

These assumptions are reasonable based on the experimental design. The constant temperature is justified due to the small size of the sample. The sample is 100 mg and the temperature diffusion within the sample is negligible. The experiment has a continuous inflow of air. We assume this inflow replaces oxygen consumed during the reaction. If we removed any assumptions then it adds more complexity into the model. We could measure the oxygen concentration at the outflow point in order to determine the validity of this assumption.

We still need to determine an appropriate set of units. From dimensional analysis we have,

$$\frac{[M]}{[t]} = [M]\,[OA]\,. \tag{3.11}$$

Since $[M]$ appears on both sides of the equation then its units does not matter for the TGA experiment. For the TGA model we use mass for the units of M. We choose to use mass since it is easy to convert the equation into a mass density equation for models that consider spatial dimensions. Since we can scale M by a constant then if we assume that the volume is kept constant and the mass doesn't change spatially then the equation remains true. When mass does change spatially then the change in mass density follows the mass change equation in Equations (3.9) and (3.10).

We now consider appropriate units for oxygen. In the context of the TGA experiment it does not matter as oxygen is kept constant. In the large stockpiles the concentration is the most natural quantity to consider. When we include consumption of iron, our reaction rate is expressed as the rate at which a mass of iron is consumed. In order to determine how much oxygen is consumed at the same time it is easiest to work in terms of a mass of oxygen rather than a concentration. We are still able to convert mass density into a concentration and this is done after the simulation is complete. The pre-exponential factor, A, subsequently has units, $[A] = \mathrm{m}^3/\mathrm{kg}/\mathrm{s}$. We may scale this constant to have different units of time depending upon the context.

In order to model the TGA experiment conducted by Longbottom et al. [1], we consider a two reaction system. We only consider simulating the region that is considered from $130-600^oC$. This is where the bulk of the self heating is expected to occur in the large stockpiles. Prior to this region, moisture is being removed from the sample. Although there is a possibility that more reactions occur in this region, we expect two reactions are sufficient to simulate the experiment. The TGA experiment measures the total mass of the sample. Our model equations are for the mass of the individual components being consumed. This means that we need an equation for the sample mass. The sample mass is given by,

$$\frac{dM_t}{dt} = w_{c,1}M_1A_1\exp\left(\frac{-E_1}{RT}\right) + w_{c,2}M_2A_2\exp\left(\frac{-E_2}{RT}\right), \tag{3.12}$$

where the subscripts 1,2 refers to the first and second reaction respectively, M_1, M_2 follow the system given by equation 3.9, and w_c refers to the change in sample mass per M consumed.

This now forms a complete system that can be used to simulate the TGA experiment. In the next chapter we use the TGA experiment to estimate the parameters in our model in order to translate this into the larger stockpiles.

Differential Scanning Calorimetry

During the TGA experiment, Differential Scanning Calorimetry (DSC) can be conducted. This provides additional data that we can use to estimate the kinetic parameters. In order to do this we need to simulate the DSC data. We can do this using the equation,

$$H = Q_1\frac{dM_1}{dt} + Q_2\frac{dM_2}{dt}. \tag{3.13}$$

This equation introduces a heat of reaction parameter Q, for each reaction. This is the same parameter that is included in our stockpile model. We can rewrite Equation (3.13) by substituting the Arrhenius reaction rates. This form allows a simple calculation to be made once we have solved the system of equations.

In this chapter we have established the different models we consider in this thesis. We have also stated the necessary assumptions that underpin each of the models. These assumptions all simplify the model down. We conduct analysis on each of these models and present different alterations to these. It is also important to consider these models together to ensure the parameters between the models are consistent in order to effectively model the stockpiles onsite at BlueScope Steel.

Chapter 4

Bayesian Inference

In this chapter we present Markov Chain Monte Carlo (MCMC) algorithms to estimate the Bayesian posterior distribution of the parameters in our Thermogravemetric Analysis (TGA) model. A more general introduction to MCMC algorithms is presented in Appendix A. We begin by investigating the effectiveness of this method on a simulated single reaction TGA. We then apply this method to a simulated experiment with two parallel reactions, before applying the algorithm to experimental data. At each stage we look for issues that arise with the algorithm and how to refine it for our application in thermogravemetric modelling.

The key advantage of the simulated data is that we know the underlying parameters and the model predicts exactly the behaviour of our system. For the experimental data we also face the potential issue of model miss-specification [123, 124]

4.1 Bayesian Framework

For each application we apply the same statistical model. Let $\mathbf{F}(\theta_M)$ denote the vector corresponding to the solution of the ODE, evaluated at times, $\mathbf{t}$. This solution is obtained numerically using a Runge-Kutta algorithm, hence we approximate $\mathbf{F}(\theta_M)$ by $\hat{\mathbf{F}}(\theta_M)$. The symbol θ_M denotes the parameters required for the differential equation. The simulated data, $\mathbf{y}$, is generated by adding Gaussian noise to the simulation data. The likelihood function is,

$$\mathbf{y} \sim N\left(\hat{\mathbf{F}}(\theta_M), \sigma^2\mathbf{I}\right). \tag{4.1}$$

To evaluate this likelihood we are required to solve the ODE. The posterior distribution follows the equation

$$P(\theta|\mathbf{y}) = \frac{P(\mathbf{y}|\theta)P(\theta)}{\int P(\mathbf{y}|\theta)P(\theta)d\theta}, \tag{4.2}$$

where: $\theta = (\theta_M, \sigma)$ represents our parameters, $P(\mathbf{y}|\theta)$ follows the likelihood function in Equation 4.1, and $P(\theta)$ is the prior for our parameters which is generally a uniform dis-

tribution. We have a normalising constant, $P(\mathbf{y}) = \int P(\mathbf{y}|\theta)P(\theta)d\theta$, sometimes referred to as the model evidence. Since the likelihood (4.1) requires the solution to the ODE, $P(\mathbf{y})$ is unavailable in closed form. Instead we use MCMC methods to simulate from the posterior $P(\theta|\mathbf{y})$.

Algorithm 1: Base Metropolis Hastings Algorithm used to sample the posterior.

Result: Write here the result

Initialise θ by sampling from the prior;

while *sample_size* $<$ *required sample* **do**

  Sample $\theta^* \sim Q(\theta^*|\theta)$;

  Set $\alpha = \frac{P(\mathbf{y}|\theta^*)Q(\theta_{t-1}|\theta^*)P(\theta^*)}{P(\mathbf{y}|\theta)Q(\theta^*|\theta_{t-1})P(\theta_{t-1})}$;

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

  if $\alpha > b$ **then**

    Set $\theta_t = \theta^*$

  else

    Set $\theta_t = \theta_{t-1}$

  end

end

The MCMC algorithm we use is a Metropolis-Hastings Algorithm outlined in Algorithm 1. Our algorithm proposes new parameters, θ^*, using a Gaussian random walk proposal, centered on the previous parameters. This proposal is symmetric, $Q(\theta|\theta^*) = Q(\theta^*|\theta)$. This has the advantage that it reduces the amount of computations that are required when determining the likelihood, α, in Algorithm 1. The typical chain length used in our analysis is 20000 draws per chain with the first 10000 draws were discarded as burn-in. In some instances 30000 draws were sampled with the first 10000 or 15000 discarded as burn-in dependent upon the chains convergence speed.

We have not yet defined the variance of the Gaussian proposal distribution. When selecting the variance it is important to consider what acceptance rate is acceptable. If the acceptance rate is too large, then the algorithm may not have explored enough of the parameter space. A low acceptance rate increases the number of iterations required to generate our sample. There are numerous ways to consider the optimal acceptance rate. Both Roberts et al. [125] and Yang et al. [126] indicate that an acceptance rate of 0.234 can be obtained under some conditions. Achieving a rate similar to this over the course of the algorithm is unlikely. At different stages of the algorithm we want different acceptance rates. In an ideal case, the early stages of the algorithm quickly locates the maximum likelihood point. To achieve this a large variance is considered. Once the algorithm is near the local maximum smaller steps are required to generate a sample from the posterior distribution. We vary the variance to allow greater step sizes to occur initially and smaller step sizes to occur later in the algorithm.

The initial values proposed in the algorithm are heavily dependant upon the starting

point. At this stage, known as the transition phase, the algorithm has yet to converge and is not approximating the posterior distribution. These values are considered the burn-in for the algorithm and must be discarded. Our algorithm is set up so that a fixed number of points are proposed, then we discard the first half of the chain as burn-in.

The algorithm is dependant upon the starting values. We simulate four Markov Chains, each with different starting values, with each chain following Algorithm 1. This is particularly important when the posterior distribution has multiple modes. This is a key advantage of using the MCMC approach as traditional optimisation based methods can become stuck in the sub-optimal modes.

It is important to examine whether the Markov Chains have converged to their stationary distribution which is our posterior distribution. Given that we have multiple chains we can use the Gelman Split-$\widehat{R}$ statistic [8]. This in conjunction with examining the trace plots of the Markov Chains allows us to determine whether the Markov Chains have converged. In addition we can calculate an effective sample size (ESS) to evaluate the quality of the sample. The Split-$\widehat{R}$ statistic for each parameter is determined using the within chain variance and the between chain variance, whilst the ESS uses an estimate for the auto-correlation within the sample. Samples that have high auto-correlation yield a low ESS, whilst samples that are completely independent will have ESS equal to the sample size. The Markov Chains form a dependant sample but by understanding the ESS parameter and with the option of thinning the sample we can produce a large enough independent sample.

To calculate the Split-$\widehat{R}$ statistic, let ψ denote the parameter to be estimated, with ψ_{ij} denote the ith observation from the jth chain. We compute the between chain variance B using,

$$B = \frac{n}{m-1} \sum_{j=1}^{m} \left(\bar{\psi}_{\cdot j} = \bar{\psi}_{\cdot\cdot} \right),$$

where $\bar{\psi}_{\cdot j}$ is the mean of the jth chain and $\bar{\psi}_{\cdot\cdot}$ is the mean across the sample. We then compute the within chain variance, W by,

$$W = \frac{1}{M} \sum_{j=1}^{m} s_j^2,$$

where s_j^2 is the sample variance of the jth chain. These are combined to estimate the marginal posterior variance, $\text{var}(\psi \mid y)$,

$$\widehat{\text{var}}^{+} (\psi \mid y) = \frac{n-1}{n} W + \frac{1}{n} B$$

which overestimates the variance, $\text{var}(\psi \mid y)$, of the marginal posterior distribution. The

$\hat{R}$ diagnostic is calculated as,

$$\widehat{R} = \sqrt{\frac{\widehat{\text{var}}^{+}(\psi \mid y)}{W}}.$$

4.2 Single Reaction

To develop the MCMC algorithm, we need to define our model and model parameters. We use the model

$$\frac{dM}{dt} = -MOA\exp\left(\frac{-E}{RT}\right), \tag{4.3}$$

$$\frac{dT}{dt} = \alpha, \tag{4.4}$$

derived in section 3.2. In this application we assume the oxygen concentration, O is constant. This constant is absorbed into the pre-exponential factor A.

We have physical parameters A and E. We also consider the parameter σ that represents the variance of the noise in our data. We set the true parameters to be $A_{\text{true}} = 3.17 \times 10^{6}\text{min}^{-1}$, $E_{\text{true}} = 8.616 \times 10^{4}\text{J/mol}$, and $\sigma_{\text{true}} = 0.05$. For our analysis we set the remaining parameters to correspond with the experimental conditions with the initial mass for one reaction $M(0) = 0.45\text{g}$. The specific values are listed in Appendix C. Since the parameters A and E can vary significantly, we explore the parameter space using the parameters, $\widetilde{E} = \log_{10} E$ and $\widetilde{A} = \log_{10} A$. This is a key transformation allowing the algorithm to explore many orders of magnitude much more easily than if we used the original variables. This can reduce the burn-in time and provide greater precision. Our vector of parameters, θ, in Algorithm 1 becomes $\theta = \left(\widetilde{E}, \widetilde{A}, \sigma\right)$.

We require a prior and proposal distribution for our parameters. In the absence of any additional information, we implement a weakly informative guassian distribution over the parameters $\widetilde{A}$ and $\widetilde{E}$. As we are using the parameter space of these variable in our algorithm this reduces some of the calculations. This choice of prior equates to a log-normal distribution over the variables A and E. As stated in at the beginning of this chapter, we use a Gaussian proposal distribution around the previous estimates which reduces the number of calculations required. For the noise parameter which has the constraint, $\sigma > 0$, we transform the variables to a logarithmic scale for the random walk. This prevents the algorithm from proposing values less than zero. Specification of these distributions are provided in Appendix C.2. Simulating the single reaction using the true parameters and adding some noise we obtain the mass curve in Figure 4.1. For the single reaction we limit our discussions to include only the mass of the reactant, rather than a sample mass. Converting to a sample mass is a simple manipulation of the same curve.

The random sample of parameter estimates obtained through the MCMC algorithm are displayed in Figures 4.2a and 4.2b. These histograms are indicative of the posterior

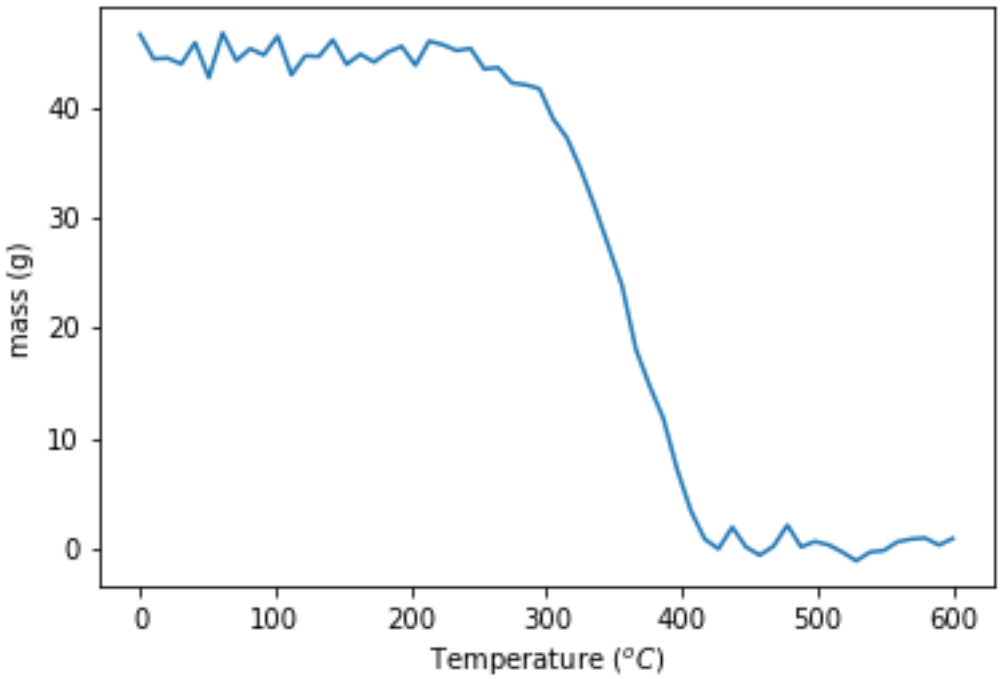


Figure 4.1: Simulated experimental data with noise. Parameter Values: $A_{\text{true}} = 3.17 \times 10^6 \text{min}^{-1}$, $E_{\text{true}} = 8.616 \times 10^4 \text{J/mol}$, $\alpha = 10$ and $\sigma_{\text{true}} = 0.05$.

distribution. We can use this data to analyse the posterior distribution to obtain point estimates and confidence intervals for the parameters. As mentioned in Appendix A, we can apply a functional transformation of our sampled data allowing us to utilise these estimates for $\widetilde{A}$ and $\widetilde{E}$ to determine the critical stockpile lengths required for stockpile ignition. This is a significant advantage of the Bayesian approach.

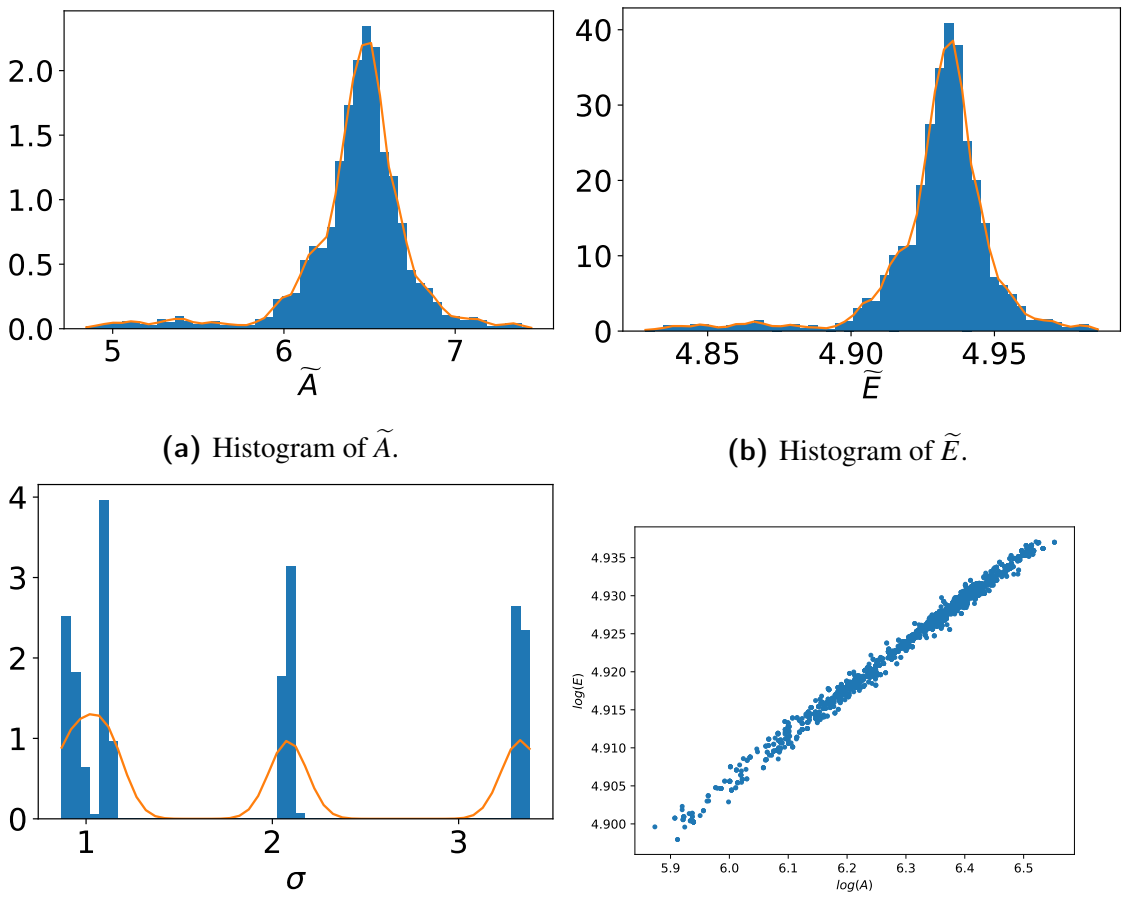
Figure 4.2 displays the sample density that our Metropolis-Hastings Algorithm generates. In each Histogram we capture the true value within the bounds of our sample. We can construct confidence intervals for the true parameters conditioned upon our samples. Given we have the true values we can calculate the percentile of each value and these are stated in Table 4.1. Various summary statistics can be found in Appendix C.3.

Parameter	True Value	Percentile
$\widetilde{E}$	4.94	89.7
$\widetilde{A}$	6.5	89.5
σ	0.05	0

Table 4.1: The percentile scores from the Metropolis-Hastings Algorithm.

Figure 4.2c indicates that the four chains have not converged and that no chain has estimated the true noise. Our diagnostic statistics, Split-$\widehat{R}$ and ESS, displayed in table 4.2, also indicate there are major issues within our algorithm. The ESS sizes are extremely low rendering such samples useless and the Split-$\widehat{R}$ is not close to one for the noise parameter.

The figure indicates that some chains take significantly longer to converge. This causes



(a) Histogram of $\widetilde{A}$.

(b) Histogram of $\widetilde{E}$.

(c) Histogram of σ.

(d) Joint distribution of $\widetilde{A}$ and $\widetilde{E}$.

Figure 4.2: Output from the Metropolis-Hastings algorithm for a simulated single reaction.

issues in determining the appropriate number of samples to accept before terminating the algorithm. This is also a reason why the sigma values are comparatively large. If the sampled points produce a FWC curve that is not sufficiently close to the data, then increasing the noise parameter increases the likelihood function. Additional figures for the trace plots of each parameter are included in Appendix C.3.

	Split-$\widehat{R}$	ESS
$\widetilde{A}$	1.036	32
$\widetilde{E}$	1.038	32
σ	66.136	4

Table 4.2: Diagnostic statistics for the single reaction MCMC.

Figure 4.2d plots the joint distribution of the parameters $\widetilde{E}$ and $\widetilde{A}$. This indicates that the two variables are highly correlated, with correlation coefficient 0.99. The fact that the variables are highly correlated causes inefficiencies in the MCMC algorithm. Given the significance of this curve in accepting new data points it is useful to design an algorithm

that samples along this curve. When we propose $\widetilde{A}$ and $\widetilde{E}$ independently, our proposed values are unlikely to remain on this curve.

Consider Equation 4.5,

$$A \exp\left(\frac{-E}{RT_m}\right) = \frac{E}{RT_m^2}\frac{dT}{dt}, \tag{4.5}$$

derived in section 3.2. This equation relates the pre-exponential factor, A, and activation energy, E, to the temperature at which the reaction rate is maximised, T_m. Figure 4.3 compares the points sampled by our original algorithm with the theoretical true curve with $T_{m,\text{true}} = 635$K. We observe that the Markov Chains follow this curve. It would be more beneficial to propose new parameters along these curves. However, we may not be able to determine T_m, to sample from this curve.

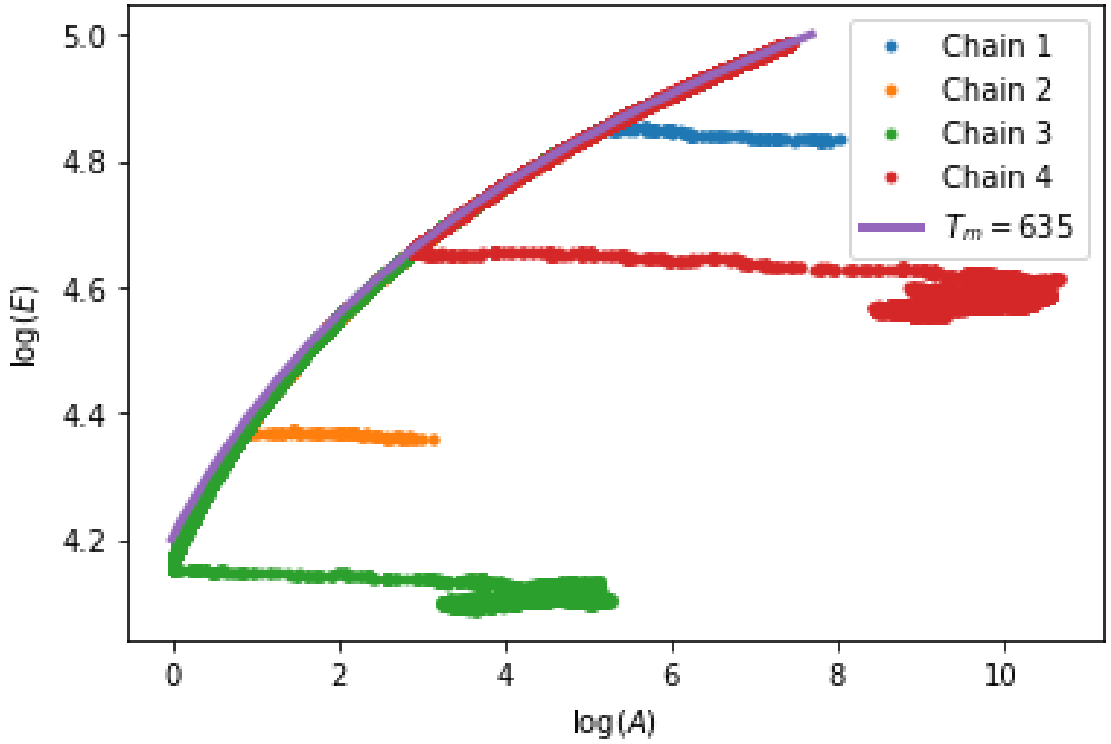


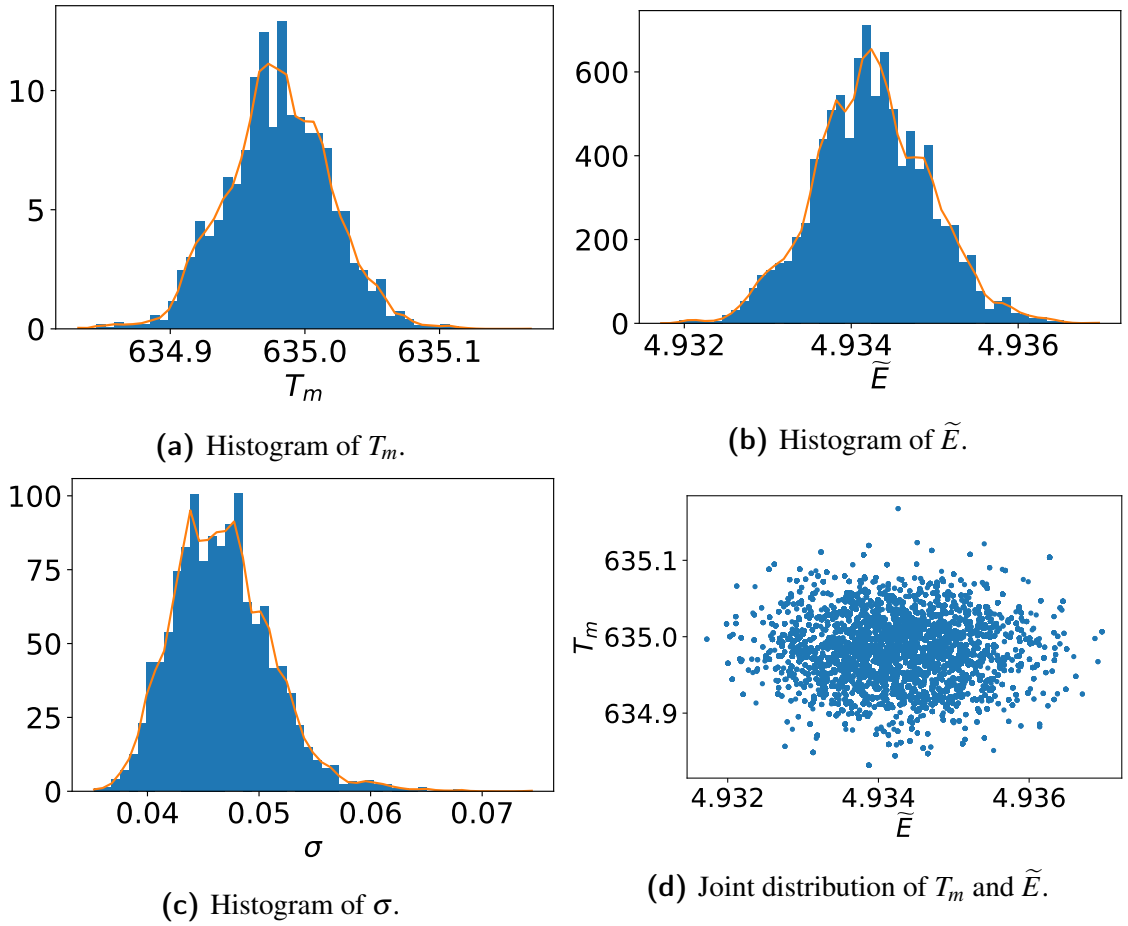
Figure 4.3: A comparison of the sampled points and the curve in equation 4.5.

Using Equation 4.5 alleviates another issue with our sampling method; we do not have good prior information on either the pre-exponential factor or the activation energy. This equation allows us to replace the parameter $\widetilde{A}$ from our sample space with the parameter T_m. We propose T_m and $\widetilde{E}$ values, and then using Equation 4.5 we determine the corresponding $\widetilde{A}$ value. In the case where we have one reaction, we can usually measure T_m directly from the experimental data. For this simulation we do not wish to use this explicit value as we intend to apply this to a scenario with two reactions where this parameter cannot be determined prior to conducting inference. Even without knowing the exact value, we are able to obtain a useful prior distribution for T_m. We use a uniform prior with a limited range. By changing the parameters our MCMC algortihm has converged as indicated

by the Split-$\widehat{R}$ values in Table 4.3

	Split-$\widehat{R}$	ESS
$\widetilde{E}$	1.007	1224
T_m	1.005	807
σ	1.002	977

Table 4.3: Diagnostic Parameters for proposals with T_m.



(a) Histogram of T_m.

(b) Histogram of $\widetilde{E}$.

(c) Histogram of σ.

(d) Joint distribution of T_m and $\widetilde{E}$.

Figure 4.4: Output from the Metropolis-Hastings algorithm for a single simulated reacion using the parameters (T_m, E).

Using the parameter T_m, we obtain very similar graphs for A and E. The plot for E against T_m is displayed in Figure 4.4d. This figure indicates that there is no significant correlation between the two parameters. This is useful in our algorithm as it improves the acceptance rate with similar step sizes. This is further supported by the trace plot in Figure 4.5.

One of the key metrics we are interested in is the effect that this variation has on the predicted critical length of the stockpile. This is a major benefit of the MCMC approach as we can conduct a functional transform from our sample to the critical length using the

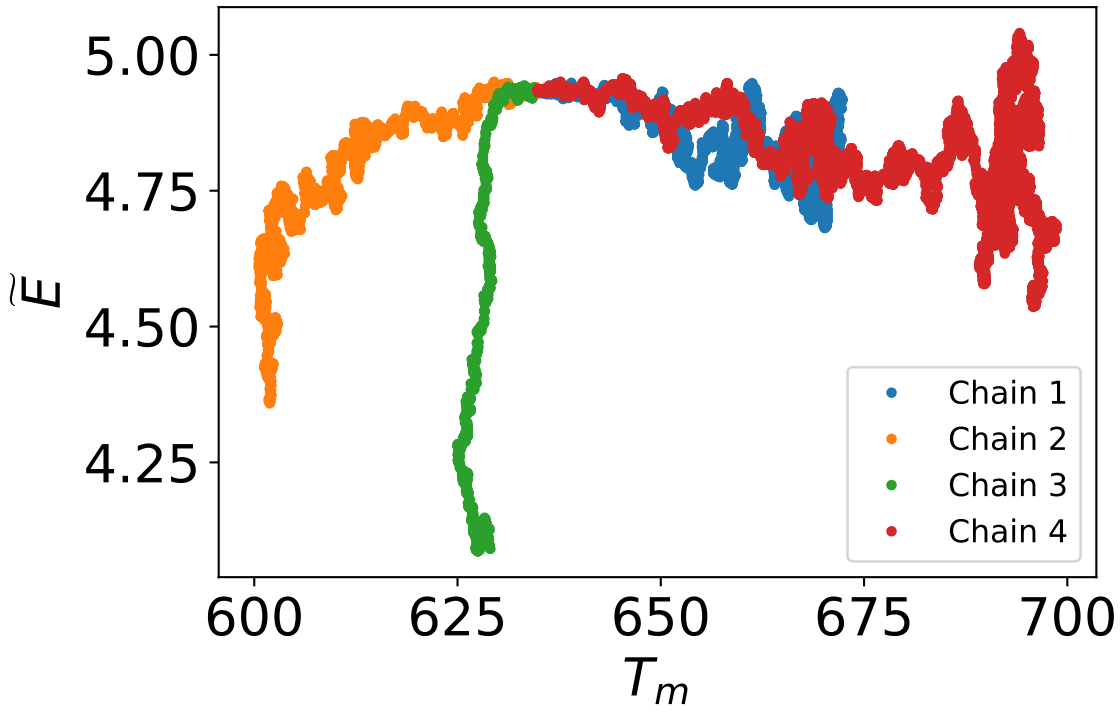


Figure 4.5: Trace plots of each of the Markov-Chain including burn-in of the parameters T_m and $\widetilde{E}$.

formula,

$$L_{cr} = K\sqrt{\frac{\exp\left(\frac{E}{RT_a^2}\right)}{A}\frac{RT_a^2}{E}}, \tag{4.6}$$

where K is some constant that depends upon various constants in our stockpile model. Equation (4.6) was derived in Section 2.2. We have taken this constant such that the minimum critical length is 1m. Figure 4.6 indicates that the critical length of the stockpile can be determined to within some reasonable range. The values in the histogram are indicative of the critical lengths and are dependant on the thermal conductivity and the heat of reaction; these parameters are not included in our TGA analysis. Using Equation 4.6 and taking logarithms we obtain,

$$\log_{10}(L_{cr}) = \log_{10}(K) + \log_{10}\left(\frac{RT_a^2}{EA}\right) + \frac{E}{RT_a^2}\log_{10}(e). \tag{4.7}$$

This form indicates that the parameter, K shifts the values of $\log_{10} L_{cr}$ and does not affect the spread.

4.2.1 Selecting Appropriate Parameters

This section has shown that utilising a different set of parameters can improve the efficiency of the MCMC algorithm. It is useful to consider why this is the case.

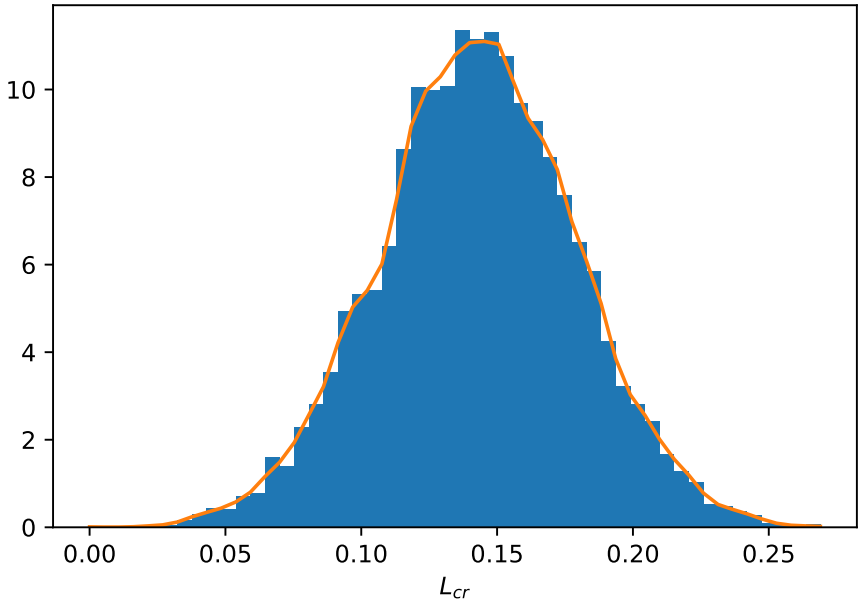
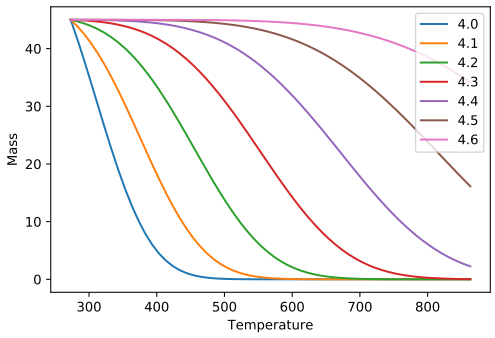
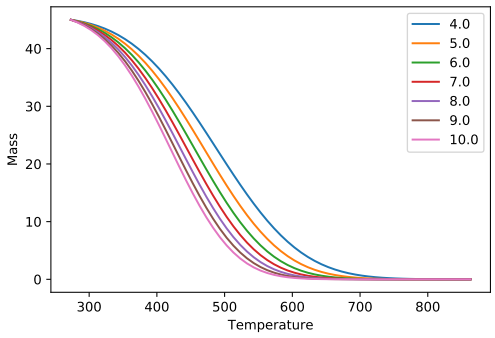


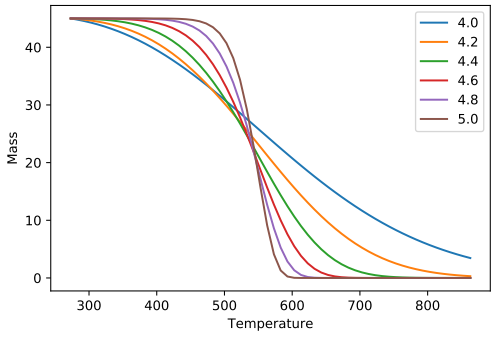
Figure 4.6: Histogram of the sample obtained for $\log_{10}(L_{\text{cr}})$ through the MCMC algorithm.



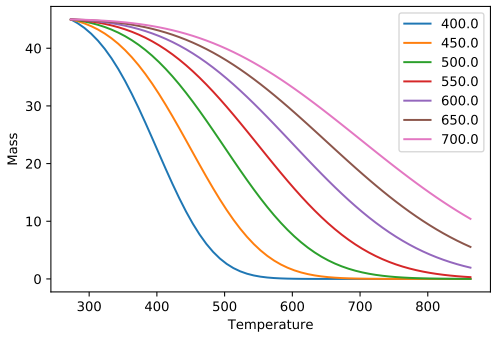
(a) The effect of $\widetilde{E}$ with $\widetilde{A} = 6$.



(b) The effect of $\widetilde{A}$ with $\widetilde{E} = 4.2$.



(c) The effect of $\widetilde{E}$ with $T_m = 550K$.



(d) The effect of T_m, in units Kelvin, with $\widetilde{E} = 4.2$.

Figure 4.7: Comparison of the effects of each parameter pairing on the FWC curve.

Figure 4.7 demonstrates the effect that altering each parameter has on the FWC curve. We observe in Figures 4.7a and 4.7b, that the $\widetilde{E}$ parameter has a large effect on both when the reaction occurs and its duration; small changes in $\widetilde{E}$ cause a large shift in the FWC curve. In contrast the affect of changing $\widetilde{A}$ appears to be minimal. Dramatically changing the order of magnitude of the parameter has little effect on FWC curve. This is in contrast to the $\left(\widetilde{E}, T_m\right)$ pairing displayed in figures 4.7c and 4.7d. In this set up, T_m is the sole indicator of the reaction's location. As T_m is increased, the duration of the reaction increases. The key difference is that as we vary $\widetilde{E}$ the location of the reaction remains the same; only the duration of the reaction changes. It is important to note the the temperature ranges from 400 to 700 K, which is far greater than any prior information will deem necessary. This is a key observation in increasing the efficiency as we are able to start the algorithm with a useful prior for T_m and small changes in T_m do not cause large changes in the duration of the reaction. This enables the parameter $\widetilde{E}$ to control the duration whilst T_m refines the location. Figure 4.7 shows why the parameter pairing $\left(\widetilde{E}, T_m\right)$ is a better basis than the parameter pairing $\left(\widetilde{E}, \widetilde{A}\right)$.

4.2.2 Application to experimental data

We now apply this algorithm to an experimental data set where only a single reaction occurs. This simple scenario serves to test the algorithm on experimental data where the underlying parameters are unknown. We model our reactant using the equation,

$$\frac{dM}{dt} = -MOA \exp\left(\frac{-E}{RT}\right),$$
$$\frac{dT}{dt} = \alpha.$$

This models the mass of the reactant, whereas the experimental data is in terms of a mass percentage. Since our model is invariant for mass we can scale this by the initial mass, without affecting any of the other parameters. We determine the mass percentage using the equation,

$$SM = M + \frac{M_f}{M_0}\left(M_0 - M\right), \tag{4.8}$$

where M_f, is the final mass, M_0, is the initial mass, and SM is the sample mass as a percentage of the initial mass. This equation is a sum of the mass of reactant, M, and the product mass $\left(M_f/M_0\right)\left(M_0 - M\right)$. This approach assumes that sample is pure and the reaction is complete at the end of the experiment; this is valid with our data. The fractional weight change data is presented in Figure 4.8. For our analysis, we restrict the data to the window where the temperature is between 250^oC and 1250^oC. This eliminates some complications, such as the early stages of the experiment where the sample is not heated at a constant rate, and removes the back end of the reaction where the mass change

is negligible.

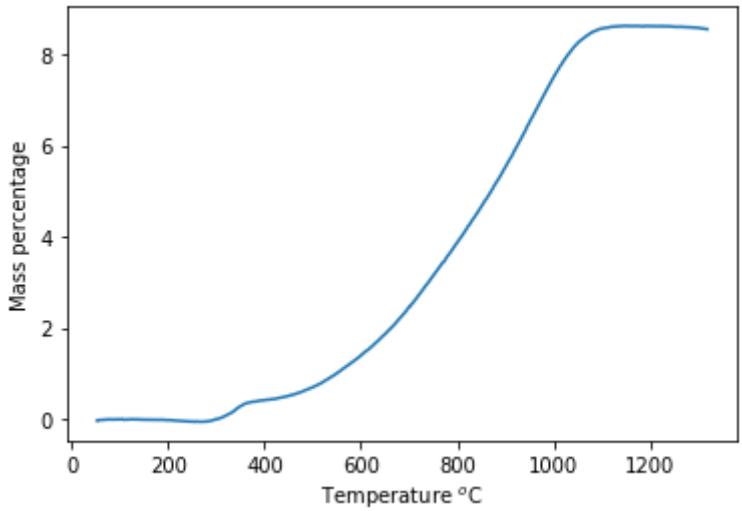


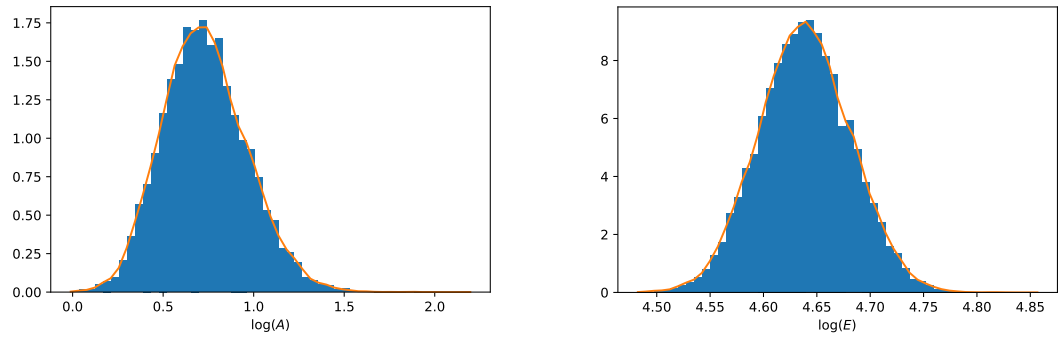
Figure 4.8: The FWC curve for a simple experiment where one reaction is occurring.

We introduce a variable proposal distribution into our algorithm. The prior information we have in determining the temperature where the reaction rate is maximised, is less precise than in our simulated examples. It requires a larger step size initially to explore this space, though we still require a narrower proposal once the chains have converged to efficiently sample the posterior. To implement this our proposal is $N\left(\theta_{t-1}, s/\sqrt{d}\right)$, where, s is our fixed proposal standard deviation, and d is our control parameter. We simulate for 20000 iterations total and after we reach 10000 iterations we increase the parameter d, narrowing the proposal distribution. This method requires manual tuning and a implementation of an adaptive method such as Adaptive Metropolis-Hastings [127] or Delayed-rejection adaptive Metropolis (DRAM) algorithm [128] would reduce the time required to tune the algorithm. These methods adaptively change the proposal algorithm improving overall convergence.

We check whether the Markov Chains have converged using the diagnostic statistics in Table 4.2. We observe that the Split-$\widehat{R}$ statistic is close to one, providing strong evidence that the chains have mixed and converged to the stationary distribution. The ESS is relatively large indicating that there is not too much auto-correlation occurring within the data. These statistics indicate that our algorithm has provided a sufficient sample of the posterior distribution.

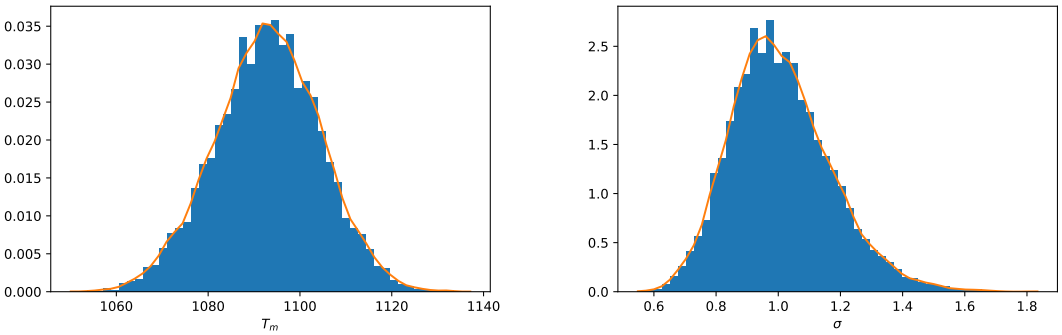
	Split-$\widehat{R}$	ESS
$\widetilde{E}_1$	1.0016	2249
T_m	1.0023	1662
σ	1.0006	3479

Table 4.4: Diagnostic statistics for the experimental data modelled with one reaction.



(a) Histogram of $\widetilde{A}$.

(b) Histogram of $\widetilde{E}$.



(c) Histogram of T_m.

(d) Histogram of σ.

Figure 4.9: Histograms of sampled parameters of the posterior distribution for the experimental data.

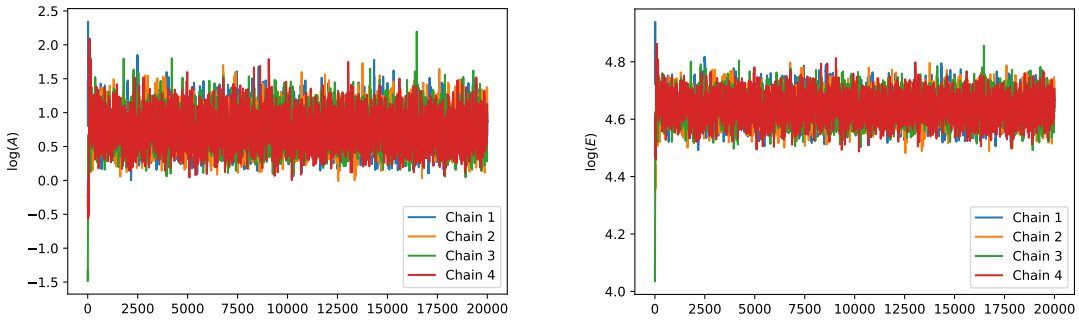
Figure 4.9 displays the samples from the posterior distribution. This figure indicates that we have a large uncertainty in our parameters. Additionally, we can consider the summary statistics for these parameters in Table 4.5 to describe the distribution.

	Mean	Minimum	Median	Maximum	90% CI
$\widetilde{A}$	0.73	-0.01	0.72	2.19	(0.36, 1.12)
$\widetilde{E}$	4.63	4.48	4.63	4.85	(4.56, 4.71)
σ	1.00	0.54	0.99	1.83	(0.77, 1.30)
T_m	1092	1050	1093	1137	(1073, 1111)

Table 4.5: Summary statistics for the experimental data.

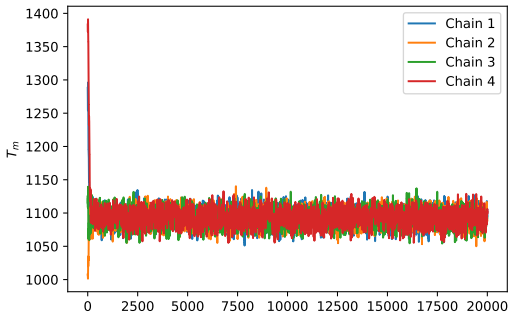
When examining the trace plots in Figure 4.10, the algorithm behaves as we expect. These trace plots include the first 10000 points sampled which are discarded as burn-in. We also observe that the variance in the noise σ, is much higher than in our simulated examples and is higher than we expect for the experimental data. This suggests that there may be an issue with model miss-specification.

To examine the hypothesis that the model does not accurately represent the data, we plot some of the sampled curves. We randomly select 10 data points, simulate the experiment,

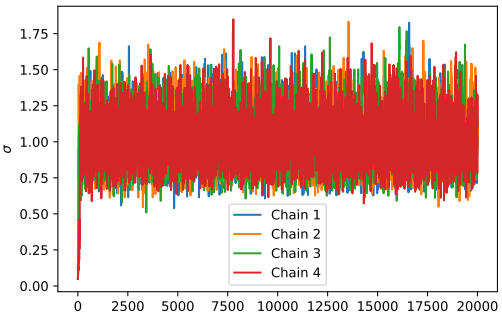


(a) Trace plot of $\widetilde{A}$.

(b) Trace plot of $\widetilde{E}$.



(c) Trace plot of T_m.



(d) Trace plot of σ.

Figure 4.10: Trace plots for each of the parameters of our MCMC algorithm applied to experimental data.

and plot the results against the FWC data. We observe in Figure 4.11 that the FWC curves do not completely agree with the data. It is worth noting that it is difficult to determine which curve best fits the data visually. This results is expected in our algorithm. The data has an increase in mass early in the reaction that does not appear to be captured by our model and this is likely causing the issues in our estimates.

Persisting with this model we are able to sample from the posterior of the critical length of a stockpile built with this material. Figure 4.12 indicates that the critical length varies over many orders of magnitude. This indicates that using any of the acceptable curves in Figure 4.11 in isolation provides a poor estimate of the critical length needed for ignition. This highlights the need for consideration of the parameter uncertainty in assessing models. This analysis indicates that if we wish to approximate the reaction using this model, then we have a large uncertainty within our parameters which causes a large uncertainty in the critical length of the stockpile. To avoid this issue, we need either additional data or an alternative model that better predicts the data. This example also indicates the dangers we can face if we use a simple one reaction model as an approximation for a more complicated reaction scheme.

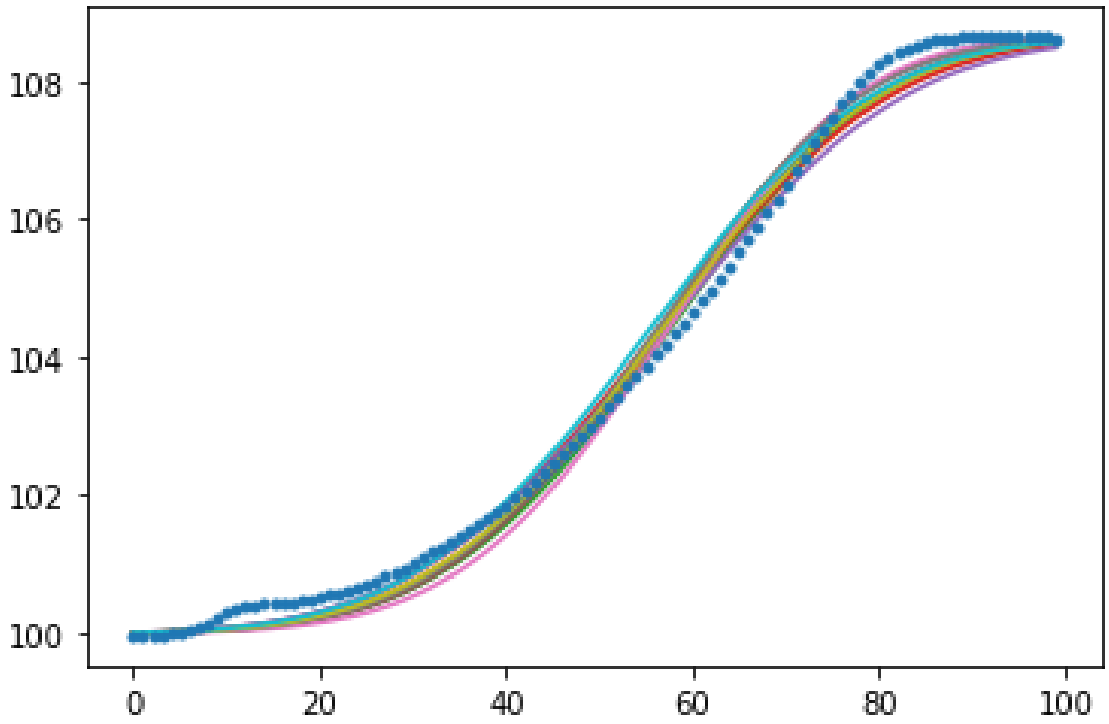


Figure 4.11: A random sample of FWC curves against the experimental data.

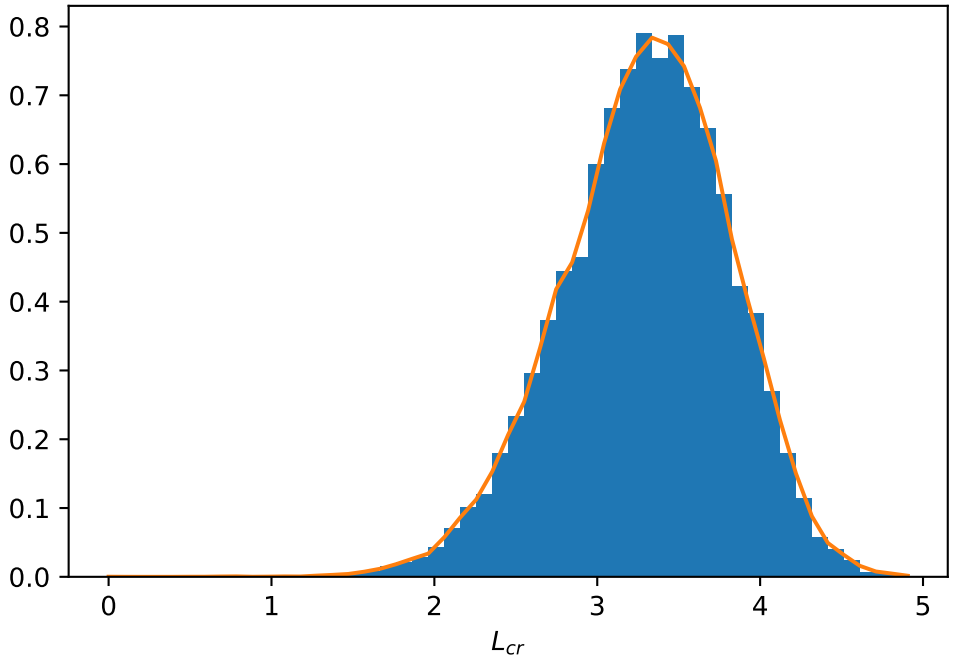


Figure 4.12: Histogram of the logarithm of the critical length of the stockpile made from the material in the experiment.

4.3 Two Reactions

We extend our algorithm from a single reaction to two parallel reactions. For the identifiability of the parameters we need to ensure the reactions have a defined order. Experimentally we can determine which of the two reactions is occurring first, the simulations need to reflect this.

To model the experiment we use equation (3.12),

$$\frac{dM_t}{dt} = w_{c,1} M_1 A_1 \exp\left(\frac{-E_1}{RT}\right) + w_{c,2} M_2 A_2 \exp\left(\frac{-E_2}{RT}\right). \tag{4.9}$$

Using equation (4.3) for $M_1$and M_2, we integrate Equation (4.9), simplifying to the equation,

$$M_t = M_{t_0} + w_{c,1}\left(M_1 - M_{1_0}\right)) + w_{c,2}\left(M_2 - M_{2_0}\right)), \tag{4.10}$$

where the 0 subscript denotes the initial mass. The masses, M_1 and M_2 follow Equation 4.3. For the statistical model, Equation 3.12, is used to determine the theoretical solution, which we discretise at specified times, $\mathbf{t}$. We then add Gaussian noise to our simulated solution. As our results for the one reaction case indicate that it is better to estimate T_m. Hence we explore the parameter space of $(\widetilde{E}, T_m)$, for the two reaction model. We initially consider a scheme with well separated reactions. The true parameters for our parameter space are defined in table 4.6.

Parameter	Value
E_1	8.6×10^4 J/mol
E_2	2.6×10^5 J/mol
$T_{m,1}$	560 K
$T_{m,2}$	750 K
A_1	3.5×10^7 m^3/kg/s
A_2	7.1×10^{17} m^3/kg/s
σ	0.05

Table 4.6: Parameter values for the simulation.

For the two reaction model we adjust our algorithm. As we introduce more complex behaviour into our model efficiency becomes more of an issue. We observed in the one reaction model that the algorithms can take a significant time to converge. This is computationally inefficient. We also still want our algorithm to have an acceptance rate of approximately 0.234. As such we divide our algorithm into three stages. Initially, our proposal distribution is a Gaussian distribution with high variance. In this stage we anticipate the acceptance rate can vary dramatically, but once the optimal node is found, we will have a low acceptance rate. The intermediate stage has a more narrow Gaussian proposal distribution where the algorithm refines this node further. In these stages we are not overly concerned about the acceptance rates, and these are discarded as burn-in.

The final stage is where we obtain our sample. In this region the algorithm is expected to have converged, so the proposal distribution can be altered to achieve an acceptance rate around the optimal value.

With two reactions occurring it is likely that a sub-optimal mode within the posterior distribution exists where the order of the reactions is reversed. In practical applications the reaction order can often be determined through additional experimentation. To mimic this we add a constraint into our prior information that $T_{m,1} < T_{m,2}$. The simplest way to introduce this is to use weakly-informative uniform priors for each of the parameters such that the intervals the parameters are distributed over are disjoint. This eliminates any suboptimal mode that may have occurred with the reaction order changed.

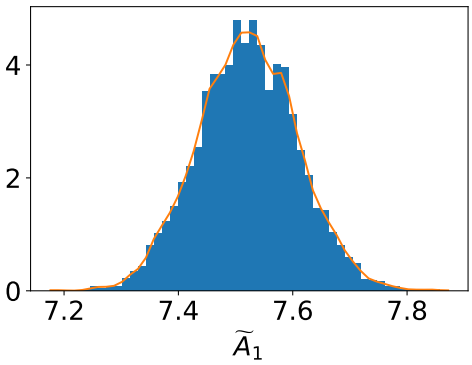
Using Algorithm 1, we generate the posterior samples in Figure 4.13. Checking the diagnostic statistics in Table 4.7, the chains have converged and the ESS of each parameter is reasonable. We find that our posterior distribution contains each of the true values, indicating that the posterior distribution we defined can be used within the TGA context and extended into multiple dimensions.

	Split-$\widehat{R}$	ESS
$\widetilde{E}_1$	1.0028	1442
$T_{m,1}$	1.0017	2267
$\widetilde{E}_2$	1.0019	1634
$T_{m,2}$	1.0025	1650
σ	1.0016	3024

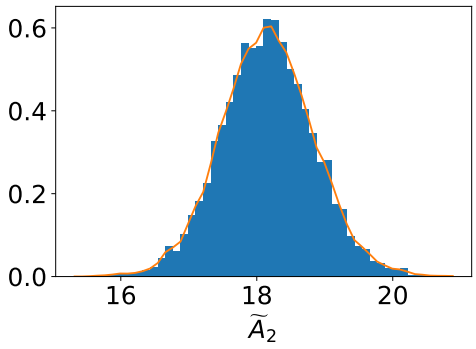
Table 4.7: Diagnostic Statistics for the 2 reaction model.

We observed in the single reaction case that there was significant correlation between the parameters A and E, so it would be beneficial to determine if adding a second reaction introduces additional correlations. Table 4.8 indicates the only significant correlation between any of the parameters in our example are the two kinetic parameters of the second reaction. Having minimal correlations increasing the efficiency of the random walk MH algorithm as the auto-correlation is reduced. In the case the two reactions are well separated, with one reaction going to complete before the other begins, this is ideal for using this MH algorithm.

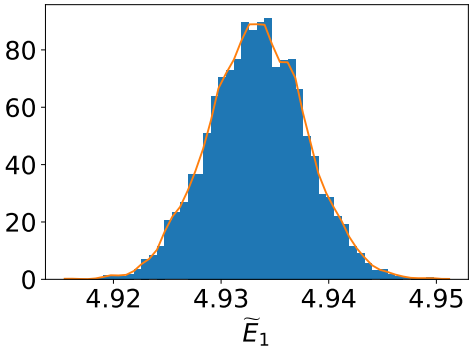
For well separated reactions, the algorithm has no issues sampling from the posterior distribution. This presents an ideal scenario and a necessary step to show the algorithm is sufficient. In our experimental data, the reactions are not as well separated. We therefore simulate some new data using altered parameters. We change the parameters $T_{m,1}$ and $T_{m,2}$ so that, $T_{m,1} = 620$K and $T_{m,2} = 700$K. Given these values are much closer, it is harder to distinguish the two peaks in the simulated data. The non-overlapping uniform prior is much more difficult to implement in this scenario since the we do not have a clear division between the two values. Since these two parameters are linked through our constraint, we



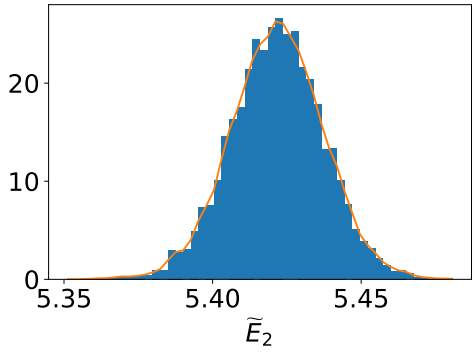
(a) Histogram of $\widetilde{A_1}$.



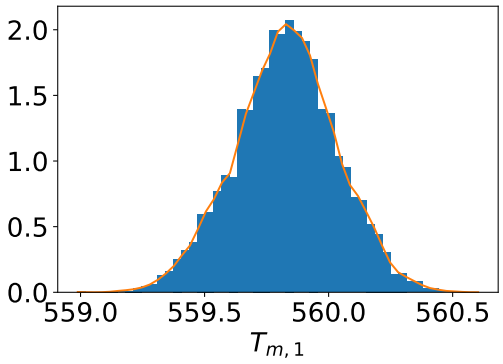
(b) Histogram of $\widetilde{A_2}$.



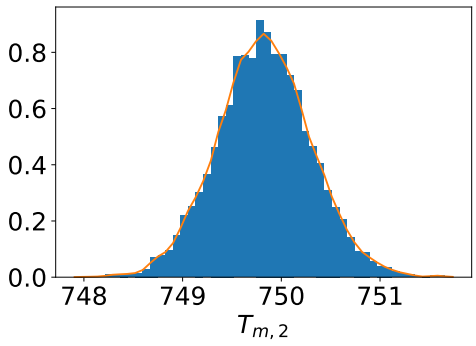
(c) Histogram of $\widetilde{E_1}$.



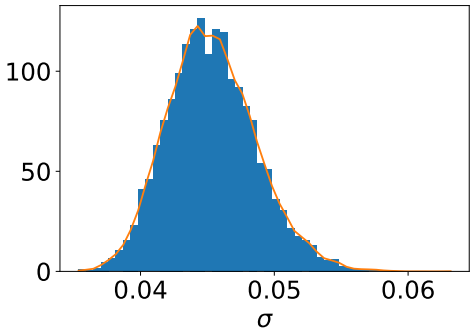
(d) Histogram of $\widetilde{E_2}$.



(e) Histogram of $T_{m,1}$.



(f) Histogram of $T_{m,2}$.



(g) Histogram of σ.

Figure 4.13: Histograms of the posterior sample generated from the Metropolis-Hastings for a simulated experiment with two reactions.

	$\widetilde{E}_1$	$T_{m,1}$	$\widetilde{E}_2$	$T_{m,2}$	σ
$\widetilde{E}_1$	1.00	-0.00	0.03	-0.00	0.00
$T_{m,1}$	-0.00	1.00	0.01	-0.02	-0.00
$\widetilde{E}_2$	0.03	0.01	1.00	-0.19	-0.00
$T_{m,2}$	-0.00	-0.02	-0.19	1.00	-0.03
σ	0.00	-0.00	-0.00	-0.03	1.00

Table 4.8: Correlations between the parameters in the two reaction model.

cannot consider an independent prior distribution for these parameters Instead we consider a joint uniform prior over the region, $550 < T_{m,1} < T_{m,2} < 750$. This prior enforces our identifiability constraint, $T_{m,1} < T_{m,2}$, and bounds each value. It is possible to have other priors that fit these constraints, we use this prior for its simplicity.

	Split-$\widehat{R}$	ESS
$\widetilde{E}_1$	1.003	577
$T_{m,1}$	1.001	1163
$\widetilde{E}_2$	1.014	148
$T_{m,2}$	1.002	1782
σ	1.003	865

Table 4.9: Diagnostic statistics for the two reaction MH algorithm.

Using the results presented in Table 4.9, we can reasonably believe the Markov Chains have converged. An interesting point of note in these statistics is that the ESS is much lower than what we would expect in this case. This suggests that there is significant auto-correlation within each Markov Chain. We attribute this to the correlation between some of the parameters that we observe displayed in Table 4.10. As we observed with our algorithm for the single reaction before we changed the parameters, highly correlated parameter values made the random walk algorithm inefficient. These correlations are interesting as they did not appear in the well separated reactions. As the reactions are closer together then it is unsurprising that correlations appear. With the two activation energies controlling the speed of the reaction, it makes sense that if the first reaction happens quickly, it cannot be used to explain any of the change in mass in the later stage of the experiment making the second activation energy larger causing a quicker reaction.

Conducting the inference we obtain the samples displayed in Figure 4.14. In these samples we correctly obtain the true parameter values for each parameter. We are able to successfully implement our algorithm to estimate the parameters when there are two parallel reactions. It is reasonable to expect that we can apply this algorithm to include additional parallel reactions, though this requires further investigation. There are limitations in that the reactions have to be separated in such a way that we can impose useful prior information.

	$\widetilde{E}_1$	$T_{m,1}$	$\widetilde{E}_2$	$T_{m,2}$	σ
$\widetilde{E}_1$	1.00	-0.12	0.40	0.10	-0.03
$T_{m,1}$	-0.12	1.00	-0.32	-0.04	0.02
$\widetilde{E}_2$	0.40	-0.32	1.00	-0.11	-0.05
$T_{m,2}$	0.10	-0.04	-0.11	1.00	0.02
σ	-0.03	0.02	-0.05	0.02	1.00

Table 4.10: The correlation between the sampled parameters.

Multi-Modality

The algorithm that we use does not always work without issues. We found that when using the following model,

$$\begin{aligned}\frac{dM_1}{dt} &= -M_1 OA_1 \exp\left(\frac{-E_1}{RT}\right),\\ \frac{dM_2}{dt} &= -M_2^3 OA_2 \exp\left(\frac{-E_2}{RT}\right),\\ \frac{dT}{dt} &= \alpha,\end{aligned}$$

our posterior distribution identified different modes. This model uses for a third order Arrhenious reaction for the second reaction rather than a first order. In future work we may consider different models, though our current analysis focuses upon first order reactions.

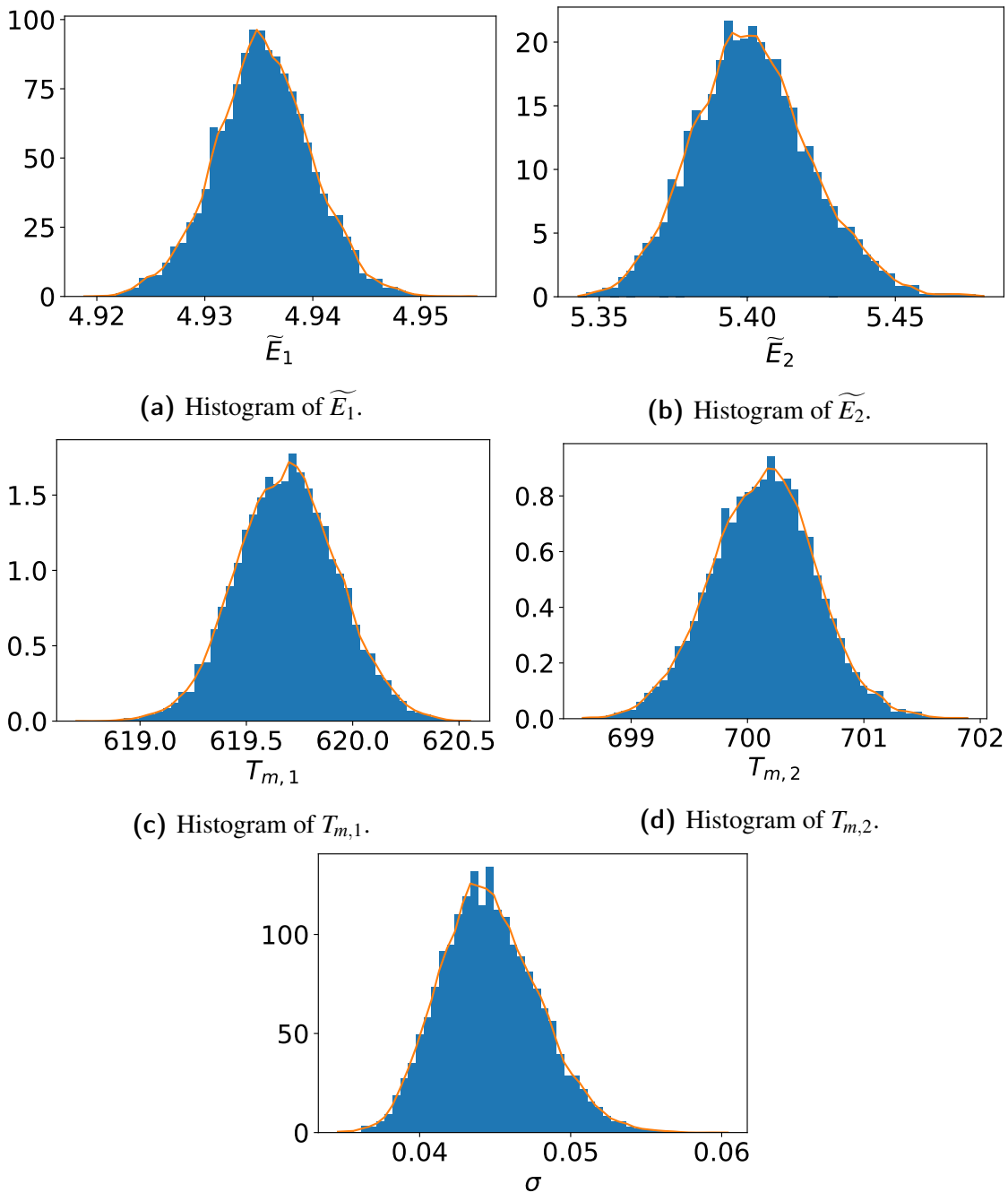
When using higher order reactions Equation 2.22, becomes,

$$A \exp\left(\frac{-E}{RT_m}\right) n M_m^{n-1} = \frac{E}{RT_m^2}\frac{dT}{dt}, \tag{4.11}$$

where n, is the reaction order and M_m is the mass when the reaction rate is maximised. We cannot use this equation to propose new values of A, as the mass term M_m is dependant upon the parameters (A, E, T_m). This yields an implicitly defined function for A, which is unlikely to yield any significant gains in computation speed. We can still use Equation (2.22) as a means to propose new A values, though the T_m parameter in this case does not have the same meaning. This still reduces some of the issues we found when using the parameters (A, E), though the prior information is not as useful.

We produce a simulated sample using the same parameters specified in Table 4.6. These parameters use Equation (2.22) to relate the pre-exponential factor to the Activation Energy, and temperature T_m.

We examine the behaviour of four chains as they move through the parameter space. Figure 4.15 indicates that the value that the mode each chain samples around is dependant upon the initial proposal for the sampling algorithm. We observe three chains (red, orange and green) sample the true parameters in our model, whilst the other chain samples a much



(a) Histogram of $\widetilde{E_1}$.

(b) Histogram of $\widetilde{E_2}$.

(c) Histogram of $T_{m,1}$.

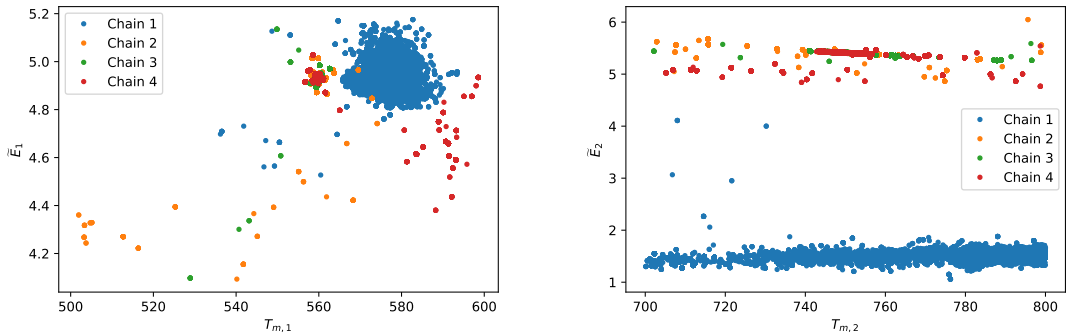
(d) Histogram of $T_{m,2}$.

(e) Histogram of σ.

Figure 4.14: Output from the Metropolis-Hastings for an experiment with two reactions, close together. True values: $T_{m,1} = 620$K and $T_{m,2} = 700$K.

larger region.

We claimed earlier that by proposing new values of T_m, we are able to eliminate some of the issues associated with proposing the parameters,(A,E). Figure 4.16 indicates that the sample from the chains that converged around the true parameters are uncorrelated for the first reaction. For the second reaction there is a correlation between the two parameters that will affect the convergence. This correlation is much less than that between E and A,

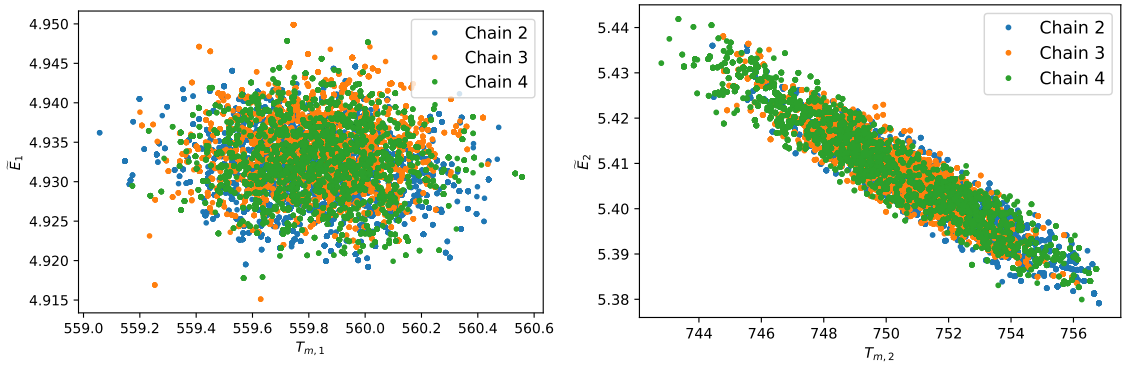


(a) Random walks for the first reaction.

(b) Random walks for the second reaction.

Figure 4.15: Random walks for the maximal temperature and activation energy for the two reactions in a simulated system.

and will allow the algorithm to be more efficient.



(a) Sample for the first reaction.

(b) Sample for the second reaction.

Figure 4.16: Samples of the converged chain for the maximal temperature and activation energy for the two reactions in a simulated system.

Figure 4.17 plots example FWC curves within each mode. We see that mode 1, indicated by the orange line, is able to approximate the data well whilst mode 2, indicated by the green line, seems to approximate the two reactions by a single reaction. This could potentially arise from the fact we are not using good information about the T_m parameter. Our current implementation of the algorithm is not able to identify the reaction order and this results in this mode occurring. Upon further investigation of the log-probability plots (Figure 4.18), the secondary chain is sampling around a local maxima within the posterior distribution and presenting as a secondary mode. In the limit where the chains run for an infinitely long amount of time, the Markov-chain will leave this local maxima, however in practice this is not possible. This highlights the need for careful consideration of the optimal values of the posterior distribution and also highlights some of the pitfalls of relying on simple optimisation based methods that may identify sub-optimal modes.

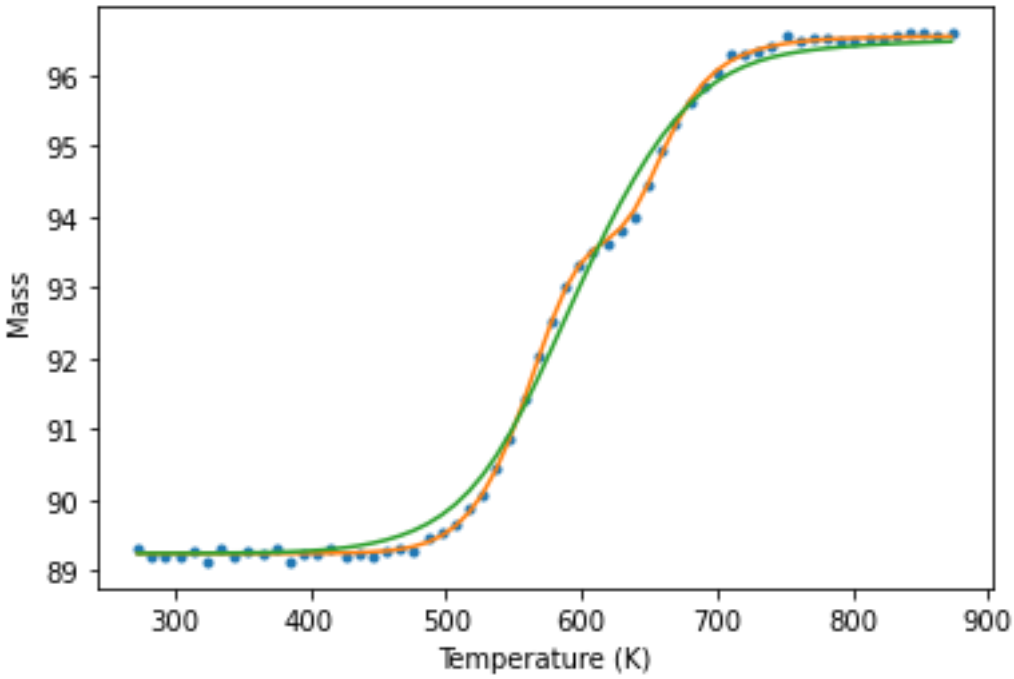


Figure 4.17: Comparison of two sampled parameter values for the two reaction case.

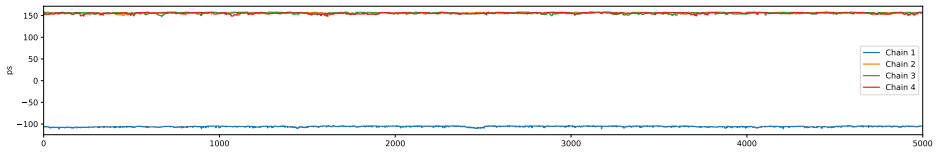


Figure 4.18: Trace plot of the log-probability function of the apparent multi-modal behaviour.

4.3.1 DSC

With the issues of multi-modality we opt to use more of the experimental data. It is hoped that it will address some of the correlations that we found in the two reaction data. Given that the DSC data can be split into a contribution from each reaction, proportional to the reaction rate, then we obtain additional information about these reaction rates. The mass data can be too noisy, to obtain this information about the reaction rates from the mass data. The addition of DSC enables this. We also obtain an improvement in our prior knowledge of the temperature at which the reaction rate is maximised as these correspond to peaks in the DSC data.

Adding the DSC data into our model requires the additional equation,

$$D = Q_1 A_1 M_1 \exp\left(\frac{-E_1}{RT}\right) + Q_2 A_2 M_2 \exp\left(\frac{-E_2}{RT}\right). \tag{4.12}$$

This equation can be evaluated after the differential equation for the mass has been solved. This introduces two new parameters, Q_1 and Q_2 into our model.

The DSC equation is evaluated at discrete time intervals to produce the vector $\mathbf{D}_\tau(\theta_m)$. Since we have introduced additional experimental data we change our notation slightly. We let $\mathbf{M}_\tau(\theta_m)$ denote the mass data. Our vector of simulated values $\mathbf{F}_\tau(\theta_m)$ contains both the mass data $\mathbf{M}_\tau(\theta_m)$ and the DSC data $\mathbf{D}_\tau(\theta_m)$. The experimental data $\mathbf{y}$, contains the experimental mass data $\mathbf{m}$ and DSC data $\mathbf{d}$. The posterior distribution has the same form as before,

$$P(\theta|\mathbf{y}) = \frac{P(\mathbf{y}|\theta)P(\theta)}{\int P(\mathbf{y}|\theta)P(\theta)d\theta}. \tag{4.13}$$

In this instance the likelihood function is different as we have introduced new data. We consider the experimental data to be normally distributed around the simulated data from the model with some noise parameter to be inferred. The residuals $\mathbf{e_M} = \mathbf{M}_\tau(\theta_m) - \mathbf{m}$ and $\mathbf{e_D} = \mathbf{D}_\tau(\theta_m) - \mathbf{d}$ are independent of each other. Within each residual vector the residuals are independent and identically distributed. Thus we have he following distributions,

$$\begin{aligned} \mathbf{e_M} &\sim \mathcal{N}(0, \sigma_M), \\ \mathbf{e_M} &\sim \mathcal{N}(0, \sigma_D), \end{aligned}$$

representing the likelihood function $p(\mathbf{y}|\theta)$.

This introduces a second noise parameter for the DSC data. Because of this we change our approach to estimating the noise parameters. When we just use the mass data we observed that the during the transitory phase the noise parameters increase dramatically. This is not surprising as during the transitory phase the residuals are much larger than when the chains are stationary. However, this complicates the algorithm and reduces the efficiency which becomes more important as we include more parameters. To account for this we change the prior and proposal distribution for the noise parameters. Rather than using a simply prior we use an inverse gamma distribution for the variance σ^2. The major benefit here is that the inverse gamma distribution is the conjugate prior distribution for the component of the likelihood from the noise parameters. Subsequently we use an inverse gamma distribution as the proposal distribution. This is a non-symmetric distribution and we cannot now ignore the proposal distribution in calculating the acceptance probability. However, the component of our acceptance probability π that is dependant upon the noise parameters is now one. We synthesise this into Algorithm 2.

Algorithm 2: Metropolis Hastings Algorithm for determining the posterior distribution with FWC and DSC data.

Result:

Initialise θ_0 by sampling from the prior;

for $t = 0$:*sample_size* **do**

  Sample $\theta^* \sim Q(\theta^*|\theta)$;

  Evaluate $\mathbf{M}_\tau(\theta^*)$ # solve differential equation;

  Evaluate $\mathbf{D}_\tau(\theta^*)$ # calculate the DSC;

  Set $\alpha = \frac{P(\mathbf{y}|\theta^*)Q(\theta_{t-1}|\theta^*)P(\theta^*)}{P(\mathbf{y}|\theta)Q(\theta^*|\theta_{t-1})P(\theta_{t-1})}$;

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

  if $\alpha > b$ **then**

    Set $\theta_t = \theta^*$

  else

    Set $\theta_t = \theta_{t-1}$

  end

end

With this algorithm we consider two cases; Fixed heat parameters, where we prescribe the values of Q_1 and Q_2, and Inferred Parameters, where we conduct inference on the parameters Q_1 and Q_2. The fixed heat parameters can be useful when we have a controlled substance and a controlled experiment with well established heat of reaction terms. In our experimental application, this is not the case and inferring these parameters is preferred.

Fixed Heat Parameters

The first case is where we consider fixed heat parameters. This is a simple case where we can examine the effects of the addition of the DSC data without increasing the dimension of our problem. In a practical application we may not be able to prescribe these heat parameters, this limits the practical implementation of this method.

	Split-$\widehat{R}$	ESS
$\widetilde{E}_1$	1.0009	4928
$T_{m,1}$	1.0018	3059
$\widetilde{E}_2$	1.0009	5655
$T_{m,2}$	1.0013	3267
σ_M	1.0002	17122
σ_D	1.0001	15918

Table 4.11: Diagnostic statistics for the two reaction MH algorithm for the DSC and FWC data sets and fixed heat parameters.

Checking the diagnositic statistics in Table 4.11 for this method we find that we have greatly improved the ESS by including the DSC data. One particular point of interest

is how large the ESS is for the two noise parameters. This large ESS is due to much lower auto-correlation for these parameters. We attribute this to the choice of proposal distribution where we use a Gibbs sampler for these two parameters. Removing this parameter from our random walk improves the remaining parameters convergence as well. All of the Split-$\widehat{R}$ values are very close to one, indicating that the chains have mixed very well. The correlation between parameters in Table 4.12 indicate that the correlation between some of the parameters is still affecting the results.

	$\widetilde{E}_1$	$T_{m,1}$	$\widetilde{E}_2$	$T_{m,2}$	σ_M	σ_D
$\widetilde{E}_1$	1.00	0.22	0.06	-0.13	-0.03	0.02
$T_{m,1}$	0.22	1.00	0.21	0.25	0.06	-0.07
$\widetilde{E}_2$	0.06	0.21	1.00	0.28	0.02	-0.02
$T_{m,2}$	-0.13	0.25	0.28	1.00	0.03	-0.03
σ_M	-0.03	0.06	0.02	0.03	1.00	-0.00
σ_D	0.02	-0.07	-0.02	-0.03	-0.00	1.00

Table 4.12: Correlations for the sampled parameters with fixed heat parameters.

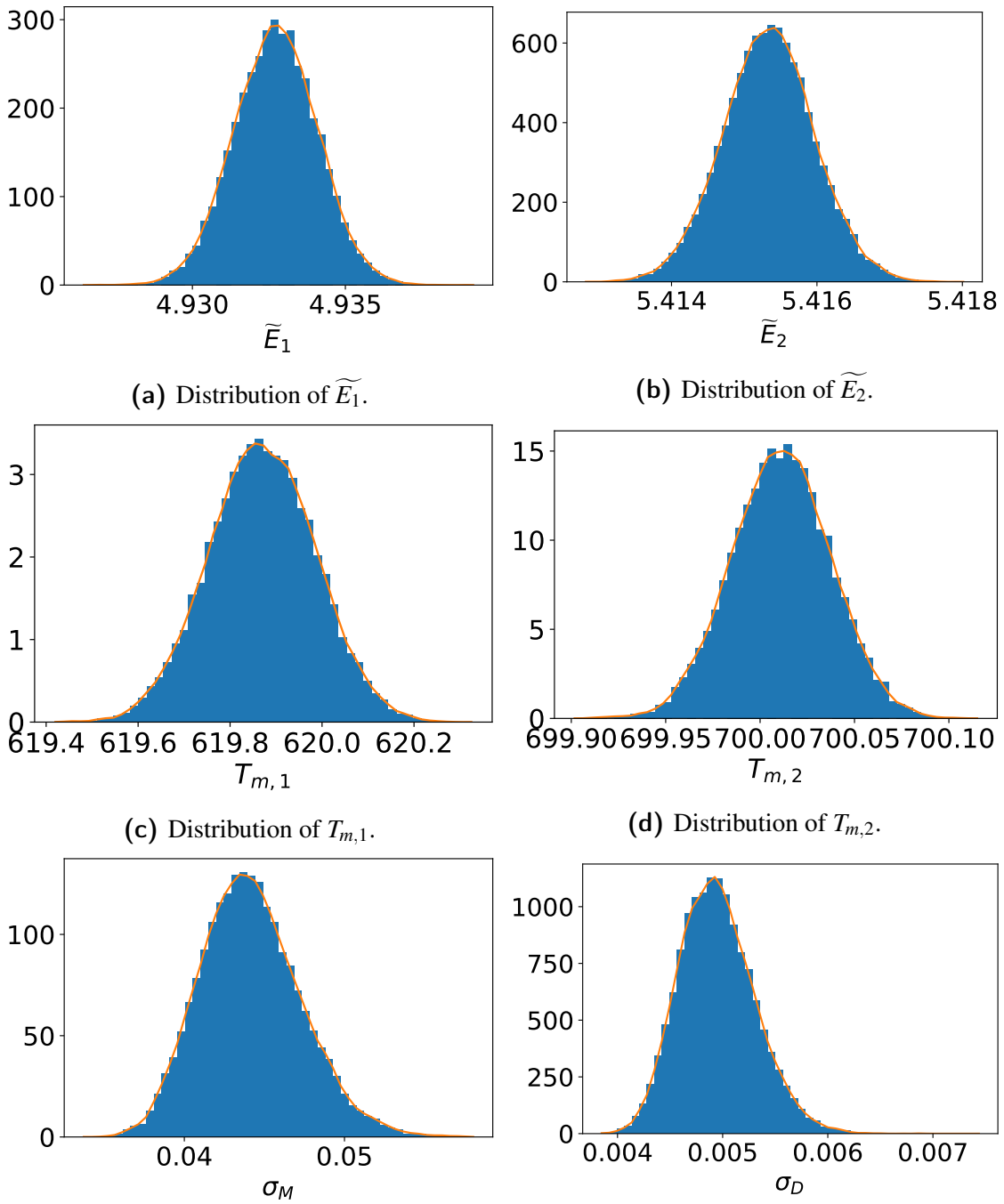
Figure 4.19 displays the posterior distributions obtained from the simulated data. When we compare this to the data posterior samples in Figure 4.14, we observe that the histograms are much narrower, whilst containing the true values. Adding the DSC data into the simulations has decreased the uncertainty in the posterior distributions.

Using fixed parameter estimates is not always possible in practical applications, since we do not know the heat of reaction parameters. This is particularly important when we are not certain of the reaction schemes. If we assume the heat of reactions are fixed with uncertainty over the reactions taking place then the result can be a poor fit to the data. In our application multiple reaction schemes have been proposed and subsequently using inferred parameter estimates is preferred.

Inferred Parameter

The alternative method for including these heat parameters is to conduct inference on them. This setting is more useful for problems where these parameters are unknown. As mentioned before introducing new parameters into the inference increases the dimensionality of the problem, possibly leading to more multi-modal behaviour.

We introduce these parameters in the log form i.e.$\widetilde{Q}_i = \log(Q_i)$. This has the same benefit that we observe with the activation energy. The log scale allows the algorithm to move efficiently through may different orders of magnitude whilst maintaining the precision necessary. In introducing this parameter we use the weakly informative uniform distribution for the parameter $\widetilde{Q}$. This concentrates the mass equally between orders of magnitude meaning that more of the probability mass is located at smaller values. The samples from the posterior distribution are displayed in Figure 4.20



(a) Distribution of $\widetilde{E_1}$.

(b) Distribution of $\widetilde{E_2}$.

(c) Distribution of $T_{m,1}$.

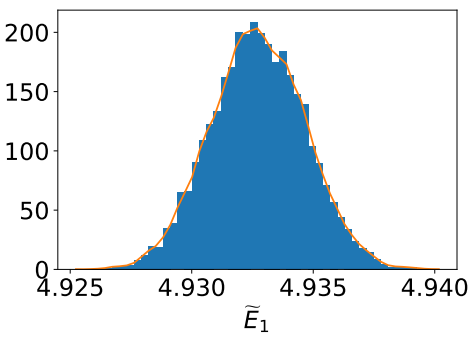
(d) Distribution of $T_{m,2}$.

(e) Distribution of σ_M.

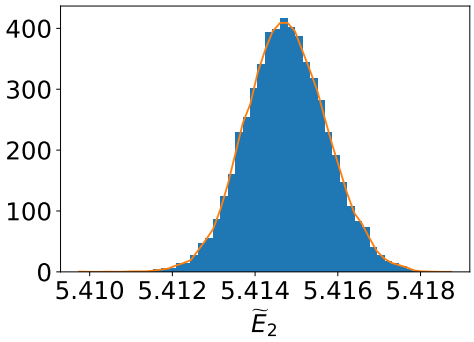
(f) Distribution of σ_D.

Figure 4.19: Histograms of the posterior samples generated from the MH algorithm for a simulated experiment with both TGA and DSC data.

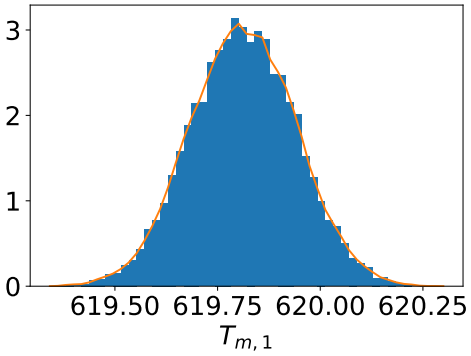
When we compare the histograms in Figure 4.20 to those in Figure 4.19 we observe that the variance of the posterior distributions is much greater when we infer the heat parameters. This is not surprising as we increase the number of unknowns in the model. Figure 4.20 also indicates that we have successfully captured the true parameters within our posterior distributions. We can confirm that the Markov Chains have converged to the stationary distribution from the Split-$\widehat{R}$ values in Table 4.13. The significantly larger ESS



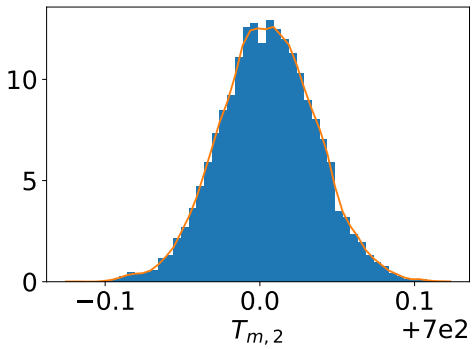
(a) Distribution of $\widetilde{E_1}$.



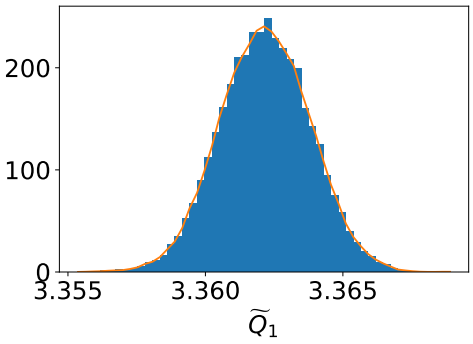
(b) Distribution of $\widetilde{E_2}$.



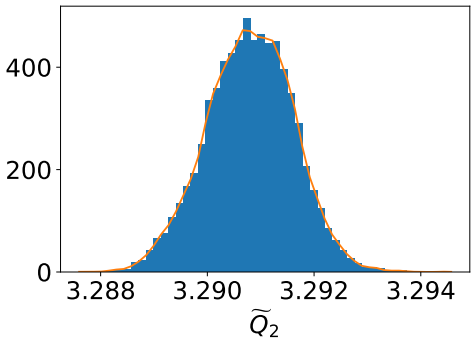
(c) Distribution of $T_{m,1}$.



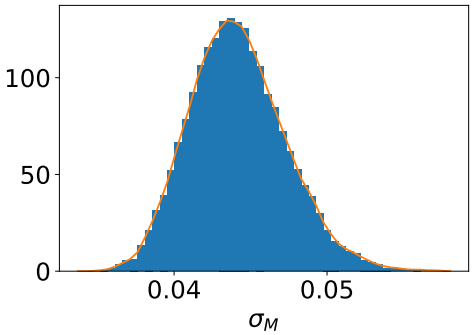
(d) Distribution of $T_{m,2}$.



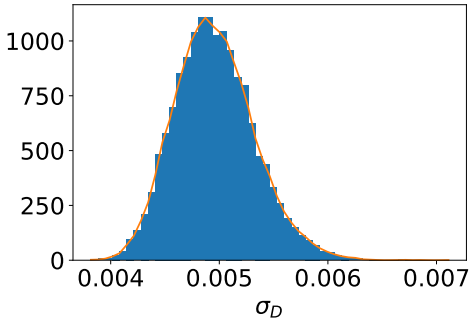
(e) Distribution of $\widetilde{Q_1}$.



(f) Distribution of $\widetilde{Q_2}$.



(g) Distribution of σ_M.



(h) Distribution of σ_D.

Figure 4.20: Histograms of the posterior samples including inference for the heat of reaction parameters, generated from the MH algorithm for a simulated experiment with both TGA and DSC data.

of the two noise parameters and lower Split-$\widehat{R}$ values highlight the benefits of using the conjugate prior distribution.

	Split-$\widehat{R}$	ESS
$\widetilde{E}_1$	1.00173	1232
$T_{m,1}$	1.00057	1483
$\widetilde{E}_2$	1.00188	993
$T_{m,2}$	1.00255	1416
$\widetilde{Q}_1$	1.00139	1216
$\widetilde{Q}_2$	1.00267	974
σ_M	1.00002	10891
σ_D	1.00019	8793

Table 4.13: Convergence Statistics for the simulated data with DSC.

The correlation coefficients between the posterior sample displayed in Table 4.14 indicate there is significant correlation between the different parameters. The posterior samples are displayed in Figure 4.21 indicating these are approximately linear correlations. This correlation contributes to the lower ESS sizes we see in Table 4.13. These correlations are much more significant than those in Table 4.12. Adding the additional parameters into the model has increased the complexity of the posterior distribution. Consequently, the simpler the model that can be used the more efficient the MCMC algorithm is at sampling the posterior distribution.

	$\widetilde{E}_1$	$T_{m,1}$	$\widetilde{Q}_1$	$\widetilde{E}_2$	$T_{m,2}$	$\widetilde{Q}_2$	σ_M	σ_D
$\widetilde{E}_1$	1.00	-0.12	-0.70	-0.45	-0.46	0.52	-0.09	0.08
$T_{m,1}$	-0.12	1.00	0.28	0.49	0.43	-0.48	0.08	-0.09
$\widetilde{Q}_1$	-0.70	0.28	1.00	0.54	0.49	-0.57	0.08	-0.08
$\widetilde{E}_2$	-0.45	0.49	0.55	1.00	0.58	-0.77	0.08	-0.10
$T_{m,2}$	-0.46	0.43	0.49	0.58	1.00	-0.51	0.06	-0.09
$\widetilde{Q}_2$	0.53	-0.49	-0.57	-0.77	-0.51	1.00	-0.07	0.09
σ_M	-0.09	0.08	0.08	0.08	0.06	-0.07	1.00	0.01
σ_D	0.08	-0.09	-0.08	-0.10	-0.09	0.09	0.01	1.00

Table 4.14: Sample correlations between each parameter.

We have established that our MH algorithm produces a sample of the posterior distribution of our parameter estimates. In each successive simulation we have made some adjustments as our data and algorithm becomes more complicated. What remains is to apply this to the experimental data.

4.3.2 Additional Parameterisation

Another common parameterisation for the Arrhenious reaction rate centers the reaction upon some reference temperature T_{ref} [129, 130]. Using this reference temperature, we

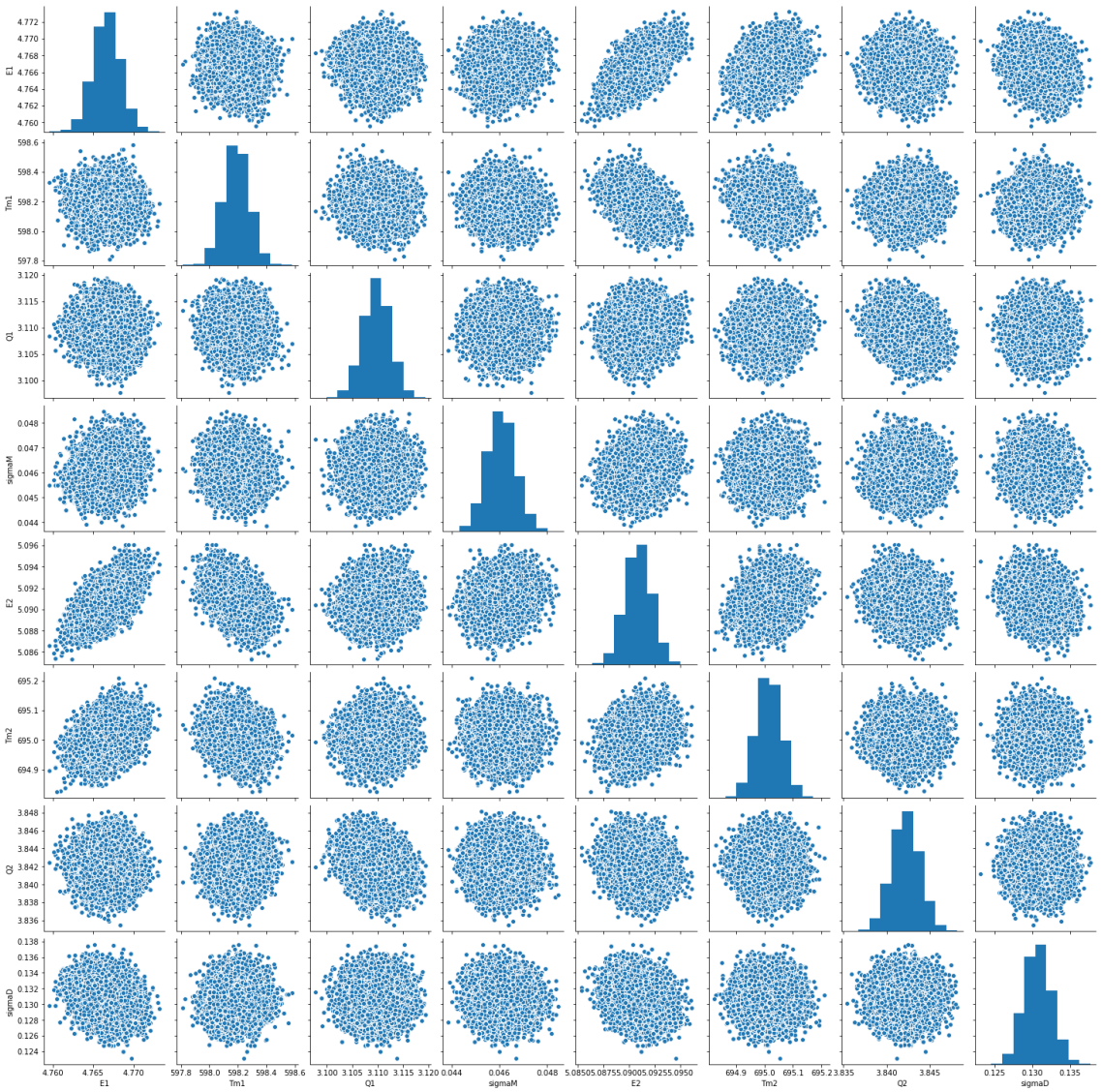


Figure 4.21: Pair plot of all the parameters estimated from the MCMC algorithm identifying any possible correlations.

rewrite the Arrhenious reaction rate as,

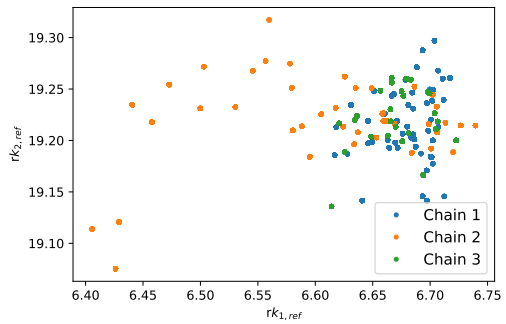
$$k(T) = k_{\text{ref}} \exp\left(\frac{-E}{R}\left(\frac{1}{T} - \frac{1}{T_{\text{ref}}}\right)\right). \tag{4.14}$$

In this parameterisation, k_{ref} is the reaction rate at the temperature T_{ref}. To convert to the original parameterisation,

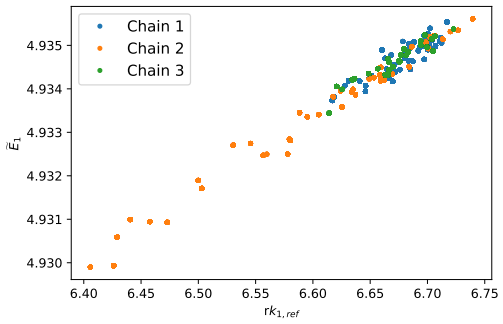
$$A = k_{\text{ref}} \exp\left(\frac{-E}{RT_{\text{ref}}}\right). \tag{4.15}$$

This relationship reduces the correlation between the inferred parameters. In our application to two reactions occurring within a sample we found that whilst the correlation had reduced, significant correlations persisted within our analysis as indicated by Figure 4.22. Using alternative methods for sampling from the posterior distribution such as adaptive

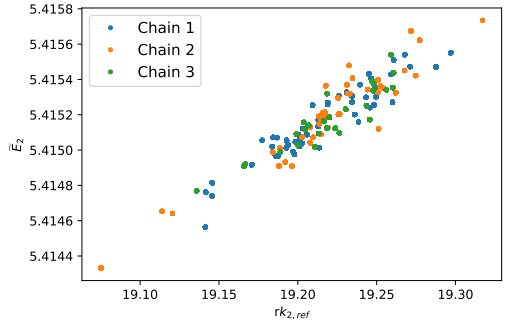
methods may lessen the impact of this correlation on the efficiency of the MH algorithm [128]. This parameterisation may become most useful when considering experiments using different heating rates. In these applications, the parameter T_m changes, however using this alternate parameterisations, this problem is not an issue. Our analysis suggests that the selection of the reference temperature is important in decoupling the inferred parameters and this may have an impact when multiple experiments are considered.



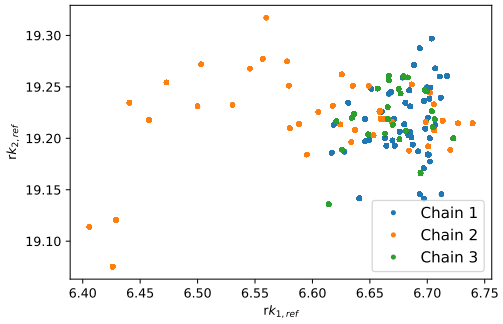
(a) Random sample for $k_{\text{ref},1}$ and $k_{\text{ref},2}$



(b) Random sample for $k_{\text{ref},1}$ and $\widetilde{E_1}$



(c) Random sample for $\widetilde{E_2}$ and $k_{\text{ref},2}$



(d) Random sample for $\widetilde{E_2}$ and $\widetilde{E_1}$

Figure 4.22: Posterior sample for the $k_{\text{ref},1}$, $k_{\text{ref},2}$, $\widetilde{E_1}$ and $\widetilde{E_2}$ parameters.

4.4 Application to Experimental Data

Now that we have established the algorithm is able to infer information about our parameters in simulated examples we apply it to experimental data. The experimental data is more complex and there are issues regarding possible instances of model miss-

specification. We use the model,

$$\frac{dM_1}{dt} = M_1 A_1 \exp\left(\frac{-E_1}{RT}\right), \tag{4.16}$$

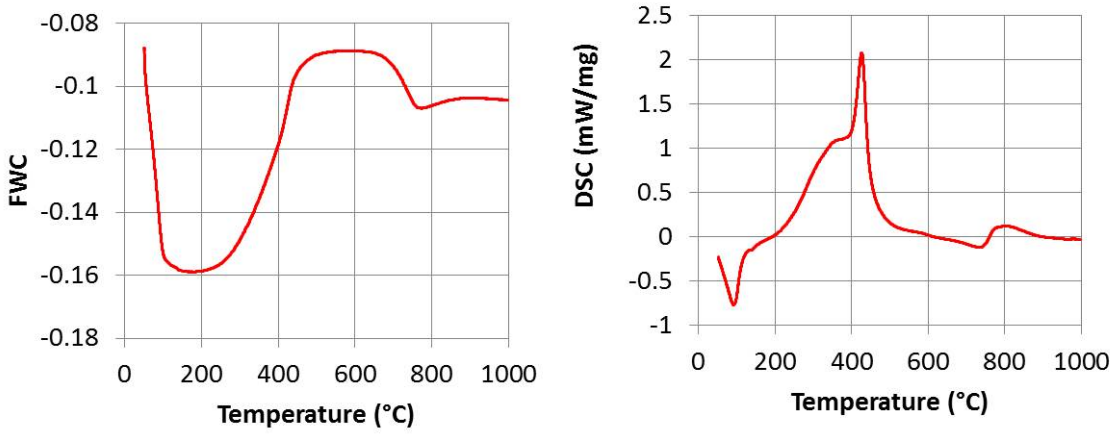
$$\frac{dM_2}{dt} = M_2 A_2 \exp\left(\frac{-E_2}{RT}\right), \tag{4.17}$$

$$\frac{dT}{dt} = \alpha, \tag{4.18}$$

$$\frac{dM_t}{dt} = w_{c,1} M_1 A_1 \exp\left(\frac{-E_1}{RT}\right) + w_{c,2} M_2 A_2 \exp\left(\frac{-E_2}{RT}\right), \tag{4.19}$$

$$H = Q_1 M_1 A_1 \exp\left(\frac{-E_1}{RT}\right) + Q_2 M_2 A_2 \exp\left(\frac{-E_2}{RT}\right), \tag{4.20}$$

that we derived in Section 3.2. This first order Arrhenious equation model links into our large stockpile models and thus it is imperative that we conduct inference on these parameters. The experimental data is obtained from the work done by Longbottom et al [1] and is displayed in Figure 4.23.

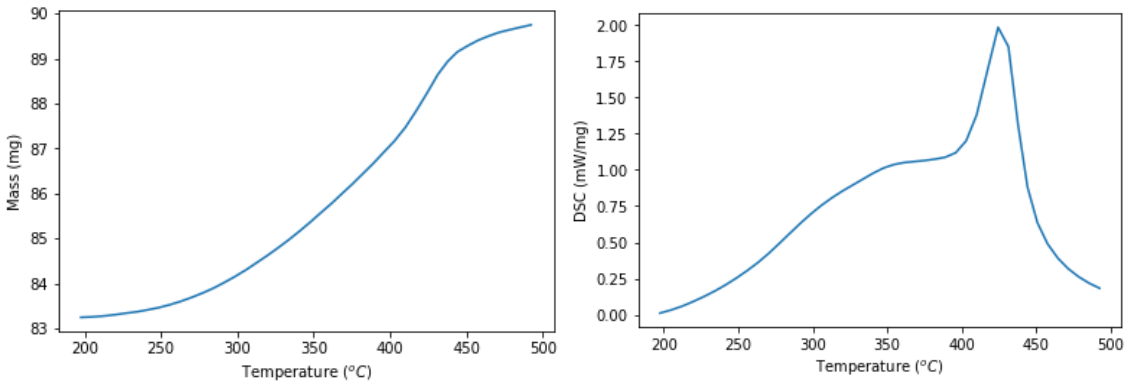


(a) Fractional weight change curve. (b) Differential Scanning Calorimetry data.

Figure 4.23: Experimental data on the Filter cake [1].

This data includes all the reactions within the stockpile. If we were to use all of the data we would need at least five reactions, of which many are unlikely to be identified through experimentation. We focus our analysis on the region from 200^oC to 500^oC. This region was identified by Longbottom et al. [1] as the critical region where the reactions that are the primary drivers of self-heating occurs. The restricted data is displayed in Figure 4.24, for clarity.

When we apply the algorithm to experimental data it is important to consider the effect of the parameters we are not conducting the inference on. The specific parameters of interest are the initial mass concentrations, denoted $M_{1,0}$ and $M_{2,0}$, and the weight change constants, $w_{c,1}$ and $w_{c,2}$. We determine the weight change constants by prescribing the



(a) Restricted FWC curve.

(b) Restricted DSC data.

Figure 4.24: Experimental data on the Filter cake restricted to the critical region [1].

reaction scheme,

$$6\mathrm{FeO} + \mathrm{O}_2 \rightarrow 2\mathrm{Fe}_3\mathrm{O}_4,$$
$$3\mathrm{Fe} + 2\mathrm{O}_2 \rightarrow \mathrm{Fe}_3\mathrm{O}_4,$$

stated in Section 3.1. Using this reaction scheme we can determine the amount of mass that is added to the sample per gram of Wüstite and Iron consumed. These weight change values are $w_{c,1} = -0.074$ and $w_{c,2} = -0.429$, respectively.

The initial masses are more challenging. The filter cake is known to be highly heterogeneous and as a result the initial masses in the sample can vary from the measured values determined by Longbottom et al [1]. It is important to have accurate estimates for these parameters as they can introduce error into the model and the estimates of the reaction parameters can have an unknown bias. We control for this by ensuring that the simulated sample mass at the beginning and end of the critical region closely match the sample mass in the experiment. To calculate this we impose an assumption that there is negligible mass change from these reactions outside of this region. Under this assumption we also assume that the moisture-less mass is the minimum sample mass. This results in the equation,

$$M_{1,0}w_{c,1} + M_{2,0}w_{c,2} = M_{t,0} - M_{t,f}. \tag{4.21}$$

The initial concentrations we select, minimise the distance, $d = |\left(M_{1,0} - \widetilde{M_1}, M_{2,0} - \widetilde{M_2}\right|$, subject to the constraint 4.21, where $\widetilde{M_1}$ and $\widetilde{M_2}$ are the masses calculated using the percentage of moistureless mass determined by Longbottom et al [1]. We calculate the proportion of moistureless mass of Wüstite, M_1, and Iron, M_2, to be 42.8% and 10.9%. Both of these values are less than those measured by Longbottom et al [1], which are 41.2% and 10.6%. One of the key aspects of Bayesian Inference is that the posterior distributions

are all conditional upon the experimental data. In our instance this also means that the estimates are conditional on our model choice and subsequently the model parameters. Changing the reaction scheme affects the posterior distribution for all the parameters. This framework allows useful forward predictions that are conditionally upon the chosen reactions though we find that if we prescribe a reaction scheme that is deemed to be inaccurate the analysis has to be conducted again with the new reaction scheme. This Bayesian approach we are using on the parameter estimation also enables future work to compare reaction schemes, though this is beyond the scope of this thesis.

When we apply our algorithm to the experimental data, there is additional information that can be used to inform our prior distributions. Additional experimental data by Longbottom et al. [7], provides a region where the reactions are occurring. Using this data we are able to use a Guassian distribution centered on this range as our prior distribution. We apply this distribution for the T_m parameters of each reaction. When this prior is applied we no longer have the identifiability condition $T_{m,1} < T_{m,2}$. This can be added in as a condition though we choose to exclude this in our analysis. As there is some uncertainty regarding the choice of reaction schemes then allowing the order of reactions to change can be helpful. If we observe that the posterior distribution has significant mass in this secondary mode, then that lends credibility to the idea that we have not selected the correct reaction scheme.

	$\widehat{R}$	ESS
$\widetilde{E}_1$	1.0016	1829
$T_{m,1}$	1.0022	2815
Q_1	1.0024	1911
E_2	1.0027	1937
$T_{m,2}$	1.0003	2639
Q_2	1.0022	1463
σ_M	1.0004	8842
σ_D	1.0004	8436

Table 4.15: Diagnostic Statistics for the experimental data.

With all these assumptions we conduct our inference on the data. Before we consider the samples from the posterior posterior distribution we want to assess if there are any issues within our algorithm. Considering the diagnostic statistics in Table 4.15, there are no obvious signs of concern. The Split-$\widehat{R}$ values are all close to one, indicating that the chains are well mixed and have reached stationarity. Additionally the ESS size for each parameter is good. Once again the noise parameters have a significantly larger sample size than the other parameters and we attribute this to the choice of proposal distribution.

The samples obtained through our algorithm are displayed in Figure 4.25. These Histograms appear nicely behaved and indicate a slightly greater spread of values than what we experience in the simulated studies. This is expected as we expect that some model

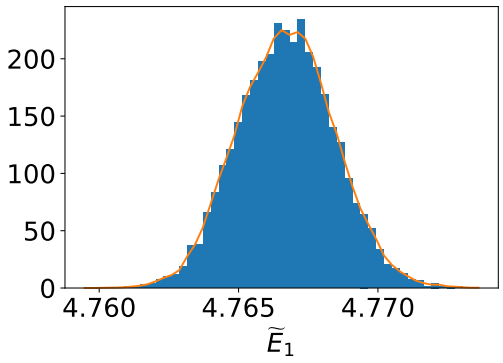
miss-specification is likely. The spread of these histograms are quite similar to those achieved when we consider two reactions close together only using the FWC data presented in Figure 4.20. The parameter with the greatest uncertainty in comparison to the results with two reactions is the temperature at which the first reaction is maximised, $T_{m,1}$.

When we apply this algorithm to experimental data it is important to carefully examine the convergence of each of the four chains in our algorithm. These investigations can be carried out with the our simulated data, though given we know the true values we can reasonably determine whether they have converged using the sample. If the sample from the posterior distribution contains the true parameters this adds weight to the argument that the Markov Chains have converged and reached stationarity. For the experimental data we do not know these values and have to rely upon the diagnostics statistics in Table 4.15 and investigating the trace plots. The diagnostic statistics indicated that the chains have mixed and reached stationarity through the Split-$\widehat{R}$ statistic and the ESS indicates that the auto-correlation is low.

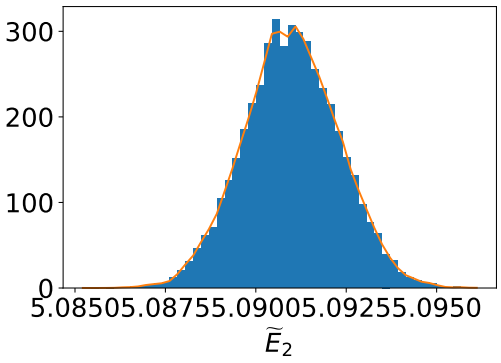
The trace plots in Figure 4.26 indicate that the chains have converged together from very different starting point. The Markov chains in these figures include the burn in period which is discarded before considering the sample. We observe an initial period where we have a lower acceptance rate. The low acceptance rate is paired with a large step size in the proposal distribution allowing the algorithm to explore greater regions of the parameter space. By having a variable proposal distribution in the initial stages of the algorithm we can have a lower acceptance rate as we attempt to explore more of the parameter space. Once the Markov Chains begin to converge onto the posterior distribution, we adjust the proposal so that we can sample the distribution efficiently.

The trace plots after the burn in is discarded are displayed in Figure 4.27. By focussing on the sample itself we observe that the chains are mixing well and the auto-correlation is low. This corroborates the diagnostic statistics in Table 4.15 and thus we conclude that the Markov Chains have converged and the sample we have obtained is from the posterior distribution of our parameters.

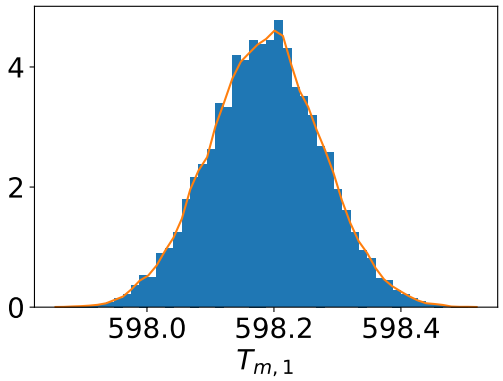
Now that we have established that the sample we have obtained from the MCMC algorithm is from the posterior distribution we can begin to analyse this sample. The correlation between the parameters is of particular interest and is displayed in Table 4.16. We observe that several parameters are highly correlated, in particular the two activation energy parameters, $\widetilde{E}_1$ and $\widetilde{E}_2$ have a correlation coefficient of 0.63. In our model the activation energy controls the duration of the reaction when the temperature that the reaction rate is maximised is fixed. This therefore indicates that the spread of the two reactions are linked. Another correlation that is of interest is with the noise parameters and the activation energy. In all of the simulated cases the noise parameters were uncorrelated to all of the parameters. When the activation energy is increased this seems to decrease the noise in the DSC and increase the noise in the FWC data. This suggests that the



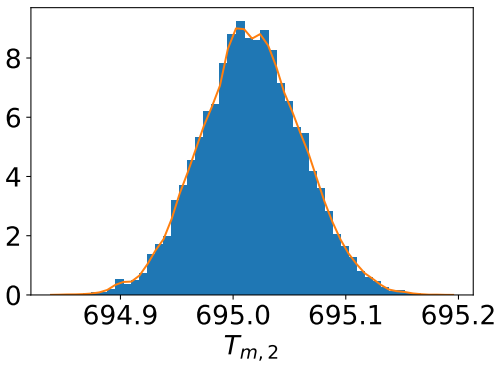
(a) Distribution of $\widetilde{E_1}$.



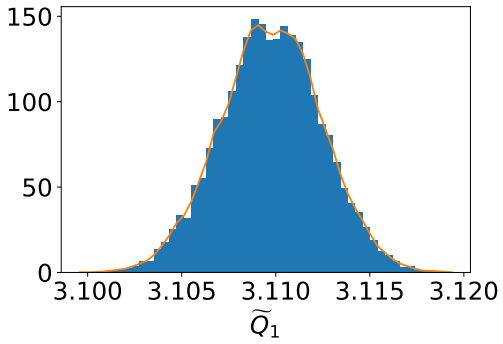
(b) Distribution of $\widetilde{E_2}$.



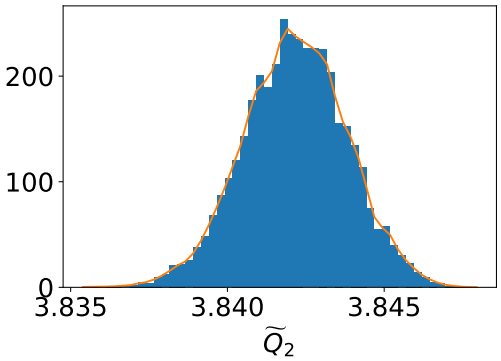
(c) Distribution of $T_{m,1}$.



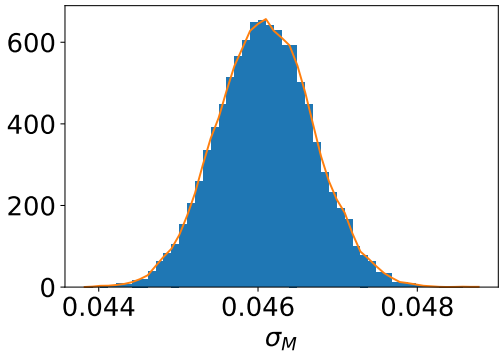
(d) Distribution of $T_{m,2}$.



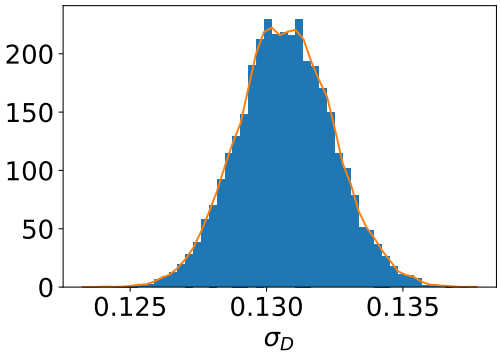
(e) Distribution of $\widetilde{Q_1}$.



(f) Distribution of $\widetilde{Q_2}$.

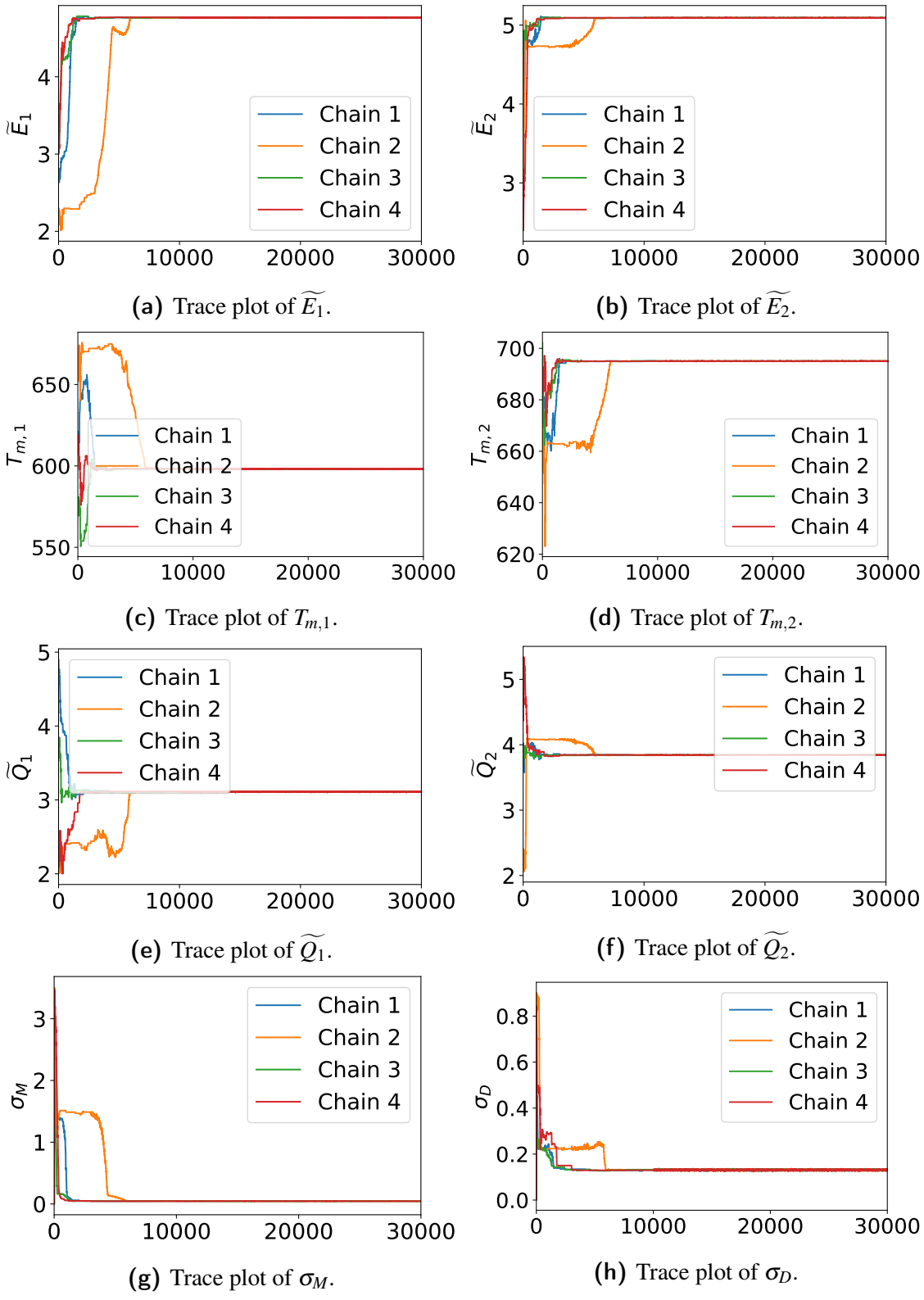


(g) Distribution of σ_M.



(h) Distribution of σ_D.

Figure 4.25: Posterior samples of the estimated parameters from the experimental data.



(a) Trace plot of $\widetilde{E_1}$.

(b) Trace plot of $\widetilde{E_2}$.

(c) Trace plot of $T_{m,1}$.

(d) Trace plot of $T_{m,2}$.

(e) Trace plot of $\widetilde{Q_1}$.

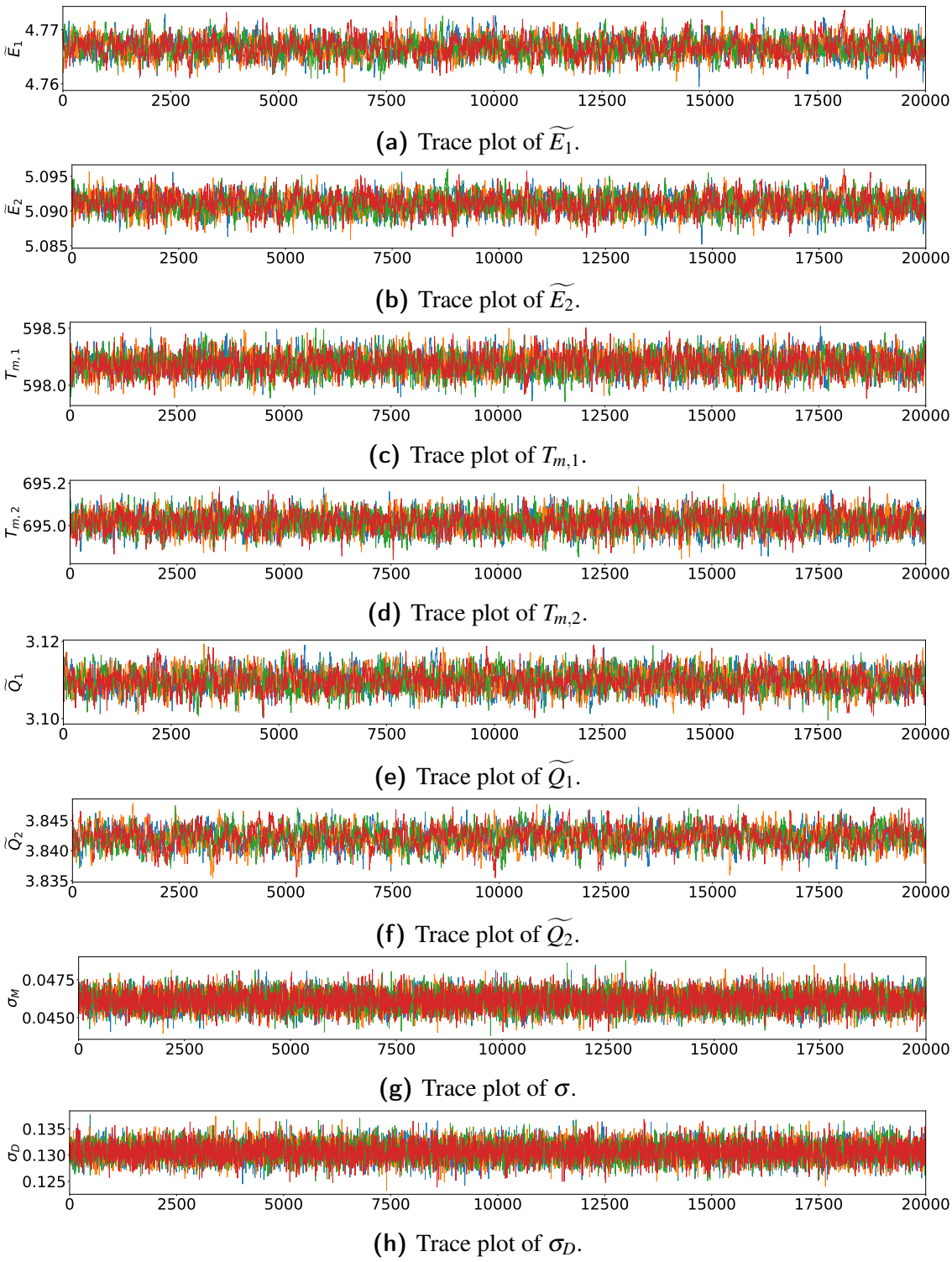
(f) Trace plot of $\widetilde{Q_2}$.

(g) Trace plot of σ_M.

(h) Trace plot of σ_D.

Figure 4.26: Trace plots of the estimated parameters from the experimental data.

DSC data indicates that the reaction rate is greater than what is indicated by the FWC data. This suggests that an additional heat producing reaction is required that does not affect the mass data. Further development of the MCMC algorithm would be required to



(a) Trace plot of $\widetilde{E_1}$.

(b) Trace plot of $\widetilde{E_2}$.

(c) Trace plot of $T_{m,1}$.

(d) Trace plot of $T_{m,2}$.

(e) Trace plot of $\widetilde{Q_1}$.

(f) Trace plot of $\widetilde{Q_2}$.

(g) Trace plot of σ.

(h) Trace plot of σ_D.

Figure 4.27: Trace plots of the estimated parameters from the experimental data after the burn-in period is removed.

thoroughly investigate this hypothesis.

We can analyse the sample to produce some point estimates for the distribution. The mean and median values are two key estimates along with the minimum and maximum sampled values that we use to understand the posterior distribution. These values along

	$\widetilde{E}_1$	$T_{m,1}$	$\widetilde{Q}_1$	$\widetilde{E}_2$	$T_{m,2}$	$\widetilde{Q}_2$	σ_M	σ_D
$\widetilde{E}_1$	1.00	-0.06	-0.12	0.63	0.42	-0.00	0.19	-0.19
$T_{m,1}$	-0.06	1.00	-0.06	-0.46	-0.18	0.04	-0.06	0.07
$\widetilde{Q}_1$	-0.12	-0.06	1.00	0.05	-0.00	-0.28	-0.01	-0.00
$\widetilde{E}_2$	0.63	-0.46	0.05	1.00	0.28	-0.19	0.21	-0.21
$T_{m,2}$	0.42	-0.18	-0.00	0.28	1.00	0.00	0.03	-0.05
$\widetilde{Q}_2$	-0.00	0.04	-0.28	-0.19	0.00	1.00	-0.03	0.02
σ_M	0.19	-0.06	-0.01	0.21	0.03	-0.03	1.00	-0.04
σ_D	-0.19	0.07	-0.00	-0.22	-0.05	0.02	-0.04	1.00

Table 4.16: Table of the correlations between each parameter for the experimental data.

with a 90% credible interval (CI) are stated in table 4.17. One thing that is important to consider is that whilst the width of the credible intervals seem small for both the activation energy and heat of reaction parameters, these are both specified on a logarithmic scale. When converted to the true scale these correspond to standard confidence intervals.

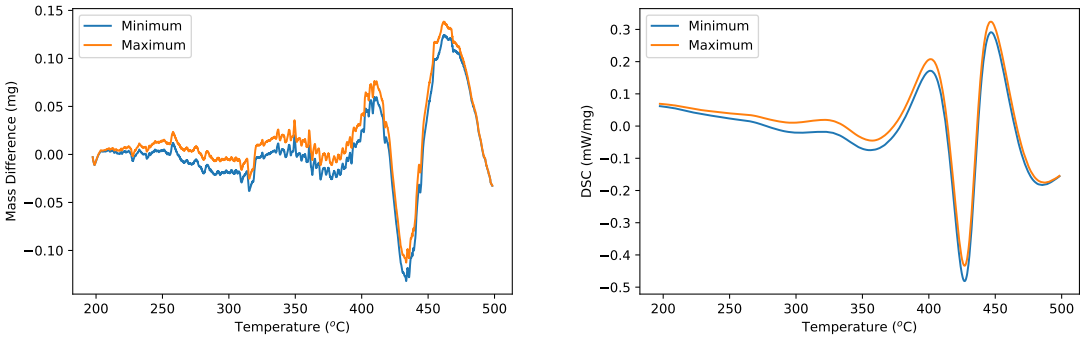
	Mean	Minimum	Median	Maximum	90% CI
$\widetilde{E}_1$	4.766	4.759	4.767	4.774	(4.763, 4.770)
$T_{m,1}$	598.1	597.8	598.2	598.5	(598.0, 598.3)
$\widetilde{Q}_1$	3.109	3.100	3.110	3.120	(3.105, 3.114)
$\widetilde{E}_2$	5.090	5.085	5.091	5.096	(5.088, 5.093)
$T_{m,2}$	695.01	694.83	695.02	695.19	(694.94, 695.09)
$\widetilde{Q}_2$	3.842	3.835	3.842	3.848	(3.839, 3.845)
σ_M	0.0461	0.0438	0.0461	0.0488	(0.0451, 0.0471)
σ_D	0.1307	0.1232	0.1307	0.1377	(0.1278, 0.1336)

Table 4.17: The summary statistics for the parameters in the experimental model.

One thing that remains is to compare the simulated experiment using the posterior sample to the experimental data. The differences between the simulated sample parameters and the experimental data is displayed in Figure 4.28. These figures indicate that there is some discrepancy between the model and the experimental data in the temperature range 400-500^{o}C. This difference is low and we are still able to use our model to explain the data. This discrepancy indicates that a different reaction scheme could be implemented with a new reaction occurring in this region.

4.5 Summary

The MCMC algorithm that we have developed in this chapter has demonstrated that we can obtain meaningful parameter estimates from the posterior distribution. We have tested this algorithm in multiple simulated cases for which we were able to obtain samples from the posterior distribution that included our true parameters. The Bayesian framework



(a) Fractional weight change curve.

(b) Differential Scanning Calorimetry data.

Figure 4.28: Experimental data on the Filter cake [1], compared to the posterior sampled values.

allows us to utilise this sampled parameter distribution to investigate the behaviour of large stockpiles and easily identify the uncertainty in key predictive parameters such as the critical length and ignition times. We have been able to generate meaningful estimates of the kinetic parameters from the experimental data which previous parameter estimation tools were insufficient for. In addition to obtaining these estimates, our algorithm has identified that there are aspects of the model that do not explain the data completely. This suggests that an additional reaction may need to be included in order to obtain a reaction scheme that best considers the data. This concept along with the inclusion of multiple sets of experimental data is considered in the next chapter.

Chapter 5

Inference using Multiple Experiments

In the previous chapter we utilised a Metropolis-Hastings algorithm to sample from the posterior distribution of our parameters. These posteriors are dependant upon the observations of a single experiment. In practice these experiments are often repeated, using different heating rates. This additional data should be included in our parameter inference model. One possible approach to do this is to simulate each experiment using the respective heating rates within our previous MCMC algorithm. This is expected to result in a more complex likelihood; additional modes may appear causing issues with convergence. Simulating multiple experiments is also a computationally expensive task. To avoid this, we develop a Sequential Monte Carlo (SMC) method that uses importance sampling (IS) to sample from the posterior distribution given the parameters in each of the experiments.

In practical applications the experiment can be either repeated under the same conditions, or a fixed control parameter can vary. In the work conducted by Longbottom et al [1], the TGA experiment was conducted four times with heating rates α of 5, 10, 10, 20 K/min. In this example the 10K/min was repeated and two new heating rates were conducted. The set-up that we introduce in this chapter can theoretically handle any arbitrary amount of repeated experiments, though as this amount increases the computation required will also increase.

5.1 The Posterior Distribution

In order to sample from the posterior distribution, we must first define it. To begin, we must first make assumptions about the data. We consider each experiment, inclusive of both FWC and DSC data, independent of each other. We cannot consider the FWC and DSC data within one experiment as independent of each other. We assume that the noise across multiple heating rates remains constant across each experiment. These assumptions

enable us to calculate the posterior distribution as,

$$p(\theta \mid \mathbf{y}) = \prod_{i=0}^{n} p(\mathbf{y}_i \mid \theta) \frac{p(\theta)}{\prod p(\mathbf{y}_i)}, \tag{5.1}$$

where $\mathbf{y}_i$ is the data from the ith experiment and $p(\mathbf{y}_i|\theta)$ is the corresponding likelihood function. The probability functions $p(\mathbf{y}_i|\theta)$ are calculated by assuming that the residuals between the solution vector $\mathbf{M}_t$ and the data $\mathbf{y}$ is normally distributed with mean 0 and variance σ_M^2. The term $\prod p(\mathbf{y}_i)$ is a normalising constant. We use this product term as it allows naturally to define this posterior distribution recursively by,

$$p(\theta \mid \mathbf{y}_{1:i}) = \frac{p(\mathbf{y}_i \mid \theta) p(\theta \mid \mathbf{y}_{1:i-1})}{p(\mathbf{y}_i)}, \tag{5.2}$$

$$p(\theta \mid \mathbf{y}_1) = \frac{p(\mathbf{y}_1 \mid \theta) p(\theta)}{p(\mathbf{y}_1)}, \tag{5.3}$$

where $\mathbf{y}_{1:i}$, includes the data from the first to the ith experiment. This form provides an intuitive basis for the remainder of the chapter as we use the previous posterior distribution as our prior information for each additional experiment.

5.2 Sequential Monte Carlo Methods

When sampling from the posterior distribution, we can utilise sequential Monte Carlo (SMC) techniques. In this instance we apply the information from each experiment sequentially, updating our original MCMC sample with each new set of experimental data. We conduct importance sampling to re-weight our original sample and then we can implement an MCMC step within our algorithm to handle any redundancies.

5.2.1 Importance Sampling

Importance sampling is a re-sampling technique, where given a sample $\mathbf{X}$, then for each point $\mathbf{x}$ in our sample a weight w is calculated [131]. These weights are calculated through the likelihood function with respect to the second experiment. Consider a distribution $p(x)$, then the expectation $E[f(x)]$ is defined as,

$$E[f(x)] = \int f(x) p(x) dx.$$

This definition implies that if we have a distribution $h(x) = w(x)p(x)$, then we can compute the expectations of this distribution as,

$$E[f(x)] = \int f(x) w(x) p(x) dx.$$

Importance sampling uses this concept where we sample from $p(x)$, and then calculate weights using the function $w(x)$. Given a sample $\{x_j\}$ drawn from a distribution $p(x)$, then the expectation $E[f(x)]$ is estimated by,

$$E[\hat{f}(x)] = \frac{1}{N}\sum_{j=1}^{N} f(x_j)$$

Assigning weights w_i to each data point such that $w_i = w(x_i)$, then the expectation can be estimated by,

$$E[\hat{f}(x)] = \frac{1}{N}\frac{\sum_{j=1}^{N} f(x_j)w_j}{\sum_{j=1}^{N}(x_j)}.$$

For normalised weights $\tilde{w}_j = w_j/\sum w_j$ this reduces to,

$$E[\hat{f}(x)] = \frac{1}{N}\sum_{j=1}^{N} f(x_j)\tilde{w}_j.$$

This estimate provides a framework of how we use importance sampling for repeated experiments. Using the recursive definition of the posterior distribution in equation (5.2), then if we obtain a sample from $p(\theta|\mathbf{y}_{1:i-1})$, then we calculate the weights by evaluating our likelihood function for the ith function $p(\mathbf{y} \mid \theta)$.

We can re-sample using these weights. This re-sampling process will result in an unweighted sample. After we conduct the re-sampling process, original data points that have higher weight appear more frequently. This can cause problems if we then want to weight the re-sampled data points as the sampled data points populate the distribution in a sparse manner. After multiple iterations this can result in a high probability mass concentrated on a few points. We counter this by introducing a Monte Carlo re-sampling technique that has the effect of perturbing these concentrated data points. The resulting data set is a weighted sample from the posterior distribution including all sets of experimental data [131]. Through the parallelisation of tasks this can increase the computation time when taking advantage of a previously obtained sample.

5.2.2 The Sequential Monte Carlo Algorithm

When we consider including multiple experiments we will often have different solution vectors for each experiment. TGA experiments are typically conducted with multiple heating rates, changing one of the parameters in our model. Additionally, for our application we implement a condition upon the initial masses of each reactant. This is a necessary step as the material is heterogeneous and we do not know the initial masses. Since we are restricting the data from the TGA to a specified time period, this adds further uncertainty to our initial masses. The result of these parameter differences between each experiment

means that each experiment has a different solution vector. Subsequently we let $\mathbf{F}_i$ denote the solution vector corresponding to the ith experiment.

We can split the Sequential Monte Carlo Algorithm up into three stages:

1. Sampling using the MCMC algorithm defined in Chapter 4.

2. Weighting the samples using importance sampling.

3. Resampling using a Metropolis-Hastings algorithm.

During the first stage we only consider the data from the first experiment which we denote $\mathbf{y}_1$. Using Algorithm 1 in Chapter 4, we obtain a sample, $\{\theta_j\}$ from the posterior distribution $p(\theta \mid \mathbf{y}_1)$. The second step is to use importance sampling to weight these particles. With $\mathbf{F}_2(\theta)$ denoting the solution vector for the second experiment, then for our likelihood, $p(\mathbf{y}_2|\theta)$, we have the data $\mathbf{y}_2$ following a Guassian distribution with mean $F_2(\theta)$ and standard deviation σ. This likelihood is the same for each experiment meaning we have y_i distributed according to,

$$\mathbf{y}_i \sim \mathcal{N}\left(\mathbf{F}_i, \sigma^2\mathbf{I}\right),$$

where $\mathbf{F}_i$ is the corresponding solution vector for the ith experiment. The notation we use is consistent with a single data set from the experiment, this is easily extended to the two data sets we observe in the TGA/DSC data.

At each iteration, each data point within the sample $\{\theta_j\}$ is assigned a weight w_j such that $w_j = p(\mathbf{y}_i \mid \theta_j)$. After the weights have been calculated then calculate the effective sample size (ESS). The ESS is calculated differently for the SMC algorithm then it was for the MCMC algorithm, though the concept is the same. We calculate ESS using the weights of the sample where,

$$\text{ESS} = \frac{(\sum_i w_j)^2}{\sum_i w_j^2}. \tag{5.4}$$

If the ESS is to low then there are degenerate point within our sample and we re-sample using a MH algorithm. After resampling we produce an unweighted sample meaning we rest our weights $w_j = 1$.

The MH algorithm we use is the same as we used in the previous chapter. The key difference is the initialisation step and the results we obtain from that. We initialise our algorithm by resampling using the normalised weights $\tilde{w}_j$. The weights are normalised such that $\tilde{w}_j = w_j/\sum_i w_i$. This allows us to re-sample the point θ_i such that, $P(\theta_i = \theta_j) = \tilde{w}_j$. This results in an unweighted sample from the posterior distribution $p(\theta \mid \mathbf{y}_{1:i})$.

We apply the algorithm differently within our SMC algorithm as we did previously. The goal of our MH algorithm here is to re-sample from the posterior to reduce the degeneracy in our previous sample. As long as the ESS isn't too low, then we still have a

sample representative of the posterior distribution. This means that initialising using this sample, then we are initialising from the posterior distribution. Since we have initialised according to the posterior distribution, then we do not have a burn-in period and this makes it computationally efficient to simulate many chains. With these multiple chains we can compute this efficiently by running our algorithm in parallel. Given that we want to generate a new independent sample to then re-weight, then we only take the last value in our small Markov Chain. If we simulate these chains for just long enough then we will obtain a sample that is independent and computationally efficient.

We build our MH algorithm using information from our sample to implement into our proposal distribution. We use the covariance matrix from the sample as the covariance matrix for the random walk proposal, that is,

$$\theta^* \sim \mathcal{N}\left(\theta_j, \Sigma\right), \tag{5.5}$$

where Σ is the sample covariance matrix. Since our sample is representative of the posterior distribution then the covariance matrix allows new proposed values to be more readily accepted and enables the algorithm to explore the parameter space more efficiently. This proposal also has the advantage that it is symmetric meaning it can be excluded from the calculation of the acceptance probability. Our acceptance probability π, includes the information from each set of experimental data that has been considered. After m experiments, the acceptance probability is calculated by,

$$\pi = \frac{\prod_{i=0}^{m} p(\mathbf{y}_i \mid \theta^*) p(\theta^*)}{\prod_{i=0}^{m} p(\mathbf{y}_i \mid \theta_j) p(\theta_j)}. \tag{5.6}$$

In this case we conduct ten iterations of this MCMC process as we only take the last sample. This method ensures that we maintain the same sample size and that after resampling we have an independent sample. This method is computationally efficient as we can run the MH algorithms in parallel.

5.2.3 Simulation Study

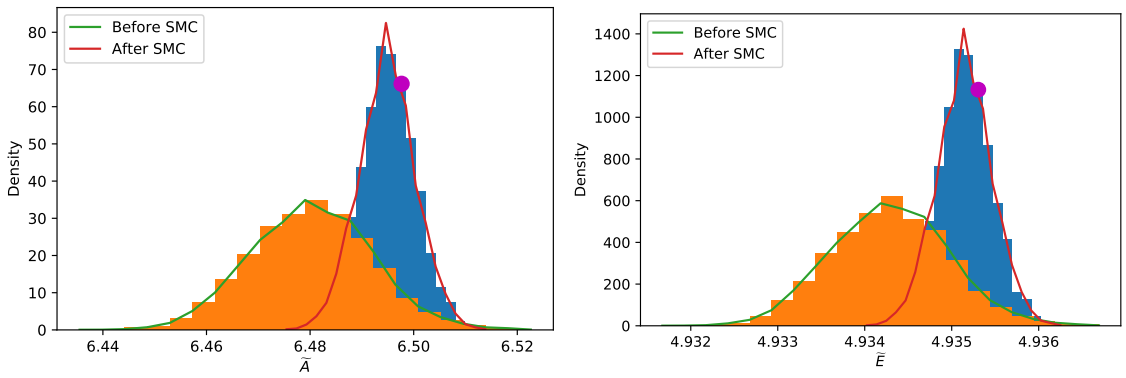
To investigate the ability for this algorithm to achieve our target goals of including as much possible data as we can and to improve the uncertainty estimation in our models, we first conduct a study on simulated data. The simulated data has known parameter values and we also have a perfect predictive model whilst our experimental data lacks this. We set up our investigations using the simplest case where we consider a one reaction model with only the mass data. We use the simulation data generated in section 4.2. We then generate additional data with heat rates of 5 and 20^{o}C/min. The result is a data set that mimics the experimental data that has been conducted.

With our simulated data there was no significant degeneracy requiring re-sampling be-

Algorithm 3: Sequential Monte Carlo Algorithm.

Initialise the sample $\{\theta_j\}$ by conducting a MH-MCMC algorithm using one Data set;
Set $w_j = 1.$;
for *i in 1:m* **do**
  $w_j = w_j \times p(\mathbf{y} \mid \theta_j)$;
  $ESS = \frac{(\Sigma w_j)^2}{\Sigma w_j^2}$;
  if $ESS < ESS_{crit}$ **then**
    Resample using a MH Algorithm;
    for *k in 1:10* **do**
      Propose $\theta^* \sim \mathcal{N}(\theta, \Sigma)$;
      Set $\pi = \frac{\prod_{i=0}^{m} p(\mathbf{y}_i|\theta^*)p(\theta^*)}{\prod_{i=0}^{m} p(\mathbf{y}_i|\theta_j)p(\theta_j)}$;
      Simulate $\beta \sim \mathcal{U}(0,1)$;
      if $\pi > \beta$ **then**
        Set $\theta_j = \theta^*$
      end
    end
  end
end
Re-sample using the weights to obtain an unweighted sample.

tween the reactions. This is ideal in that it allows for an efficient algorithm. We compare the histograms from after the MCMC process on one set of experimental data, before we apply the SMC algorithm using multiple data sets to the histograms obtained after we complete the SMC algorithm.



(a) Comparison of $\widetilde{A}$.

(b) Comparison of $\widetilde{E}$.

Figure 5.1: Comparison of the sampled values before the SMC algorithm is applied to after the SMC algorithm has been applied for the reaction Kinetic parameters.

Figure 5.1 compares the pre-exponential factor and the activation energy before and after including the additional experiments. We observe in this figure, that utilising the SMC algorithm and the additional experimental data, the posterior distributions narrows

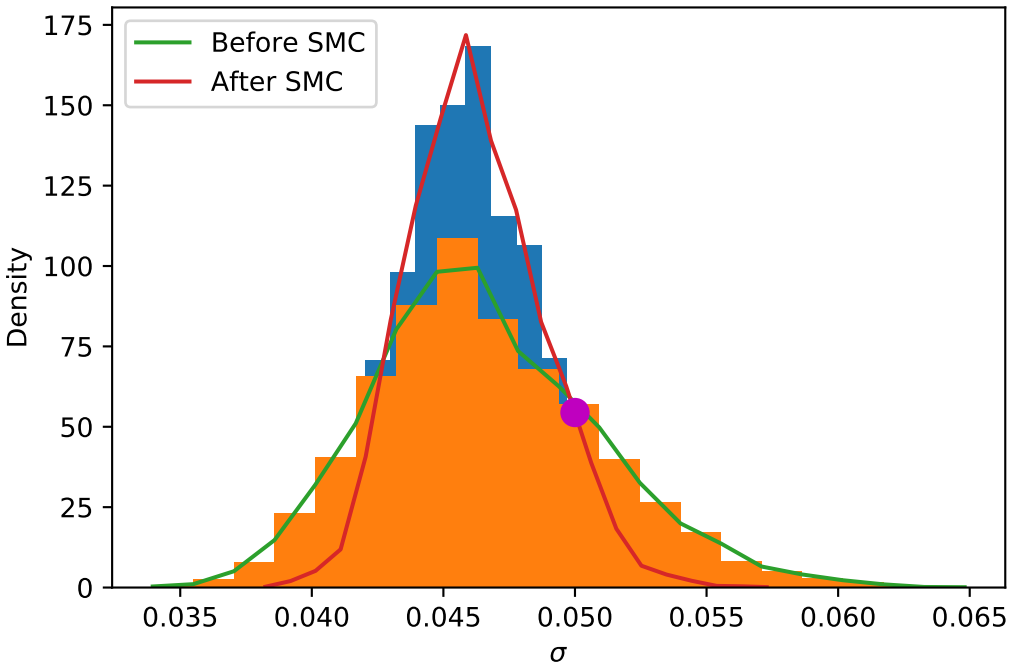


Figure 5.2: Comparison of the noise parameter σ, before and after conducting the SMC algorithm.

around the true values. This result clearly indicates that as more data is added then the uncertainty in our estimates reduces. In Figure 5.2, we observe a similar narrowing of the distribution for the noise parameter σ, although in this case the distribution under predicts the noise. The distribution still contains the noise parameter justifying the use of this algorithm.

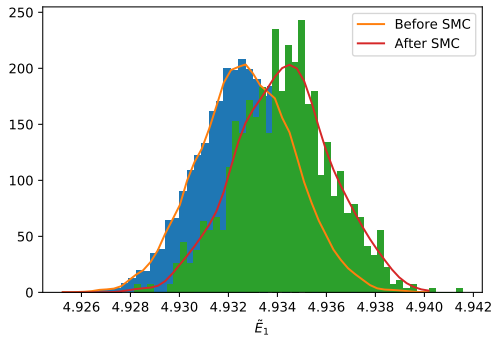
When we apply the SMC algorithm for multiple experiments, we do not use the temperature at which the reaction rate is maximised, T_m within our algorithm. This parameter is dependant upon the heating rate of each experiment and as each set of experiments have a different heating rate the T_m parameter changes. This makes the T_m parameter difficult to use in the SMC algorithm and we lose the meaning of this when applied to additional experimental data sets. The need for the T_m parameter in the proposal stage is also reduced as we have already established a sample from the posterior distribution and hence the correlations between parameters can be calculated and their effect included in the proposal distribution. This allows for a more efficient random walk proposal since each parameter is no longer proposed independently. Care needs to be taken when using the priors since we established our posterior distributions using the prior for T_m.

5.2.4 Two Reaction Model

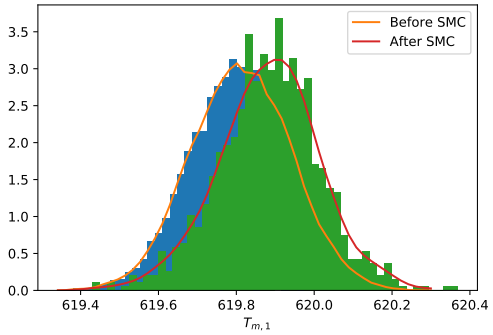
In addition to the one reaction model we test our methodology out with two reactions. We simulate the four experiments that are conducted with the filter cake; two experiments at

a heating rate of 10^{o}C/min with additional experiments at 5^{o}C/min and 20^{o}C/min. With the two reactions we found that it was necessary to conduct our MCMC step within the algorithm due to having the probability mass concentrated. In this application the ESS that we obtained was 70 after determining the weights including all reactions. This was deemed too low and subsequently we had to implement the MCMC resampling step. In this instance we resampled using 1000 points sampled according to the posterior distribution after three experiments were included as including the fourth had too low of an ESS.

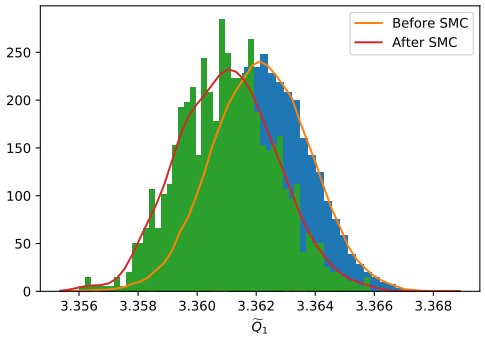
We compare the posterior distributions for the first reaction in Figure 5.3. We observe that there is a difference between the two distributions. Figure 5.3b indicates that the centre of the posterior distribution is closer to the true value, $T_{m,1} = 620$. Similarly with the activation energy and heat generation we see a marginal improvement. One key difference between our results for the two reaction case and the one reaction case is we don't observe the same narrowing of the distribution. Given we have the additional data it is still more useful to include this data. This may suggest that for our particular simulated example, the additional simulations may not be adding additional information.



(a) Distribution of $\widetilde{E}_1$.



(b) Distribution of $T_{m,1}$.

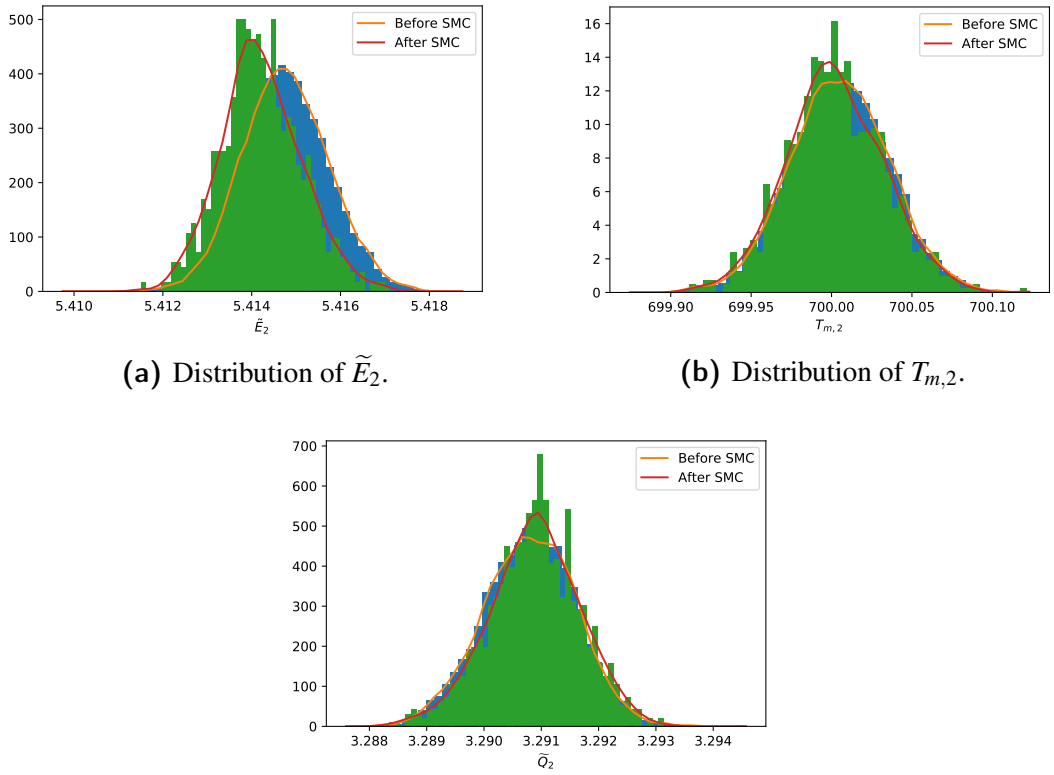


(c) Distribution of $\widetilde{Q}_1$.

Figure 5.3: Comparison of the posterior distribution after one experiment and after including multiple experiments for the first reaction.

The effects on the second reaction in Figure 5.4 are less than those we observed in with the first reaction. The distributions for $T_{m,2}$ and $\widetilde{Q}_2$ barely change and could likely be

attributed to sampling variance. There appears to be a slight narrowing of the distribution for the heat generation. These two parameters are the main cause of the large peak in the DSC data and it appears as though this can be well approximated using a single experiment. THe activation has greater affect from the additional reaction that occurs and then as the parameters for the first reaction changes then it is expected this should too.



(a) Distribution of $\widetilde{E}_2$.

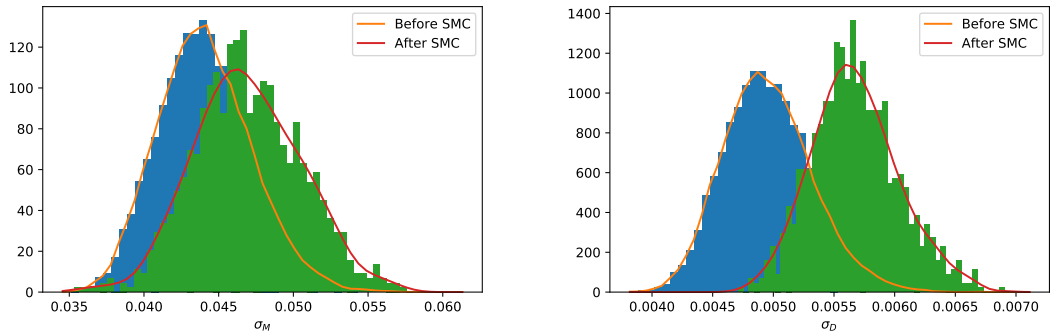
(b) Distribution of $T_{m,2}$.

(c) Distribution of $\widetilde{Q}_2$.

Figure 5.4: Comparison of the posterior distribution after one experiment and after including multiple experiments for the second reaction.

The greatest difference between the posterior distributions is observed in the noise parameters. Figure 5.5 displays just how much these parameters change. The posterior distributions for both parameters increase after the inclusion of additional experiments. This suggests that all the experiments are necessary to obtain accurate estimates as any one reaction may overfit the parameter values for that reaction and underestimate the noise parameter.

We have established that our SMC algorithm is useful for incorporating multiple sets of experimental data. In both the one and two reaction case we were able to obtain distributions that had refined our estimates of the various model parameters. These simulations have the key advantage that the model predicts the data completely. For the experimental data we do not have this luxury and various issues may arise as a result of that.



(a) Distribution of σ_M.

(b) Distribution of σ_D.

Figure 5.5: Comparison of the posterior distribution of the noise parameters between one experiment and after including multiple experiments.

5.2.5 Application to Experimental Data

When we apply our methodology onto the experimental data we have to be careful about our initial conditions. Since the composition of the filter cake varies, then the initial concentrations can vary between different samples. We resolve this by considering each experiment to have its own set of initial conditions. Our initial conditions are selected using the same methodology we used for the initial masses in Section 4.4. This method forces the initial and final sample mass to correspond. Using this method we get the initial conditions stated in table 5.1. We observe that the proposed initial conditions are quite similar and reflect the uncertainty estimates for these parameters in the work conducted by Longbottom et al. [7].

Heating Rate	% Moistureless Mass	% Wüstite	% Iron
5	85.17	43.5	11.0
10	84.64	42.8	10.8
10	84.17	41.2	10.9
20	85.54	43.3	11.0

Table 5.1: Initial Conditions for the experimental data.

As noted in the previous section that the T_m parameter is specific to the heating rate. For our application we use the temperature at which the heating rate is maximised for only the heating rate of 10^oC and this then loses its meaning when applied to the additional reactions. This is one draw back from using these parameter choices though we accept this due to the large benefit in the initial Metropolis–Hastings algorithm.

During our algorithm we re-weight the sample using the repeated experiment with $\alpha = 10$. After this re-weighting step we have the large majority of the probability mass concentrated on a single observation. After re-sampling using this weighting scheme, our sample contains only a single observation. This observation is the one whose mass noise σ_M is the greatest. This indicates that the estimates we obtain from the first experimental

data, do not describe the results of the second experimental data set well enough. Considering the experiments are conducted under the same heating rate, then theoretically if the material is identical then any difference in results are due to experimental error. Given that the filter cake is known to have a varying composition, then we conclude that the variation in initial masses is causing such a discrepancy. This is either the result of uncertainty in the underlying measurements, or it could be the result of using an incorrect reaction scheme. We observe similar behaviours when we re-weight using the other experiments.

Due to this issue with re weighting, we extend the number of iterations in the SMC resampling step. Rather than perturbing the points slightly using ten iterations we extend this to 200 iterations. This allows the algorithm to explore the joint posterior distribution using these starting points.

5.3 Alternative Approaches

Using the SMC algorithm for the experimental data has indicated that the reaction scheme we have used may not be sufficient in interpreting all of the experimental data. We examine this hypothesis by conducting inference on each experiment independently using our MCMC algorithm in the previous chapter. This will enable us to observe if there are any patterns that we observe between experiments and will enable us to identify if any of the marginal posterior distributions overlap.

Figure 5.6 clearly indicates that there is a large discrepancy between the estimated parameters within our model. This provides a strong indication that our approach is being limited in some capacity. There are four causes that we have identified to be causing these issues:

- Multi-modality in the posterior distribution.
- Sensitivity to the prescribed initial conditions.
- Miss-specified reaction Scheme.
- Variation in the parameter values between experiments.

We can investigate the effects that each of these can have and propose methods for overcoming these issues.

5.3.1 Multi-Modality

Multi-Modality is an issue that we have previously come across when estimating our parameter values. Given that our posterior distribution combines likelihoods from two sets of experimental data, it is unsurprising that there is a slight variation with the best fit for the parameter values. Our approach to potentially overcome this issue is to change

our prior for the joint distribution of multiple experiments. Rather than using our previous uninformative prior, we use the samples from the posterior distribution from each of our experiments, to form a new prior. Equation (5.1) remains the same but we use the prior,

$$\theta \sim \mathcal{N}(\mu_\theta, \Sigma_\theta),$$

where μ_θ is the sample mean of the combined sample from the previous MCMC algorithm, and Σ_θ is the combined sample Covariance matrix. With this prior we then conduct importance sampling to sample from the posterior distribution. We generate a sample θ_i from our prior distribution $\mathcal{N}(\mu_\theta, \Sigma_\theta)$. We then weight each observation using the likelihood function,

$$w_i = p(\mathbf{y}_1 \mid \theta_i) p(\mathbf{y}_2 \mid \theta_i).$$

As with our previous application of importance sample, redundancies can occur and the weights can be concentrated on a small number of particles. We monitor this by calculating an effective sample size (ESS). The ESS has the same conceptual basis as that which was used for the MCMC algorithms though the calculation is different. We calculate this with the formula,

$$\text{ESS} = \frac{(\sum_i w_j)^2}{\sum_i w_j^2}. \tag{5.7}$$

From this we observe that if all the points are equally weighted then our ESS will be the same as the original sample size. This formula also indicates the opposite where if one weight is much greater than any other, then the ESS approaches one.

We apply this joint posterior approach to two sets of experimental data. We focus on the two reactions that are conducted at the same heating rate. These experiments are selected since theoretically the only difference in the data is the slight differences in the sample of filter cake used. This will then limit the effect of a miss-match in the choice of reaction scheme. From Figure 5.6, we also observed that the samples from the two experiments with a 10^oC heating rate were consistently the closest.

Using our importance sampling algorithm we calculate ESS$=$ 87. Given that this is from an initial sample size of 100000, then we are observing many degenerate points. We can resolve this by tempering the posterior distribution. In this method a tempering parameter β is introduced. The weights are calculated as $w_i = (p(\mathbf{y}_1 \mid \theta_i) p(\mathbf{y}_2 \mid \theta_i))^\beta$. We let β vary between 0 and 1. This tempers our sample such that if $\beta = 0$, then we have the prior distribution, and $\beta = 1$ denotes the posterior distribution. This parameter can be selected adaptively by the sampling algorithm [132]. Using this we can change the parameter beta and by re-sampling at various stages, we can achieve a sample from the posterior distribution with a reasonable ESS. This tempering process will become extremely important if we look to apply this method with multiple reactions as this will

result in a broader prior distribution and can result in more redundant points sampled from the prior. Such methods have been mathemaically justified [133] and applied in previous inverse problems [134].

Persisting with our methodology, we can visualise the effect of resampling on the posterior distribution. Due to the low ESS posterior estimates are not reasonable. Figure 5.7 provides a comparison of the posterior distribution of the results from the independant MCMC algorithms and the joint posterior distribution using the SMC algorithm. This clearly indicates that the probability masses of each distribution are concentrated with no significant overlap. In this instance the marginal joint posterior distributions are between the posterior distributions from each experiment.

For the Iron reaction we observe different behaviour of the posterior distributions. Figure 5.8 indicates that the joint posterior distribution for the heat enthalpy Q_2, is not between the two posterior distributions from the MCMC algorithm. The noise parameters in Figure 5.9 is slightly different. There is significant overlap in the posterior distribution of the noise parameter for the DSC whilst the posterior distributions for the noise in the mass data is close but doesn't overlap.

It is important to note that the ESS of the joint posterior distribution is 87, and subsequently it is expected that the sample will not contain many points outside of the regions of high probability. Similarly the kernel density estimates for the distribution along the diagonal is not as smooth as the MCMC posterior distribution because we have a small ESS.

Using information from the MCMC algorithms for each independent reaction has enabled us to propose a new informative prior on the joint posterior distribution. Using a normal distribution in this way enables importance sampling to be used to weight the particles sampled from the normal distribution. We have shown the algorithm samples from the joint posterior distribution and has provided a small sample. The algorithm can be further improved through the use of tempering to increase the ESS. Whilst our algorithm provides a mechanism to adjust for these disjoint posterior distributions from the MCMC, it does not address why there is the discrepency. Given that theoretically the difference between the two experiments at a 10^oC heating rate are only dependant upon the sample, then we still want to explore why these results are different.

5.3.2 Sensitivity of Initial Conditions

As part of our analysis we must prescribe initial conditions for each of the reactants. The concentrations of the reactants have uncertainty over their values. In addition to this there is known variation of the concentration within the filter cake and this will then affect the concentration of each experiments sample. Another source for uncertainty is that we do not model the reactions from the beginning of the experiment. It may be the case where

some unknown quantity of material has previously reacted that we do not account for. It is for these reasons why we place the restriction upon our choice of initial conditions. If our method for estimating the initial concentrations are not accurate, then this could explain some of the variation in the posterior distributions from each reaction scheme. Given that the samples and their initial conditions are the main difference between the two experiments with the same heating rate, then this suggests that there is some effect from the choice of initial conditions.

We test how sensitive the posterior distributions are to the choice of initial conditions, by comparing the posterior distributions for different initial conditions. We use four different initial conditions. We use the vector of initial masses we calculated for the MCMC algorithm, denoted, $\mathbf{IC}_0$ and the initial masses determined experimentally [7], denoted, $\mathbf{IC}_3$. In addition to these two values we select the two points, $\mathbf{IC}_1$ and $\mathbf{IC}_2$, such that the restriction on the initial values is maintained, and that $||\mathbf{IC}_0 - \mathbf{IC}_3|| = ||\mathbf{IC}_0 - \mathbf{IC}_1|| = ||\mathbf{IC}_0 - \mathbf{IC}_2||$ This provides a rather broad selection of initial conditions but this is useful when investigating the sensitivity of the initial conditions.

Our results in Figure 5.10 indicate that the initial condition have a significant effect on the posterior distribution. This implies that we either require more accurate data about the initial masses of each reactant or we can change our approach to using the initial masses. Additionally it appears that the discrepancy between the posterior distribution for each of the two experiments could be attributed to the difference in initial conditions. It is logical that this would be the case since the only difference between the experiments is the sample.

With the filter cake being heterogeneous, there are significant limitations to the accuracy at which we can calculate the initial masses. As mentioned previously even if the initial sample masses would be known with a very high accuracy for each sample before the experiment is conducted, our simulated experiment has a different starting point so that we can ignore the effect of the moisture drying. Given that the heating rate is less stable at some of the lower temperatures then matching the data becomes more difficult if we begin the simulation from the beginning. This then limits the information we can obtain about the initial mass of each reactant at the point we are beginning our simulation.

The best method we have for addressing these initial conditions is to include them as parameters that need to be inferred. This allows our algorithm to infer these parameter values rather than have them prescribed. This will increase the complexity of the model and can make computation more challenging. With the experimental data available we would be able to use an informative prior for these initial masses.

5.3.3 Alternate Reaction Schemes

One possible explanation for the results we observe in figure 5.6, is that the proposed reaction scheme is incorrect. We selected the particular reaction scheme as a two-reaction scheme was the simplest way to fit the data, given the experimental data indicated at least two significant reactions were occurring. The particular reactions were selected based upon the advice of experimentalists and their results. Changing the reaction choice within the two reaction model would alter the weight change coefficients and using our method for calculating the initial conditions, they would be altered as well.

If we were to persist with a two-reaction scheme, there are three ways we propose to alter this. Firstly we could select different reaction schemes and test whether these produce better results when applied to multiple sets of data. Such a method can be computationally expensive. Secondly, we can change the reaction order of each reaction. We have assumed first order Arrhenius reactions though the experimental data may be explained by much higher order reactions. Finally, we could infer the weight change parameters. This approach would leave the reactions arbitrary and we would determine which weight change values best fit the data and then use that information to select an appropriate reaction scheme. Whilst it may produce the ideal weight change values, this may not correspond to any physical reaction and thus cannot be reconciled with the experimental data.

Alternatively to a two-reaction scheme we can extend our reaction scheme to include a third reaction. When adding an additional reaction we introduce more parameters into the model increasing its complexity. This enables our model to fit the data better but this also increases the complexity of our posterior distribution. Adding a third reaction increases the possibility of additional modes occurring in the posterior distribution. This will complicate the MCMC algorithm and can induce additional inefficiencies within the algorithm. In addition, we have reduced our prior knowledge of the T_m parameters by including an extra reaction. The experimental data defines two clear reactions but the location of the third reaction is unclear. The third reaction also decreases the confidence we have in the peak locations. Adding a third reaction also adds new initial conditions that may need to be taken into consideration. Since we consider a reaction of two solids where there is no mass change, then this condition isn't as important for a single reaction. Since we can scale the initial mass out of the mass equation, then the initial mass will be absorbed into the heat of reaction Q. We assume an Arrhenius reaction where the mass of the second reactant does not change. This is not a great assumption but due to limited information and the need for a simplified model then use it. Adding the third reaction provides a means to examine the credibility of this third reaction without putting too many limitations on the exact reaction and over-complicating the model.

We extend the model to include the equation,

$$\frac{dM_3}{dt} = -M_3 A_3 \exp\left(\frac{-E_3}{RT}\right), \tag{5.8}$$

for our new reaction. The DSC equation then becomes,

$$H = Q_1 M_1 A_1 \exp\left(\frac{-E_1}{RT}\right) + Q_2 M_2 A_2 \exp\left(\frac{-E_2}{RT}\right) + Q_3 M_3 A_3 \exp\left(\frac{-E_3}{RT}\right). \tag{5.9}$$

Such a model has greater ability to explain the data.

Using our MCMC algorithm for the thre reaction model we find that the additional reaction is having minimal contribution to the data. The diagnostic statistcs in Table 5.2 indicate that the chains have not converged. The posterior distribution is extremely complex and more tuning is required for the proposal distributions in order to obtain a high quality sample.

	Split-$\widehat{R}$	ESS
E_1	40.14674	4.0
Q_1	4.03305	4.0
$T_{m,1}$	145.63248	4.0
E_2	68.60897	4.0
$T_{m,2}$	359.18913	4.0
Q_2	14.68980	4.0
E_3	1.20088	17.0
$T_{m,3}$	2.20414	6.0
Q_3	1.35211	13.0
σ_M	3.94838	4.0
σ_D	4.38695	4.0

Table 5.2: Diagnositc Results for the MCMC algorithm with three reactions

Persisting with this model understanding the significant uncertainty we increase the number of Markov Chains that we use. This allows the algorithm to have greater scope and capture the many different modes that can be present. The trace plots in Figures figs. 5.11 to 5.14 highlight just how complex this posterior distribution is.

For the first reaction in Figure 5.11 it is clear that there are at least 3 modes. We find that our MCMC algorithm has a strong preference for the Temperature at which the reaction rate is maximised is approximately 600K. Only one chain finds each of the other two nodes. It is likely that these Markov Chains found these nodes because of particular initial values. Looking carefully there appears as though there are two nodes appearing at the lower temperature of 600K. The activation energy and subsequently the pre-exponential factor also display the same multi-modality. The heat generation parameter is slightly more interesting since we start to see some of the separation within chains between the

various chains.

The trace plots for the second reaction in Figure 5.12 correspond to what we observed with the first reaction. The temperature at which the reaction rate is maximised shows a clear mode where most of the mass is concentrated. We observe that one chain samples a mode where the two reactions merge together and another chain samples a mode where the reaction order is swapped. When we observe the activation energy and heat generation parameters we see that within this large mode for the temperature that the reaction rate is maximised this mode consists of two separate modes.

The trace plots for the third reaction in Figure 5.13 are perhaps the most interesting. The chain corresponding to the pink line appears as though it has not reached stationarity. This particular chain has such a low activation energy that the pre-exponential factor is very small. A pre-exponential factor that is very small will mean that the reaction is very slow and as such minimal heat generation will occur. This suggests that it is not influencing the model. Aside from this chain we see the emergence of three distinct modes. Given that this reaction is not affected by the order of the other two reactions, when this order swaps then we do not see an obvious change in the parameters for the third reaction.

The trace plots for the noise parameters in Figure 5.14 provide some valuable insight to what each of these modes are explaining. We can see from figure 5.14a that the grey Markov chain which combines the two reactions is associated with the largest amount of noise for the mass data. The green chain where the reaction order is reversed also produces a larger estimate for this noise parameter. The remaining chains all produce a similar noise associated with the data indicating that they all provide a similar fit for that data. The noise parameter for the DSC data sheds more light on what is occurring with these multiple modes. The pink and orange chains were identified as where the third reaction has negligible impact. We then see a split in the main modes that appeared from the $T_{m,1}$ and $T_{m,2}$ parameters. These modes correspond with those that are seen easily in Figure 5.13c. The mode that corresponds to $T_{m,3} \approx 700K$ corresponds to the lower DSC noise, whilst the mode with $T_{m,3} \approx 760K$ corresponds to the higher noise. Since in our application the noise parameter is used to consider the fit of the data, then it would be worthwhile using this posterior distribution to obtain more information about the possible reaction schemes such that we can use more informative priors. It would then be possible to use the prior distribution to eliminate these additional modes and that way the MCMC algorithm can sample this main mode more efficiently.

We have claimed that the different modes approximate the data in different ways. Figure 5.15 compares the DSC data to a small sample of possible simulated DSC curves. We observe from this figure the effects of these different modes.

We have considered one of many different ways of changing the reaction scheme. Our initial investigations indicate that this reaction scheme has the possibility to better explain the data. More information about this third reaction would be required in order to sample

effectively. When considering different reaction schemes we must consider how we select which model provides the best fit. SMC methods are able to be used for this model comparison though this is outside of the scope of this thesis.

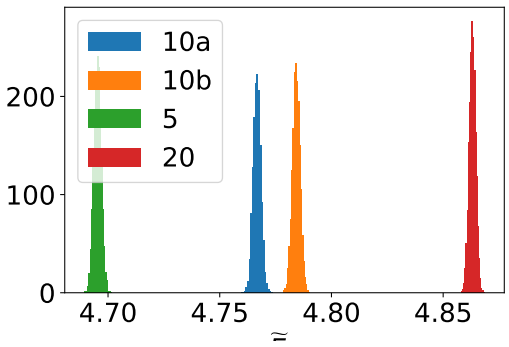
5.3.4 Hierarchical Models

One of the issues we identified was that the parameters may vary between different experiments. This has already been considered with the initial masses but what about the other parameters. The activation energy E, and heat of reaction Q are both physical parameters dependent upon the reaction and therefore it is more reasonable to fix these across the experiments rather than let them vary. However for the pre-exponential factor, A and subsequently the temperature at which the reaction is maximised T_m, we can allow this to vary between reactions. This pre-exponential factor is a fitted parameter and can be dependent upon the sample. We still expect that the pre-exponential factors for each experiment will relate to each-other, hence making a hierarchical model worth investigating. Similarly the initial conditions if we choose to infer them would be best set up in this manner as well.

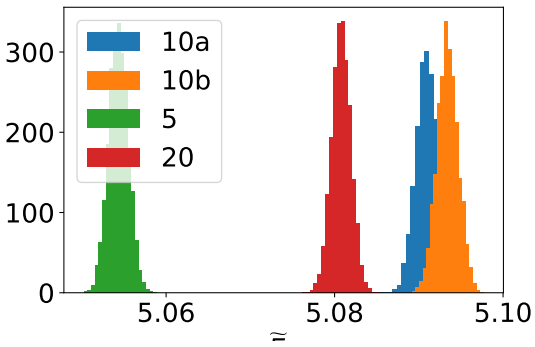
5.4 Summary

The sequential Monte Carlo Algorithm we have proposed successfully incorporates additional reactions into our model and narrows the posterior parameter distributions for simulated data. When this was applied to experimental data it concentrated around a single observation from the previous posterior distribution indicating that there was an issue with the data. Our re-sampling method produced a sample from the posterior distribution though the probability mass being heavily concentrated around values where the noise parameter was higher indicates there may be other issues with adding multiple experiments, though this isn't unexpected.

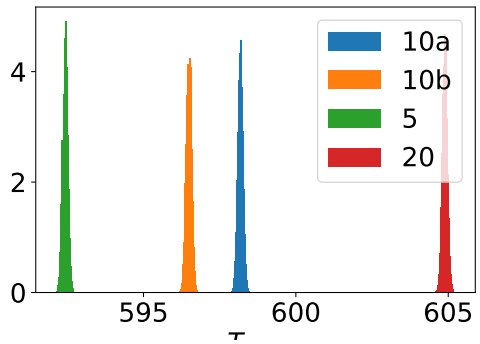
Our analysis indicated that there were issues with out application to multiple experiments as our sample would depend upon which experiment was sampled first using the MCMC method. This was demonstrated with only accounting for two experiments at the same heating rate. As such we didn't apply this method to incorporate all the experiments as we found the assumptions in our model and choice of reaction scheme produced vastly different posterior distributions for each set of experimental data. This indicates that more needs to be done to refine the choice of reaction scheme before our SMC method can be used effectively to combine the data.



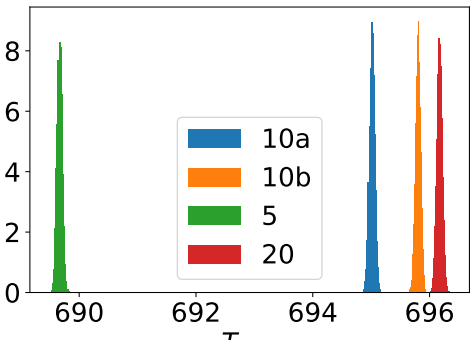
(a) Distribution of $\widetilde{E_1}$.



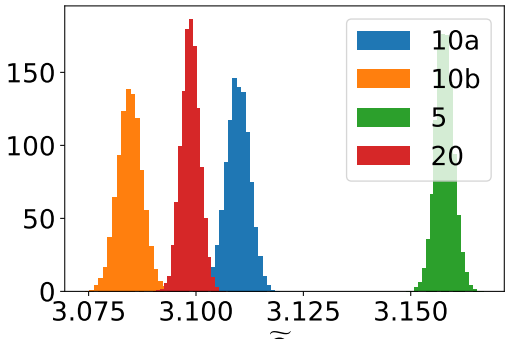
(b) Distribution of $\widetilde{E_2}$.



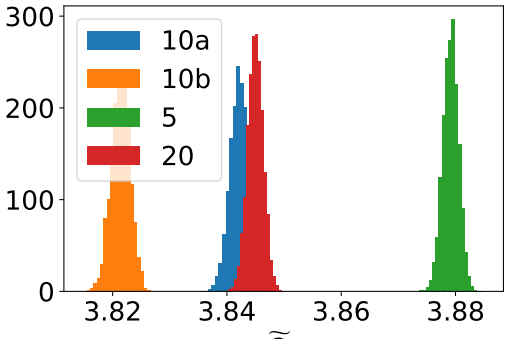
(c) Distribution of $T_{m,1}$.



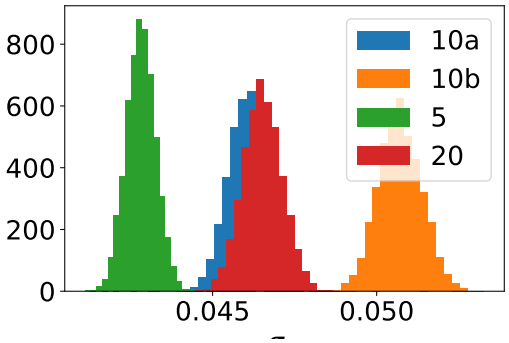
(d) Distribution of $T_{m,2}$.



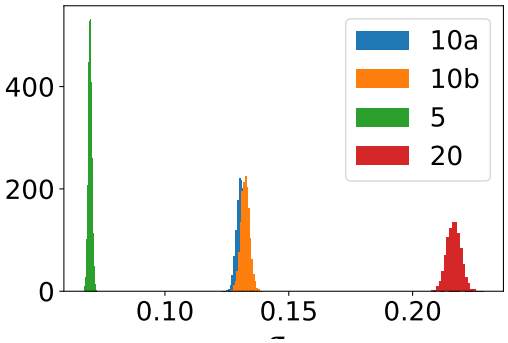
(e) Distribution of Q_1.



(f) Distribution of Q_2.



(g) Distribution of σ_M.



(h) Distribution of σ_D.

Figure 5.6: Comparison of the posterior samples using the different sets of experimental data.

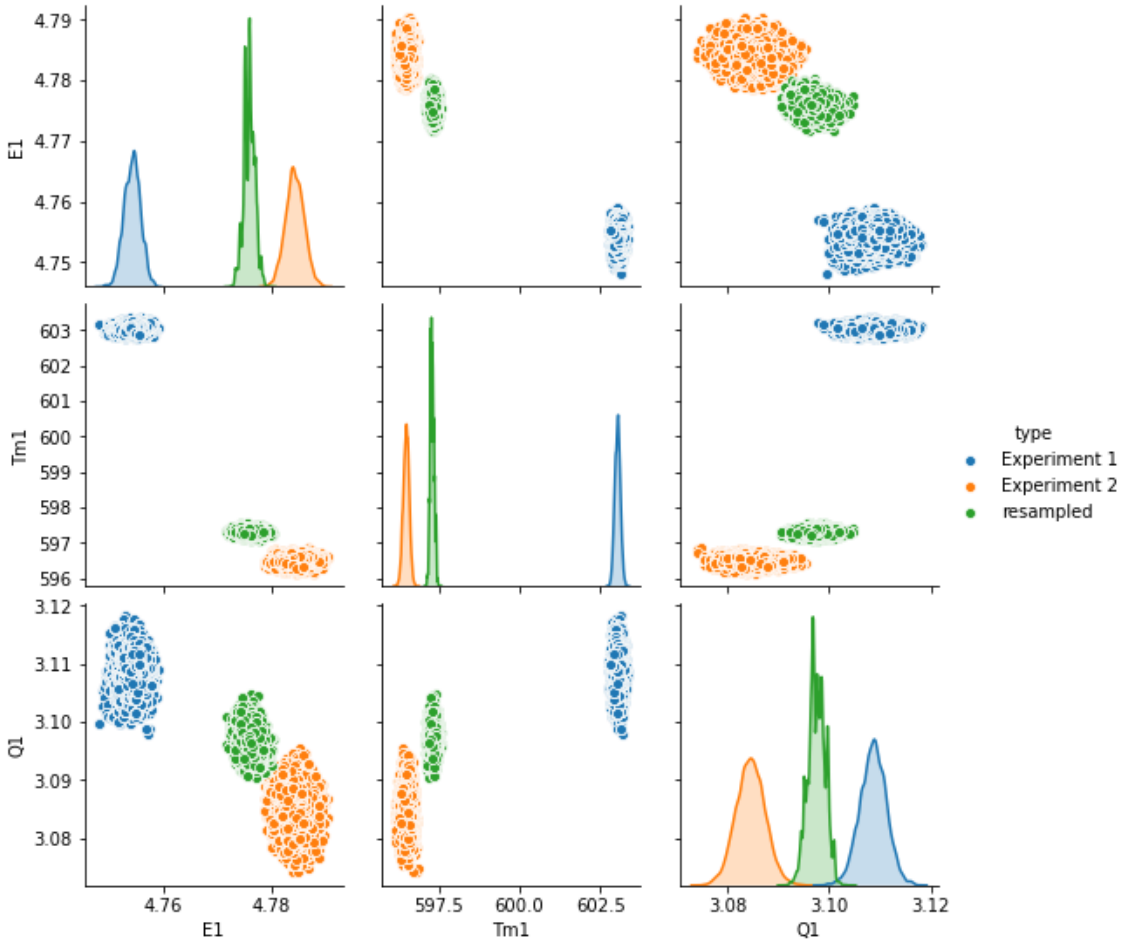


Figure 5.7: Comparison of the posterior distributions for the reaction kinetics of the wüstite reaction.

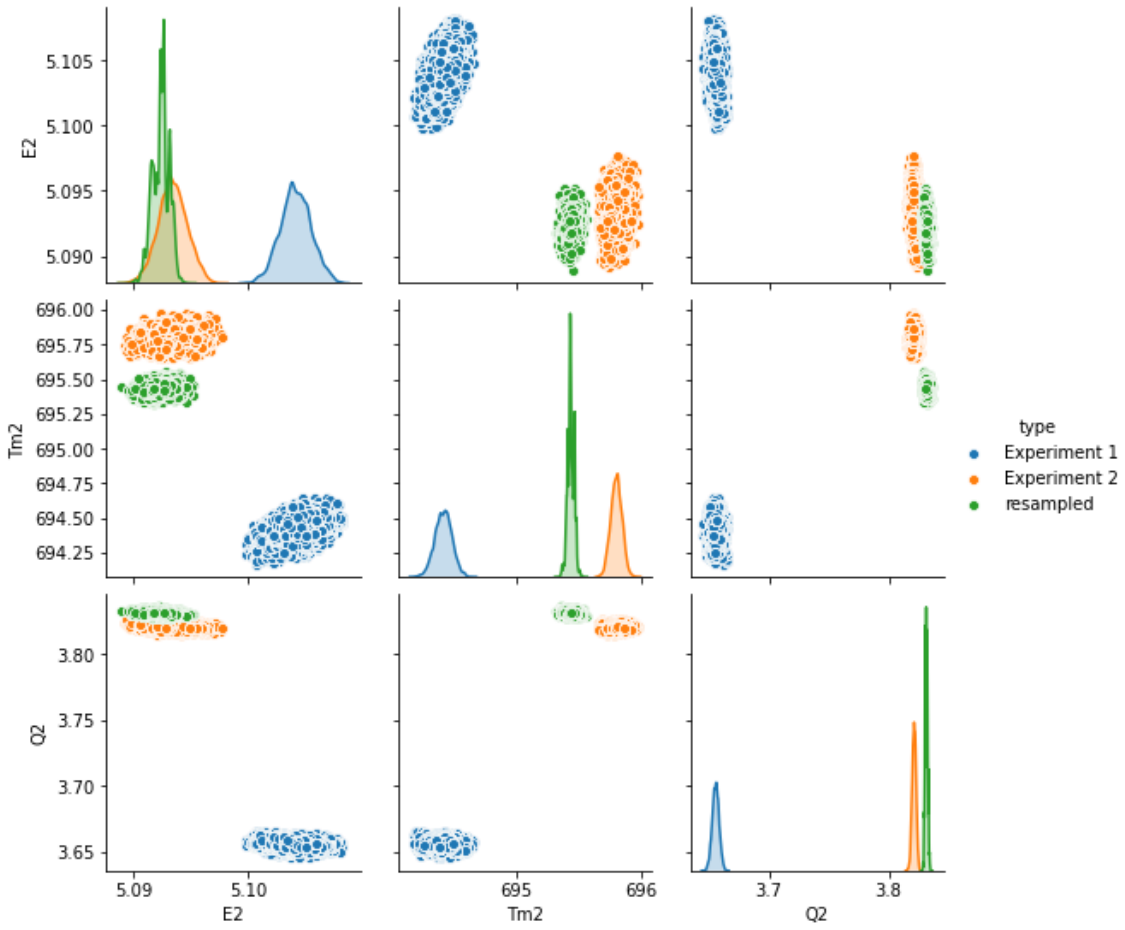


Figure 5.8: Comparison of the posterior distributions for the reaction kinetics of the Iron reaction.

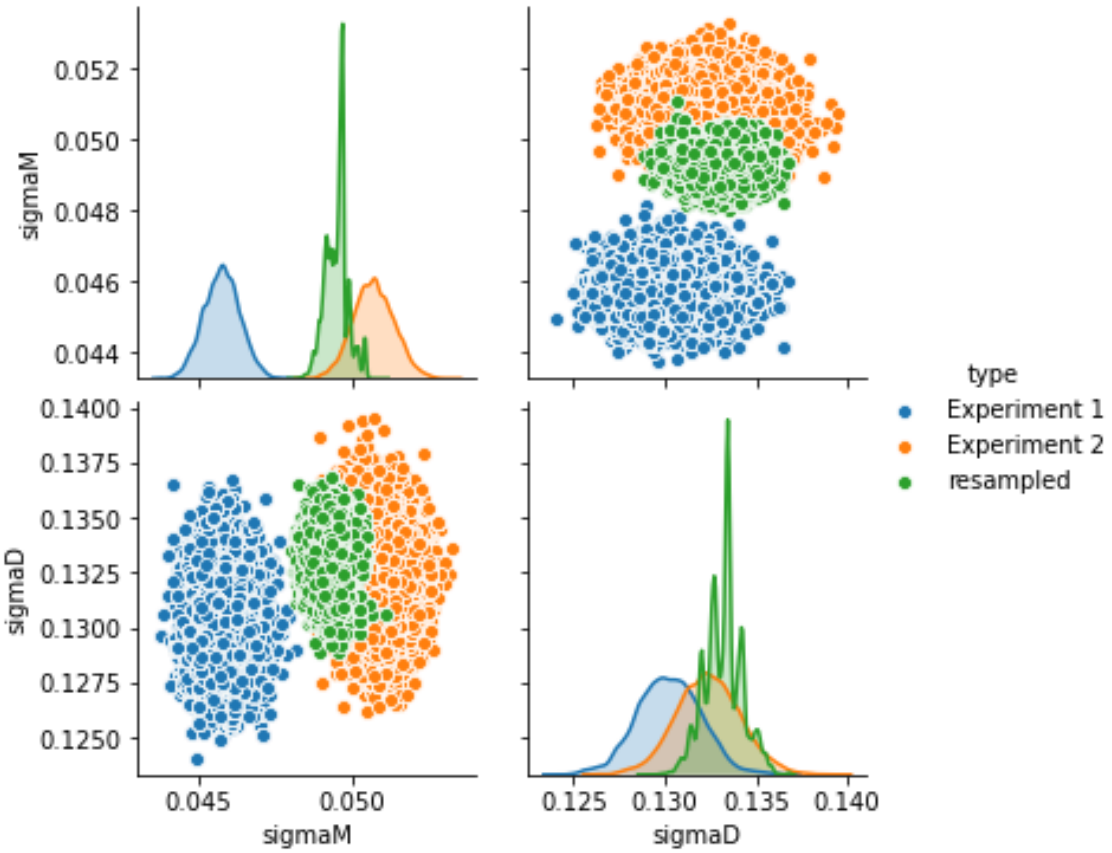
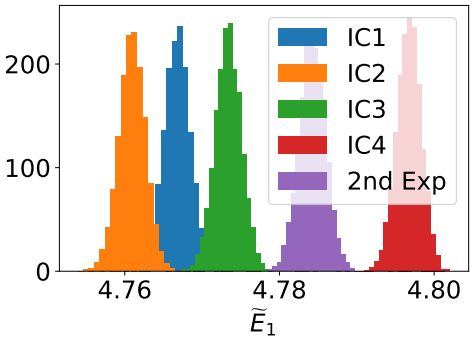
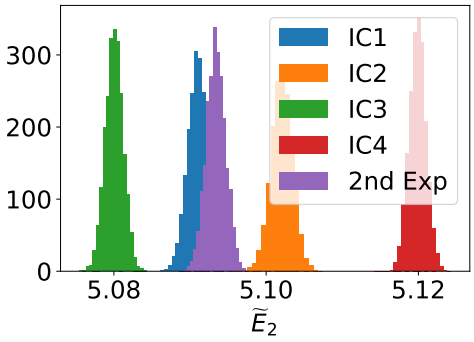


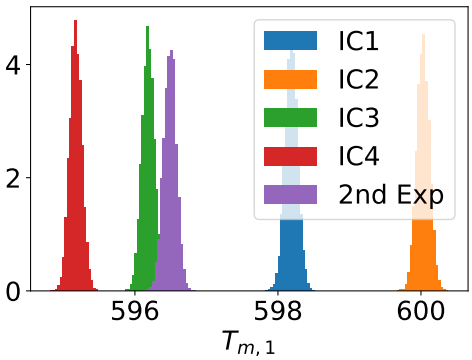
Figure 5.9: Comparison of the posterior distributions for the noise parameters of the experiments.



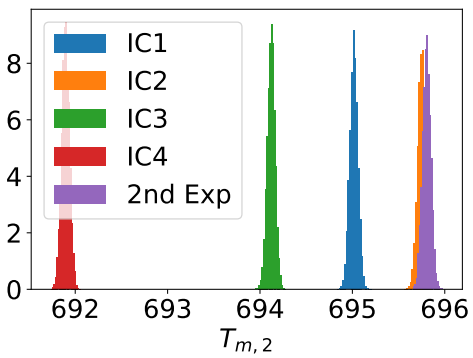
(a) Distributions of $\widetilde{E}_1$.



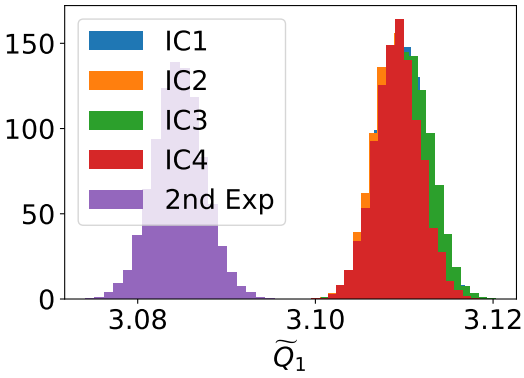
(b) Distributions of $\widetilde{E}_2$.



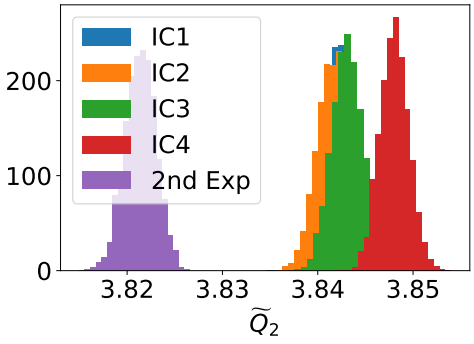
(c) Distributions of$T_{m,1}$.



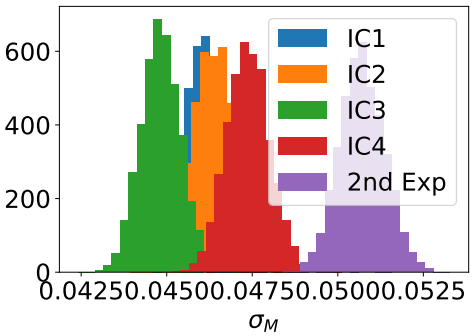
(d) Distributions of $T_{m,2}$.



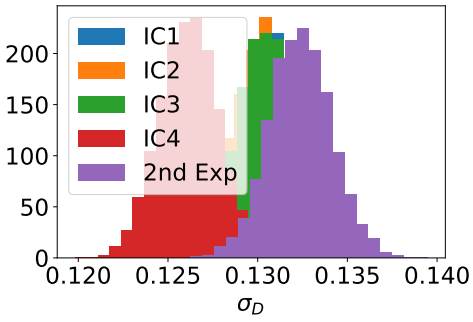
(e) Distributions of $\widetilde{Q}_1$.



(f) Distributions of $\widetilde{Q}_2$.

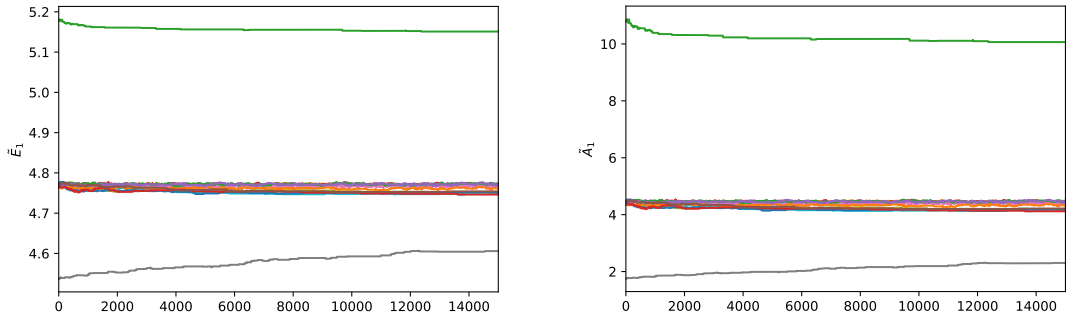


(g) Distributions of σ_M.



(h) Distributions of σ_D.

Figure 5.10: Comparison of the posterior distributions generated from each of the different initial conditions and the second experiment.



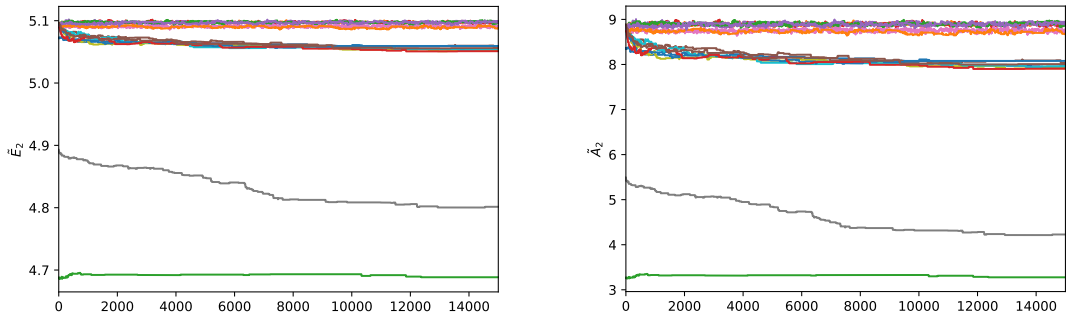
(a) Trace plot of $\widetilde{E_1}$.

(b) Trace plot of $\widetilde{A_1}$.

(c) Trace plot of $T_{m,1}$.

(d) Trace plot of $\widetilde{Q_1}$.

Figure 5.11: Trace plots for the kinetic parameters of the first reaction.



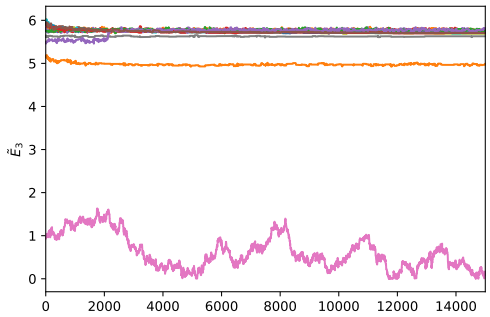
(a) Trace plot of $\widetilde{E_2}$.

(b) Trace plot of $\widetilde{A_2}$.

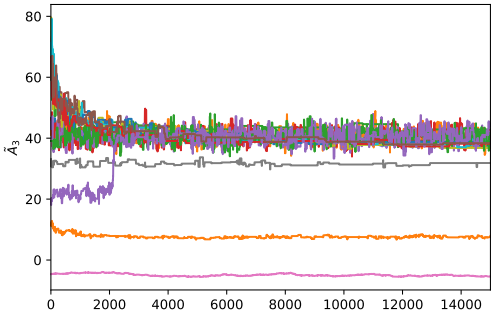
(c) Trace plot of $T_{m,2}$.

(d) Trace plot of $\widetilde{Q_2}$.

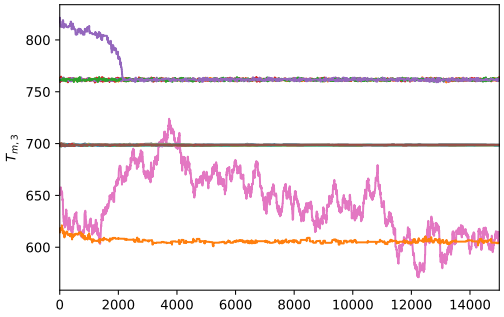
Figure 5.12: Trace plots for the kinetic parameters of the Second reaction.



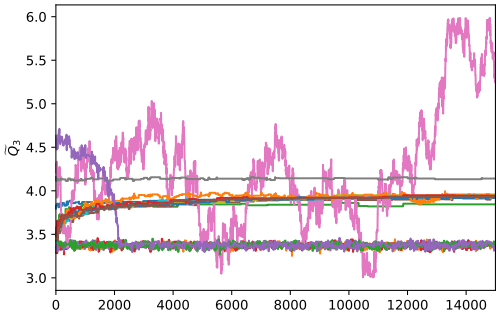
(a) Trace plot of $\log E_3$.



(b) Trace plot of $\widetilde{A_3}$.

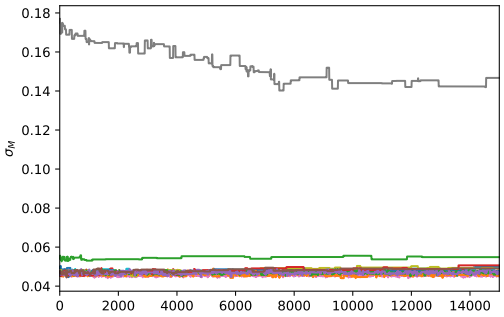


(c) Trace plot of $T_{m,3}$.

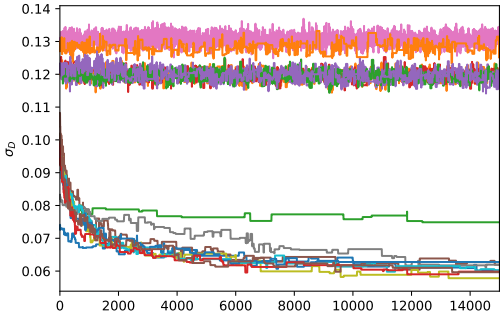


(d) Trace plot of $\widetilde{Q_3}$.

Figure 5.13: Trace plots for the kinetic parameters of the third reaction.



(a) Trace plot of σ_M.



(b) Trace plot of σ_D.

Figure 5.14: Trace plots for the kinetic parameters of the noise parameters.

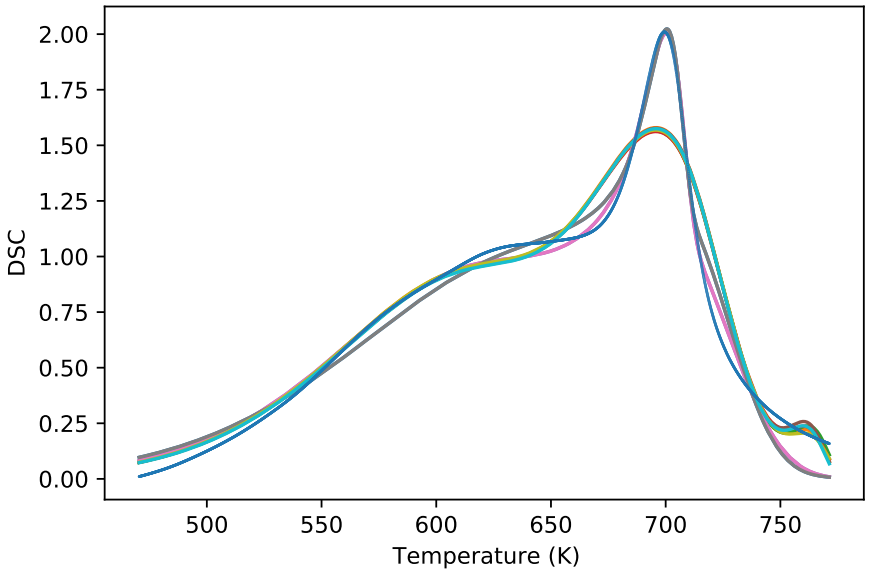


Figure 5.15: Comparison of the DSC data through experimentation with a random sample of DSC curves generated from the sampled parameter estimates to compare the model fit.

Chapter 6

The Scaled Stockpile Equation

One of the ways we can analyse the stockpile model,

$$\rho c \frac{\partial T}{\partial t} = \nabla \cdot (k \nabla T) + Q(T), \tag{6.1}$$

section 3.1, is investigate the non-dimensional problem. In this technique we scale the dimensional variables, x, t and T using the parameters that appear in the equations. This analysis allows us to determine some of the features of stockpiles without having as many parameters involved. This analysis determines a relationship for the critical length of a stockpile and the reaction kinetics. In this chapter we analyse, through non-dimensional analysis, the effects that periodic boundary condition can have on stockpile ignition and also the effects of a hotspot within the stockpile for a subcritical stockpile.

Consider a one-dimensional stockpile, of length L, with a first order Arrhenious reaction rate. We use the scalings,

$$\begin{aligned} x^* &= \frac{x}{L}, \\ t^* &= \frac{\alpha}{L^2} t, \\ u &= \frac{E}{RT_a^2} \left(T - T_a\right), \end{aligned}$$

to rewrite the equation in non-dimensional form,

$$\frac{\partial u}{\partial t^*} = \frac{\partial^2 u}{\partial x^{*2}} + \delta \exp\left(\frac{u}{1+\varepsilon u}\right), \tag{6.2}$$

where $\varepsilon = T_a R/E$ and,

$$\delta = L^2 \frac{Q \rho A}{k} \exp\left(\frac{-E}{RT_a}\right) \frac{E}{RT_a^2}. \tag{6.3}$$

The parameter δ, is known as the Frank-Kamanetskii (FK) parameter . We use the Dirichlet boundary condition, $T = Ta$, on both boundaries, which is the non-dimensional bound-

ary condition $u = 0$. The domain of the scaled equation is $x^* \in [0,1]$. With this formulation we often have $\varepsilon \ll 1$, so we consider $\varepsilon = 0$ [28]. In this case we have a critical Frank-Kamanetskii parameter such that if $\delta > \delta_{\text{cr}}$ then the stockpile ignites, and if $\delta < \delta_{\text{cr}}$ then the stockpile ignition is dependent upon the initial condition, which we consider sub-critical [28].

This non-dimensionalisation can be extended to higher spatial dimensions. If we consider three dimensions, (x,y,z), then by scaling each spatial dimension by the height, L_y, Equation 6.1, becomes,

$$\frac{\partial u}{\partial t^*} = \frac{\partial^2 u}{\partial x^{*2}} + \frac{\partial^2 u}{\partial y^{*2}} + \frac{\partial^2 u}{\partial z^{*2}} + \delta \exp\left(\frac{u}{1+\varepsilon u}\right), \tag{6.4}$$

on the domain $[0, L_x/L_y] \times [0,1] \times [0, L_z/L_y]$, where L_x, L_y, L_z are the stockpile lengths in each direction. We scale each spatial dimension by the same length to maintain the Laplacian in the equation. The results are dependant upon the shape of the domain, generally we consider nice shapes where the lengths have certain nice ratios. From here we drop the $*$ notation from the dimensionless variables.

We need to define what we mean by ignition. In the theoretical case with $\varepsilon = 0$, the temperature can increase without an upper bound [28]. We define supercritical stockpiles as those that continue heating indefinitely and sub-critical are those that do not. This phenomena cannot be determined numerically as we are restricted to finite times and temperatures. How can we define ignition in a practical sense? Within our model there are two things to consider: how hot does the stockpile need to get to to categorise the stockpile as supercritical, and how much time do we allow for the stockpile to reach this state? The answers depend upon the context of the problem that is being addressed.

For our context we consider ignition times of approximately one year as this is considered a reasonable storage time frame for the storage of the BOS stockpiles. In our application this corresponds to a final time $t_f = \omega = 0.3$. For the purposes of our analysis we will consider ignition occurring when the maximum non-dimensional temperature exceeds 100, so $u_{\max} = 100$. Using the parameters in [135], this equates to a temperature of roughly $800^o C$. The parameters used here are very similar to those we use in this thesis and as such the critical temperature is roughly $800^o C$. If we were to use a point estimate of $E = 10^{4.637} J$ for the experimental data, this would correspond to a critical temperature of $1600^o C$ which may be excessive. This highlights the need for these considerations to be made when applying the non-dimensional results to physical problems. When using the scaled equations we are able to deduce relationships between certain parameters.

6.1 Periodic Boundary Conditions

In most of the spontaneous combustion literature the boundary conditions do not depend upon time. In practice large stockpiles are subject to seasonal and diurnal variations in the ambient temperature. We only consider seasonal temperature variations. We model the seasonal temperature variations using a sine function. The boundary condition becomes,

$$T = T_a + T_o \sin\left(\frac{2\pi t}{\omega_Y} + \phi\right), \tag{6.5}$$

where T_a is the average ambient temperature, T_o is the maximum temperature oscillation, ω_Y is the oscillatory period which for seasonal temperature variations is one year, and ϕ is a phase shift. The phase shift parameter ϕ allows us to control when during the year the stockpiles are constructed. The choice $\phi = 0$ is indicative of a stockpile constructed in spring whilst $\phi = \pi/2$ corresponds to construction during summer when the ambient temperature is at its maximum. This is a simple way to model the seasonal temperature variation which allows us to examine the relationship and effect on stockpile ignition. After scaling the boundary condition we obtain,

$$u = u_o \sin\left(\frac{2\pi t}{\omega} + \phi\right), \tag{6.6}$$

where $u_o = \left(E/RT_a^2\right) T_o$, and $\omega = \alpha\omega_Y/L_y^2$.

In higher dimensions not all of the sides are exposed to ambient air. The base of the stockpile, $y = 0$ requires a separate condition. For simplicity we assume that no heat is exchanged at this boundary, that is $\partial T/\partial y = 0$. When we have this condition we are able to reflect around the plane $y = 0$ if we also reflect the boundary and initial conditions around the plane $y = 0$. This enables us to use a consistent boundary condition across the domain.

6.1.1 Dirichlet Boundary Condition

Initially we consider the simple dirichlet boundary condition for two specific stockpile geometries: $L_x = 2L_y$ and $L_x = \infty$. For the case where $L_x = 2L_y$, the reflective condition prescribed at the boundary $y = 0$ means that we can solve the problem on the square with non-dimensional length 2. The solution one the square is equal to the solution to the differential equation on the domain $[-1,1] \times [0,1]$ For the approximation $\varepsilon = 0$ and Dirichlet boundary conditions, the critical value is $\delta_{cr} = 1.7$ [28]. For $L_x \gg L_y$ we approximate the model by a one-dimensional model without diffusion in the x direction. For the approximation $\varepsilon = 0$ and static boundary condition the critical value is $\delta_{cr} = 0.88$ [28]. The value obtained from our simulations, displayed in Table 6.1 equals the value stated by Bowes [28] to the given number of significant figures.

Table 6.1: The critical Frank-Kamenetskii parameter, δ_{cr}, in a one dimensional stockpile for different cut-off times with $u_o = 0.637$, $\omega = 0.3$ and $\phi = 0$ for the cases with oscillations.

ε	Boundary Condition	$t_f = 0.3$	$t_f = 1$	$t_f = 10$	$t_f = 100$
0	Static	3.701	1.582	0.899	0.877
0	Dynamic	3.428	1.537	0.895	0.875
0.027	Static	3.910	1.658	0.927	0.904
0.027	Dynamic	3.639	1.612	0.924	0.901

Table 6.1 suggests that for the larger cut-off times ($t_f = 10, 100$) the dynamic boundary condition does not have a significant influence on the value of the critical FK parameter with a slight reduction observed, when the oscillations are added. However, there is a significant difference in the ignition times. For the case where $\varepsilon = 0$ and $\delta = 0.89$, the ignition times are $t_{\text{ig}} = 14.00$ and $t_{\text{ig}} = 12.04$, for the static and dynamic boundary conditions respectively. Similarly for the case $\varepsilon = 0.027$ and $\delta = 0.92$ the ignition times are $t_{\text{ig}} = 12.40$ and $t_{\text{ig}} = 11.67$, for the static and dynamic boundary conditions respectively.. This indicates that for a given stockpile the ignition time is different once the dynamic boundary condition is applied.

Table 6.2: The critical FK parameter δ_{cr} for a two-dimensional rectangular domain with $u_o = 0.637$, $\omega = 0.3$ and $\phi = 0$.

ε	Boundary Condition	$t_f = 0.3$	$t_f = 1$	$t_f = 10$
0	Static	4.072	2.20	1.71
0	Dynamic	3.615	2.14	1.69
0.027	Static	4.294	2.3	1.76
0.027	Dynamic	3.843	2.23	1.75

In two dimensions, the effects of the dynamic boundary condition are more prominent. For the case $\varepsilon = 0$ the oscillations induce ignition in stockpiles that have a FK parmaeter less than the established long term critical parameter 1.70 [28]. The one-dimensional stockpiles have a lower critical FK value. We anticipate that as the ratio of stockpile length to stockpile height L_x/L_y increases then the FK parameter should decrease and approach the one-dimensional critical FK parameter. This relationship is verified in Figure 6.1

So far our analysis has assumed the stockpiles are constructed at a specific time of year, specifically spring, $\phi = 0$. Figure 6.2 indicates which stockpiles ignite within a year. Certain stockpiles constructed in Spring ignite whilst identical ones constructed in Autumn do not.

The final time is now increased to $t_f = 1$. The critical Frank-Kamanetskii parameter becomes less sensitive to the construction date. Another key observation is that when $t_f = 1$ there are more values of ϕ where the critical parameter is lower than the static boundary condition. Figure 6.3 shows a shift in the oscillation peak towards $\phi = 0$.

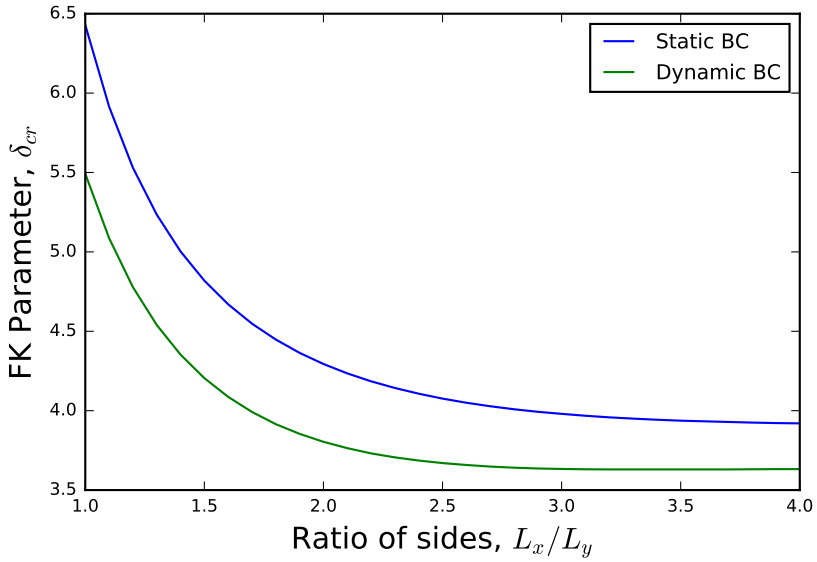


Figure 6.1: The critical FK parameter δ for ignition to occur within a year, as a function of the ratio of side length to height. Parameter values: $\varepsilon = 0.027$, $\phi = 0$, $u_o = 0.637$, $\omega = 0.3$ and $t_f = 0.3$.

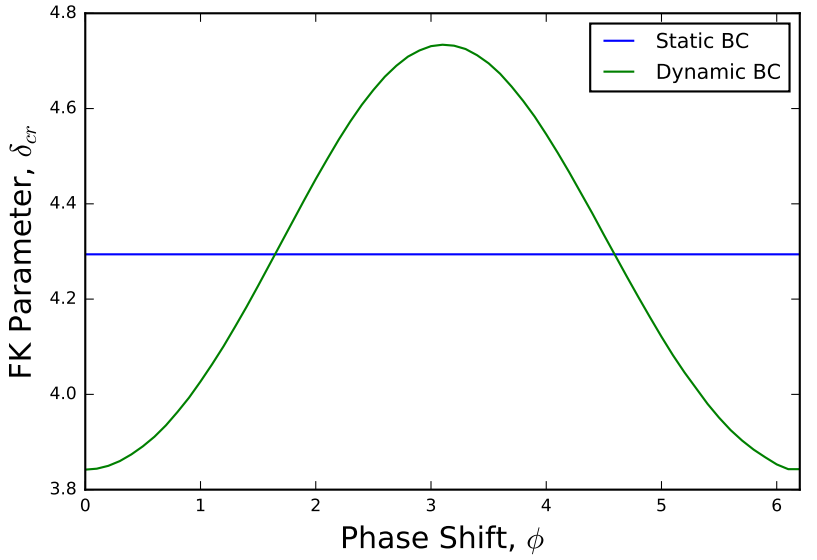


Figure 6.2: The effects of changing the stockpile construction time (ϕ) on the critical FK parameter, δ, for stockpile ignition within a year ($t_f = 0.3$). Parameter Values $\varepsilon = 0.027$, $L_x/L_y = 2$ $u_o = 0.637$ and $\omega = 0.3$.

We now consider the ignition times as the FK parameter (δ) is varied.

Figure 6.4 compares the ignition times for the static boundary condition with the dynamic boundary condition. Stockpiles constructed in spring, $\phi = 0$, are plotted along with stockpiles constructed half a year later, $\phi = \pi$. For the larger values of the FK parameter, there is a distinct difference between the ignition times in each case. For smaller values of the FK parameter the ignition times for stockpiles constructed later ($\phi = \pi$) are lower

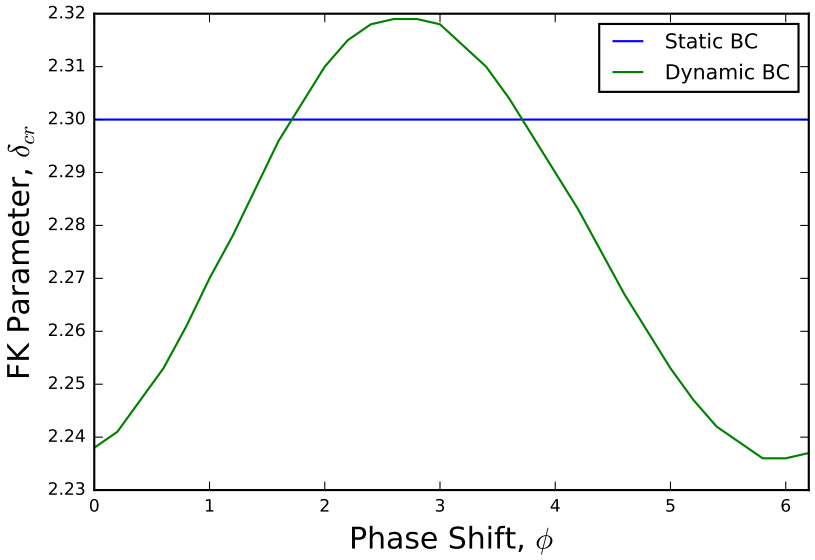


Figure 6.3: The effects of changing the construction time (ϕ) on the value of the critical FK parameter, δ, for a final time, $t_f = 1$, corresponding to just over a three year period. Parameter Values: $\varepsilon = 0.027$ and $L_x/L_y = 2$, $u_o = 0.637$ and $\omega = 0.3$.

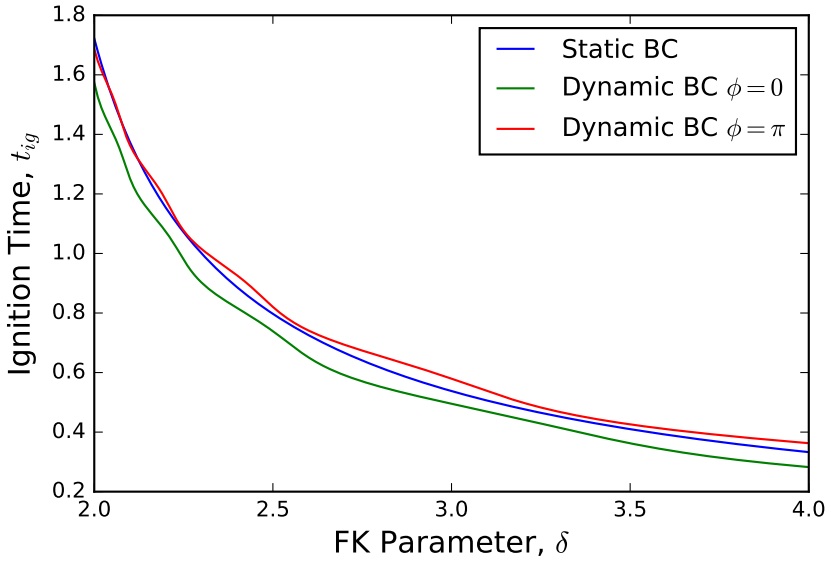


Figure 6.4: The ignition times as the Frank-Kamanetskii parameter, δ, is varied. Parameter Values: $\varepsilon = 0.027$ and $L_x/L_y = 2$.

than those with the static boundary condition. For FK parameters close to 2, there are instances where the ignition time for the dynamic boundary condition with $\phi = \pi$ is less than the ignition time for the static boundary condition. This suggests that if we consider ignition times that are longer, for example, $t_f = 1.8$, then at this point the critical FK parameter required to induce ignition for a stockpile with dynamic boundary condition will be less than the critical FK parameter for a static boundary condition for all construction times ϕ.

6.1.2 Newtonian Cooling Boundary Conditions

The alternative to using a Dirichlet boundary condition is to apply a convective heat transfer condition. We use the boundary conditions given in section 3.1.1, with the ambient temperature following the same oscillations in equation 6.5 resulting in the equation,

$$\frac{\partial T}{\partial n} = h\left(T - \left(T_a + T_o \sin\left(\frac{2\pi t}{\omega_Y} + \phi\right)\right)\right).$$

After scaling this equation we obtain,

$$\frac{\partial u}{\partial n} = \tilde{h}\left(u - u_o \sin\left(\frac{2\pi t}{\omega} + \phi\right)\right),$$

where $\tilde{h} = L_y h$ and, u_o and ω are defined as before.

For the Robin Boundary condition we focus on the two-dimensional model with domain $[-1,1] \times [0,1]$. We investigate whether the effects that we observed with the Dirichlet boundary condition correspond to those we see with the Newtonian cooling boundary condition. When examine the Newtonian cooling boundary condition a new parameter, $\tilde{h}$ referred to as the heat transfer coefficient is introduced. For our analysis here we consider ignition times with $t_f = 1$ The first relationship that we examine is the impact of this parameter on the critical FK parameter δ.

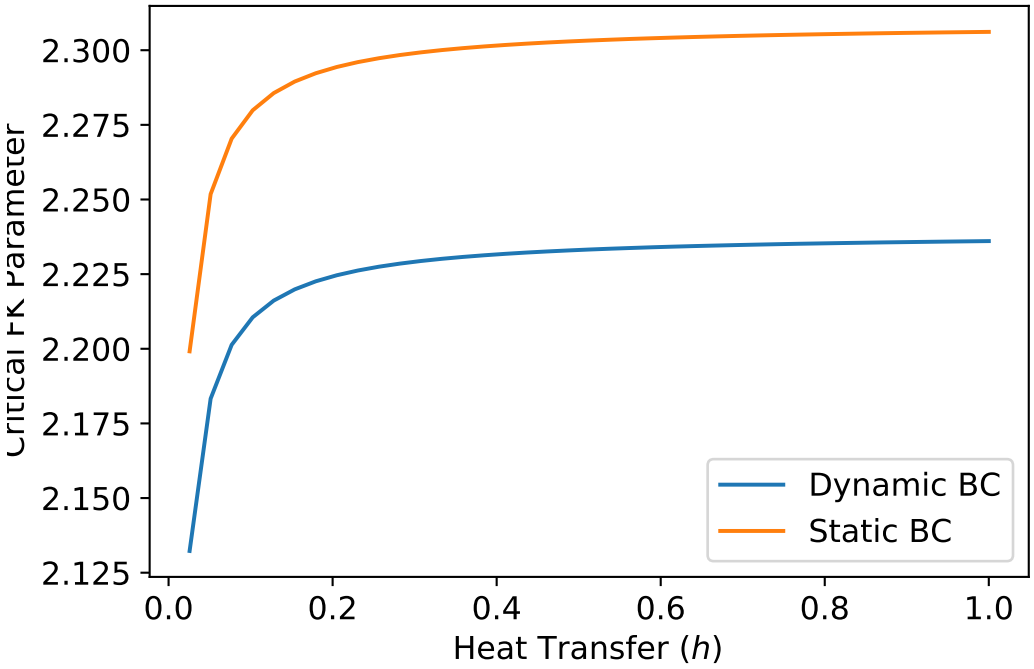


Figure 6.5: The effect of the heat transfer coefficient on a the critical FK parameter. Parameter Values: $u_o = 0.637$, $\phi = 0$.

Figure 6.5 compares the impact of the heat transfer parameter when we include periodic boundary conditions and without periodic boundary conditions. We observe that as the heat transfer coefficient increases then the critical FK parameter approaches the critical value for the Dirichlet condition, $\delta_{cr} = 2.239$, with periodic boundary conditions and $\delta_{cr} = 2.309$ without periodic boundary conditions. This is expected as the limit $h \to \infty$, then the boundary condition approaches the Dirichlet boundary condition. We observe the expected reduction in the critical FK parameter as less heat is exchanged at the boundary. When comparing the oscillating boundary condition to the static boundary condition $u_o = 0$, then we observe that the periodic boundary condition has reduced the critical FK parameter.

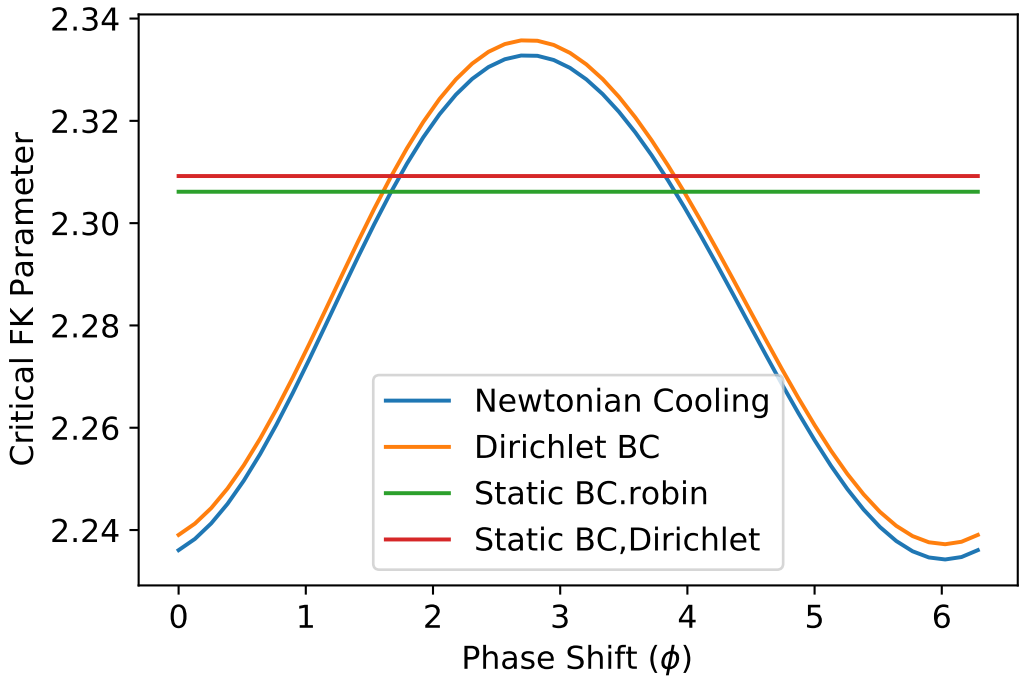


Figure 6.6: Comparison of the impact that the Phase shift has on the critical Hotspot Temperature, for the Dirichlet and Newtonian boundary conditions. Parameter Values: $u_o = 0.637$, $h = 0.5$.

Likewise with the Dirichlet boundary conditions, we need to examine the effect that the phase shift of the periodic boundary conditions has on the critical FK parameter. Figure 6.6 indicates that the effect of the phase shift is very similar with the Newtonian cooling boundary condition and the Dirichlet condition.

One of the key parameters we haven't investigated the effect of is the amplitude of the temperature oscillations. The relationship displayed by Figure 6.7 is interesting. As the oscillation temperature increases, then we find that the critical FK parameter decreases. This is intuitive, as the higher ambient temperatures add more energy to the stockpile. Where this becomes most interesting is the discrepancy between when the stockpile is

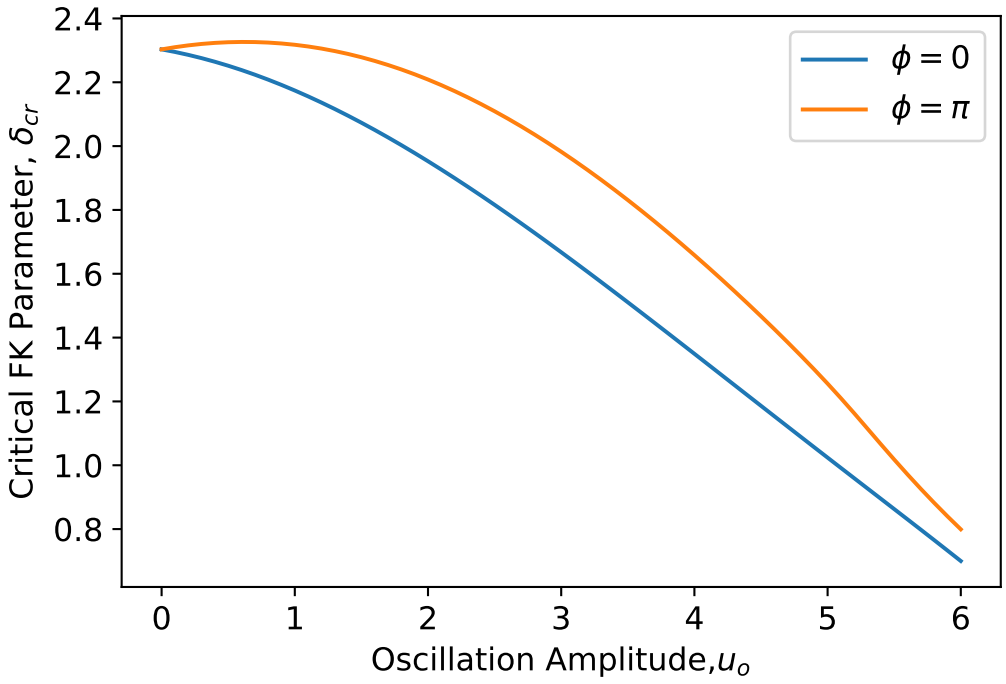


Figure 6.7: A comparison of the effect of the amplitude of temperature oscillations on the critical FK parameter.Parameter Values: $h = 0.5$.

constructed. As the oscillation temperature increases the effect of this phase shift becomes more pronounced. We observe in our results that using a Newtonian Boundary condition, as opposed to the Dirichlet Condition, has a much lower effect on the critical FK parameter. The effect of the Newtonian Cooling boundary condition also decreases at the oscillation temperature increases. Whilst these results are of interest, with our parameter values we have prescribed, the temperature oscillations, u_o, will not approach these large values.

6.2 Hotspot Ignition

One of the methods to induce ignition within large stockpiles is to take a portion of material from a supercritical stockpile and add this to a stockpile that is subcritical. From a theoretical basis, in the approximation $\varepsilon = 0$, if the temperature within a stockpile is high enough, then we obtain runaway self-heating and the stockpile becomes supercritical even for FK parameters less than the critical parameter. Since this is a possibility, we seek to identify some of the conditions that leads to ignition occurring. For $\varepsilon > 0$, we have an upper bound on the amount of energy released from the reactions. This means that the temperature does not continually increase and will reach a steady state solution that is typically much greater than the temperatures that are observed. For our investigations into the hotspots we only consider $\varepsilon = 0.027 > 0$

We introduce the hotspot by changing the initial temperature distribution. A standard initial condition is to assume that the stockpile is assembled at ambient temperature, i.e., $u = 0$. We introduce a hotspot by changing the initial condition to,

$$u = u_h \left(H(x - h_c - h_l) - H(x - h_c + h_l) \right), \tag{6.7}$$

where u_h is the dimensionless hotspot temperature, h_c is the centre location of the hotspot and h_l is the width of the hotspot, generally expressed in terms of the ratio of hotspot to stockpile size. This formulation allows us to control the location, size and temperature of the hotspot. This scaled initial condition uses the temperature scaling such that,

$$u_h = \frac{E}{RT_a^2} \left(T_h - T_a \right), \tag{6.8}$$

where T_h is the temperature of the hotspot.

6.2.1 One-Dimensional Stockpiles

We begin our analysis by considering a one-dimensional pile. These piles are the simplest and most computationally efficient to solve. This allows us to examine the dependencies efficiently before extending this to higher dimensions. The first relationship we explore is for a fixed size hotspot; how hot does this hotspot need to be in order to cause a subcritical stockpile to ignite? Figure 6.8 displays this relation and indicates what we would expect; that as we decrease the Frank-Kamanetskii parameter we require a hotter hotspot to cause ignition.

Figure 6.8 uses a hotspot that is 10% of the size of the stockpile and located on the edge. It is worth investigation what occurs when we change these parameters as we can change the size and location of the hotspot when adding one.. Figure 6.9 displays these results. We find that the size of the hotspot has a significant impact on the critical temperature. Figure 6.9a also indicates that if the whole stockpile is assembled at a temperature of 40^oC with our parameters, then this would be enough to ignite the stockpile. The location of the hotspot within the stockpile has a lower impact on the necessary hotspot temperature, though this can differ by approximately 5^oC with our parameters. We observe that hotspots near the boundary require a slightly higher temperature and attribute this to the cooling that occurs at the boundary.

In the previous section we showed that periodic boundary conditions had an impact on the critical value of the FK parameter. It would be useful to know if this has an impact on the critical hotspot temperature required for ignition. Figure 6.10 indicates that phase shift parameter ϕ has a minimal effect on the critical hotspot temperature required to induce ignition. In our application the critical temperature varies by less than a degree. We compare this to the critical hotspot temperature without periodic boundary conditions,

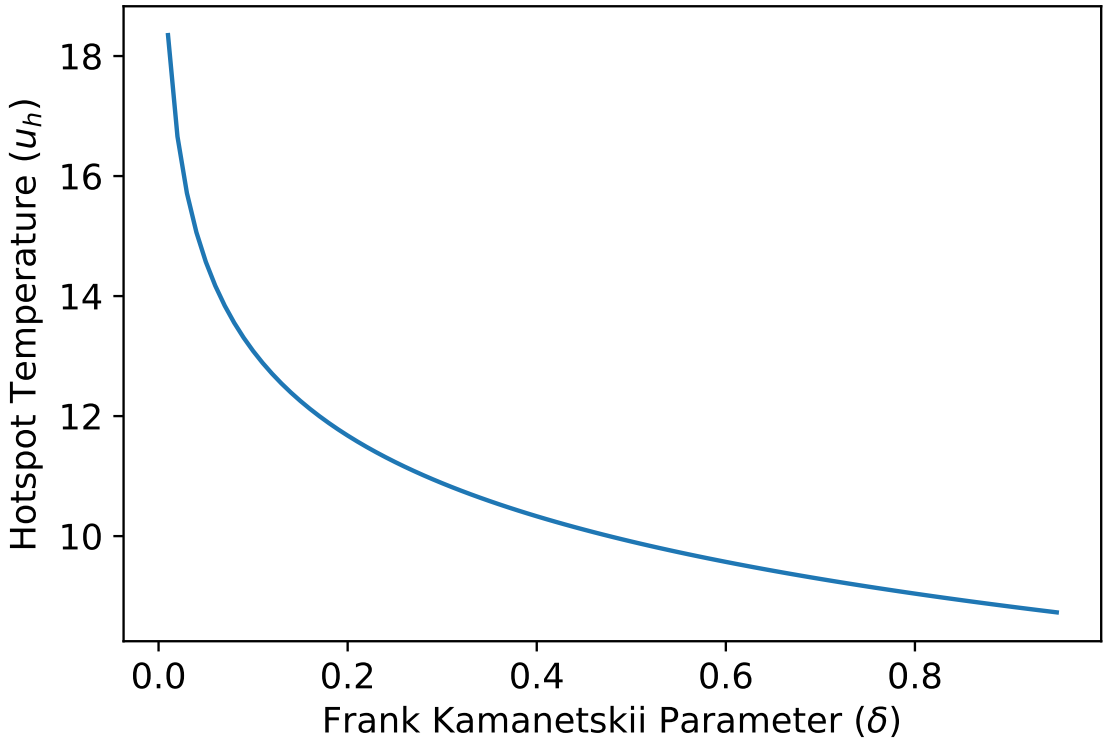
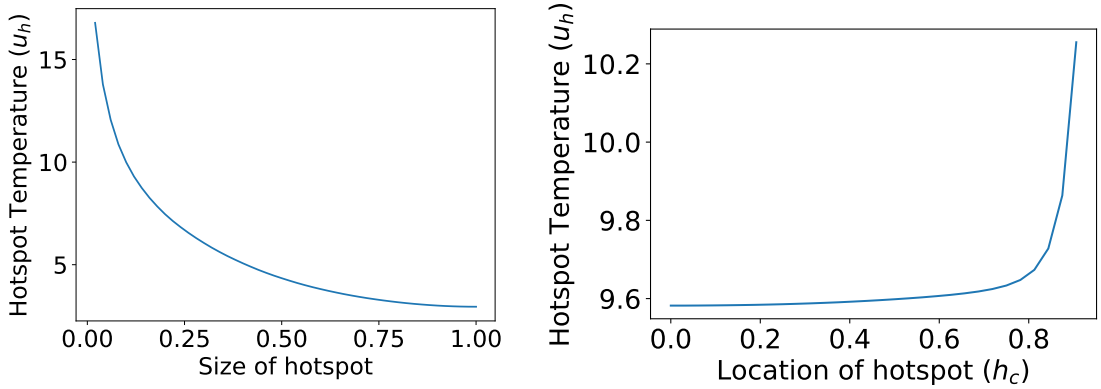


Figure 6.8: The temperature required for subcritical stockpile to ignite. Parameter Values: $h_c = 0.9$, $h_l = 0.1$



(a) Effect of hotspot size.

(b) Effect of hotspot location.

Figure 6.9: The critical temperatures needed for different hotspots in order to cause ignition of subcritical stockpiles. Parameter Values: $h_c = 0.9$, $h_l = 0.1$.

which is 7.526. This indicates that there is a slight increase in the critical temperature required, though this is not a significant difference.

To this stage our results have used an ignition condition of $u_{\max} = 100$. We now examine how this ignition criteria determines the critical temperature of the stockpile. For the range of ignition conditions we tested, $[10, 100]$, the critical hotspot temperature did not change. This suggests that setting a high enough temperature is sufficient to determine ignition in this model. To investigate why this may be the case we examine two stockpiles

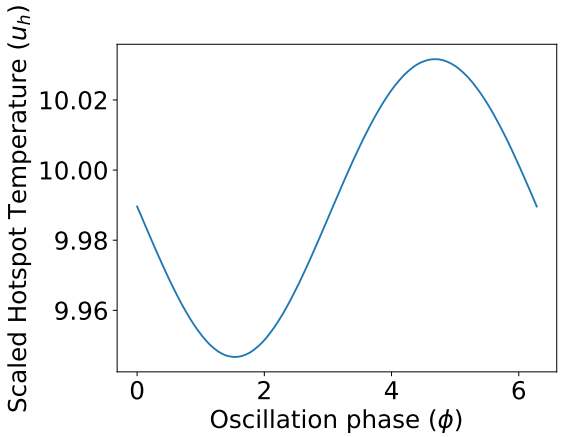
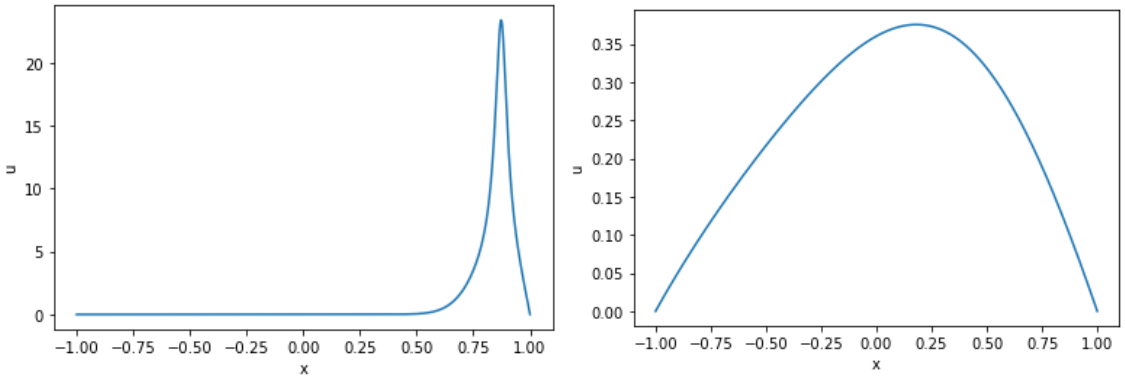


Figure 6.10: The temperature required for a subcritical stockpile to ignite as we change the time of year the hotspot is introduced.Parameter Values: $h_c = 0.9$, $h_l = 0.1$, $u_o = 0.627$, $\omega = 0.3$.

that have slightly different hotspot temperatures.



(a) Stockpile before ignition.

(b) Stockpile below critical hotspot.

Figure 6.11: Temperature profiles within two identical stockpiles with different hotspot temperatures. Parameter Values: $h_c = 0.9$, $h_l = 0.1$.

Figure 6.11 compares a stockpile with non-dimensional hotspot temperature, $u_h = 7.53$, with a stockpile with $u_h = 7.52$. This figure clearly indicates that it is the region of the hotspot itself, and the reactions occurring in that region which are causing the stockpiles to ignite. This provides further validation to the claim that as the ignition criteria is changed, the critical hotspot temperature remains the same. We observe in Figure 6.11b that the subcritical stockpile has cooled from its initial elevated hotspot. Figure 6.11a poses the question: What if the hotspot is inert?

In practical applications the hotspot contains material from a supercritical stockpile. The concentrations of the reactants can differ from those of a fresh subcritical stockpile, hence the hot-spot is unlikely to have the same reactivity. It may also be of interest to know if we are able to promote ignition using a region of heated material that does not react. To investigate this behaviour we consider the extreme case of an inert material used

in the hotspot region. In this region no reaction occurs.

The Inert Hotspot

The inert hot-spot provides an extreme case where we add material to the stockpile but this has no effect on the heat production internally. We implement this by implementing a reactant term into our equation. For the inert hotspots we also introduce an advection term into the model. Given that we have a porous material this mechanism of heat transfer occurs within the stockpiles. This becomes more important as we consider an inert hotspot since the heat transfer from the hotspot to the stockpile is extremely important. Equation (6.2) becomes,

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + v\frac{\partial u}{\partial x} + M(x,t)\delta \exp\left(\frac{u}{1+\varepsilon u}\right), \tag{6.9}$$

where $M(x,t)$ defines the mass of reactant at the point x and time t. In our previous analysis we have, $M(x,t) = 1$. For the inert hotspots, we set $M(x,t) = 1$ whenever x is outside our hotspot region and set $M(x,t) = 0$ inside the hotspot region. This function provides the desired effect of limiting heat production from the reaction to areas outside of the hotspot, congruent with the features of an inert hotspot.

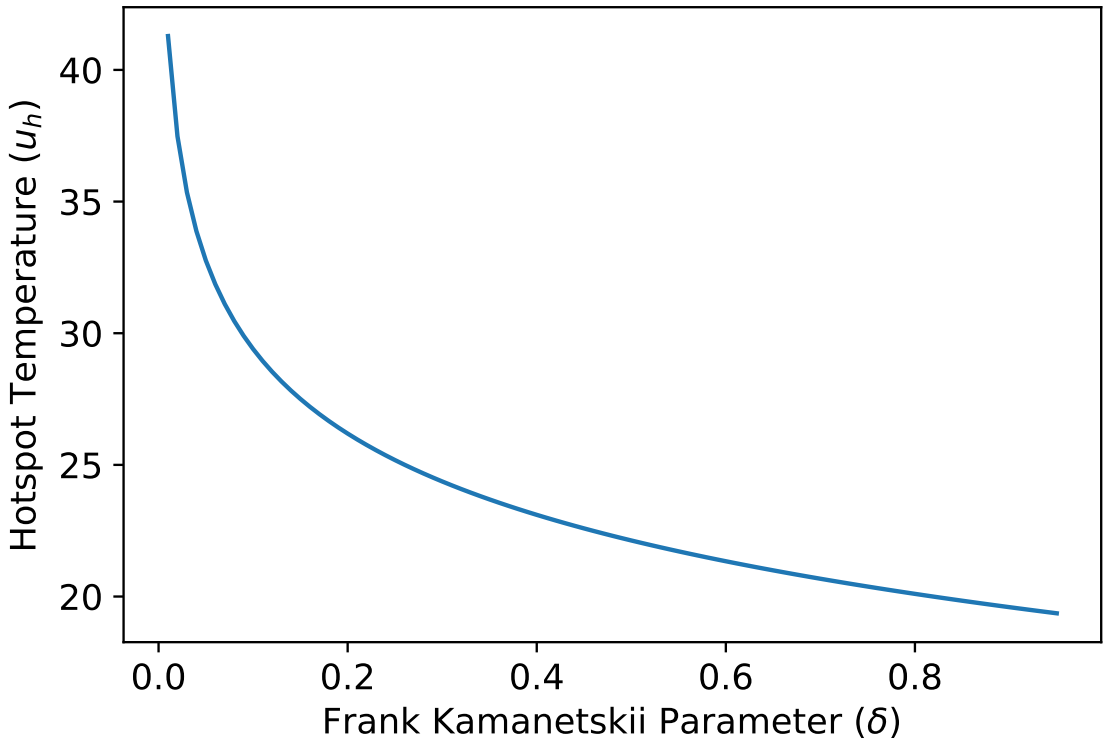
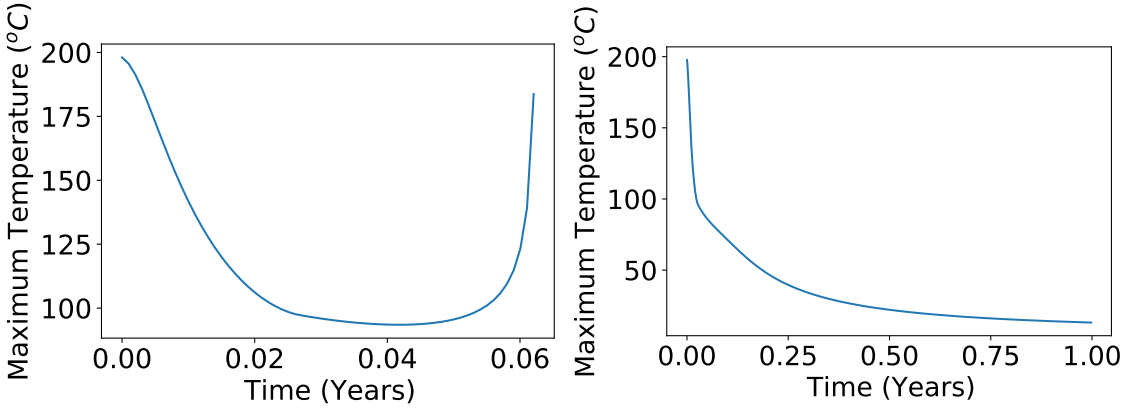


Figure 6.12: The temperature required for subcritical stockpile to ignite.Parameter Values: $h_c = 0.9$,$h_l = 0.1, v = 0$.

Figure 6.12 indicates that we require the hotspot to have a higher temperature as the FK parameter decreases. This is consistent with the reactive hotspot in Figure 6.8 but we

require a higher hotspot temperature in order to achieve ignition. This is intuitive as we need to heat the material surrounding the hotspot so that the material in those regions are heated. We now compare the temperature profiles generated from identical stockpiles that have slightly different hotspot temperatures.



(a) Stockpile just before ignition. $u_h = 22.6$.

(b) Subcritical stockpile. $u_h = 22.55$.

Figure 6.13: Temperature profiles within two identical stockpiles with different slightly different hotspot temperatures. Parameter Values: $h_c = 0.9$,$h_l = 0.1$, $v = 0$ and $\delta = 0.45$.

Figure 6.13 displays the maximum temperature profiles over the integration period. Figure 6.13a has a truncated integration period as we observe the temperature beginning to increase rapidly. This truncation has been implemented to simplify the figure so that we can see the initial cooling phase before thermal runaway occurs. The difference between these two stockpiles is that the hotspot temperature differs by less than one degree, with the ignited stockpile having a hotspot temperature just above the critical value, whilst the subcritical stockpile has a hotsopt temperature less than the critical value.

What we have discovered is that inert hotspots can induce ignition quickly. This feature also renders the integration period as rather obsolete in that the hotspot temperature breaching the critical threshold is the determining factor for ignition via hotspots. This contrasts to what is observed in the ignition of stockpiles without hotspots, where there is a slow build up of heat that eventually induces thermal runaway which was observed in Section 6.1 of this chapter in the periodic boundary conditions.

The sharp contrast in this critical value highlights the need for a statistical approach to ignition when considering uncertainty within our parameters. Given that the work in the previous chapters has highlighted some significant uncertainty in our parameter estimates, this will extend into uncertainty around the parameter estimate for the FK parameter δ. Using Equation (6.3), and our sampled parameters from Chapter 4, we can construct an estimate for our distribution of the FK parameter for a stockpile with either the Wüstite or Iron reactions that are occurring in the filter cake stockpiles. Using the distribution of the FK parameter displayed in figure 6.14, we can determine a probability of ignition occurring for any given stockpile using this distribution of parameters. It is important to

note however that the parameters will also alter the scaling for temperature, so care needs to be taken when using the hotspot temperature.

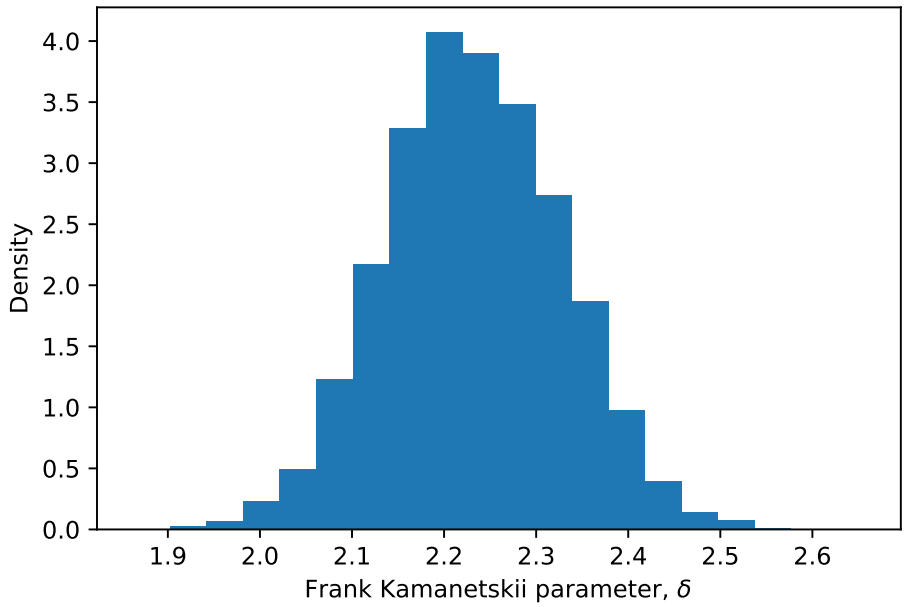


Figure 6.14: Posterior Distribution of the Frank-Kamanetskii parameter for the wüstite reaction.

Similarly to the reactive hotspot case, we can investigate the effect that location and size has on the critical temperature of the inert hotspot. These parameters are crucial as they allow us to dictate the size and position of the hotspot within the stockpile. Changing the location parameter, h_c, has the same effect as when we had a reactive hotspot within the stockpile, as displayed by Figure 6.15.

We are also able to change the size, h_l, of the stockpile. The results for an inert hotspot differ significantly from those with a reactive hotspot. We assess the impact that the size of the hostspot has for both a centrally located hotspot and an edge located hotspots wth the results displayed in Figure 6.16. We find that for small hotspots the behaviour is similar to that of reactive hotspots, though as our hotspots increase in size, larger hotspot temperatures are required. This is attributed to the fact no additional heat is generated within the hotspot so heat can only be generated in the region which there are reactants. The size parameter h_l, is indicative of the proportion of the stockpile that the hotspot takes up so a value near one means the reactive part of the stockpile is much smaller in size. We could consider a stockpile that has a fixed amount of reactive material and increase the size of the hotspot though this would change the combined length which will affect the FK parameter and the scaling for length and time, hence it was not included in our analysis.

The instance of large inert stockpiles can be considered if we are looking to add material onto the end of an existing stockpile that has completed the self-sintering process.

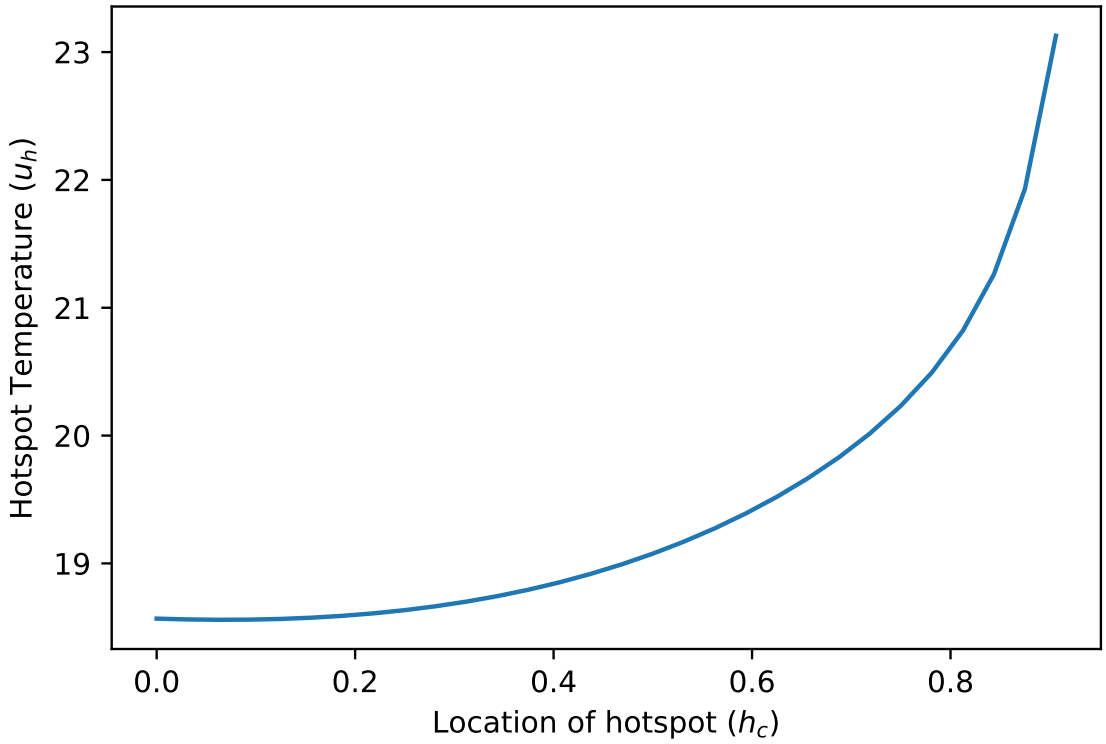
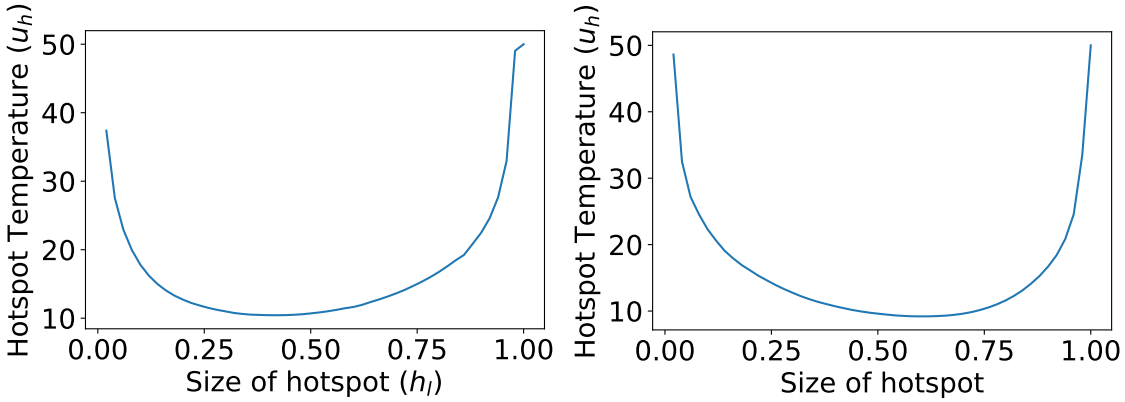


Figure 6.15: The effect that the location of the hotspot has on the critical ignition. Parameter Values: $h_l = 0.1$, $v = 0$ and $\delta = 0.45$.



(a) Centered hotspot, $h_c = 0$.

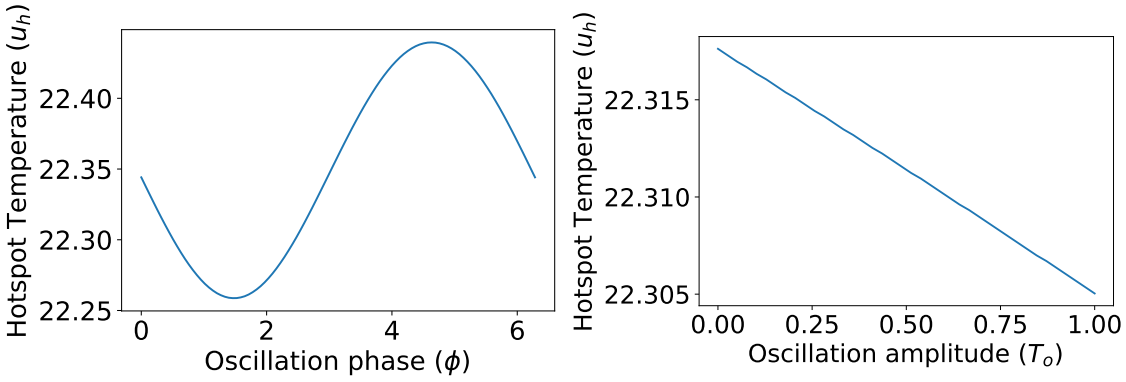
(b) Edge hotspot, $h_c = 1 - h_l$.

Figure 6.16: The effect of the hotspot size on the critical temperature of the hotspot required to induce ignition.Parameter Values: $v = 0$ and $\delta = 0.45$.

Whilst the assumptions regarding a uniform temperature do not exist in this case, having a large hotspot still equates to a fully reacted stockpile. Considering new material is consistently generated the instances of large stockpiles are worth investigating, though the hotspot temperature profile may need to resemble that of a heated stockpile rather that a flat temperature.

Additionally we investigate the influence of external factors on the critical hotspot temperatures by investigating advection through the stockpile, and the effect of oscillating

boundary conditions. Similarly to the reactive hotspot case, the boundary conditions have a minimal effect on the critical hotspot temperature. It does have a slight effect, though in a practical sense the oscillations are generally too small to cause a large difference in the result. We largely attribute this to the early ignition times that are observed with a hotspot as opposed to the slow build up we saw when we examined the effect of the boundary conditions.



(a) Oscillation Phase Shift.

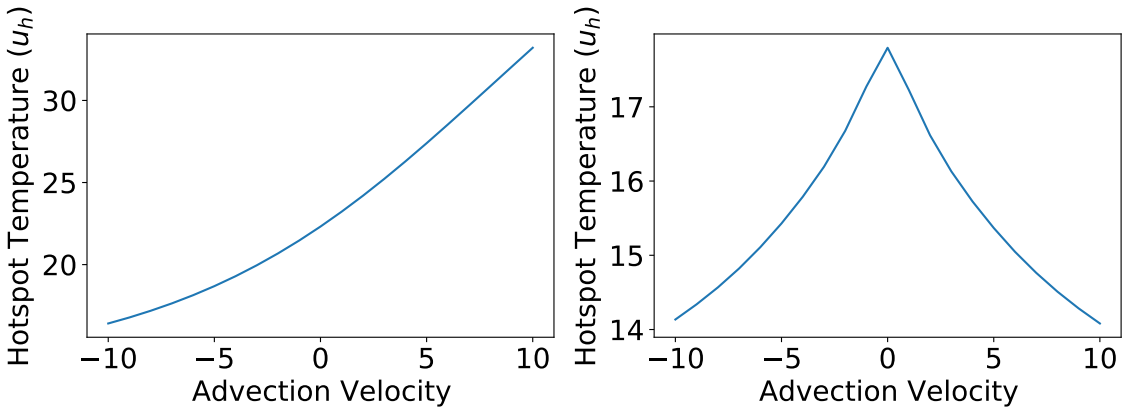
(b) Changing Oscillation Amplitude.

Figure 6.17: The effect of the seasonal temperature oscillation and the oscillation phase have on the critical hotspot temperature. Parameter Values: $h_c = 0.9$, $h_l = 0.1$, $v = 0$ and $\delta = 0.45$.

As the stockpiles are porous, we can consider the effect that the movement of air has throughout the stockpile. We find the advection can have a significant effect depending upon the advection velocity. In both the case of the edge and central hotspot we see a pronounced effect the advection velocities that we investigated have on the critical hotspot temperature. For the edge hotspot, advection in the direction towards the center of the stockpile decreases the hotspot temperature necessary for ignition whilst advection in the opposite direction increases this temperature. The behaviour of the centred stockpile is quite different in the sense that any advection decreases the necessary stockpile temperature. Our advection currents are simply in the sense the advection is maintained at a constant rate and constant direction whilst in a practical example, the advection term will constantly be changing. In addition, determining an accurate scaled advection velocity through the stockpile is quite challenging as it requires measurement of the airspeed within the stockpile. As a result the practical effect may only be small as the advection velocity my be small, though our analysis indicates that further examination of the advection term is valuable as this can have a sizable effect.

Consumption of Material

Our analysis has focused on a model where the reactant is not consumed. When we introduce consumption of the reactants we change the model and introduce a new equation



(a) Hotspot located on the edge. **(b)** Central Hotspot.

Figure 6.18: The effect of the advection on the critical hotspot temperature. Parameter Values: $h_c = 0.9$, $h_l = 0.1$, and $\delta = 0.45$.

for the reactants. The new model is given by,

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + M(x,t)\delta \exp\left(\frac{u}{1+\varepsilon u}\right). \tag{6.10}$$

The concentration of reactant $M(x,t)$ is no longer fixed. It changes according to the consumption equation,

$$\frac{\partial M}{\partial t} = -M(x,t)\delta_M \exp\left(\frac{u}{1+\varepsilon u}\right), \tag{6.11}$$

where δ_M is analagous to the FK parameter δ, with

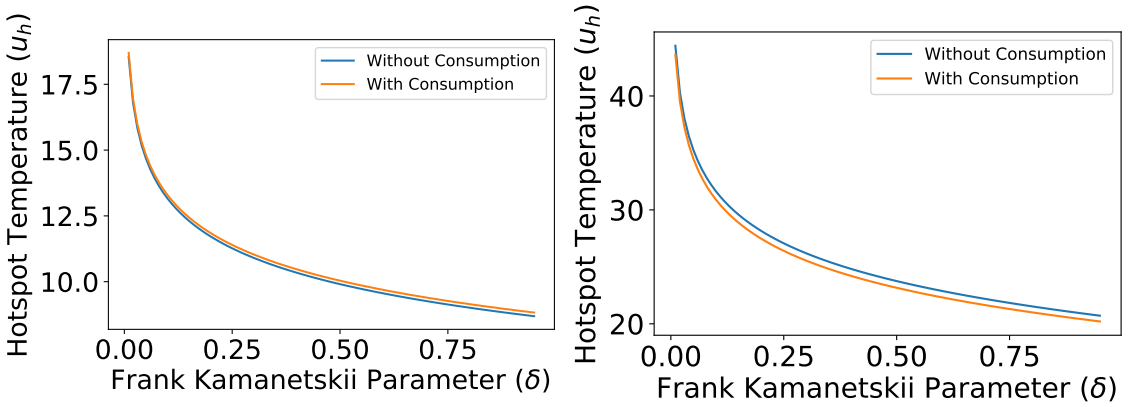
$$\delta_M = \frac{L^2}{\alpha} A \exp\left(\frac{-E}{RT_a}\right).$$

For our analysis we set $\delta_M = \delta/100$. This consumption equation is the scaled form of those used in the TGA models where the mass loss is significant. We define the initial reactant concentration, $M_0(x)$ as,

$$M_0(x) = \begin{cases} 1 - M_c & x \in [h_c - h_l, h_c + h_l] \\ 1 & x \notin [h_c - h_l, h_c + h_l] \end{cases}, \tag{6.12}$$

where M_c is the proportion of mass consumed initially within the hotspot. For an inert hotspot $M_c = 1$ meaning all the mass within the hotspot has been consumed, whilst $M_c = 0$ would be an unreacted hotspot.

The consumption of material has a small effect on the critical hotspot temperatures as the FK parameter is changed, as demonstrated by Figure 6.19. This observation applies to both the inert hostpot and the reactive hotspot. When we consider a hotspot temperature

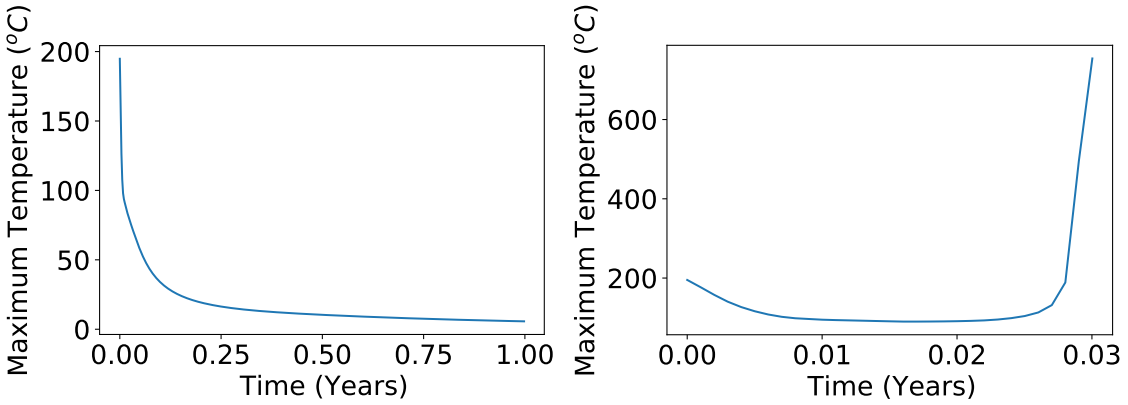


(a) Reactive Hotspot.

(b) Inert Hotspot.

Figure 6.19: Comparison of the effect that including reactant consumption in the model has for different FK parameters. Parameter values: $h_c = 0.9$, $h_l = 0.1$.

that varies slightly from the critical hotspot temperature, then we observe the same behaviour as we found in Figure 6.13, with the results of including consumption displayed in Figure 6.20.

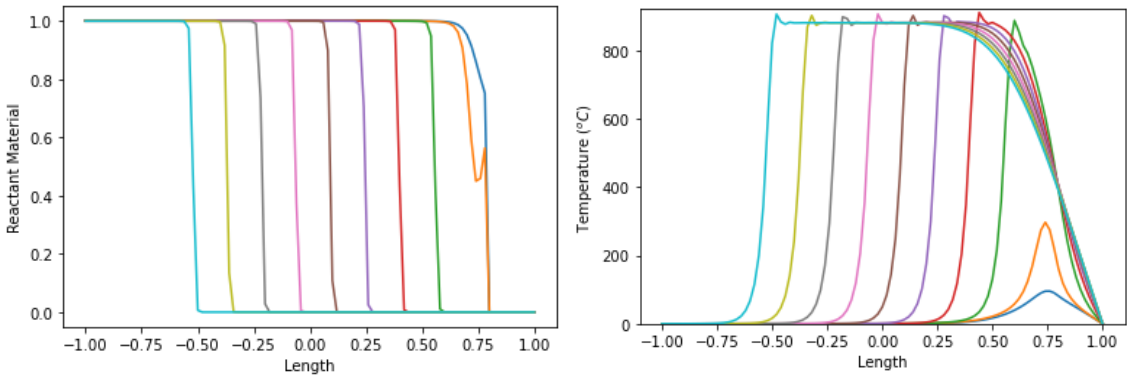


(a) Subcritical hotspot. $u_h = 23.5$.

(b) Supercritical hotspot $u_h = 23.55$.

Figure 6.20: Comparison of the thermal evolution in two two stockpiles with slightly different hotspot temperatures. The simulation includes fuel consumption. Parameter values: $h_c = 0.9$, $h_l = 0.1$, $\delta = 0.45$.

When we look at the temperature profiles of these supercritical stockpiles, we observe a similar behaviour in the temperature profiles that occurred in both the inert and reactive hotspot cases. The hotspot heats the surrounding material such that the reaction rate increases rapidly leading to ignition. When we include the consumption of materials this limits the temperature that the reactive parts of the stockpile reach, as there is only a finite amount of fuel to induce the reactions. The end result of this process is that travelling waves form within our stockpile as the region ahead of the wave is preheated by reactions occurring just behind the wave front, as displayed in Figure 6.21. Travelling wave solutions have been well studied in the combustion literature [77, 78, 79, 80, 81]



(a) Material Profile.

(b) Temperature Profile.

Figure 6.21: The travelling wave profile for a supercritical stockpile with inert hotspot. Parameter values: $h_c = 0.9$, $h_l = 0.1$, $u_h = 23.55$ and $\delta = 0.45$.

The dynamics involved in inducing ignition within the stockpile from hotspots are similar whether consumption is included or not. Including consumption has a minimal effect upon the critical hotspot temperature. It has a larger effect on the long term dynamics of the stockpile as well as the combustion evolution. One of the major considerations we need to consider when assessing whether a stockpile has ignited or not is that material consumption limits the stockpile temperature. This means that our ignition criteria, $u_{\max} > u_{\text{ig}}$ may not be a suitable criteria. However since we have introduced the reactant concentration into the equation, it is logical to use certain thresholds for the consumption of the reactants as our ignition criteria. A suitable choice will be dependant upon the particular application.

Including reactant consumption facilitates an intuitive way to link inert hotspots with the reactive hotspots. We change the initial consumption percentage of the hotspot, M_c, and allow this to range from $M_c = 0$, indicating a completely reactive hotspot to $M_c = 1$ being an inert hotspot. This provides a method to connect these two extremes and provides a more realistic framework for us to work in. Figure 6.22 indicates that as the hotspots initial mass conversion rate increases, then the hotspot temperature required to induce ignition increases. This corresponds with our previous results for the reactive and inert hotspots.

6.2.2 Two-Dimensional Stockpiles

We extend the stockpiles into two dimensions. We use Equation (6.4) as the model for the self-heating within the stockpile. The domain for the two dimensional stockpiles are $[-L_x, L_x] \times [0, L_y]$, though we restrict our analysis to those with $L_x = 2L_y$. We implement a Neumann boundary condition along y=0, meaning that no heat is exchanged between the stockpile and the ground. Since the PDE is symmetric about the y axis, then if we take the even extension of the initial and boundary conditions, our problem is equivalent to

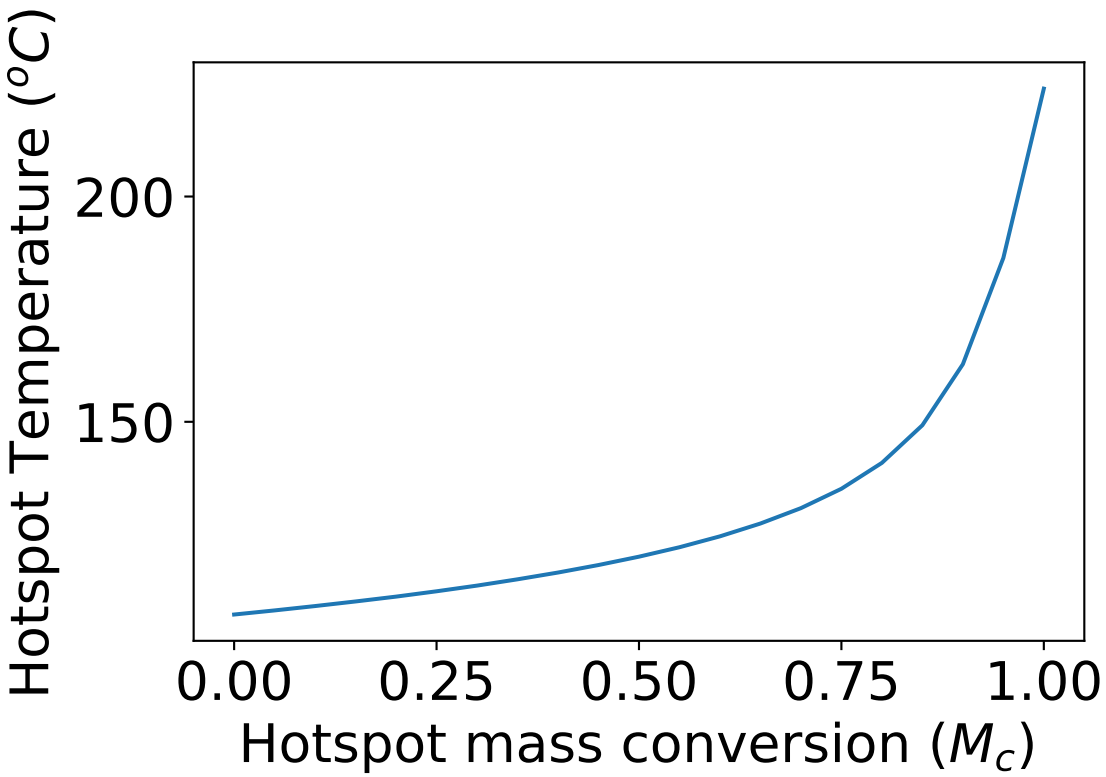


Figure 6.22: The effect of initial hot spot consumption on the critical hotspot temperature. Parameter values: $h_c = 0.9$, $h_l = 0.1$ and $\delta = 0.45$

the same PDE on a square with Dirichlet boundary conditions. This enables us to utilise known analytic results to compare to our numerical results.

For the hotspot in two dimensions, we consider the hotspot to be the region $[h_c - h_l, h_c + h_l] \times [-1/8, 1/8]$ for the square domain. In the x-direction we maintain the same hotspot that we had in the one dimensional case. We set the height of the hotspot to be $1/8$ as this aids with the computational accuracy of floating point numbers. For the two-dimensional stockpiles computation time is important to consider. A fine mesh dramatically increases the computation time and hence we trade off some accuracy for a reduced computation time. Since we are focusing on the relationship between the parameters we choose to reduce some accuracy in the values for an ease of computation. We use a 64x64 mesh over our domain and ensure that as we change the location, h_c and length, h_l of our hotspot, both of these variables can be expressed exactly in base 2. This decision ensures a consistent number of mesh points are covered by the initial condition and enables a consistent hotspot. In the one dimensional case, we found that even for much finer meshes, the influence of the size of the hotspot was strong enough that including an additional node had an effect on the critical hotspot temperature. Using numbers that can be represented exactly in base 2 prevents the rounding of floating point numbers which has an effect on the whether end nodes are included or not. We use the same initial condition as we used in the one-dimensional problem and limit our analysis of two-dimensional stockpiles to

the instances of inert hotspots.

The results that we obtain for the two-dimensional stockpiles are very similar to those we obtained in one dimension. Figure 6.23 indicates that the Frank Kamanetskii parameter has the same effect: as the FK parameter increases then the critical hotspot temperature decreases, as we saw in the one dimensional case. We find that higher critical temperatures are necessary to induce ignition within two dimensional stockpiles; the smaller hotspot size comparatively to the stockpile may be having an effect on this.

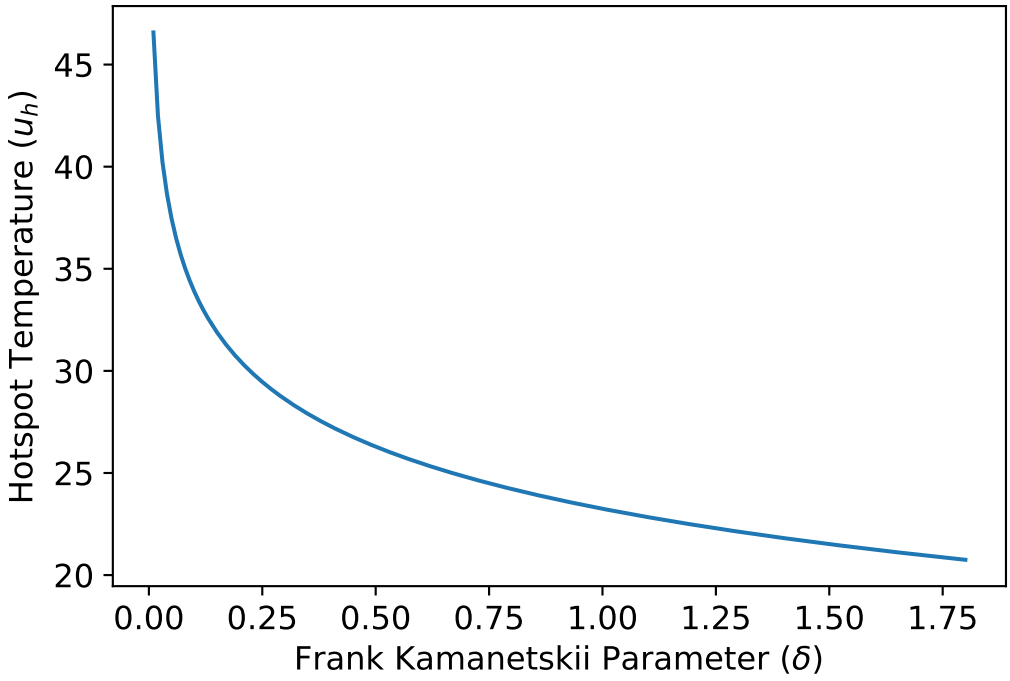
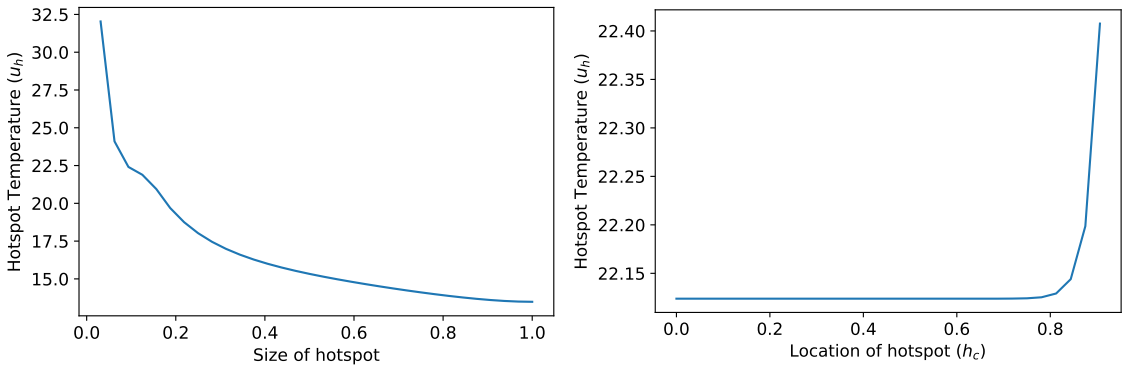


Figure 6.23: Effect of the FK parameter on the critical hotspot temperature of a two dimensional stockpile.Parameter values: $h_c = 0.9$, $h_l = 0.1$.

The size and location of the stockpile differ slightly from what we found in the one dimensional case though the overarching trends remain the same, these are displayed in figure 6.24. We observe that the location of the hotspot is much less significant in the two dimensional model. The critical hotspot temperature changes slightly when it is close to the boundary but once the hotspot is moved in slightly then any additional movement does not decrease the critical hotspot temperature. This contrasts to the one dimensional model where there was a considerable change in the critical hotspot temperature as the location of the hotspot moved closer to the centre of the stockpile. The size of the stockpile still has a considerable effect on the the critical hotspot temperature. We do not observe the same increase in critical hotspot temperature when the hotspot length, h_l approaches 1 as there is still reactive material in the stockpile.

When we consider stockpiles in two dimensions some of the behaviour of the one dimensional stockpiles changes. The location of the hotspot has less effect on the critical hotspot temperature and the size of the hotspot has a different behaviour as the length of



(a) Size of hotspot. $h_c = 0.9$.

(b) Location of hotspot. $h_l = 0.1$.

Figure 6.24: The effect of the size and location of the hotspot on the critical hotspot temperature. Parameter values:$\delta = 0.45$.

the hotspot increases. However the key behaviours of the stockpiles are still observed. As the FK parameter increases the critical hotspot temperature decreases. Larger hotspots also decrease the critical hotspot temperature. The main difference is the effect that the location of the hotspot has though it must be acknowledged that we are only shifting the hotspot in one direction and we may find additional benefits by changing the location in the other dimension.

6.3 Summary

Our analysis into the relationships between the hotspot characteristics and external factors in inducing stockpile ignition provide a basis for further commercial viability into the use of hotspots. We have considered the two extreme cases of a hotspot consisting of fresh heated material, and a hotspot consisting of a purely inert material. In practical cases where a previous stockpile has been used as the material for a new hotspot, then the likely result is that the initial concentration will be between these two bounds. On top of this there are logistical considerations that need to be made about the size of the stockpiles and hotspot that can be used as well as if there is any cost involved with heating the material. The location parameter is also crucial in a logistical sense because there may be more cost associated with placing the hotspot internally within the stockpile rather than just limiting this to the end which is simpler. Such a cost benefit analysis could be carried out though further information would be required about the specific parameters and constraints involved.

Chapter 7

Applications to Stockpile Ignition

Now that we have established the posterior distribution of the kinetic parameters and examined the non-dimensional stockpile equation, we can apply these results to the stockpiles used by BlueScope Steel. As described in Chapter 2, the self-heating behaviour of the stockpiles is irregular. Simulating the large stockpiles can develop insights into understanding this irregularity. We investigate the parameters that influence the stockpiles self-heating. Initially, we use the model

$$\rho c \frac{\partial T}{\partial t} = \nabla \cdot (k \nabla T) + Q_1 A_1 \rho_W \exp\left(\frac{-E_1}{RT}\right) + Q_2 A_2 \rho_I \exp\left(\frac{-E_2}{RT}\right), \tag{7.1}$$

introduced in Chapter 3. In this model we have the reactions,

$$6\mathrm{FeO} + \mathrm{O}_2 \rightarrow 2\mathrm{Fe}_3\mathrm{O}_4,$$
$$3\mathrm{Fe} + 2\mathrm{O}_2 \rightarrow \mathrm{Fe}_3\mathrm{O}_4.$$

In Equation (7.1), the wüstite (FeO) reaction is denoted by the subscript 1 whilst the Iron (Fe) reaction is denoted by the subscript 2 We have included the mass density of the reactants ρ_W and ρ_I. Typically, such values are absorbed into the pre-exponential factor. This changes the value of the pre-exponential factor we estimated in Chapter 4, hence the explicit inclusion of the reactant densities in our model. This is why it is important to simultaneously consider both the large stockpile model and the TGA model when defining the parameters. We change the units for time between the simulation of the stockpiles and those of the experiments. The experiments are conducted on a minute timescale whilst the stockpiles are considered across a year.

The stockpile undergoes a self-heating process. When this self-heating is rapid, the stockpile temperature rises rapidly, which we refer to as ignition. Mathematically, we need to define when ignition occurs and we do this in a similar manner to our approach in Chapter 6. We define a threshold temperature such that when the maximum temperature

within the stockpile exceeds this point, then the stockpile is determined to have ignited. If the stockpile ignites within our integration period, then we class these stockpiles as supercritical. Stockpiles that have not ignited are classed as subcritical. We consider initial temperature distribution equal to the ambient temperature. If we increase this initial temperature we can induce self-heating in these subcritical stockpiles as was demonstrated in Section 6.2. Supercritical stockpiles experience thermal runaway for our initial condition, and in the theoretical case $\varepsilon = 0$ thermal runaway occurs independently of the initial condition. Without the consumption of any reactants, the heat produced from the reactions increases as the temperature increases. Our definition of ignition is sufficient in these cases but may need to be altered when considering stockpiles where the reactant is consumed.

The FK theory predicts that there is a critical length L_{cr}, such that if the length of the stockpile L is less than the critical value, then the stockpile is subcritical. Conversely, if the length is greater than the critical value then the stockpile is supercritical. Using the posterior distribution of the parameters we defined in Chapter 4, we can calculate the posterior distribution for the critical length required for the stockpile to ignite. An analytical expression for this distribution is not possible. We need to evaluate the critical length numerically, and we do not have an analytic expression for the posterior distribution of the parameters. Since we have a sample for the posterior distribution of the parameters, we can simulate the stockpiles for each sampled point and solve for the critical length numerically. Subsequently, we obtain a sample from the posterior distribution of the critical length of the stockpile. This process can also be used to determine the posterior distribution of the ignition time required for a stockpile. The ability to apply a functional transformation of the sampled points is one of the major benefits of working within the Bayesian framework.

To simulate these stockpiles we solve Equation (7.1) numerically using the Fenics package in Python [136, 137]. We use a finite difference approximation in time whilst using Fenics which solves a spatial PDE at each time point using a finite element method. Additional details about the numerical algorithm and convergence plots are in Appendix B. Simulating the stockpiles for every sampled parameter in the posterior distribution is computationally intensive. As a result, it is more computationally effective to examine the behaviour of the stockpiles and the relationships between the parameters by using point estimates of the parameters. One option to do this would be to randomly sample from the posterior distribution. However we opt to use the medians of each of the marginal distributions. The median values are a useful point estimate for the location of the parameter values. The median parameter values along with the other model parameters are provided in Table 7.1.

Parameter		Value
Pre-Exponential Factor	A_1	2.48×10^4 min^{-1}
	A_2	1.42×10^8 min^{-1}
Activation Energy	E_1	5.84×10^4 J/mol
	E_2	1.23×10^5 J/mol
Reaction Enthalpy	Q_1	1.29×10^3 J/g
	Q_2	6.95×10^3 J/g
Thermal Conductivity	k	80 J/min/m/K
Specific Heat	c	1.6 J/g/K
Density	ρ	7.874×10^6 g/m^3
	ρ_I	0.11ρ
	ρ_W	0.42ρ
Ideal Gas Constant	R	8.314 J/mol/K
Ambient Temperature	T_a	290 K

Table 7.1: Parameter values for the stockpile simulations.

7.1 The one-dimensional model

We begin our analysis on Equation (7.1) for a one dimensional domain, $[-L,L]$ with the Dirichlet boundary condition,

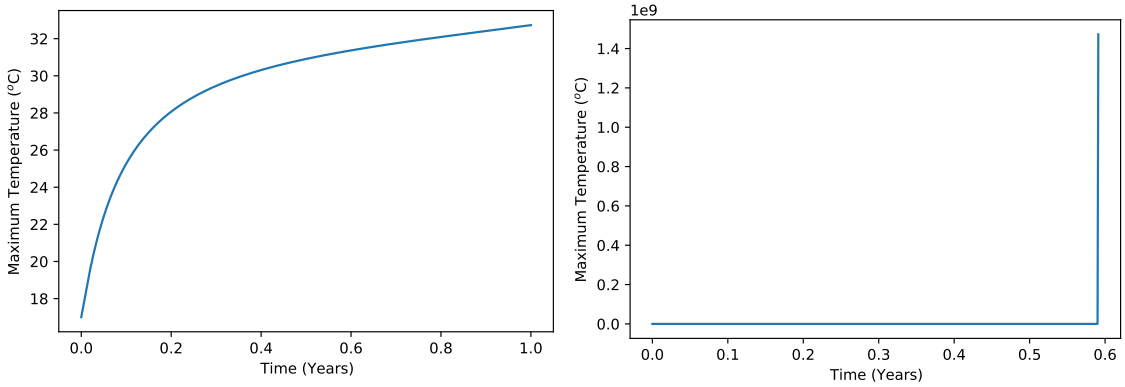
$$T = T_a \tag{7.2}$$

We begin our analysis by ignoring the consumption of any reactants. We consider the consumption of reactants in Section 7.1.2. Initially, we only consider the median values of our marginal posterior distribution; we consider the posterior predictive distributions in Section 7.3.

We simulate the stockpiles using the parameter values in Table 7.1. In order to evaluate criticality we must define our integration period and ignition temperature. We use an integration period of one year and a critical ignition temperature, $T_{\text{ig}} = 800^o$C. Using this criteria we determine the critical length, $L_{cr} = 4.13$m. This length represents the half length of the stockpile. The one-dimensional model restricts diffusion to one dimension. This approximation is valid when one dimension is much smaller than the other two [28]. Subsequently, our results indicate that if a large area is covered with the filter cake to a height of 8.26m, then this is necessary to induce ignition. If less heat is transferred to the ground, with the maximal scenario being perfect insulation, this height is reduced to the critical value $L_{\text{cr}} = 4.13$m. This one dimensional approximation is useful, particularly to examine the different behaviour that can be observed, though it is not as accurate as the higher dimensional simulations we consider in Section 7.2. The key advantage of a one dimensional approach is computational efficiency, and the temperature distribution of one dimensional stockpiles is easier to visualise. Given that we observe similar behaviours in different dimensions, we examine the relationships in one-dimension.

Now that we have evaluated the critical length of the stockpiles, we simulate sub-critical

and supercritical stockpiles and compare their behaviour. We also want to examine the effect that each reaction has on the stockpiles ignition. To do this we set the pre-exponential factor equal to zero. Alternatively the reaction enthalpy could also be set to zero. This removes any heat generation from the respective reaction.



(a) Sub-critical Stockpile $L = 4.1$m.

(b) Super-critical Stockpile $L = 4.2$m.

Figure 7.1: Comparison of the maximum temperatures within a sub-critical and supercritical stockpile in one dimension.

In Figure 7.1 we compare the maximum temperatures within the sub-critical and supercritical stockpile. We observe that there is a significant difference between the two stockpiles with a slight difference in length. This highlights how quickly the stockpiles heat without any material consumption. In practice there are several limiting factors upon the maximum temperature, such as the consumption of the reactants and additional heat transfer mechanisms such as radiation. As noted in Section 2.2.6, this form of energy transfer scales with T^4, hence at the low temperatures observed in the sub-critical cases, this can be ignored; as the temperature increases this becomes significant. This means that in reality the maximum temperature does not reach those seen figure 7.1b. The FK models are useful for determining when the reaction leads to significant self-heating but cannot determine the temperature a supercritical stockpile reaches.

The most interesting stockpiles to examine are those that have a length just below our critical value. Many of these stockpiles would reach our ignition criteria with a longer integration period and as such we can see the beginning of the self-heating process. Figure 7.2 shows the temperature of a stockpile throughout the integration period. The stockpile is classed as sub-critical since the temperature does not exceed our ignition temperature within the integration period. In these plots we scale the length so that the plotted length is 10. This allows us to compare stockpiles on similar plots when we change the length of the stockpile. Figure 7.3 plots the maximum temperature of the stockpile and indicates that there is an inflexion point where the rate at which the stockpile temperature is increasing begins to increase. This inflexion point could be used as a means of determining ignition, however, this does not significantly affect the critical length. The sub-critical stockpile in

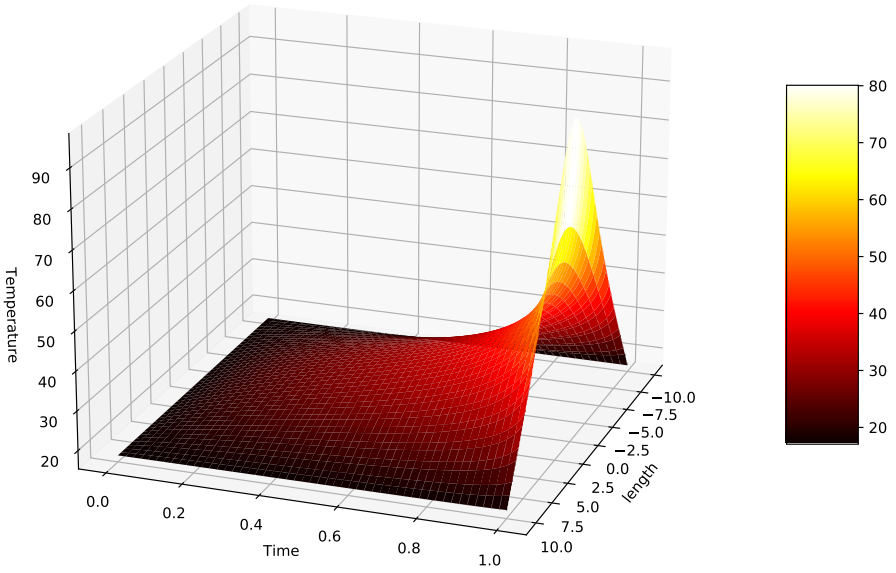


Figure 7.2: Simulation of a sub-critical stockpile with $L = 4.128$m, with the length scaled.

Figure 7.1a does not have an obvious inflection point and as such it may take a significant amount of time for this stockpile to reach the ignition temperature if it does. Considering these stockpiles differ by less than 0.1m then it is possible this is less than the variance in the critical lengths due to variance in the parameters.

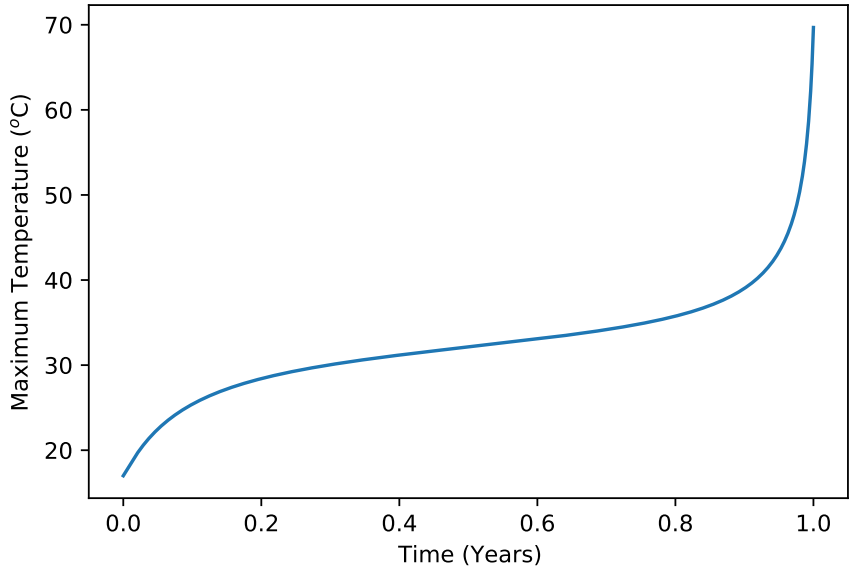
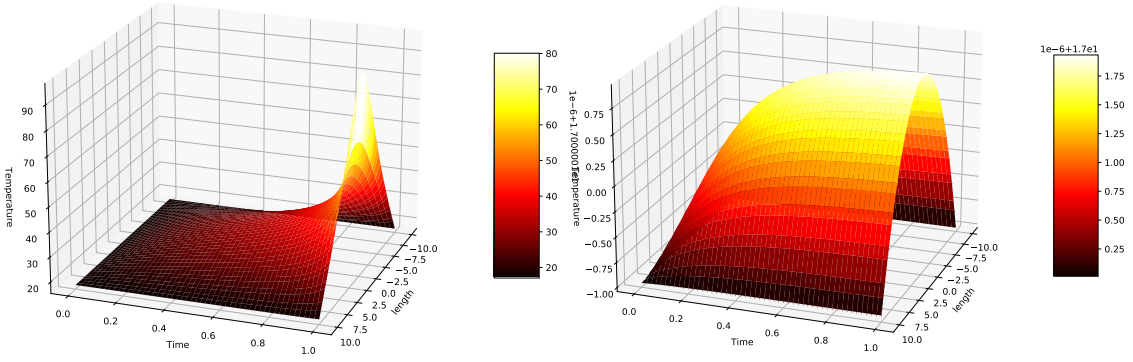


Figure 7.3: Maximum temperature for a stockpile with $L = 4.128$m.

In order to understand the effect that each of the reactions has on the stockpiles we



(a) Wüstite reaction only.

(b) Iron reaction only.

Figure 7.4: Comparison of the stockpiles that only have a single reaction occurring, $L = 4.128$m.

compare this result to stockpiles that only have one reaction. Figure 7.4 compares the stockpiles with only having one reaction. What we notice is that the stockpile with the wüstite reaction has an almost identical temperature profile to the stockpile containing both of the reactants. The stockpile with the iron reaction experiences negligible self-heating as the change in temperature is on a scale of 10^{-6}K. This implies that we can assess criticality of the stockpiles using only the wüstite reaction.

To investigate this further we compare the heat generation terms, $QA\exp(-E/RT)$, for both of the reactions. Figure 7.5 verifies that at low temperatures the heat produced from the iron reaction is many orders of magnitude less than the heat produced from the wüstite reaction. For higher temperatures this is reversed and as such if we consider high temperature simulation of the stockpiles after considering other limiting factors then we would observe that the stockpiles are affected by both reactions.

7.1.1 Oscillating boundary conditions

In Chapter 6 we found that oscillating boundary conditions have a significant effect on the critical length. Using this knowledge we investigate the effect of periodic boundary conditions for the steel stockpiles. We implement the Dirichlet boundary condition,

$$T = T_a + T_o \sin\left(\frac{2\pi t}{\omega} + \phi\right). \tag{7.3}$$

By setting the parameter $T_o = 0$ we recover the static boundary condition. We simulate the stockpiles using the boundary condition equation (7.3) with oscillation temperature $T_o = 8^o$C, oscillation period $\omega = 1$year, and phase shift $\phi = 0$. As with section 6.1, the phase shift corresponds to the time of construction with $\phi = 0$ occurring in spring.

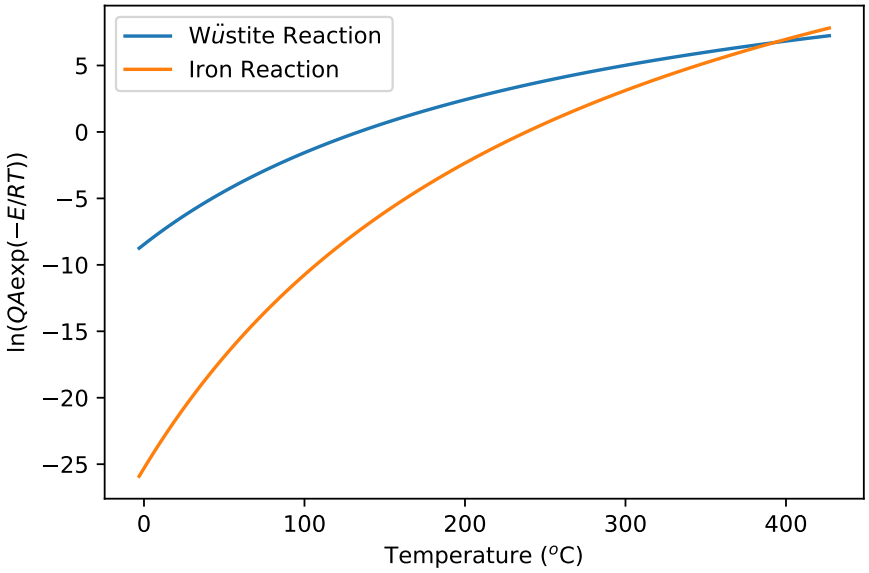


Figure 7.5: Comparison for the heat generation terms for the two reactions.

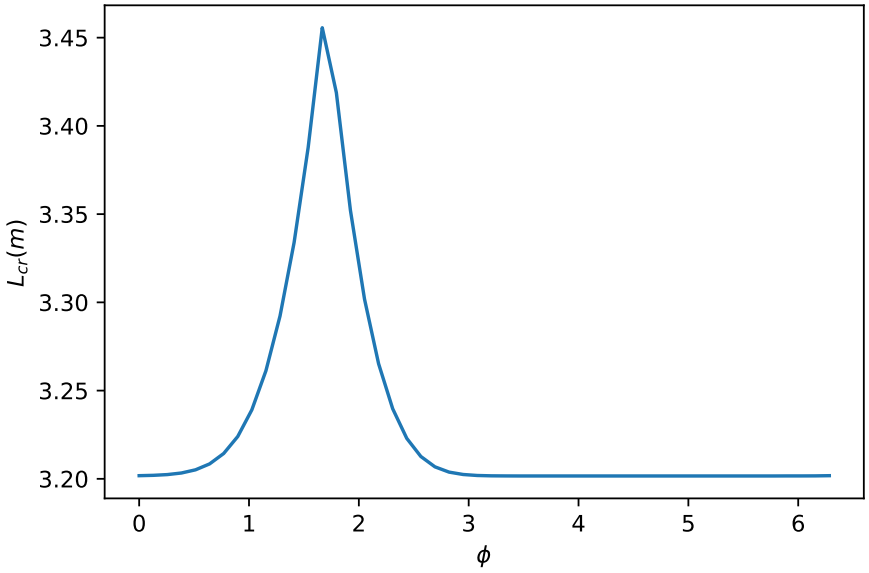


Figure 7.6: Critical stockpile lengths for different times of construction. $T_o = 8$K.

Figure 7.6 indicates that for different construction times the difference in critical length required for ignition is not significant. This appears to contradict the results that we obtained in Section 6.1. The difference between the two sets of analysis is that we fixed the length of the stockpile meaning that changing the FK paramter δ was changing the reaction kinetics. However for our analysis in this section, the reaction kinetics are fixed and the length is changing. When the length changes this also affects the scaling of the integration period which causes the discrepency between the results. Our examination

of the one reaction problem in Section 6.1 is still relevant as this needs to be considered when we view the parameters as coming from the posterior distribution. For a fixed length stockpile, the variation in the posterior distribution changes the FK parameter which affects the probability that the stockpiles will ignite given the model and the experimental data.

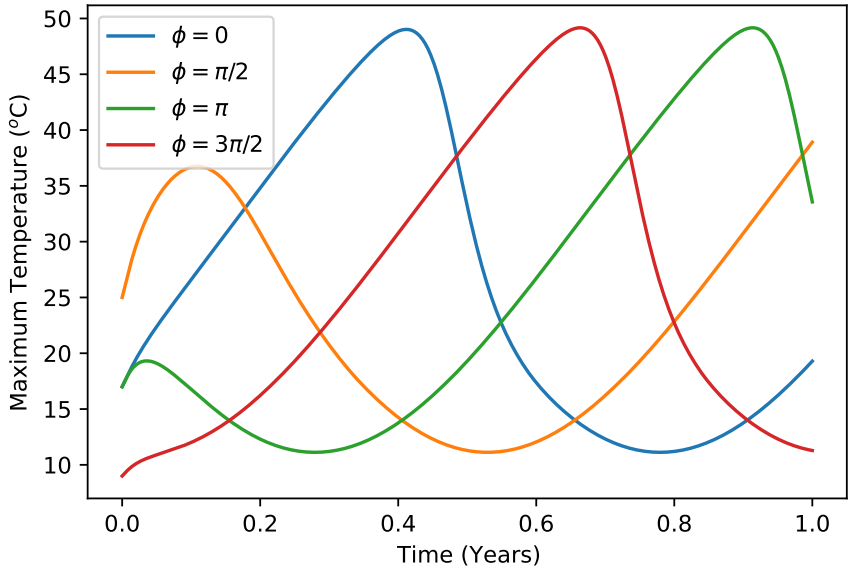


Figure 7.7: Comparison of the maximum temperature of stockpiles with different construction times, $T_o = 8$K.

Figure 7.7 indicates that for sub-critical stockpiles the variation in construction time has minimal effect. We observe that the maximum temperatures follow the same pattern with time shifted. This explains why the difference in critical length was negligible for many of the construction times. We provide further explanation to this by observing the ignition times in Figure 7.8. We select a length for which most of the stockpiles ignite and calculate the time passed before we reach ignition. There is an almost linear relationship between the phase shift and the ignition time for a range of values. The gaps indicate stockpiles that are sub-critical. It appears that although the time of construction has minimal effect on the necessary length required to induce ignition, it does significantly affect the ignition time.

In addition to the phase shift we also investigate the effect of the oscillation temperature. The amplitude of the oscillations reflect the difference in the maximum variation in the seasonal temperature variations. This will be dependant upon the location of the stockpiles. Figure 7.9 indicates that the critical length decreases as the oscillation temperature increases. This result is similar to what we found in Chapter 6.

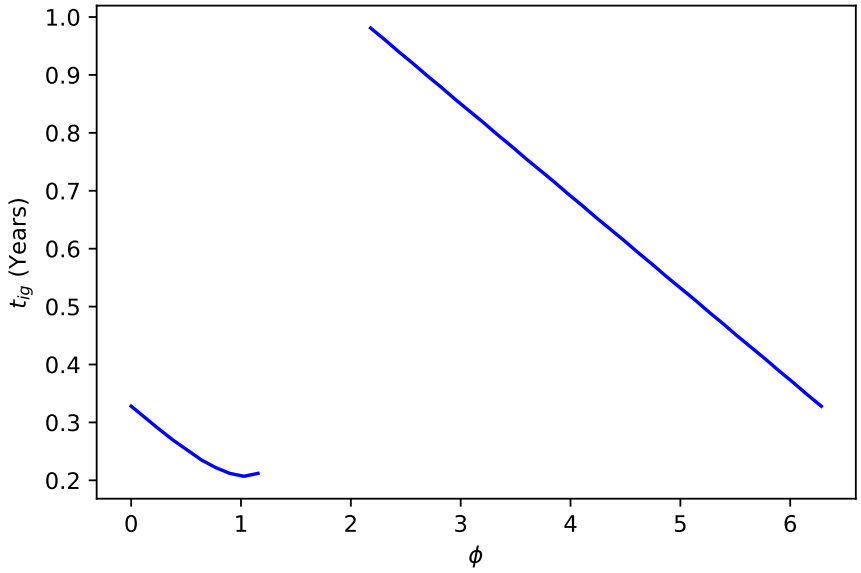


Figure 7.8: Comparison of the ignition times of stockpiles with different construction times. $T_o = 8$K, $L = 3.3$m.

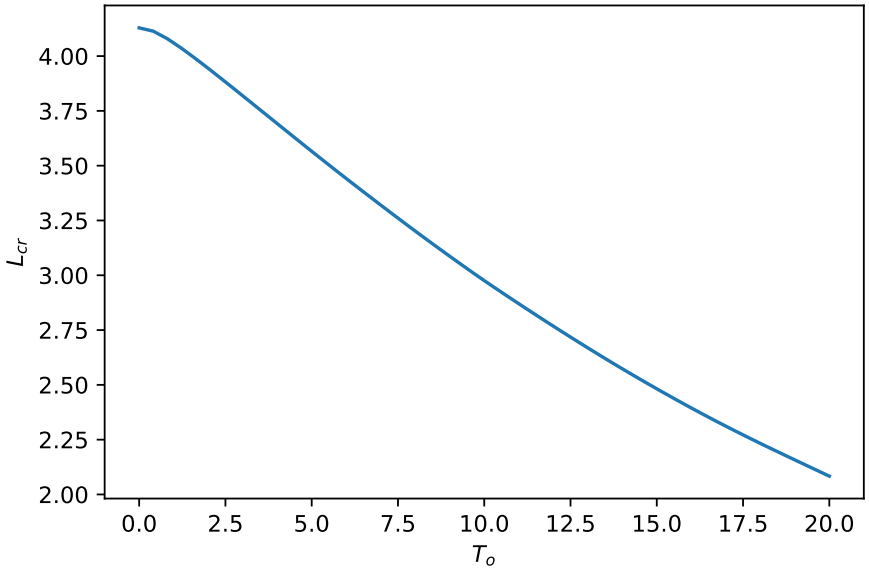


Figure 7.9: Critical stockpile lengths for different oscillation temperatures $\phi = 0$.

7.1.2 Addition of Reactants

One point of interest regarding the stockpiles at BlueScope steel is that the material within the stockpile becomes stronger after the self-sintering process which occurs at the elevated temperatures. As such, modelling how the material is consumed in the stockpile is a valuable process. Including the consumption of reactants into the stockpile models requires consideration of the units involved. This is why it was so critical that we were mindful

of units when establishing the model for the TGA/DSC reactions in Chapter 3, so that the estimated parameters are transferable to the larger stockpile model. When including reactant consumption equation (7.1), becomes,

$$\rho c \frac{\partial T}{\partial t} = \nabla \cdot (k \nabla T) + Q_1 A_1 \rho_W W \exp\left(\frac{-E_1}{RT}\right) + Q_2 A_2 \rho_I I \exp\left(\frac{-E_2}{RT}\right), \tag{7.4}$$

where W and I represent the mass fraction of wüstite and iron respectively. The equations for the reactants are,

$$\frac{\partial W}{\partial t} = -A_1 W \exp\left(\frac{-E_1}{RT}\right), \tag{7.5}$$

$$\frac{\partial I}{\partial t} = -A_2 I \exp\left(\frac{-E_2}{RT}\right). \tag{7.6}$$

These equations represent the decrease in mass fraction of the wüstite and iron. The initial value of wüstite and iron are $W_0 = 1$ and $I_0 = 1$. In this set up the mass densities ρ_W and ρ_I are interpreted as the initial mass densities of wüstite and iron respectively. This provides consistency with our TGA/DSC models that we used to estimate the kinetic parameters and allows us to use these results in this stockpile model. It is important to note that there is no diffusion with the reactants as they are solid. The variables W and I vary spatially since the rate of consumption is higher in locations with a higher temperature.

When reactant consumption is included we have to reconsider the ignition criteria. For small stockpiles it may be the case where there is insufficient material to cause the stockpile to reach certain temperatures. In some cases it may be more relevant to categorise ignition based upon the consumption of material. We therefor classify the stockpile into four categories:

- Subcritical- Stockpiles that have consumed insufficient material.
- Partially Consumed - Stockpiles where the wüstite has been consumed.
- Completely Consumed - Stockpiles where both wüstite and iron have been consumed.
- Supercritical - Stockpiles that reach the ignition temperature.

We classify consumption when at least 50% of the respective material has been depleted. Defining criticality based off reaching a specified temperature becomes less important when we consider the reactant consumption. This is mainly useful if we require a certain temperature to be reached, however most of the time the interest will be in when the reactants are consumed.

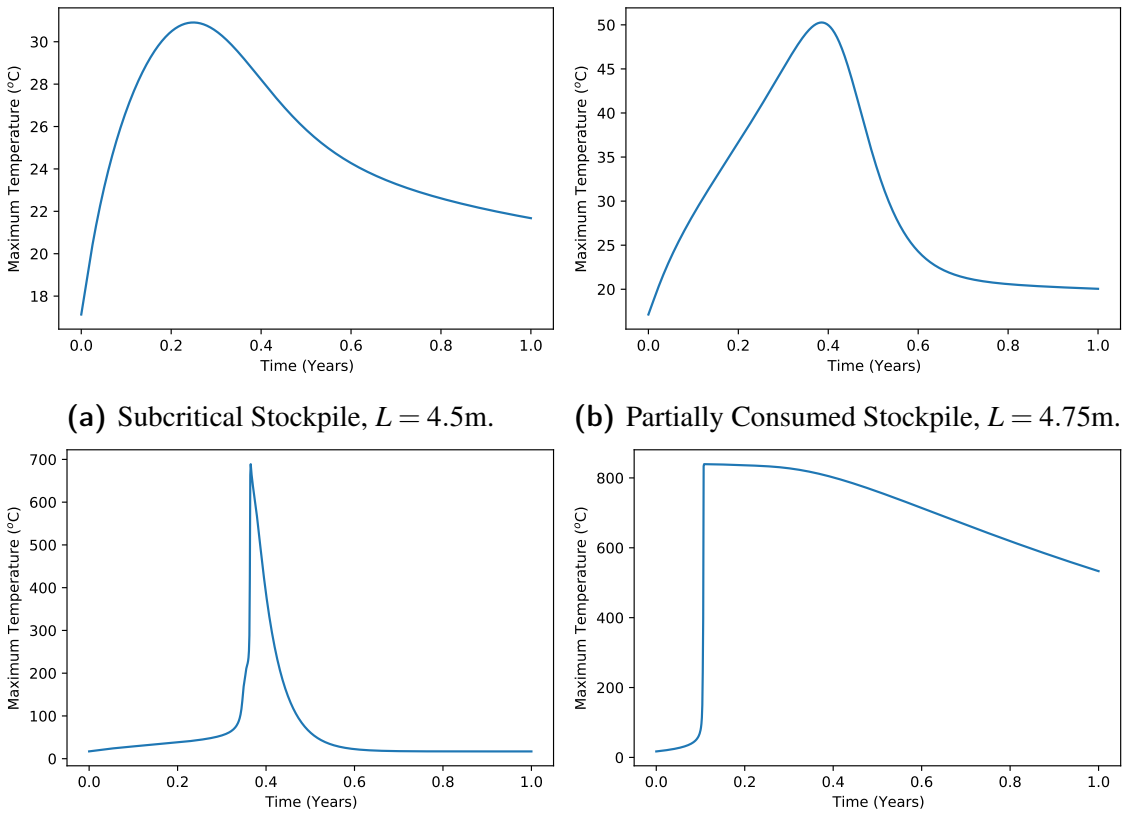
From our analysis of a single reaction model, for each reactant without consumption, we observed that the wüstite reaction occurred before the iron reaction. In simulating the stockpiles with reactant consumption we observe that the wüstite is consumed before the iron. This is expected based upon our estimates for the parameter values and the stockpile simulations correspond with the TGA data.

Our investigations indicate that these criticality classifications, (excluding subcritical) are subsets for our parameters. We do not have a supercritical stockpile that does not meet the criteria for partial of complete consumption. We have not investigated the full scope of possibly parameters that can be used for the model and thus cannot state whether this is universal. In instances where the heat of reaction Q and activation energy is large, or where we define ignition using a low temperature, than supercritical stockpiles, by our definition, may not reach complete or partial consumption. Because of these instances, we focus on the consumption conditions as the points of interest in our investigation. Completely consumed stockpiles are a subset of the partially consumed stockpiles by definition. Furthermore, we find that for fixed parameters this classification can be determined based upon the length of the stockpile.

For each classification we can define a critical length such that if the stockpile length is greater than the critical value then the stockpile is within that classification. We denote the critical lengths L_P, L_C and L_I, for the critical lengths necessary for partial consumption, complete consumption and supercritical respectively. These critical lengths satisfy the condition, $0 < L_P < L_C < L_I$. For the median values of our parameter estimates we calculate the critical values as, $L_P = 4.54$m, $L_C = 4.79$m and $L_I = 24.8$m. These critical lengths are higher than the critical length we found for stockpiles where the reactant consumption is ignored. The critical length for ignition L_I is high due to the choice of parameters that we have. For our parameter values there is not enough material in the smaller stockpile to cause the temperature to reach our ignition threshold. It is worth noting that from our analysis of the experimental data it appeared that there may be a third reaction and that our parameter estimates using a two reaction model underestimate the amount of heat being generated. A revised reaction scheme is therefore likely to produce different results.

Our parameters create an interesting region where only the wüstite has been consumed whilst the iron has not. Understanding the temperature profile in this region may shed the most light on the self-heating behaviour. The existence of such a region provides a clear indication that the wüstite reaction drives the self-heating process. We compare a partially consumed and completely consumed stockpile in Figure 7.10. One of the most significant differences is the maximum temperature that the stockpile reaches. When both reactants are consumed then the stockpiles reach a much higher temperature.

We have identified that the wüstite reaction drives the self-heating within these stockpiles, so it is useful to compare the temperature that the stockpiles reach if only the wüstite



(a) Subcritical Stockpile, $L = 4.5$m. **(b)** Partially Consumed Stockpile, $L = 4.75$m.

(c) Completely Consumed Stockpile, $L = 4.8$m. **(d)** Supercritical Stockpile, $L = 25$m.

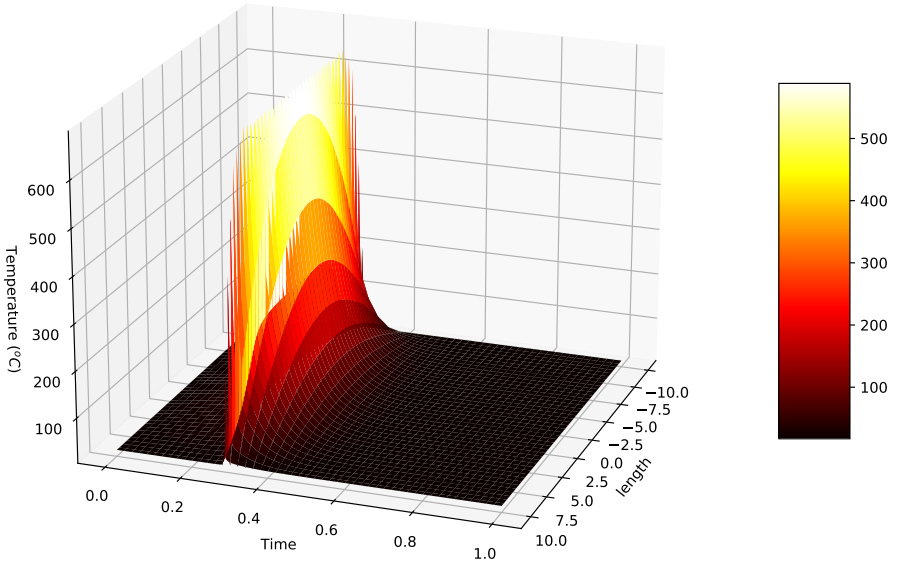
Figure 7.10: Comparison of the maximum temperature between stockpiles in different ignition classes.

reaction was occurring. We set the pre-exponential factor for the iron reaction to be zero. Figure 7.11 indicates there is a significant temperature difference between the two stockpiles. Subsequently this temperature difference is attributed to the effect of the iron reaction. It seems as though the wüstite reaction is initiating the self-heating process but the iron is enabling high temperatures to be observed.

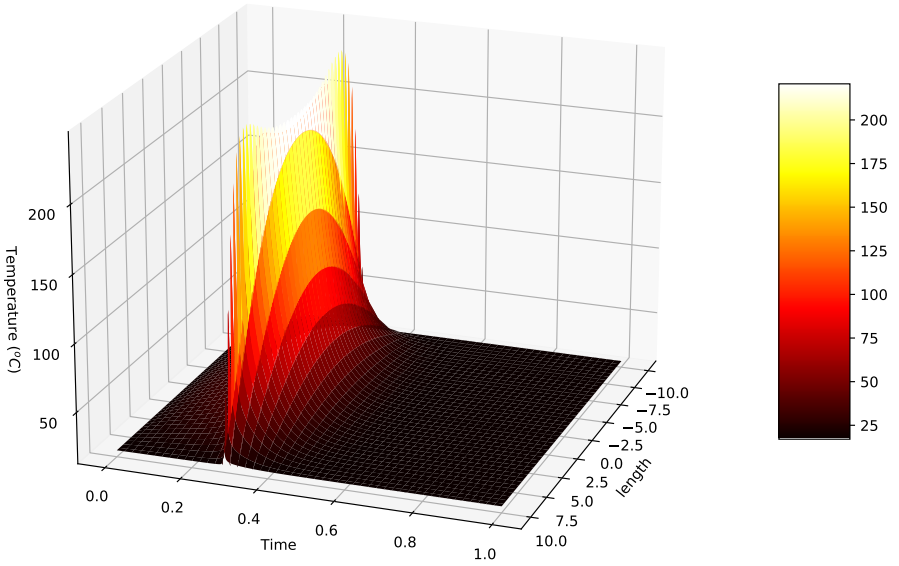
Addition of Hotspots

The addition of heated material has been applied to the stockpile stored onsite with various levels of success. Ideally we would like to determine a bound for the size and temperature of a hotspot that would induce ignition within these large stockpiles. In the previous chapter, we have demonstrated that when hotpsots are added into subcritical stockpiles that this can induce the rapid self-heating that causes the filter cake to self-sinter. Here we introduce hotspots into the larger stockpile model.

We add hotspots into the model using the same mechanism as in the previous chapter. We consider the hotspot H to be the region $H = [L(h_c - h_l), L(h_c + h_l)]$ with the restriction $-1 \leq h_c - h_l < h_c + h_l \leq 1$. We define our hotspot by changing the initial condition of the stockpile. We begin by considering a region of heated material that has not reacted.



(a) Standard stockpile.



(b) Wüstite only.

Figure 7.11: Comparison of the temperature of a stockpile with both reactants against a stockpile with only wüstite. The Temperature axis has a different scale on both to demonstrate a similar profile but different temperatures are reached. $L = 5$m.

In later analysis we use inert hotspots. We set the initial temperature T_0 as,

$$T_0(x) = \begin{cases} T_h + T_a & x \in H \\ T_a & x \notin H \end{cases}, \tag{7.7}$$

with the initial mass fractions given by,

$$W_0(x) = \begin{cases} W_h & x \in H \\ 1 & x \notin H \end{cases}, \qquad I_0(x) = \begin{cases} I_h & x \in H \\ 1 & x \notin H \end{cases}, \tag{7.8}$$

where W_h and I_h are the initial mass fractions inside the hotspot for wüstite and iron respectively. For unreacted material these values are $W_h = I_h = 1$ whilst for the inert hotspots $W_h = I_h = 0$. This hotspot condition is the same as was used in the previous chapter where we examined the relationships between the critical hotspot temperature and the size and location of the hotspot.

When investigating the influence of hotspots on ignition there are several important relationships to consider. Introducing hotspots into the problem adds additional parameters that can be controlled. Previously we were limited to changing the length of the stockpile. We want to examine how the critical length of the stockpile changes as we introduce different temperatures for the hotspot. We showed in our previous analysis in Chapter 6 hotspots are particularly useful for inducing ignition in stockpiles. We need to know what hotspot temperatures are needed for a hotspot of specified size in order to induce ignition.

When we include hotspots we observe that these have minimal effect on the maximum temperature that the stockpile reaches. In the cases where all the material is consumed, we notice that this self-heating process is quick and reaches a peak before cooling gradually as heat diffuses through the stockpile and is exchanged at the boundary. For a fixed length stockpile there is only a limited amount of reactants that can be consumed. This limits the amount of heat that can be generated and subsequently the maximum temperature of the stockpile. This explains why the hotspots have minimal effect on inducing this critical condition. In circumstances where there is more heat generated within the stockpile it is likely this will have an impact though for our parameter values we have no effect.

Figure 7.12 displays the relationship between the critical length of the stockpile and the temperature of the hotspot. As the temperature of the hotspot increases then the critical length decreases. This is consistent with our previous analysis looking at the effect of the hotspot temperature on the FK parameter. We find that for low hotspot regions these is a much larger region where only one reaction will occur. In addition to looking at the critical lengths we want to know if there are critical hotspot temperatures that induce our different ignition stages. We define these critical hotspot temperatures such that if $T_h > T_P$, then the stockpile reaches partial consumption and if $T_h > T_C$ then the stockpile reaches complete consumption. Our analysis indicates that we do not have a critical temperature to induce our critical ignition temperature as we are considering stockpile where the length is much less than the critical length for ignition L_I.

Figure 7.13 indicates that the shorter the stockpile, the higher the hotspot temperature necessary to induce ignition. This agrees with our previous results looking at ignition

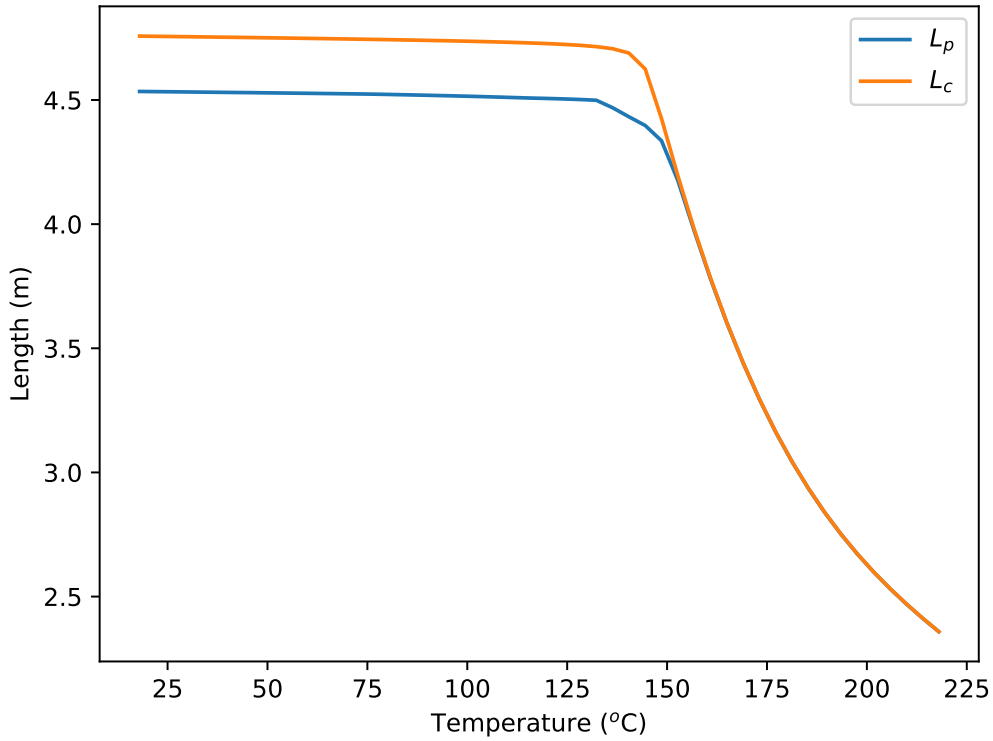


Figure 7.12: The critical length for a stockpile with a hotspot. Parameter Values: $h_c = 0.9$, $h_l = 0.1$.

from hotspots. We find that the length of the stockpile has a significant effect in the critical hotspot temperature. Additionally, for stockpile near the maximum length we consider, the difference between the critical temperature for one reaction occurring and two reactions becomes much greater. For stockpiles that are considerably smaller than critical, then a high hotspot temperature is necessary to induce ignition.

There are two ways in which we control the geometry of the hotspot. We can change the size, h_l or the location h_c. Investigating the effect that each of these parameters have on the critical hotspot temperature can influence how the hotspots are constructed. Figure 7.14 indicates that we have a similar behaviour to that we saw with the non-dimensionalised FK model, as the size of the hotspot increases we reduce the critical hotspot temperature. In this instance, there is no significant difference between one or two reactions occurring, relative to the change in hotspot width. Similarly, figure 7.12 displays that locating the hotspot at the center of the stockpile is the most conducive to inducing reactant consumption. Although these results were obtained in Chapter 6, it is useful to demonstrate that similar effects are seen when the estimated parameter values are used. This also gives greater insight into the real temperatures that are required.

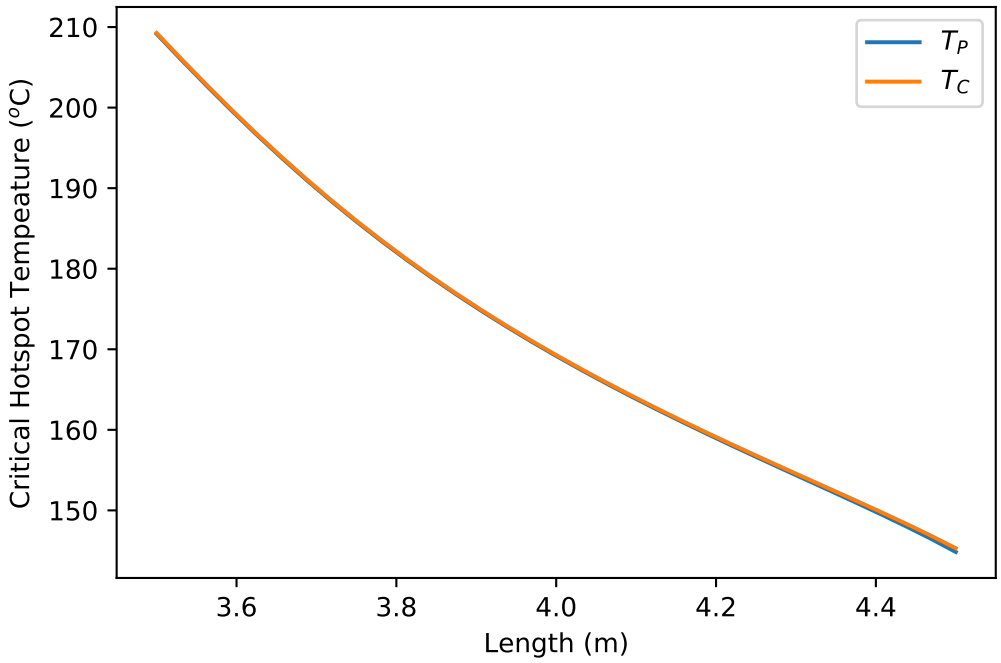


Figure 7.13: Critical hotspot temperature for different length stockpiles. Parameter Values: $h_c = 0.9$, $h_l = 0.1$.

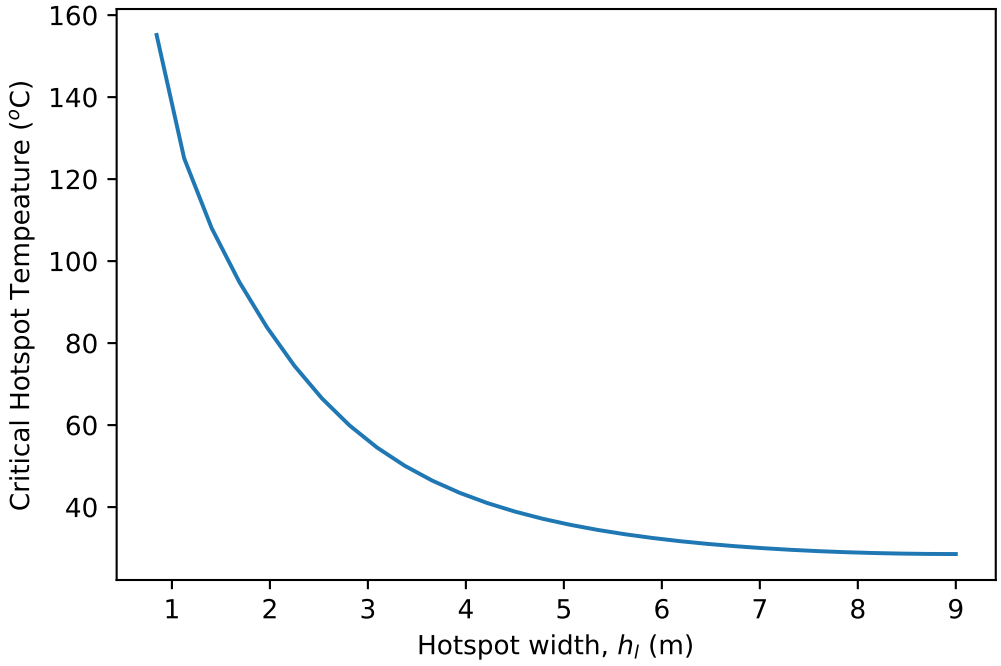


Figure 7.14: The effect of the hotspot size on the critical hostpot temperature. Parameter values: $L = 4.5$ $h_c = 1 - h_l$.

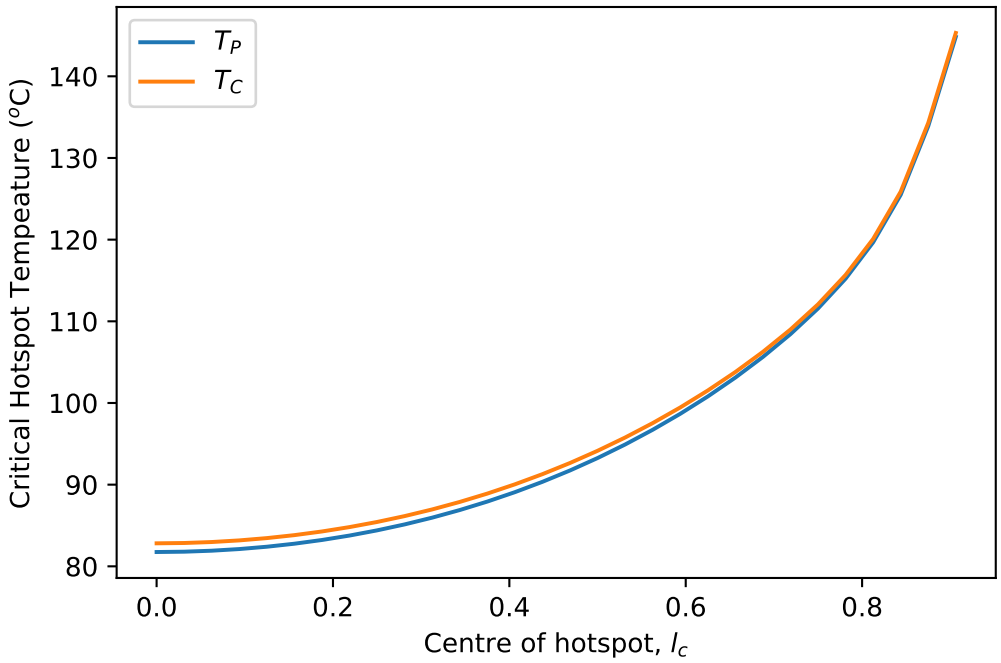


Figure 7.15: The effect the hotspot location has on the critical hostpot temperature. Parameter values: $L = 4.5$m, $h_l = 0.1$.

7.2 Higher Dimensions

One-dimensional stockpile models are very useful for investigating the relationships between parameters. The lower the dimension the lower the computation time needed to numerically solve the PDE. In practice, stockpiles are three dimensional objects. As such it is useful to simulate some stockpiles that are higher dimension. For the higher dimension stockpiles we consider the reactions occurring with consumption of material. That is, we have the model,

$$\begin{aligned}
\rho c \frac{\partial T}{\partial t} &= \nabla \cdot (k \nabla T) + Q_1 A_1 \rho_W W \exp\left(\frac{-E_1}{RT}\right) + Q_2 A_2 \rho_I I \exp\left(\frac{-E_2}{RT}\right), \\
\frac{\partial W}{\partial t} &= -A_1 W \exp\left(\frac{-E_1}{RT}\right), \\
\frac{\partial I}{\partial t} &= -A_2 I \exp\left(\frac{-E_2}{RT}\right),
\end{aligned}$$

with the Dirichlet boundary condition $T = T_a$.

7.2.1 Two Dimensions

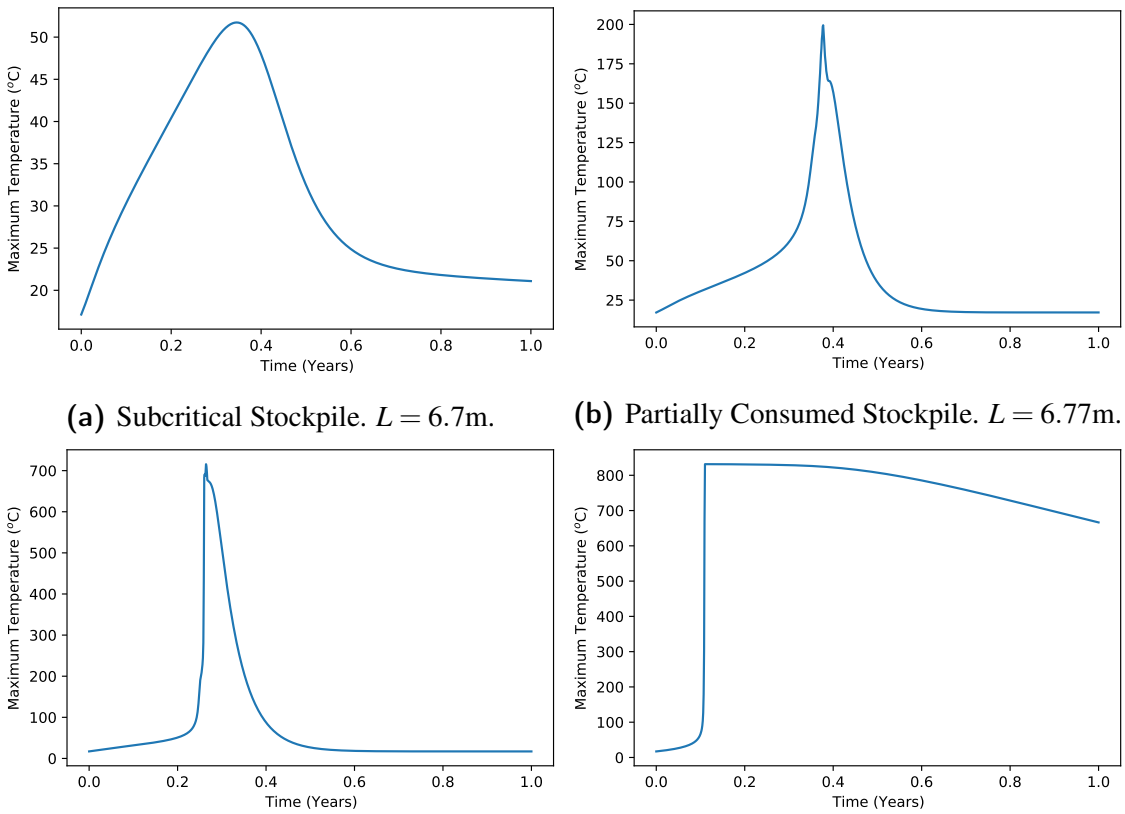
To begin with we investigate two dimensional stockpiles. The two dimensional geometry provides a useful approximation to stockpiles where one dimension is much longer than

the other two. In two dimensions we consider the PDE over a square domain. A square provides a simple geometry that can be described by only one length. Long rectangular stockpiles can be meaningfully described using these models, by making the assumption that diffusion in the long direction is negligible. The temperature variation in this long direction is much lower because the boundary effects where the heat is exchanged is limited in the centre of these stockpiles. Later we consider an example of a three dimensional stockpile where one dimension is much larger than the other two and demonstrate that this is well approximated in terms of critical length by the two dimensional model. In two dimensional models, the height is usually one of the dimensions considered. This means we must question whether the same boundary condition is appropriate on the base, which is exposed to the ground rather than the air.

The two main cases that are considered in the literature discussed in Chapter 2 are a fixed Dirichlet condition $T = T_g$ where this temperature can differ from the ambient temperature T_a, and a Neumann Boundary condition $\frac{\partial T}{\partial n} = 0$. We consider that the ground temperature is unlikely to be fixed and differ from the ambient significantly and hence we can consider the ground temperature to be the same as ambient. This may need to be revisited if we consider Newtonian cooling boundary conditions or periodic boundary conditions. For the Neumann boundary condition, if the initial condition is symmetric about the x-axis (with the y-axis corresponding to height), then this problem can be considered on an extended domain, that is reflected about the x-axis. This allows us to use a Dirichlet boundary condition over this extended domain. For a Neumann bounday condition the stockpile height will be L rather than $2L$ which it is for the Dirichlet boundary condition. These two boundary conditions provide the general extremes of perfect heat transfer, with the Dirichlet condition, and perfect insulation, with the Neumann boundary condition. In practice the value is between these two cases and more investigation would be necessary to determine this.

For the PDE problem on the square domain we calculate the critical lengths necessary to induce each stage of ignition. We calculate that $L_P = 6.75$, $L_C = 6.78$, $L_T = 30.8$. These parameters follow a similar pattern as we saw with the one dimensional case. The length required to cause the stockpile to reach our ignition temperature is much higher than the other critical lengths. These parameters indicate that if we consider perfect insulation with the ground, then in a stockpile with width 14m and height of 7m, all the reactants are consumed. Similarly to the one-dimensional case, we find there is a small region where the only the wüstite reaction occurs.

Simulating stockpiles that are in the different ignition classes, Figure 7.16 shows there is similar behaviour across the different stockpiles. In each case we see a phase of rapid heating until all the reactants are consumed. This is then followed by a phase where the stockpile cools. We see very similar behaviour to what we observed in the one dimensional stockpile.



(a) Subcritical Stockpile. $L = 6.7$m.

(b) Partially Consumed Stockpile. $L = 6.77$m.

(c) Completely Consumed Stockpile. $L = 7$m.

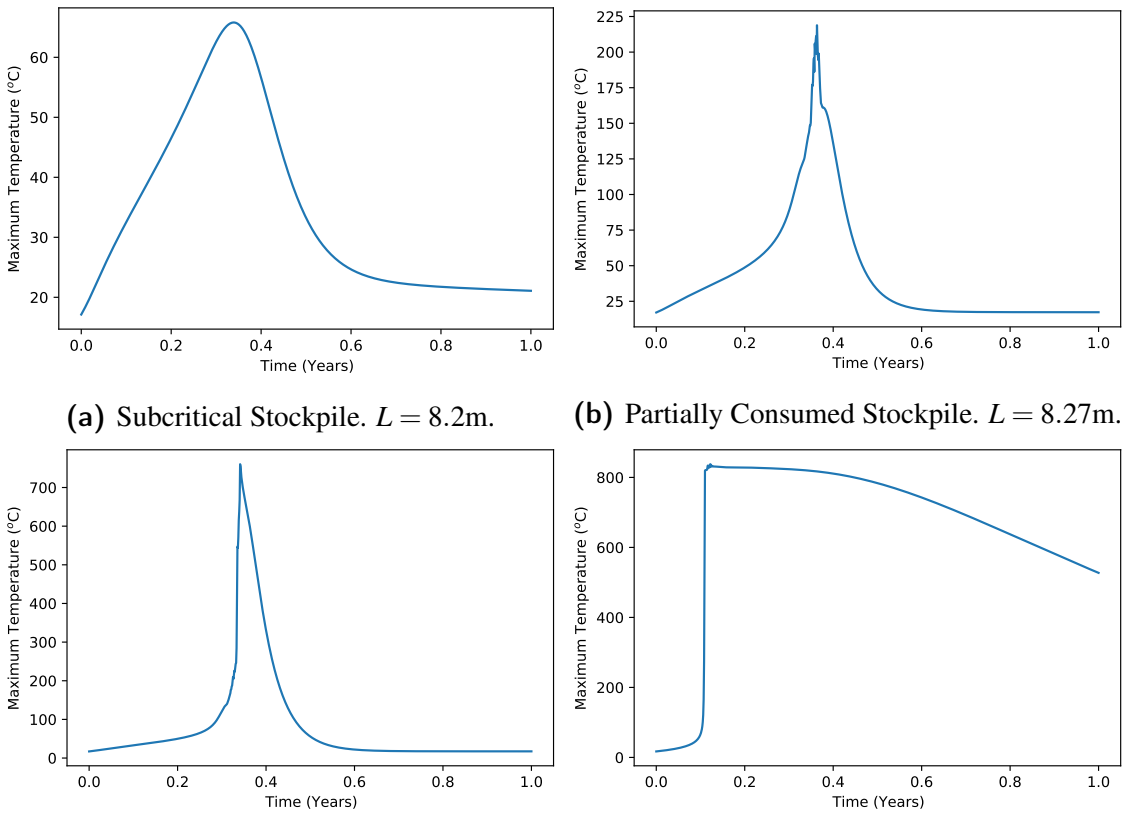
(d) Super Stockpile. $L = 31$m.

Figure 7.16: Comparison of the maximum temperature between two dimensional stockpiles in different ignition classes.

7.2.2 Three Dimensions

We can construct stockpiles that better reflect the geometry of those that are used by BlueScope Steel by extending our work further into three dimensions. The three dimensions add additional computation time and complexity into our model. To keep things simple we consider domains that are cubic. This keeps some consistency with our two dimensional approach and there is a critical FK parameter, $\delta_{\text{cr}} = 2.71$, for the one reaction case where the approximation $\varepsilon = 0$ is made [28].

When calculating the critical lengths necessary for each class of ignition the results are particularly interesting. The critical length for partial consumption, $L_p = 8.257$m, and complete consumption, $L_C = 8.289$m are consistent with what we would anticipate from our previous results. This suggests that for the higher dimensional stockpiles, the critical lengths for the two reactions merge. This is not surprising since the volume of the stockpile scales by L^d where d is the dimension. We also calculate the length for ignition, $L_I = 34$m, which is consistent with our previous investigations.



(a) Subcritical Stockpile. $L = 8.2$m.

(b) Partially Consumed Stockpile. $L = 8.27$m.

(c) Completely Consumed Stockpile. $L = 8.3$m.

(d) Super Stockpile. $L = 35$m.

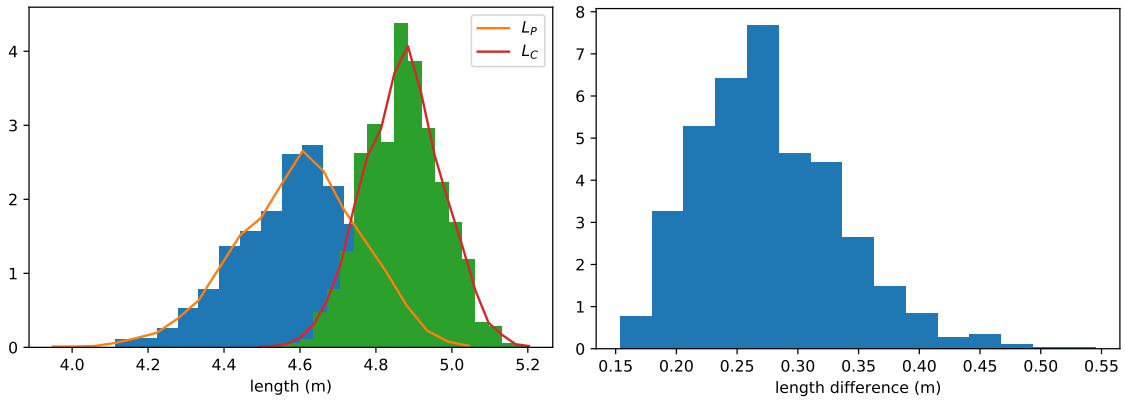
Figure 7.17: Comparison of the maximum temperature between three dimensional stockpiles in different ignition classes.

7.3 Using the Sampled values

Now that we have established some of the behaviour of the stockpiles using the median values from each of our parameters, then it is prudent to consider the posterior distribution of our parameters. Through our Bayesian approach to parameter estimation we are not limited to point estimates for the kinetic parameters. We can use the posterior distribution to determine the posterior distribution for the critical lengths of the stockpiles and determine credible intervals for the critical lengths. Additionally, we can asses the ignition time of the stockpiles and use the posterior distribution of our parameter estimates to determine the posterior distribution of these ignition times. This application of the Bayesian posterior distribution of our parameter estimates provides a valuable tool for assessing the uncertainty in our parameter estimates from the experimental data.

We opt to use a small random sample from this posterior distribution. This reduces the overall computation time whilst allowing us to explore the posterior distribution for the critical lengths and also the ignition times. For the critical lengths we examine the stockpiles including consumption and focus on the two consumption classes, partial and complete. In Figure 7.18 we observe that there is a significant difference between the critical lengths for each randomly sampled point. However we find that from a distribu-

tional instance, the posterior distributions have considerable overlap. We also find that the posterior distribution of the critical length for complete consumption is narrower than the distribution for partial consumption. The skewness in the distribution of the difference between critical lengths suggests that the samples with a low critical length for partial consumption are more likely to be those points with a large difference.



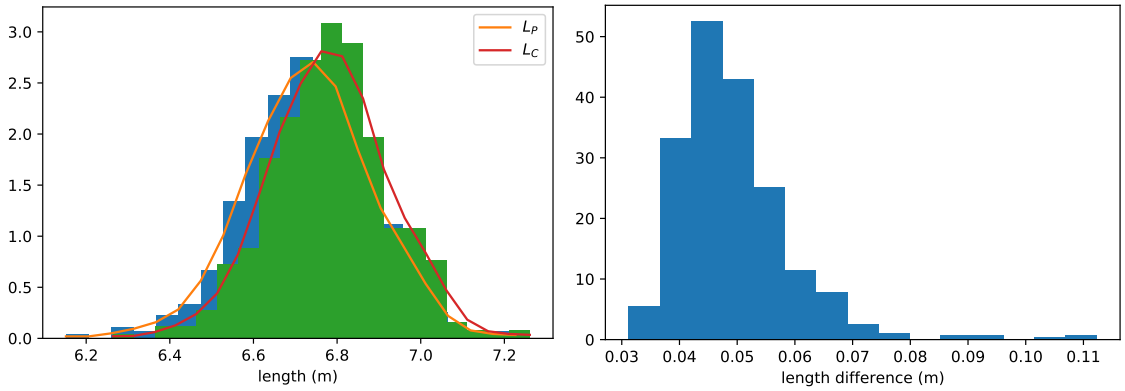
(a) Critical lengths. (b) Critical lengths difference.

Figure 7.18: The posterior distribution for the critical lengths of a one dimensional stockpile and the difference between the two critical lengths.

We can construct credible intervals using the samples from the posterior distribution of critical lengths. In addition to credible intervals we can use this sample to construct a probability that specific stockpiles undergo the different stages of consumption. For the critical length a 90% credible interval is $[4.37, 4.86]$ and $[4.63, 5.05]$ for L_P and L_C respectively. For example if a stockpile has length $L = 5$, then this stockpile is almost certain to reach partial consumption conditioned upon the data and the models that we have used. Additionally we can construct a probability of the stockpiles reaching complete consumption by considering the proportion of sampled critical lengths less than the stockpile length. When we compare the critical lengths calculated for the median parameters estimates, $L_P = 4.54$ and $L_C = 4.79$, we observe that these parameters are not centred within the posterior distribution, with the percentiles being 30 and 20 respectively. This highlights the importance of considering the posterior sample rather than limiting ourselves to just point estimates.

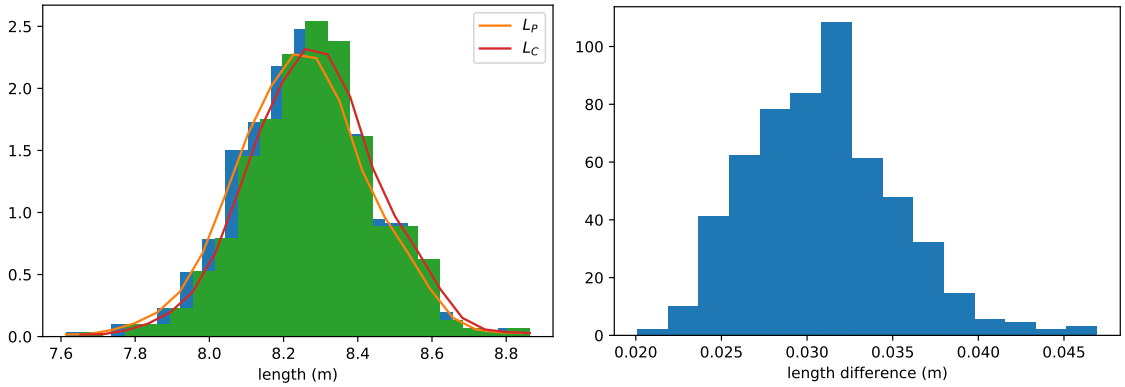
Similarly, we can determine the critical lengths in higher dimensions as well. Calculating these we find very similar results in the two dimensional case in Figure 7.19 and also in the three dimensional case in Figure 7.20. In the higher dimensions there is significant overlap between the two posterior distributions. We observe that as the dimensionality increases the difference between the critical lengths decrease. This is consistent with the results that we observed using the median values. The spread of the critical lengths for the higher dimensions is similar to the spread in one dimension.

In the previous section we claimed that using a two dimensional stockpile is a good



(a) Critical lengths. (b) Critical lengths difference.

Figure 7.19: The posterior distribution for the critical lengths of a two dimensional stockpile and the difference between the two critical lengths.



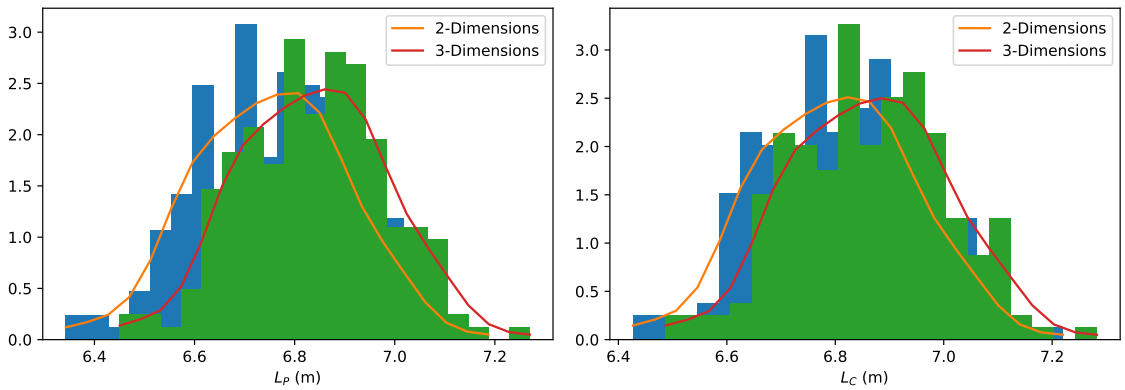
(a) Critical lengths. (b) Critical lengths difference.

Figure 7.20: The posterior distribution for the critical lengths of a three dimensional stockpile and the difference between the two critical lengths.

approximation for when the third dimension is large. We can validate this by comparing the critical length of a stockpile with lengths, $L_x = L$, $L_y = L$ and $L_z = 10L$. Given a stockpile of these proportions we compute the critical length for L which causes the stockpile to ignite. Using our sample, figure 7.21, we observe that the values are very similar. The difference between the two critical lengths for each sample are in Figure 7.22, which highlights that these differences are small. Hence we can show that the two dimensional model can be used effectively for this geometry.

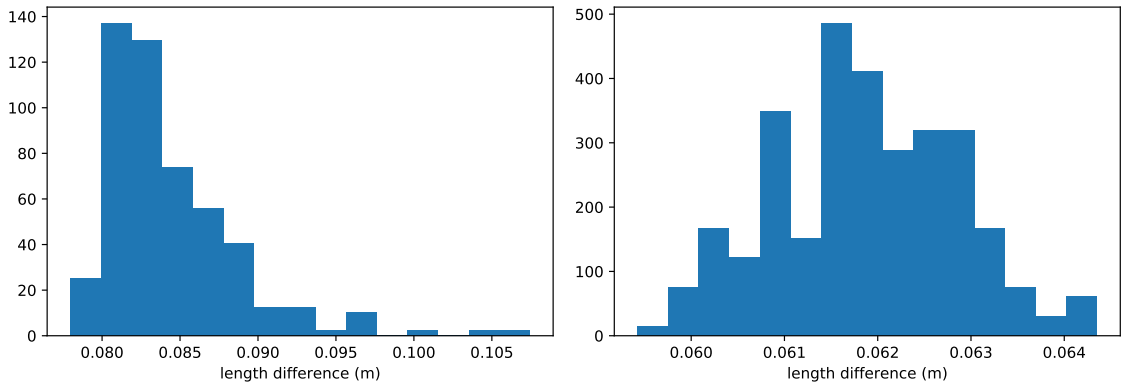
We have determined the critical lengths for these stockpiles in each dimension. It is also worthwhile investigating the effects that the uncertainty in our parameter estimates have on the ignition times. Whilst we utilised the consumption model for the critical lengths we implement the non-consumption model for the ignition times. We use this model as it is simpler to define an ignition time.

Figure 7.23 displays an estimate of the posterior distribution of the ignition times for a one dimensional stockpile. In this figure we observe that there is a spread amongst the



(a) Critical lengths L_P. (b) Critical lengths L_C.

Figure 7.21: A comparison of the critical lengths between a two dimensional stockpile and a three dimensional stockpile with the third dimension having length $10L$.



(a) Critical lengths difference L_P. (b) Critical lengths difference L_C.

Figure 7.22: The difference between the critical lengths between a two dimensional stockpile and a three dimensional stockpile with the third dimension having length $10L$.

ignition times and subsequently we have uncertainty about the ignition times. A similar distribution is observed when we look at higher dimensional stockpiles, see Figures 7.24 and 7.25. There is a slight difference in the spread of the ignition times between each dimension, but this also corresponds with a difference in the median values. The standard deviation of the distribution increases as the median increases. Because of this it is difficult to determine whether this spread is related to dimension or the median value of the ignition time. Given that we consider the same length stockpile for the two and three dimensional stockpiles, then comparing the ignition times of the two we observe that increasing the length in the third dimension reduces the ignition time.

7.4 Summary

Throughout this chapter we have explored the behaviour of large stockpiles using our estimates for the reaction kinetics. We focused mainly on the influence of external factors

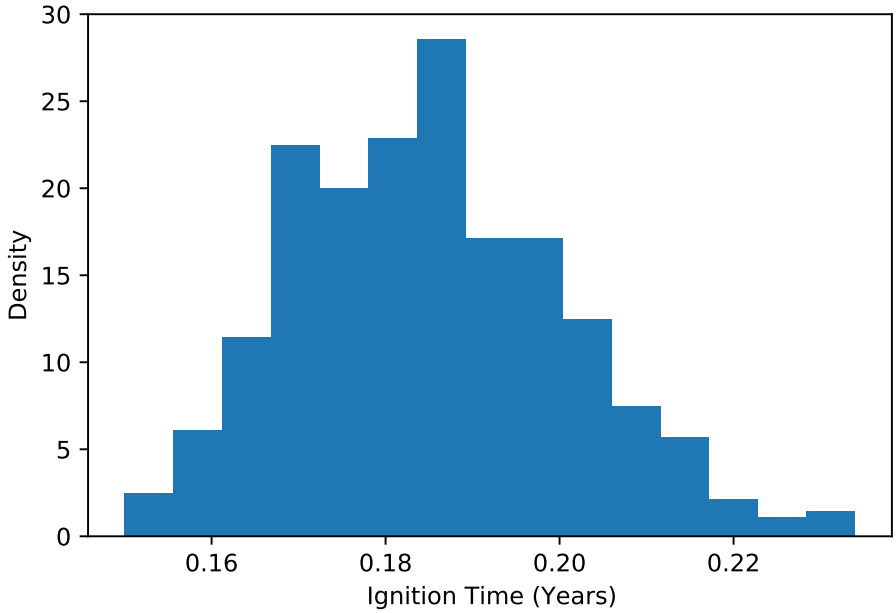


Figure 7.23: Posterior distribution of the ignition times for a one dimensional stockpile with $L = 5.2$m.

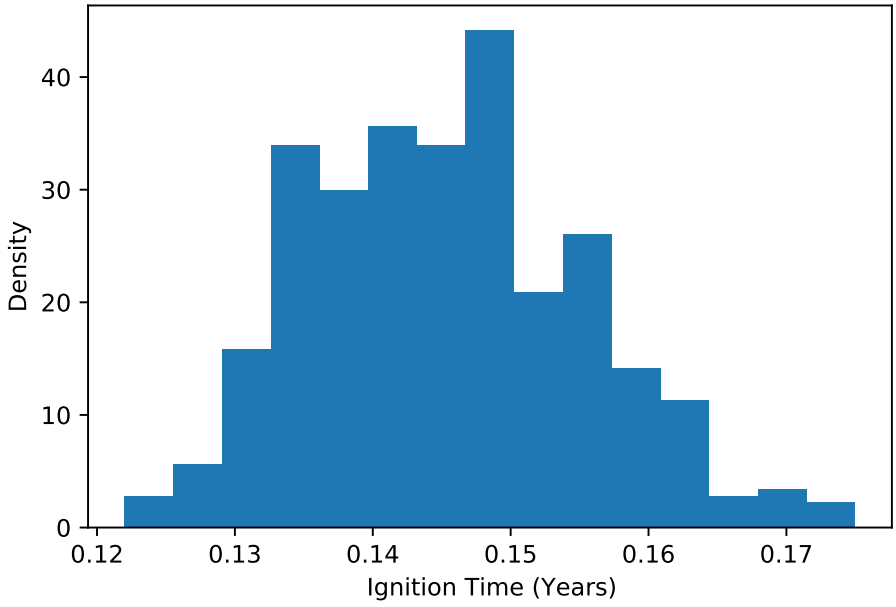


Figure 7.24: Posterior distribution of the ignition times for a two dimensional stockpile with $L = 8$m.

like temperature variation in the boundary condition and how hotspots can be used to induce ignition within subcritical stockpiles. Most of our work used the median values of each parameter as these provide a rough idea of the behaviours of the stockpile and is a useful point estimator of each parameter. We found that when we consider periodic

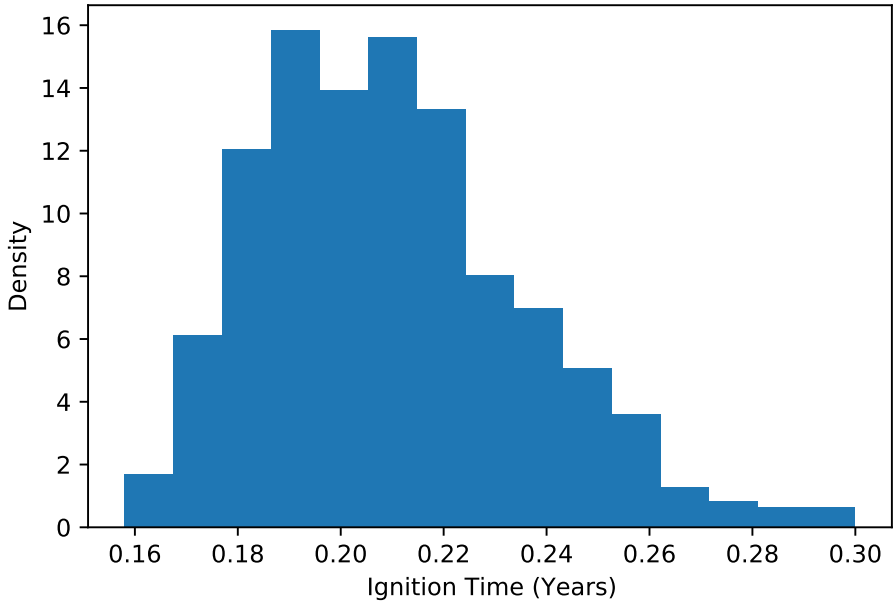


Figure 7.25: Posterior distribution of the ignition times for a three dimensional stockpile with $L = 8$m.

boundary conditions, the time when stockpile is constructed has minimal influence upon the critical length of the stockpile. Our result in Chapter 6 indicated this affects the critical FK parameter, however the influence from stockpile length is not significant. Using our particular parameter values we found that the temperatures that are necessary for the hotspot to induce ignition are higher than we anticipated when we scaled the one reaction model in Chapter 6. These two results highlight the need for further investigation into the stockpile models using the parameters that we obtain from the experimental data.

In addition to these results in one dimension we briefly examined the behaviour of the stockpiles in two and three dimensions. We found the same behaviour with a difference in critical lengths in each dimension. This is useful as it means we can focus our investigations to one dimension to determine the contributing factors to ignition. This reduces the computation time required, and the effects that are deemed significant, warrant more thorough investigation in higher dimensions.

Furthermore, we began exploring the posterior distributions of the critical lengths and ignition times. These posterior distributions are particularly important when considering decision making about the construction of these stockpiles. These also highlight the change of approach to simulating stockpiles where we consider a more probabilistic approach to determining critical conditions rather than relying on point estimators. Our posterior distributions show that the median values of each parameter does not result in the median of the critical length posterior distribution.

Chapter 8

Conclusion

Throughout this thesis I have presented various novel approaches that can be used to understand the reactions occurring within the filter cake and their application to the self-sintering that occurs within large stockpiles. A core model was developed that was useful to apply seamlessly between simulating the experimental data and simulating large scale stockpiles. This model was developed using the the core ideas of Frank–Kamentskii (FK) theory, taken from similar applications, and applied to both the stockpiles and to the thermogravemetric analysis (TGA). Within our application of FK theory we apply one of the reaction schemes proposed by Longbottom et al. [7], to our model.

In order to use this kinetic model in stockpile simulations we need to estimate the kinetic parameters of the reactions. Due to the complexity of the material and the uncertainty of the reaction schemes, identifying these parameters using a classical method was not possible. These previous methods were also not able to capture the uncertainty in the parameter estimates as effectively as we needed. Hence we proposed a novel application of existing MCMC methods for parameter estimation. Our method of inference in Chapter 4 uses a Bayesian framework to consider a posterior distribution of the parameter estimates rather than point estimates with errors attached. This new framework integrates particularly well with the forward modelling of the stockpiles as we are able to consider distributions of the critical conditions that influence the stockpiles behaviour. This shifts our perspective on stockpile ignition from a deterministic viewpoint as to whether certain stockpiles ignite to a probabilistic interpretation.

Our Bayesian framework allows us to define a posterior distribution for the kinetic parameters. This distribution is complex and cannot be obtained analytically or sampled from directly. We constructed a Metropolis–Hastings (MH) random walk algorithm to sample from this distribution. This allows us to gather information about the posterior distribution and generates a sample of our parameter estimates that we can use in the forward stockpile modelling.

Our algorithm conducts inference using one set of experimental data. However, we want to incorporate all the available data into our estimates. In Chapter 5 we include all

the experimental data by extending our algorithm using Sequential Monte Carlo (SMC) methods. Whilst we were able to apply this successfully to simulated data, the application to experimental data identified some issues. Our SMC algorithm suggested that, whilst the chosen reaction scheme was sufficient to explain any one set of experimental data, the parameter estimates were not consistent amongst the four sets of experimental data. This method provides a means to test the robustness of any reaction scheme coupled with the mathematical model. The SMC algorithm highlights the importance that conducting repeated experiments have when trying to estimate the reaction kinetics. This has particular importance when planning experiments as repeated experiments are valuable, particularly with different heating rates.

After establishing estimates of the parameters, we simulate the stockpiles. Using the FK theory as our framework, we consider a single reaction model. Non-dimensionalising the problem reduces the number of model parameters, simplifying the model. Using this framework we investigate the effect of external temperature variation and the use of hotspots to induce ignition. For small temperature oscillations we found that there was an effect on the critical FK parameters which is important to consider when simulating the larger stockpiles.

Stockpiles can be constructed with material from previously reacted stockpiles embedded in them. This increases the energy level within the stockpile which has the effect of possibly inducing ignition in stockpiles that would have been subcritical. This technique has been employed onsite at BlueScope Steel with mixed results. We set about examining whether this construction technique can be used to induce ignition. Our results indicate that this is a very effective method to induce ignition within the stockpiles. We determine numerical some relationships between the necessary temperature of the hotspot and the hotspot characteristics such as size and location.

In Chapter 7 we use our posterior samples of the parameters from Chapter 4 to simulate stockpiles that are constructed onsite. Using the median values of the marginal distributions, we determined the critical length that the stockpiles need to be greater than in order to cause ignition. We then simulated subcritical and supercritical stockpiles to investigate how the temperature within these changed over time. Simulating the stockpiles in two and three dimensions provided consistent results. Further to this, we simulated stockpiles using a sample from the posterior distribution so as to obtain a sample of the posterior distribution of critical lengths. Using this sample we generate credible intervals for the critical lengths. Such plots enable us to consider the probability of a certain stockpile igniting based on our estimates of the parameters.

8.1 Discussion

Many aspects of this thesis lends itself to further investigation. Each chapter presents new work that can be extended upon. The models that I have implemented have been developed to reflect fundamental aspects of the stockpiles used by BlueScope Steel. These models can be extended further to account for matters such as the influence of oxygen concentration on the reactions. Additionally, we are interested in the behaviour of the stockpiles after ignition has occurred, where factors such as radiative heat transfer become much more important. Including such terms is one of the next stages in the model development.

The estimation of the kinetic parameters is an important aspect of simulating the stockpiles. It is necessary to know the variation amongst these parameters since they affect the simulations of the stockpiles and subsequently any further decisions derived from them. The MH algorithm that we developed can be applied to more complicated reaction schemes. We modelled a two-reaction scheme where the reactions were distinct. More complicated reaction schemes introduce more complexities into the posterior distribution and subsequently influence the efficiency of the MH algorithm. In Chapter 5 we observed this complexity as we trialled a 3 reaction scheme for the experimental data. For well separated reactions we anticipate that the algorithm will work effectively with appropriate conditioning upon the priors for the T_m parameters; this still requires investigation though.

In addition to adding more reactions, there are reaction schemes where the products of one reaction react with other materials in the sample. This results in reactions occurring in series rather than the parallel reactions that we modelled. This is highly likely for complex materials such as the filter cake and multiple schemes of this nature have been proposed by Longbottom et al. [7]. Our formulation of the TGA data allows these schemes to be introduced easily into the model. However, this introduces additional parameters, which can result in additional modes in the posterior distribution. These reaction schemes remain to be investigated and tested with simulated data. They form a very natural extension of this thesis.

Our algorithms to conduct inference can be further refined. The Metropolis-Hastings algorithm we employed required manual tuning. This process is time consuming and can be automated using AMH or DRAM [127, 128]. Additionally our SMC algorithm developed in Chapter 5 can be further refined. Our initial MCMC algorithm incorporates a Runge-Kutta solver for the differential equation. This numerical method approximates the solution to the differential equation with increasing precision as the step-size parameter is reduced. Subsequently, multilevel Monte-Carlo methods provide a natural framework to account for this approximation [138]. Additionally, when initialising the SMC algorithm we use samples from the previous experiments to initialise the SMC algorithm. This may

introduce minor biases and subsequently a sequence of distributions that converge to the initial posterior distribution as described in Beskos et al.[133] is more appropriate.

When we applied our SMC method to the experimental data, we had issues using multiple TGA experiments. Specifically the posterior mass after one experiment was not concentrated in the same region as the posterior mass after two reactions. We attribute this discrepancy to the model miss-specification for the experimental data. The points at which the posterior concentrates is affected by the order in which the experiments are selected highlighting the issues with the model. Tempering the distribution is also expected to mitigate any effects of the order of experiments may have on the posterior sample [132]. In an idealised system where the model perfectly explains the data, such as a simulated set of experiments, the order of experiments does not have an impact [131]. In particular, we highlighted that the model selection process needs to be refined. In Chapter 5, we proposed various methods to examine the choice of model used in our inference. A critical finding was that any future work in this area needs to consider the effect of the initial conditions on the posterior distribution. It is likely that the initial concentrations need to be inferred from the experimental data and cannot be prescribed at the start.

One of the important questions that came out of our inference work was how to select the reaction scheme that best fits the data. We showed that our two reaction model may not be the best fit due to the discrepancy in the parameter values when we consider each experiment separately. Other reaction schemes should be considered, but then it remains how do we determine which model provides the best fit. There are methods that can be used for such model selection such as using the Bayes factor [8]. This will form an important aspect of any future work that is undertaken for these algorithms.

In our stockpile models we investigated the effects of periodic boundary conditions and hotspots. The periodic boudary conditions provide an approximation for the seasonal temperature variation in the ambient air temperature. This work can be extended to three dimensional problems. As we found the results are consistent in one and two dimensions, we anticipate that results from three dimensions will also be consistent. For the periodic boundary conditions there are two main ways we would look to extend this model. The first is to incorporate diurnal oscillations and determine if these are significant. More importantly, we can incorporate weather data to improve our model for the ambient temperature. We chose to model temperature variation using a simplifying approximation to examine if this has any effects. Reported weather data is unlikely to follow this sinusoidal function and a more representative function needs to be identified. With the Covid-19 pandemic causing issues with obtaining localised weather data, this was not possible.

Our analysis into the effect of hotspots on stockpile ignition mainly considered one spatial dimension. To simulate stockpiles that are used by BlueScope Steel we need to consider three dimensions. Additionally, we need more information on the characteristics of the hotspots used by BlueScope within the stockpiles to provide greater accuracy. This

is particularly important when we consider simulations in two and three dimensions.

The final problem that we examined was to simulate the stockpiles using the estimates from our inference. This produced some interesting results where for specific lengths, only one of the reactions would occur. These are all conditional upon the parameter estimates and the reaction scheme that we choose. As we refine the parameter estimation method to allow us to consider different reaction schemes, we need to re-simulate the stockpiles to check if they have a similar behaviour. We found that for the kinetic parameters we used that the wüstite reaction initiates the self-heating process whilst the iron reaction mainly contributes to the stockpiles reaching higher temperatures. This is of important practical use to BlueScope Steel, as we observe which reaction is the most useful for promoting ignition. This is a useful observation that will need to be checked when new reaction schemes are considered.

We limited ourselves to what we investigated using point estimates of the sampled parameter values when investigating the periodic boundary conditions and hotspot temperatures. We considered how the various model parameters change this posterior distribution of critical lengths. Similarly we could use the posterior distribution to examine the distribution of critical hotspot temperature required to induce stockpile ignition. There is significant value in using a sample from the posterior distribution rather than focusing on point estimators.

When we simulated the stockpiles, we had model parameters such as the diffusivity that were not inferred using the filter cake. Given that the filter cake is complex the model parameters we have used will differ to those of the filter cake. It would be valuable to undertake additional experimental work to determine what these parameter values are. Having greater consultation with BlueScope Steel would allow these models to be compared against observations of stockpiles at the Port Kembla facility, although the Covid-19 pandemic prevented that possibility. This would provide further justification regarding the importance of model selection for the reaction schemes and provides an additional method to justify this choice.

There are some aims that we set out in Chapter 1 that would still provide valuable insights into the self-heating of stockpiles. The addition of moisture is complex as the affect this has on the wüstite and iron reactions needs to be defined. Using radiation as a form of heat transfer is particularly important for high temperature stockpiles and if we wish to simulate these after ignition. These are two ways to extend the model to ensure that the simulated stockpiles better reflect reality. We have also not considered air flow through the stockpile. The filter cake is a porous material and advection occurs as the air moves through the stockpile. We considered advection breifly in the dimensionless hotspot problem in Chapter 6, but have not considered this in our stockpile simulations. This is an important concept that needs to be addressed in future work. Due to the complexity we discovered in order to estimate the kinetic parameters, we were unable to consider these

aims within this thesis.

8.2 Summary

I have considered many novel approaches to addressing the self-ignition process of steel stockpiles. I have considered a new approach to estimating parameters using TGA data that utilises a Bayesian framework where I consider a distribution of parameters rather than point estimates with associated errors. This approach utilises a Metropolis–Hastings random walk algorithm that needed to be calibrated to efficiently sample this posterior distribution. I have incorporated a sequential Monte–Carlo algorithm to utilise the data from multiple experiments, which was used to indicate that our proposed reaction scheme does not fully explain the experimental data. In addition to the Bayesian Inference, I have examined the effects that seasonal variations in ambient conditions have on stockpile ignition. These results can be considered in many types of stockpiles such as coal and compost. Furthermore, I investigated the ability of hotspots to be used to induce ignition. These novel approaches came together to simulate the large stockpiles that are stored by BlueScope Steel in Chapter 7.

This thesis poses some research questions that require further investigation. The most pressing is how best to incorporate data from multiple experiments, whether that be from reducing the restriction on the initial conditions or upon implementing a hierarchical model. This leads to the important question of how do we select different models? Within this thesis we suggested a different reaction scheme that suggested a very complex posterior distribution. The more complex the posterior distribution, the more challenging it is to sample effectively. The Metropolis–Hastings algorithm needs to be calibrated further, requiring additional work. Once results are obtained then it is important to determine how we select which model has the best fit for the data.

Bibliography

[1] Raymond James Longbottom, Brian Joseph Monaghan, Guangqing Zhang, David John Pinson, and Sheng Jason Chew. Self-sintering of bos filter cake for improved recyclability. *ISIJ International*, 59(3):432–441, 2019.

[2] R.L. Nyirenda. The processing of steelmaking flue-dust: A review. *Minerals Engineering*, 4(7):1003–1025, 1991.

[3] Hesham Mohamed Ahmed, Amanda Persson, Lena Sundqvist Ökvist, and Bo Björkman. Reduction behaviour of self-reducing blends of in-plant fines in inert atmosphere. *ISIJ International*, 55(10):2082–2089, 2015.

[4] Lotta Nedar. Dust formation in a bof converter. *Steel research*, 67(8):320–327, 1996.

[5] Raymond Longbottom, Brian J Monaghan, Guangqing Zhang, Sheng Chew, and David J Pinson. Characterisation of steelplant by-products to realise the value of fe and zn. *Characterisation of Steelplant By-Products to Realise the Value of Fe and Zn. Proc. 7th European Coke and Ironmaking Congress*, 2016.

[6] Raymond J Longbottom, Brian J Monaghan, David J Pinson, and Sheng J Chew. Understanding the self-sintering process of bos filter cake for improving its recyclability. *Journal of Sustainable Metallurgy*, 5(3):429–441, 2019.

[7] Raymond James Longbottom, Brian Joseph Monaghan, David John Pinson, Nathan Alan Stewart Webster, and Sheng Jason Chew. In situ phase analysis during self-sintering of bos filter cake for improved recycling. *ISIJ International*, pages ISIJINT–2020, 2020.

[8] Andrew Gelman, John B. Carlin, Hal S. Stren, David B. Dunsun, Aki Vehtari, and Donald B. Rubin. *Bayesian Data Analysis*. Chapman and Hall/CRC, New York, 2013.

[9] Sara J Palmer and Ray L Frost. Characterisation of bauxite and seawater neutralised bauxite residue using xrd and vibrational spectroscopic techniques. *Journal of materials science*, 44(1):55–63, 2009.

[10] Zhanping You, Julian Mills-Beale, Justin M Foley, Samit Roy, Gregory M Odegard, Qingli Dai, and Shu Wei Goh. Nanoclay-modified asphalt materials: Preparation and characterization. *Construction and building materials*, 25(2):1072–1078, 2011.

[11] RST Manhães, LT Auler, MS Sthel, J Alexandre, MSO Massunaga, JG Carrió, DR Dos Santos, EC Da Silva, A Garcia-Quiroz, and H Vargas. Soil characterisation using x-ray diffraction, photoacoustic spectroscopy and electron paramagnetic resonance. *Applied Clay Science*, 21(5-6):303–311, 2002.

[12] Anna Mazzinghi, Chiara Ruberto, Lisa Castelli, Caroline Czelusniak, Lorenzo Giuntini, Pier Andrea Mandò, and Francesco Taccetti. Ma-xrf for the characterisation of the painting materials and technique of the entombment of christ by rogier van der weyden. *Applied Sciences*, 11(13):6151, 2021.

[13] ZHI Sun, Y Xiao, J Sietsma, H Agterhuis, G Visser, and Y Yang. Characterisation of metals in the electronic waste of complex mixtures of end-of-life ict products for development of cleaner recovery technology. *Waste management*, 35:227–235, 2015.

[14] Rufino Lozano and Juan Pablo Bernal. Characterization of a new set of eight geochemical reference materials for xrf major and trace element analysis. *Revista mexicana de ciencias geológicas*, 22(3):329–344, 2005.

[15] T. Hatakeyama and F.X. Quinn. *Thermal analysis: fundamentals and applications to polymer science*. Wiley, 1999.

[16] Junmeng Cai and Sibinuer Alimujiang. Kinetic analysis of wheat straw oxidative pyrolysis using thermogravimetric analysis: Statistical description and isoconversional kinetic analysis. *Industrial & Engineering Chemistry Research*, 48(2):619–624, 2009.

[17] Hsisheng Teng, Hung-Chi Lin, and Jui-An Ho. Thermogravimetric analysis on global mass loss kinetics of rice hull pyrolysis. *Industrial & Engineering Chemistry Research*, 36(9):3974–3977, 1997.

[18] Morten Gunnar Grønli, Gábor Várhegyi, and Colomba Di Blasi. Thermogravimetric analysis and devolatilization kinetics of wood. *Industrial & Engineering Chemistry Research*, 41(17):4201–4208, 2002.

[19] SA Hadigheh and S Kashi. Characterisation of carbon fibre recovered by pyrolysis using thermal gravimetric analysis (tga). In *The 14th International Symposium on Fiber-Reinforced Polymer Reinforcement of Concrete Structures (FRPRCS-14), Northern Ireland, United Kingdom*, 2019.

[20] Xiang Gou, Xuan Zhao, Surjit Singh, and Da Qiao. Tri-pyrolysis: A thermo-kinetic characterisation of polyethylene, cornstalk, and anthracite coal using tga-ftir analysis. *Fuel*, 252:393–402, 2019.

[21] Niroj Kumar Mohalik, Somu Mandal, Santosh Kumar Ray, Asfar Mobin Khan, Debashish Mishra, and Jai Krishna Pandey. Tga/dsc study to characterise and classify coal seams conforming to susceptibility towards spontaneous combustion. *International Journal of Mining Science and Technology*, 32(1):75–88, 2022.

[22] A Skreiberg, O Skreiberg, J Sandquist, and L Sørum. Tga and macro-tga characterisation of biomass fuels and fuel mixtures. *Fuel*, 90(6):2182–2197, 2011.

[23] Siti Shawalliah Idris, Norazah Abd Rahman, and Khudzir Ismail. Combustion characteristics of malaysian oil palm biomass, sub-bituminous coal and their respective blends via thermogravimetric analysis (tga). *Bioresource technology*, 123:581–591, 2012.

[24] Aneta Magdziarz and Małgorzata Wilk. Thermogravimetric study of biomass, sewage sludge and coal combustion. *Energy Conversion and Management*, 75:425–430, 2013.

[25] Yogeshwari Pooja Rago, François-Xavier Collard, Johann F Görgens, Dinesh Surroop, and Romeela Mohee. Co-combustion of torrefied biomass-plastic waste blends with coal through tga: Influence of synergistic behaviour. *Energy*, 239:121859, 2022.

[26] Bojan Janković, Nebojša Manić, and Dragoslava Stojiljković. The gaseous products characterization of the pyrolysis process of various agricultural residues using tga–dsc–ms techniques. *Journal of Thermal Analysis and Calorimetry*, 139(5):3091–3106, 2020.

[27] OA Pinchuk, SR Aubuchon, C Marks, R Dominey, F Dundar, OF Deniz, A Ata, and KJ Wynne. Thermally pretreated 46% pt/vulcan xc72: Characterisation by tga/dsc/tem and cyclic voltammetry. *Fuel cells*, 9(5):554–561, 2009.

[28] PC Bowes. *Self-heating; evaluating and controlling the hazards*. Elsivier, 1984.

[29] Kevin Brooks, Vemuri Balakotaiah, and Dan Luss. Effect of natural convection on spontaneous combustion of coal stockpiles. *AIChE journal*, 34(3):353–365, 1988.

[30] Brian F Gray, John F Griffiths, and Stephen M Hasko. Spontaneous ignition hazards in stockpiles of cellulosic materials: criteria for safe storage. *Journal of chemical technology and biotechnology. Chemical technology*, 34(8):453–463, 1984.

[31] Nathasak Boonmee and Pongsak Pongsamana. Spontaneous ignition of bagasse stockpiles in thailand: a fire safety concern. *Engineering Journal*, 21(3):37–50, 2017.

[32] M.I. Nelson, E. Balakrishnan, and X.D. Chen. A semenov model of self-heating in compost piles. *Process Safety and Environmental Protection*, 81(5):375 – 383, 2003. Sustainable Development and Technology.

[33] Jian Zhang, Ting Ren, Yuntao Liang, and Zhongwei Wang. A review on numerical solutions to self-heating of coal stockpile: Mechanism, theoretical basis, and variable study. *Fuel*, 182:80 – 109, 2016.

[34] JM Ashman, JM Jones, and A Williams. Some characteristics of the self-heating of the large scale storage of biomass. *Fuel Processing Technology*, 174:1–8, 2018.

[35] Francesco Restuccia, Nicolas Ptak, and Guillermo Rein. Self-heating behavior and ignition of shale rock. *Combustion and Flame*, 176:213 – 219, 2017.

[36] Andrew G Salinger, Rutherford Aris, and Jeffrey J Derby. Modeling the spontaneous ignition of coal stockpiles. *AIChE Journal*, 40(6):991–1004, 1994.

[37] Supan Wang, Haixiang Chen, and Naian Liu. Ignition of expandable polystyrene foam by a hot particle: An experimental and numerical study. *Journal of hazardous materials*, 283:536–543, 2015.

[38] Joseph H Flynn and Leo A Wall. General treatment of the thermogravimetry of polymers. *J Res Nat Bur Stand*, 70(6):487–523, 1966.

[39] John Henry Sharp and Sally A Wentworth. Kinetic analysis of thermogravimetric data. *Analytical chemistry*, 41(14):2060–2062, 1969.

[40] Takeo Ozawa. A new method of analyzing thermogravimetric data. *Bulletin of the chemical society of Japan*, 38(11):1881–1886, 1965.

[41] F Carrasco. The evaluation of kinetic parameters from thermogravimetric data: comparison between established methods and the general analytical equation. *Thermochimica Acta*, 213:115–134, 1993.

[42] Homer E Kissinger. Variation of peak temperature with heating rate in differential thermal analysis. *Journal of research of the National Bureau of Standards*, 57(4):217–221, 1956.

[43] N Sbirrazzuoli, Y Girault, and L Elégant. Simulations for evaluation of kinetic methods in differential scanning calorimetry. part 3—peak maximum evolution

methods and isoconversional methods. *Thermochimica Acta*, 293(1-2):25–37, 1997.

[44] Gerhard Anton Wilhelm Van Doornum. *The spontaneous heating of coal*. PhD thesis, University of Natal, Durban, 1960.

[45] Dick Schmal, Jan H Duyzer, and Jan Willem van Heuven. A model for the spontaneous heating of coal. *Fuel*, 64(7):963–972, 1985.

[46] BD Young, DF Williams, and AW Bryson. Two-dimensional natural convection and conduction in a packed bed containing a hot spot and its relevance to the transport of air in a coal dump. *International journal of heat and mass transfer*, 29(2):331–336, 1986.

[47] Harvinder S Sidhu, Mark I Nelson, NA Ansari, Xiao Dong Chen, et al. Mathematical analysis of self-heating in compost piles. *Chemeca 2006: Knowledge and Innovation*, page 897, 2006.

[48] Mark I. Nelson, Xiao Dong Chen, and Harvinder S. Sidhu. A spatially uniform model of oxidative self-heating in compost piles. *Chemeca:Australasian Conference on Chemical Engineering*, 2007.

[49] H. S. Sidhu, M. I. Nelson, and X. D. Chen. A simple spatial model for self-heating compost piles. In Wayne Read and A. J. Roberts, editors, *Proceedings of the 13th Biennial Computational Techniques and Applications Conference, CTAC-2006*, volume 48 of *ANZIAM J.*, pages C135–C150, May 2007. `http://anziamj.austms.org.au/ojs/index.php/ANZIAMJ/article/view/86` [May 20, 2007].

[50] Peter Nordon, Brian C Young, and Norman W Bainbridge. The rate of oxidation of char and coal in relation to their tendency to self-heat. *Fuel*, 58(6):443–449, 1979.

[51] T Luangwilai, HS Sidhu, MI Nelson, XD Chen, et al. Biological self-heating of compost piles with air flow. *Engineering Our Future: Are We up to the Challenge?: 27-30 September 2009, Burswood Entertainment Complex*, page 2683, 2009.

[52] T Luangwilai, HS Sidhu, MI Nelson, and XD Chen. Modelling air flow and ambient temperature effects on the biological self-heating of compost piles. *Asia-Pacific Journal of Chemical Engineering*, 5(4):609–618, 2010.

[53] Thiansiri Luangwilai and Harvinder Sidhu. Determining critical conditions for two dimensional compost piles with air flow via numerical simulations. *ANZIAM Journal*, 52:463–481, 2010.

[54] T Luangwilai, HS Sidhu, MI Nelson, XD Chen, et al. Modelling the effects of air flow, ambient temperature and radiative boundary conditions in compost piles. *Chemeca 2010: Engineering at the Edge; 26-29 September 2010, Hilton Adelaide, South Australia*, page 3585, 2010.

[55] Rachael Aganetti, Aymeric Lamorlette, E Guilbert, Dominique Morvan, and GR Thorpe. Advection and the self-heating of organic porous media. *International Journal of Heat and Mass Transfer*, 93:1150–1158, 2016.

[56] R. Aganetti, A. Lamorlette, and G.R. Thorpe. The relationship between external and internal flow in a porous body using the penalisation method. *International Journal of Heat and Fluid Flow*, 66:185 – 196, 2017.

[57] C.E. Zambra, N.O. Moraga, and M. Escudey. Heat and mass transfer in unsaturated porous media: Moisture effects in compost piles self-heating. *International Journal of Heat and Mass Transfer*, 54(13):2801 – 2810, 2011.

[58] Thiansiri Luangwilai, Harvinder S Sidhu, and MI Nelson. One-dimensional spatial model for self-heating in compost piles: Investigating effects of moisture and air flow. *Food and Bioproducts Processing*, 108:18–26, 2018.

[59] Thiansiri Luangwilai, Harvinder Sidhu, and Mark Nelson. A two dimensional, reaction-diffusion model of compost piles. *ANZIAM Journal*, 53:34–52, 2011.

[60] T Luangwilai, HS Sidhu, MI Nelson, et al. Understanding effects of ambient humidity on self-heating of compost piles. *Chemeca 2018*, page 68, 2018.

[61] T Luangwilai, HS Sidhu, MI Nelson, XD Chen, et al. Modelling the effects of air flow, ambient temperature and radiative boundary conditions in compost piles. *Chemeca 2010: Engineering at the Edge; 26-29 September 2010, Hilton Adelaide, South Australia*, page 3585, 2010.

[62] N.O. Moraga, F. Corvalán, M. Escudey, A. Arias, and C.E. Zambra. Unsteady 2d coupled heat and mass transfer in porous media with biological and chemical heat generations. *International Journal of Heat and Mass Transfer*, 52(25):5841 – 5848, 2009.

[63] Mark I Nelson and Xiao Dong Chen. Survey of experimental work on the self-heating and spontaneous combustion of coal. *Reviews in Engineering Geology*, 18(1):1831–1883, 2007.

[64] Moshood Onifade and Bekir Genc. A review of research on spontaneous combustion of coal. *International Journal of Mining Science and Technology*, 30(3):303–311, 2020.

[65] Naphat Nopsanti and Nathasak Boonmee. A numerical simulation of spontaneous ignition of bagasse. *Engineering and Applied Science Research*, 47(1):27–35, 2020.

[66] Numan Luthfi, Tappei Ohkoshi, Yutaka Tamaru, Takashi Fukushima, and Kenji Takisawa. Investigation into the combustion kinetics and spontaneous ignition of sweet sorghum as energy resource. *Bioresources and Bioprocessing*, 9(1):1–12, 2022.

[67] Francesco Restuccia, Nieves Fernandez-Anez, and Guillermo Rein. Experimental measurement of particle size effects on the self-heating ignition of biomass piles: Homogeneous samples of dust and pellets. *Fuel*, 256:115838, 2019.

[68] Lars Schwarzer, Peter Arendt Jensen, Stig Wedel, Peter Glarborg, Oskar Karlström, Jens Kai Holm, and Kim Dam-Johansen. Self-heating and thermal runaway of biomass–lab-scale experiments and modeling for conditions resembling power plant mills. *Fuel*, 294:120281, 2021.

[69] AS Shteinberg and SI Khudyaev. Thermal explosion under oscillations of the external temperature. In *Doklady Physical Chemistry*, volume 403, pages 120–123. Springer, 2005.

[70] VA Gorel'skii, AA Yashchuk, and AS Shteinberg. Thermal explosion of a reactive mixture under conditions of oscillating ambient temperature. *Russian Journal of Physical Chemistry B*, 4(5):755–759, 2010.

[71] Vasily Novozhilov. Thermal explosion in oscillating ambient conditions. *Scientific reports*, 6:29730, 2016.

[72] Vasily Novozhilov. Kinetic effects in thermal explosion with oscillating ambient conditions. *Scientific reports*, 8(1):4030, 2018.

[73] Nepal C Roy. Convection characteristics in a closed vessel in the presence of exothermic combustion and ambient temperature oscillations. *International Journal of Heat and Mass Transfer*, 116:655–666, 2018.

[74] B.F. Gray and G.C. Wake. Critical initial conditions for thermal ignition. *Mathematical and Computer Modelling*, 18(8):65–75, 1993.

[75] R O Weber, E Balakrishnan, and G C Wake. Critical initial conditions for spontaneous thermal ignition. *Journal of the Chemical Society, Faraday Transactions*, 94(24):3613–3617, 1998.

[76] J Brindley, JF Griffiths, and AC McIntosh. Ignition phenomenology and criteria associated with hotspots embedded in a reactive material. *Chemical engineering science*, 56(6):2037–2046, 2001.

[77] AG Merzhanov and BI Khaikin. Theory of combustion waves in homogeneous media. *Progress in energy and combustion science*, 14(1):1–98, 1988.

[78] VV Gubernov, JJ Sharples, HS Sidhu, AC McIntosh, and J Brindley. Properties of combustion waves in the model with competitive exo-and endothermic reactions. *Journal of mathematical chemistry*, 50(8):2130–2140, 2012.

[79] AC McIntosh, RO Weber, and Geoffry N Mercer. Non-adiabatic combustion waves for general lewis numbers: wave speed and extinction conditions. *The ANZIAM Journal*, 46(1):1–16, 2004.

[80] RO Weber and SD Watt. Combustion waves. *The ANZIAM Journal*, 38(4):464–476, 1997.

[81] GN Mercer, RO Weber, BF Gray, and SD Watt. Combustion pseudo-waves in a system with reactant consumption and heat loss. *Mathematical and Computer Modelling*, 24(8):29–38, 1996.

[82] AC McIntosh, J Brindley, and JF Griffiths. An approximate model for the ignition of reactive materials by a hot spot with reactant depletion. *Mathematical and computer modelling*, 36(3):293–306, 2002.

[83] T. L. Jackson, A. K. Kapila, and D. S. Stewart. Evolution of a reaction center in an explosive material. *SIAM Journal on Applied Mathematics*, 49(2):432–458, 1989.

[84] AK Kapila. Evolution of deflagration in a cold combustible subjected to a uniform energy flux. *International Journal of Engineering Science*, 19(4):495–509, 1981.

[85] AA Shah, J Brindley, AC McIntosh, and J Griffiths. Ignition and combustion of low-exothermicity porous materials by a local hotspot. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 463(2081):1287–1305, 2007.

[86] Dmitry Olegovich Glushkov, Geniy Vladimirovich Kuznetsov, and Pavel Alexandrovich Strizhak. Numerical simulation of solid-phase ignition of metallized condensed matter by a particle heated to a high temperature. *Russian journal of physical chemistry B*, 5(6):1000–1006, 2011.

[87] Peter James Caine, SJ Puttick, J Brindley, AC McIntosh, JF Griffiths, and R Mullins. *Ignition of bulk solid materials by a localised hotspot*. PhD thesis, University of Leeds, 2010.

[88] T Boddington, Peter Gray, and GC Wake. Theory of thermal explosions with simultaneous parallel reactions i. foundations and the one-dimensional case. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 393(1804):85–100, 1984.

[89] JG Graham-Eagle and GC Wake. Theory of thermal explosions with simultaneous parallel reactions. ii. the two-and three-dimensional cases and the variational method. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 401(1820):195–202, 1985.

[90] JG Graham-Eagle and GC Wake. The theory of thermal explosions with simultaneous parallel reactions. iii. disappearance of critical behaviour with one exothermic and one endothermic reaction. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 407(1832):183–198, 1986.

[91] GC Wake, M Sleeman, XD Chen, and JC Jones. Theory and applications of ignition with variable activation energy. *Journal of Thermal Science*, 1(3):208–212, 1992.

[92] S Pushpavanam and R Narayanan. Ignition and extinction in a model problem with parallel endothermic and exothermic reactions. *Chemical engineering science*, 44(11):2611–2618, 1989.

[93] SO Ajadi and V Gol'Dshtein. Critical behaviour in a three-step reaction kinetics model. *Combustion Theory and Modelling*, 13(1):1–16, 2009.

[94] JC Jones, H Rahmati, and GC Wake. The unpiloted ignition of eucalyptus leaves treated as a parallel reaction system. *Journal of fire sciences*, 9(4):311–329, 1991.

[95] Ernst L Back. Auto-ignition in hygroscopic, organic materials-especially forest products-as initiated by moisture absorption from the ambient atmosphere. *Fire safety journal*, 4(3):185–196, 1981.

[96] Ian K Walker. The role of water in spontaneous combustion of solids. In *Fire Research Abstracts and Reviews*, volume 9. Committee on Fire Research of the Division of Engineering, National Research ..., 1967.

[97] Christian Lohrer, Martin Schmidt, and Ulrich Krause. A study on the influence of liquid water and water vapour on the self-ignition of lignite coal-experiments and numerical simulations. *Journal of Loss Prevention in the Process Industries*, 18(3):167–177, 2005.

[98] BF Gray and GC Wake. The ignition of hygroscopic combustible materials by water. *Combustion and Flame*, 79(1):2–6, 1990.

[99] Xiao Dong Chen. On the fundamentals of diffusive self-heating in water containing combustible materials. *Chemical Engineering and Processing: Process Intensification*, 37(5):367–378, 1998.

[100] T Luangwilai, HS Sidhu, MI Nelson, and Xiao Dong Chen. Modelling the effects of moisture content in compost piles. *CHEMECA 2011: Australian Chemical Engineering Conference Australia: Engineers Australia*, 2011.

[101] GN Mercer and RO Weber. Radiation enhanced combustion wave speeds. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 453(1962):1543–1549, 1997.

[102] M.N. Özışık and M Necati. *Radiative Transfer and Interactions with Conduction and Convection*. A Wiley-Interscience publication. Wiley, 1973.

[103] Guy Joulin. On reaction-radiation equilibria in slab geometry. *Combustion science and technology*, 50(4-6):205–216, 1986.

[104] Esmail R. Monazam, Ronald W. Breault, Ranjani Siriwardane, and Duane D. Miller. Thermogravimetric analysis of modified hematite by methane (ch4) for chemical-looping combustion: A global kinetics mechanism. *Industrial & Engineering Chemistry Research*, 52(42):14808–14816, 2013.

[105] Quang-Vu Bach and Wei-Hsin Chen. Pyrolysis characteristics and kinetics of microalgae via thermogravimetric analysis (tga): A state-of-the-art review. *Bioresource Technology*, 246:88–100, 2017. Special Issue on Biochar: Production, Characterization and Applications – Beyond Soil Applications.

[106] Nicholas S. Reding and Mark B. Shiflett. Characterization of thermal stability and heat absorption for suppressant agent/combustible dust mixtures via thermogravimetric analysis/differential scanning calorimetry. *Industrial & Engineering Chemistry Research*, 58(11):4674–4687, 2019.

[107] Dario C. Lewczyk, Jesse W. Cohan, Melanie L. Goetz, Brendan L. Trafford, Robert L. Fuller, and Justin R. Sparks. Kinetic treatment of evaporation via thermogravimetric analysis: The case of d-limonene. *Industrial & Engineering Chemistry Research*, 59(33):15069–15074, 2020.

[108] T Akahira and T Sunose. Method of determining activation deterioration constant of electrical insulating materials. *Res Rep Chiba Inst Technol (Sci Technol)*, 16(1971):22–31, 1971.

[109] Joseph H. Flynn and Leo A. Wall. A quick, direct method for the determination of activation energy from thermogravimetric data. *Journal of Polymer Science Part B: Polymer Letters*, 4(5):323–328, 1966.

[110] Jianbiao Liu, Xuya Jiang, Hongzhen Cai, and Feng Gao. Study of combustion characteristics and kinetics of agriculture briquette using thermogravimetric analysis. *ACS Omega*, 6(24):15827–15833, 2021. PMID: 34179626.

[111] John N Carras and Brian C Young. Self-heating of coal and related materials: models, application and test methods. *Progress in Energy and Combustion Science*, 20(1):1–15, 1994.

[112] JC Jones. Calculation of the frank–kamenetskii critical parameter for a cubic reactant shape from experimental results on bituminous coals. *Fuel*, 78(1):89–91, 1999.

[113] J Clifford Jones and Graeme C Wake. Measured activation energies of ignition of solid materials. *Journal of Chemical Technology & Biotechnology*, 48(2):209–216, 1990.

[114] JC Jones. A novel approach to the estimation of combustion rates of solid fuels. *Journal of Chemical Technology & Biotechnology*, 53(1):29–32, 1992.

[115] JC Jones. Difficulties with standard tests to predict shipping hazards with coals and carbons. *Journal of fire sciences*, 15(3):175–179, 1997.

[116] Lionel Elliott, Derek B Ingham, Adrian G Kyne, Nicolae S Mera, Mohamed Pourkashanian, and Christopher W Wilson. The use of ignition delay time in genetic algorithms optimisation of chemical kinetics reaction mechanisms. *Engineering Applications of Artificial Intelligence*, 18(7):825–831, 2005.

[117] Marco AB Zanoni, Henrique Massard, Marcio F Martins, and Sylvain Salvador. Application of inverse problem and thermogravimetry to determine the kinetics of oil shale pyrolysist. *High Temperatures–High Pressures*, 41(3), 2012.

[118] Cédric Reverte, Jean-Louis Dirion, and Michel Cabassud. Kinetic model identification and parameters estimation from tga experiments. *Journal of Analytical and Applied Pyrolysis*, 79(1-2):297–305, 2007.

[119] Musa Buyukada. Uncertainty estimation by bayesian approach in thermochemical conversion of walnut hull and lignite coal blends. *Bioresource technology*, 232:87–92, 2017.

[120] Musa Buyukada. Probabilistic uncertainty analysis based on monte carlo simulations of co-combustion of hazelnut hull and coal blends: Data-driven modeling and response surface optimization. *Bioresource technology*, 225:106–112, 2017.

[121] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953.

[122] W. Keith Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.

[123] Marc C. Kennedy and Anthony O'Hagan. Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3):425–464, 2001.

[124] Jenný Brynjarsdóttir and Anthony O'Hagan. Learning about physical parameters: the importance of model discrepancy. *Inverse Problems*, 30(11):114007, oct 2014.

[125] Andrew Gelman, Walter R Gilks, and Gareth O Roberts. Weak convergence and optimal scaling of random walk metropolis algorithms. *The annals of applied probability*, 7(1):110–120, 1997.

[126] Jun Yang, Gareth O. Roberts, and Jeffrey S. Rosenthal. Optimal scaling of random-walk metropolis algorithms on general target distributions. *Stochastic Processes and their Applications*, 130(10):6094–6132, 2020.

[127] Heikki Haario, Eero Saksman, and Johanna Tamminen. An adaptive metropolis algorithm. *Bernoulli*, pages 223–242, 2001.

[128] Heikki Haario, Marko Laine, Antonietta Mira, and Eero Saksman. Dram: efficient adaptive mcmc. *Statistics and computing*, 16:339–354, 2006.

[129] T. Salmi, J.P. Wärnå, J.R.H. Carucci, and C.A. de Araújo Filho. *Chemical Reaction Engineering: A Computer-aided Approach*. De Gruyter Textbook Series. De Gruyter, 2020.

[130] Livingstone Murray Rose et al. *Chemical reactor design in practice*. Elsevier Scientific Pub. Co., 1981.

[131] Simo Särkkä. *Bayesian filtering and smoothing*. Cambridge university press, 2013.

[132] Ajay Jasra, David A Stephens, Arnaud Doucet, and Theodoros Tsagaris. Inference for lévy-driven stochastic volatility models via adaptive sequential monte carlo. *Scandinavian Journal of Statistics*, 38(1):1–22, 2011.

[133] Alexandros Beskos, Ajay Jasra, Nikolas Kantas, and Alexandre Thiery. On the convergence of adaptive sequential monte carlo methods. *The Annals of Applied Probability*, 26(2):1111–1146, 2016.

[134] Nikolas Kantas, Alexandros Beskos, and Ajay Jasra. Sequential monte carlo methods for high-dimensional inverse problems: A case study for the navier–stokes equations. *SIAM/ASA Journal on Uncertainty Quantification*, 2(1):464–489, 2014.

[135] Matthew Berry, Mark Nelson, Brian Monaghan, and Ben Whale. The effects of oscillating boundary conditions on thermal ignition. *ANZIAM Journal*, 61:C45–C58, 2019.

[136] G. N. Wells A. Logg and J. Hake. DOLFIN: a C++/Python finite element library. In K.-A. Mardal A. Logg and G. N. Wells, editors, *Automated Solution of Differential Equations by the Finite Element Method*, volume 84 of *Lecture Notes in Computational Science and Engineering*, chapter 10. Springer, 2012.

[137] A. Logg and G. N. Wells. DOLFIN: automated finite element computing. *ACM Transactions on Mathematical Software*, 37, 2010.

[138] Alexandros Beskos, Ajay Jasra, Kody Law, Raul Tempone, and Yan Zhou. Multilevel sequential monte carlo samplers. *Stochastic Processes and their Applications*, 127(5):1417–1440, 2017.

[139] Albert Tarantola. *Inverse problem theory and methods for model parameter estimation*. SIAM, 2005.

[140] Gareth O. Roberts and Jeffrey S. Rosenthal. General state space Markov chains and MCMC algorithms. *Probability Surveys*, 1:20–71, 2004.

[141] A. Gelman, W. R. Gilks, and G. O. Roberts. Weak convergence and optimal scaling of random walk Metropolis algorithms. *The Annals of Applied Probability*, 7(1):110 – 120, 1997.

[142] Alan E. Gelfand and Adrian F. M. Smith. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85(410):398–409, 1990.

[143] Aki Vehtari, Andrew Gelman, Daniel Simpson, Bob Carpenter, and Paul-Christian Bürkner. Rank-normalization, folding, and localization: an improved $\widehat{R}$ for assessing convergence of MCMC (with discussion). *Bayesian Analysis*, 16(2):667–718, 2021.

[144] Dootika Vats, James M. Flegal, and Galin L. Jones. Multivariate output analysis for Markov chain Monte Carlo. *Biometrika*, 106(2):321–337, 2019.

[145] John Charles Butcher. *The numerical analysis of ordinary differential equations: Runge-Kutta and general linear methods*. Wiley-Interscience, 1987.

[146] John C Strikwerda. *Finite Difference Schemes and Partial Differential Equations*. Chapman and Hall, 1989.

[147] Albert C Reynolds, Jr. Convergent finite difference schemes for nonlinear parabolic equations. *SIAM Journal on Numerical Analysis*, 9(4):523–533, 1972.

[148] Lawrence C Evans. *Partial differential equations*, volume 19. American Mathematical Soc., 2010.

Appendix A

Inverse Problems and Markov Chain Monte Carlo

In this appendix we will look at the theory behind the MCMC algorithms used in the thesis. To begin we briefly look at some basic probability and inverse problem theory, before moving on to MCMC algorithms. The material from this appendix is predominantly sourced from Tarantola [139] and Gelman et al [8].

A.1 Inverse Problems

In order to look at inverse problems we first have to address the concept of a model and model parameters. Let $\mathscr{O}$ be a physical system being studied; in our case this is the TGA experiment. We then determine a minimal set of parameters that can be used to characterise the experiment. These are our model parameters. We can then have two types of modelling; forward modelling takes inputs of activation energy and the pre-exponential factor and determines the fractional weight change; inverse modelling takes the measurements obtained in the experiment and then infers the values of the model parameters.

The model parameters are not always unique. We are able to use two different sets of model parameters; we can use the temperature at which the reaction rate is maximised rather than the pre-exponential factor. There are advantages of using certain parameterisation methods. In our case we found that using one set of parameters were heavily correlated causing the Markov chain to experience a high level of auto-correlation. This is just one of the benefits that can come from reparameterisation.

A.1.1 Bayesian Framework

We approach the inverse problem using a Bayesian framework. Bayesian inference consists of a sampling distribution and a prior distribution and uses Bayes Rule to connect

these together. One of the major aspects of Bayesian Inference is the parameters θ are all explicitly conditional upon the data, y. Bayes Rule states,

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)}. \tag{A.1}$$

In the bayesian framework $p(y|\theta)$ is the sampling distribution, $p(\theta)$ is our prior distribution, and $p(y)$ is the model evidence, which is a normalising constant in inverse problems. The concept follows that we update our distribution of the parameters θ from our prior distribution $p(\theta)$ using the sampling distribution $p(y|\theta)$ to get a posterior distribution $p(\theta|y)$ that is conditional upon the experimental data y.

This framework allows us to take the observed measurements y and infer the model parameters θ. Such a framework in the sampling distribution exploits the relative simplicity of the forward model. Our sampling distribution for the data, dependent upon the the model parameters, allows us to incorporate the much simpler forward model into our solution. In our case the simplest sampling distribution we can consider is that our measurements y are normally distributed around the forward model predictions $F(\theta)$, with some unknown sampling variance,

$$y \sim \mathcal{N}(F(\theta), \sigma).$$

The Prior Distribution

In order to solve the inverse problem we need to place some prior information on the state of the model parameters. This can be done in a number of ways. One of the most basic ways to assign a prior is to use the homogeneous distribution over the parameter space, which in our space is just a uniform distribution. This is a non-informative prior.

The next prior distribution that we consider is a weakly-informative prior distribution. These prior distributions will provide a little more information than the homogeneous distribution. One such way to do this is to restrict the prior distribution to some interval within the parameter space. This is quite common as we can place physical limits upon our parameters in many cases. This can often be reduced further into a plausible range based upon some additional knowledge. In the context of our TGA experiments, our data indicates that the reaction is completed within our experimental time-frame. This means that the temperature at which the reaction rate is maximised must be restricted to within the temperature range of our experiment.

Non-informative priors such as the homogeneous distribution is not proper on an infinite parameter space. This may cause some complications in setting up our inverse problem. The weakly-informative distribution is proper and does not suffer from such drawbacks.

Alternatively to these two types of prior distributions we can implement a conjugate

distribution. In these instances the prior distribution is selected based upon the choice of sampling distribution. A conjugate prior distribution is one that, when combined with the sampling distribution produces a posterior distribution that is of the same form as the prior distribution. Formally, let $\mathscr{F}$ be a class of sampling distributions, $(y|\theta)$ and $\mathscr{P}$ be a class of prior distributions $p(\theta)$, then the class $\mathscr{P}$ is conjugate to $\mathscr{F}$ if

$$p(\theta|y) \in \mathscr{P} \text{ for all } p(y|\theta) \in \mathscr{F} \text{ and } p(\theta) \in \mathscr{P} \tag{A.2}$$

Example A.1.1. Consider a binomial sampling distribution such that,

$$p(y|\theta) = \text{Bin}(\theta, n) \propto \theta^y (1-\theta)^{n-y}$$

In this instance if we consider a prior distribution of the form

$$p(\theta) \propto \theta^{\alpha-1}(1-\theta)^{\beta-1},$$

which is a beta distribution with parameter α and β: $\theta \sim \text{Beta}(\alpha, \beta)$. We can then write the posterior distribution as,

$$\begin{aligned} p(\theta|y) &\propto \theta^y (1-\theta)^{n-y} \theta^{\alpha-1}(1-\theta)^{\beta-1} \\ &= \theta^{y+\alpha-1}(1-\theta)^{n-y+\beta-1} \\ &= \text{Beta}(\theta|\alpha+y, \beta+n-y), \end{aligned}$$

which is of the same parametric form as the prior distribution.

This conjugate prior allows us to write the posterior distribution is a nice simple form. In our application we use a normal sampling distribution, and for the standard deviation of that normal distribution, the conjugate prior is the inverse gamma distribution. Let $p(\mathbf{y}|\sigma^2)$ be a Gaussian sampling distribution with mean μ and standard deviation σ,

$$p(y|\sigma^2) \propto \frac{1}{\sigma^n} \exp\left(\frac{-1}{2}\left(\frac{(\mathbf{y}-\mu)^T(\mathbf{y}-\mu)}{\sigma^2}\right)\right),$$

where n is the dimension of the data y. For simplicity let $\mathbf{e} = \mathbf{y} - \mu$. Consider σ^2 to be distributed according to the inverse gamma distribution with shape parameter α and rate parameter β, that is,

$$p(\sigma^2) \propto \frac{1}{\sigma^{2\alpha+1}} \exp\left(\frac{-\beta}{\sigma^2}\right).$$

This then leads to the posterior distribution,

$$p(\sigma^2|y) \propto \frac{1}{\sigma^n}\exp\left(\frac{-1}{2}\left(\frac{(\mathbf{y}-\mu)^T(\mathbf{y}-\mu)}{\sigma^2}\right)\right)\frac{1}{\sigma^{\alpha+1}}\exp\left(\frac{-\beta}{\sigma}\right)$$
$$\frac{1}{\sigma^{2n+\alpha+1}}\exp\left(\frac{-\left(\beta+\frac{\mathbf{e}^T\mathbf{e}}{2}\right)}{\sigma^2}\right).$$

This leaves the posterior distribution $p(\sigma^2|y) = \text{IG}\left(\alpha+n/2, \beta+\mathbf{e}^T\mathbf{e}/2\right)$. Therefore we can state that the Inverse Gamma distribution is the conjugate prior distribution to the Gaussian sampling distribution. If μ is known and fixed then this means we can get an analytic expression for the posterior distribution. In our application μ is dependant upon the model parameters and hence we cannot obtain this marginal distribution easily. If we condition upon the model parameters we do in fact get an analytic distribution.

A.2 MCMC Algorithms

In many contexts the posterior distribution cannot be determined analytically. In these cases it can be useful to sample from the distribution and use this sample to infer information about the model parameters. Sampling from this distribution is not straight-forward and this can be done using MCMC algorithms. The basis of these methods is to take a random walk through the parameter space and accepting new points on this walk using a likelihood function. We choose to use the Metropolis-Hastings algorithm outlined in 4.

Algorithm 4: Metropolis Hastings Algorithm.

Initialise θ_0 by sampling from the prior $p(\theta)$;
for $j = 1,\ldots,J$ **do**
  Sample $\theta^* \sim Q(\theta^*|\theta_{j-1})$;
  Set $\rho = \frac{p(y|\theta^*)q(\theta^*|\theta_{j-1})p(\theta^*)}{p(y|\theta_{j-1})q(\theta_{j-1}|\theta^*)p(\theta_{j-1})}$;
  Sample $b \sim U(0,1)$ **if** $b \leq \rho_j$ **then**
    Set $\theta_j = \theta^*$
  else
    Set $\theta_j = \theta_{j-1}$
  end
end

Algorithm 4 defines a Markov chain with stationary distribution given by the Bayesian posterior in Equation (A.1) [8]. What this means is that the dependence upon the initial parameter values θ_0 decreases after every iteration until the posterior samples are eventually independent. This distinguishes this approach from deterministic optimisation

algorithms such as Levenberg–Marquadt which depend upon these starting values. This phase of the algorithm is referred to as the transitory phase. This transitory phase is discarded as burn-in, leaving a Matrix Θ consisting of the random sample from the posterior distribution.

This sample is then used to calculate posterior expectations and estimate credible intervals for the parameters For any measurable function $f(\theta)$,

$$\mathbb{E}[f(\theta) \mid \mathbf{y}] = \int f(\theta)\, \mathrm{d}p(\theta \mid \mathbf{y}) \tag{A.3}$$

$$\approx \frac{1}{J-B} \sum_{j=B+1}^{J} f(\theta_j) \tag{A.4}$$

where (A.3) is a Lebesgue integral with respect to the posterior probability measure and B is the number of MCMC samples that were discarded as burn-in. Equation (A.4) is the Monte–Carlo estimate of this integral [140]. This then allows us to consider posterior expectations of any functional transform of our parameters. We make considerable use of this in our application to stockpile ignition, since the critical length is a function of the model parameters. In particular this function is not analytic and has to be solved for numerically, which is simple to apply to the sampled parameters.

A.2.1 The Proposal Distribution

The proposal distribution forms an integral part in defining the MCMC algorithm. The proposal distribution can be independent from the previous parameter. An example of such a proposal would be to use the prior distribution as the proposal distribution. This can work effectively in instances where the prior distribution is already close to the posterior distribution. In applications where the prior is non-informative or weakly informative this becomes a very computationally inefficient algorithm.

The type of proposal distribution we implement is a random walk. A random walk proposal consists of using a multivariate Gaussian distribution centred on the previous parameter values, as the proposal. One of the major benefits in this approach is that this distribution is symmetric, that is $q(\theta_{j-1} \mid \theta^*) = q(\theta^* \mid \theta_{j-1})$. This simplifies our calculation for the acceptance probability ρ. This random walk can be tuned using the covariance matrix in order to achieve an acceptance rate close to the optimal value for random walk MH of 0.234 [141, 126].

Random walk models work well for our model parameters though these can be improved further. In instances where we have a conjugate prior for one of the parameters, we can use Gibbs sampling [142]. Using our example with the noise parameter σ^2, with prior $p(\sigma^2) = \mathrm{IG}(\alpha_0, \beta_0)$ and sampling distribution $p(y \mid \sigma^2, \theta_m) = \mathrm{N}(F(\theta_m), \sigma)$, has the

conditional distribution

$$p(\sigma^2 \mid \theta_m, y) = \text{Beta}(\alpha_0 + n/2, \beta_0 + \mathbf{e}^T\mathbf{e}/2),$$

where n is number of measurements and $\mathbf{e}$ is the residual $y - F(\theta_m)$. We use this conditional posterior distribution as the proposal distribution for for the sigma parameters. The advantage of this choice of proposal is that acceptance probability $\rho_\sigma = 1$ for this parameter. To calculate the acceptance probability ρ, $q(\sigma^* \mid \sigma_{j-1}) = p(y \mid \sigma)p(\sigma_{j-1})$, which reduces the acceptance probability to 1. This would present the ideal scenario as the each new proposal would be accepted allowing an efficient algorithm. Given that we have other parameters involved in our model we do not have an acceptance of 1, though the acceptance is dependant only upon the model parameters.

A.2.2 Convergence of MCMC algorithm

Determining whether the Markov chain has converged to its stationary distribution is an integral part to MCMC algorithms. In this thesis we used two key diagnostic statistics to evaluate the convergence; the Split-$\widehat{R}$ and the effective sample size (ESS). To assist with this it is useful to run multiple Markov chains. A simple indication that the chains have reached stationarity is that the samples from each chain all agree. This simple indication is sufficient to determine if the chains have not converged but there can still be issues in determining convergence purely from these samples. This can be quantified using the Split-$\widehat{R}$ diagnostic [143].

Consider M Markov Chains of length N. The calculation for this statistic splits each chain in half to create $2M$ chains each of length $N/2$. The statistic compares the variation within the chains to the variation between the chains. If the ratio of these two variations Split-$\widehat{R} \approx 1$, than this indicates the chains have converged whilst if the variation between the chains is much larger than the variation within each chain, Split-$\widehat{R}$ is significantly greater than one, this suggests that the chains have not reached stationarity. One of the key benefits of splitting each chain is that the chains may have converged together but not reached the stationary distribution, and by splitting the chains this case will be identified by the Split-$\widehat{R}$ diagnostic. Calculating the Split-$\widehat{R}$ is simply splitting the chains and calculating the $\widehat{R}$ diagnostic after it has been split. In this instance we have $m = 2M$ chains each with sample size $n = N/2$.

To calculate the $\widehat{R}$, let ψ denote the parameter to be estimated, with ψ_{ij} denote the ith observation from the jth chain. We compute the between chain variance B using,

$$B = \frac{n}{m-1} \sum_{j=1}^{m} \left(\bar{\psi}_{\cdot j} - \bar{\psi}_{\cdot\cdot}\right)^2,$$

where $\bar{\psi}_{.j}$ is the mean of the jth chain and $\bar{\psi}_{..}$ is the mean across the sample. We then compute the within chain variance, W by,

$$W = \frac{1}{M}\sum_{j=1}^{m} s_j^2,$$

where s_j^2 is the sample variance of the jth chain. These are combined to estimate the marginal posterior variance, $\text{var}(\psi \mid y)$,

$$\widehat{\text{var}}^{+}(\psi \mid y) = \frac{n-1}{n}W + \frac{1}{n}B$$

which overestimates the marginal posterior distribution. The $\hat{R}$ diagnostic is then calculated as,

$$\widehat{R} = \sqrt{\frac{\widehat{\text{var}}^{+}(\psi \mid y)}{W}}.$$

The ESS is calculated using the emperical autocorrelation of the MCMC samples [144]. If the ESS is too low then this indicates that the markov chains have high autocorrelation. This limits how how useful the samples are in evaluating posterior expectations. If successive draws of the Markov chain were independent then the autocorrelation would be negligible and the ESS would be approximately the sample size of the Markov chains MN. To calculate the ESS we fist define the *variogram* V_t at each lag t:

$$Vt = \frac{1}{m(n-t)}\sum_{j=1}^{m}\sum_{i=t+1}^{n}\left(\psi_{i,j} - \psi_{i-t,j}\right).$$

We then estimate the autocorrelations,

$$\widehat{\rho}_t = 1 - \frac{V_t}{2\widehat{\text{var}}^{+}}.$$

Since this empirical autocorrelation is noisy for large values of t then these are summed over until two successive lags $\hat{\rho}_{2t} + \hat{\rho}_{2t+1}$ is negative. This yields the estimate,

$$ESS = \frac{mn}{1 + 2\sum_{t=1}^{T}\widehat{\rho}},$$

where T is the first odd positive integer where $\hat{\rho}_{T+1} + \hat{\rho}_{T+2}$ is negative.

These calculations for the diagnostic statistics are performed on the sample after the burn-in has been discarded. If these statistics indicate that the sample is not satisfactory then we can simple restart the algorithm and increase the number of iterations within each Markov Chain. To compliment these two statistics the trace plots of the Markov Chains can be examined. The trace plot is simply plotting the sampled parameter values of each chain for each iteration. These plots give a valuable insight into whether the chains have

converged or not and can be used in conjunction with these diagnostic statistics. One of the main issues that causes autocorrelation within a random walk MH algorithm is correlation between the parameters. The random walk proposes parameters independently of each other and hence the probability of the proposal distribution proposing a point within this correlated posterior distribution is low. Instead a narrower proposal is necessary, but this introduces autocorrelation into the Markov Chains. The best way to avoid this is by using alternative parameters to infer or to use a dependent random walk proposal distribution.

Appendix B

Numerical Techniques

A number of numerical techniques can be used to solve heat transfer in stockpiles. Initially we use finite differences. This method is useful for simple domains, but for more complex geometries it becomes more challenging to implement. The finite element method is another technique that can be used. One of the key benefits about this method is that the generalisation to complex domains is simple. There are numerous software packages developed for finite elements. We used Fenics [137, 136].

B.1 Theoretical Properties

In this section we examine the theory behind our numerical methods. We break this down into the Runge–Kutta method and the two methods we use to solve PDEs in this thesis: finite differences and finite element methods.

B.1.1 Runge-Kutta Methods

Runge–Kutta methods are used to solve first order ordinary differential equations (ODEs). We will consider the explicit Runge–Kutta methods [145] that we use to solve initial value problems. An initial value problem has the form,

$$\frac{dy}{dt} = f(t,y) \qquad\qquad y(0) = y_0. \tag{B.1}$$

The general form for a Runge–Kutta method is,

$$\begin{aligned} y_{n+1} &= y_n + h\sum_{i=1}^{s} b_i k_i, \\ t_{n+1} &= t_n + h, \\ k_1 &= f(t_n, y_n), \\ k_i &= f\left(t_n + c_i h, y_n + h\sum_{j=1}^{i-1} a_{ij} k_k\right). \end{aligned}$$

This method requires specification of the number of stages s, and the coefficients, a_{ij}, for $1 \leq j < i \leq s$, b_i for $i = 1,2,...,s$ and c_j for $j = 2,3,...,s$. The forward Euler method is the simplest Runge–Kutta method as it consists of only one stage. This means we only have $k_1 = f(t_n, y_n$ and therefore $y_{n+1} = y_n + hf(t_n, y_n)$ The coefficients can be written compactly in a *Butcher Tableau*,

0					
c_2	a_{21}				
c_3	a_{31}	a_{32}			
$\vdots$	$\vdots$		$\ddots$		
c_s	a_{s1}	a_{s2}	$\dots$	$a_{s,s-1}$	
	b_1	b_2	$\dots$	b_{s-1}	b_s

For a Runge–Kutta method to be consistent, the coefficients b_j must satisfy the equation $\sum_{i=1}^{s} b_i = 1$. The general form of the Runge–Kutta method allows for many stages. Increasing the number of stages used can increase the order of the method. For a given method the local tuncation error is the difference between the true solution and the numerical solution. For a Runge–Kutta scheme to have order p, then the local truncation error is $\mathcal{O}(h^{p+1})$.

B.1.2 Finite Difference Methods

The following theory for finite difference methods has been sourced from [146]. For many differential equations, solutions cannot be found analytically. As a result, numerical schemes are required. It is useful to obtain an estimate of how accurate the numerical solution is. To make the analysis easier it is useful to introduce some notation.

Let P be a differential operator such that,

$$Pu = 0, \tag{B.2}$$

where $u : (0,\infty) \times \Omega \subset \mathbb{R}^n \to \mathbb{R}$. Initially we consider P to be a linear differential operator.

The Finite Difference Operator

To introduce the finite difference operator we need to consider the space of functions that it operates on. To think about this we first consider the domain $\Omega = \mathbb{R}$. The standard discretisation of this domain is into a grid of equally spaced nodes with spacing h. This can be denoted by,

$$h\mathbb{Z} = \{hz : z \in \mathbb{Z}\}.$$

A similar discretisation is required for the time component as well. We define,

$$k\mathbb{N}_0 = \{kn : n \in \mathbb{N} \cup \{0\}\},$$

where k is the spacing of each nodes. For practicality we consider $k, h > 0$.

Consider a function $v : k\mathbb{N}_0 \times h\mathbb{Z} \to \mathbb{R}$ and let v_m^n denote the value of the function at the point $(n, m) \in k\mathbb{N}_0 \times h\mathbb{Z}$. We seek a difference operator, $P_{h,k}$ such that solutions to the equation,

$$P_{k,h}v = 0, \tag{B.3}$$

can be used to approximate the solutions to equation B.2. A common example of a finite difference operator is the forward difference operator, δ_+, given by,

$$\delta_+ v_m = \frac{v_{m+1} - v_m}{h}.$$

This is often denoted as δ_{x+} as it is with respect to the x variable. Similarly we can define $\delta_t +$ as,

$$\delta_{t+} v^n = \frac{v^{n+1} - v^n}{k}.$$

We also have a backward and central difference operators δ_- and δ_0, defined by,

$$\delta_- v_m = \frac{v_m - v_{m-1}}{h},$$

$$\delta_0 v_m = \frac{v_{m+1} - v_{m-1}}{2h}.$$

In each of these cases h, k represents the grid spacing of the function v_m^n. The central difference can also be generated by the average of the forward and backward differences.

These differences are used to approximate first derivatives. We can find a similar finite difference for higher order derivatives. The central difference operator for the second derivative is $\delta_+\delta_-$ which is denoted by δ^2.

Convergence

It would be good to use these finite difference operators to produce approximate solutions, though we have not set a definition for convergence.

Definition 1. A finite difference scheme approximating a differential operator is convergent, if for a solution $u(t,x)$ of equation B.2 and the solution, v_m^n, to equation B.3 such that v_m^0 converges to $u(0,x)$ as mh approaches x, then v_m^n converges to $u(t,x)$ as (nk,mh) converges to (t,x) as $h,k \to 0$ [146].

This definition does not provide a norm for the convergence and the functions, $u(t_n,x)$ and v_m^n are in different spaces so it is not clear as to how to define convergence.

For finite difference schemes it is useful to to consider consistency. Consistency is the idea that the differential operator is well approximated by the difference operator.

Definition 2. A finite difference operator $P_{k,h}$ is consistent with a differential operator P if for any smooth function ϕ, $||P_{k,h}\phi - P\phi|| \to 0$ as $h,k \to 0$, where the convergence is point-wise at each grid point.

Example B.1.1. We consider an example where $P = \frac{\partial}{\partial t} - \frac{\partial^2}{\partial x^2} - F$, and F is a function of the argument. The finite difference operator $P_{k,h} = \delta_{t+} - \delta_x^2 - F$. We wish to show consistency of these operators.

To determine consistency we need only to show that the finite difference approximations are consistent with the respective differential forms since,

$$|P\phi - P_{k,h}\phi| \leq \left| \frac{\partial \phi}{\partial t} - \delta_{t+}\phi \right| + \left| \frac{\partial^2 \phi}{\partial x^2} - \delta_x^2 \phi \right|. \tag{B.4}$$

The functional terms $F(\phi)$ drop out of the equation since the difference operator just evaluates the function at that point.

For the forward difference, since ϕ is smooth,

$$\delta_{t+}\phi(t_n,x_m) = \frac{\phi(t_n+k,x_m) - \phi(t_n,x_m)}{k}.$$

For simplicity we drop the x_m variable as this is just an equation for each x_m. We expand $\phi(t_n+k)$ by it's Taylor series about t_n,

$$\phi(t_n+k) = \phi(t_n) + k\frac{d\phi}{dt}\bigg|_{t=t_n} + \mathscr{O}(k^2)$$

Rearranging this equation, we get,

$$\frac{\phi(t_n+k) - \phi(t_n)}{k} = \frac{d\phi}{dt}\bigg|_{t=t_n} + \mathscr{O}(k). \tag{B.5}$$

Similarly we can show for the central difference that,

$$\delta_x^2\phi - \frac{\partial^2 \phi}{\partial x^2} = \mathscr{O}(h^2). \tag{B.6}$$

Combining these two components, then,

$$|P\phi - P_{k,h}\phi| \leq \mathscr{O}(k) + \mathscr{O}(h^2). \tag{B.7}$$

This approaches 0 as h,k approach 0.

With that example in mind it is useful to introduce the notion of order of accuracy.

Definition 3. A finite difference scheme $P_{k,h}v = 0$ that is consistent with the differential equation $Pu = 0$ is accurate of order p in time and q in space, if for any smooth function ϕ,

$$|P\phi - P_{k,h}\phi| = \mathscr{O}(k^p) + \mathscr{O}(h^q).$$

The next concept that needs to be defined is stability.

Definition 4. A linear finite difference scheme $P_{k,h}v_m^n = 0$ is stable in a stability region Λ if there exists an integer J such that for any positive time T, there exists a constant C_T such that,

$$\sum_{m=-\infty}^{\infty} |v_m^n|^2 \leq C_T \sum_{j=0}^{J} \sum_{m=-\infty}^{\infty} |v_m^j|^2, \tag{B.8}$$

for all $nk \leq T$ for $(h,k) \in \Lambda$.

In general for single-step time schemes, the constant J is taken as 0.

Theorem 1 (Lax-Richtmyer Equivalence Theorem)**.** *A finite difference scheme consistent with the differential equation, for which the initial value problem is well-posed, is convergent if and only if it is stable [146].*

The proof of this theorem requires the use of Fourier transforms and also a more appropriate sense of convergence.

To discuss the convergence of the finite difference schemes, we need to consider functions in the same space. To do this we introduce two operators.

Definition 5. The truncation operator T maps functions in $L^2(\mathbb{R})$ to functions in $L^2(h\mathbb{Z})$. Given $u \in L^2(\mathbb{R})$, such that,

$$u(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ix\xi} \hat{u}(\xi) d\xi. \tag{B.9}$$

We define Tu as,

$$Tu_m = \frac{1}{\sqrt{2\pi}} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} e^{imh\xi} \hat{u}(\xi) d\xi, \tag{B.10}$$

for each grid point $mh \in h\mathbb{Z}$.
When we consider the discrete Fourier transform of Tu,

$$\widehat{Tu}(\xi) = \hat{u}(\xi), \qquad |\xi| \leq \frac{\pi}{h}. \tag{B.11}$$

Definition 6. The interpolation operator S maps functions from $L^2(h\mathbb{Z})$ to functions in $L^2(\mathbb{R})$. Given $v \in L^2(h\mathbb{Z})$, with

$$v_m = \frac{1}{\sqrt{2\pi}} \int_{\frac{-\pi}{h}}^{\frac{\pi}{h}} e^{imh\xi} \hat{v}(x),$$

then Sv is defined as,

$$Sv(x) = \frac{1}{\sqrt{2\pi}} \int_{\frac{-\pi}{h}}^{\frac{\pi}{h}} e^{ix\xi} \hat{v}(x).$$

These operators are also defined in terms of the grid spacing, h, though this is left out of the notation. Using these two operators we can then define convergence in terms of the interpolation operator.

Definition 7. A finite difference scheme approximating the homogeneous initial value problem for a partial differential equation is a convergent scheme if Sv^n converges to $u(t_n,\cdot)$ in $L^2(\mathbb{R})$, where $t_n = nk$, for every solution u to the differential equation and every set of solutions to the difference scheme v depending on h, and k, for which Sv^0 converges to $u(0,\cdot)$ in $L^2(\mathbb{R})$ as h and k tend to 0

Non-linear Convergence

The convergence theory in the previous section related to linear problems rather than non-linear. Our problem we consider is non-linear which requires more specific techniques to deal with convergence. A simple theorem cannot provide convergence for all types of non-linear problems. In this section we examine a convergence theorem for a general parabolic function on a 1-dimensional domain with Dirichlet boundary conditions. The convergence results presented are from Reynolds [147].

Consider the Partial differential equation given by,

$$u_t = f(t,x,u,u_x,u_{xx}) \quad \text{in} \quad [0,T]\times(a,b), \tag{B.12}$$

$$u(0,x) = \psi(x), \quad u(t,a) = \psi_0(t), \quad \text{and} \quad u(t,b) = \psi_1(t), \tag{B.13}$$

where $\psi(a) = \psi_0(0)$ and $\psi(b) = \psi_1(0)$.
We consider a discrete mesh over the domain $[0,T]\times(a,b)$ defined by the set of points

(t_i, x_j) such that,

$$\begin{aligned} h &= \frac{b-a}{m+1}, \\ x_j &= a + ih, \\ \Delta t &= \frac{T}{n}, \\ t_i &= i\Delta t, \end{aligned}$$

for $i = 1, 2, ..., n$ and $j = 0, 1, ..., m+1$.

We can also define functions on this mesh. Let $v(t,x)$, be a function defined on the domain $[0,T] \times (a,b)$. Then the function defined on the mesh $v_j^i = v(t_i, x_j)$, for $i = 1, 2, ..., n$ and $j = 0, 1, ..., m+1$. The notation used here is to be consistent with the previous sections.

We use the finite differences that were defined in section B.1.2 to replace our derivatives. We shall also simplify our notation by denoting $f(t_i, x_j, v_j^i, \delta_0 v_j^i, \delta^2 v_j^i)$ by $f(v_j^i)$, where the differences are taken in space. For $i = 1, 2, ..., n$ and j=$0, 1, ..., m+1$ let $\alpha_j(t_i)$ and $\beta_j(t_i)$ be defined by $u_x(t_i, x_j) = \delta_0(u_j^i) + \alpha_j(t_i)$ and $u_{xx}(t_i, x_j) = \delta^2 u_j^i + \beta_j(t_i)$, where we have assumed that u(t,x) is the unique solution to B.12-B.13. Assuming that u has sufficient regularity then then α_j and β_j are $\mathcal{O}(h^2)$. By replacing the derivatives in equation B.12, we can consider the discrete problem given by,

$$\delta_{t-} v_j^i = \theta f(v_j^i) + (1-\theta) f(v_j^{i-1}) \quad \text{for } 1 \le i \le n \text{ and } 1 \le j \le m, \tag{B.14}$$

$$v_0^i = \psi_0(t_i), \quad v_{n+1}^i = \psi_1(t_i) \quad \text{and} \quad v_j^0 = \psi(x_j). \tag{B.15}$$

In order examine the convergence theorem we wish to place some restrictions on the function f. Suppose there exists constants $\alpha, A, B \ge 0$ and constants C and C' such that f satisfies,

$$\alpha(\bar{r} - r) \le f(t,x,z,p,\bar{r}) - f(t,x,z,p,r) \le A(\bar{r} - r), \quad \bar{r} \ge r, \tag{B.16}$$

$$|f(t,x,z,\bar{p},r) - f(t,x,z,p,r)| \le B|\bar{p} - p|, \tag{B.17}$$

$$-C(\bar{z} - z) \le f(t,x,\bar{z},p,r) - f(t,x,z,p,r) \le C'(\bar{z} - z). \tag{B.18}$$

These restrictions are similar to Lipschitz conditions, where condition Equation B.17 is exactly a Lipschitz condition. The value of these conditions is that we can bound the non-linear terms in the function by linear terms.

There are also some restrictions that are placed on the grid size. Suppose the grid

parameters, Δt and h, satisfy,

$$\theta \Delta t C' \leq 1, \tag{B.19}$$

$$\alpha - \frac{hB}{2} \geq 0, \tag{B.20}$$

$$(1-\theta)^2 \left(A \frac{\Delta t}{h^2} \right) \leq 1 + (1-\theta)(-C). \tag{B.21}$$

Given these restrictions then the main theorem can be stated.

Theorem 2. *Let u denote the unique solution to Equations* (B.12) *and* (B.13)*, and have bounded derivatives, u_{tt} and u_{xxxx}. Suppose f satisfies Equations* (B.16)*,* (B.17) *and* (B.18) *and Equations* (B.19)*,*(B.20) *and* (B.21)*.*

Suppose there exists a function ω defined on the domain $[0,T] \times \mathbb{R}^3$ and a function $\rho > 0$, $\rho \in C^1([0,T])$, such that,

$$f(t,x,\bar{z},\bar{p},\bar{r}) - f(t,x,z,p,r) \leq \omega(t,\bar{z}-z,|\bar{p}-p|,|\bar{r}-r|) \quad \textit{for } \bar{z} \geq z, \tag{B.22}$$

$$\rho'(t') \geq 2\omega(\bar{t},\rho(\bar{t}),|\alpha_j|,|\beta_j|) \textit{ on } [0,T] \text{ for } i = 1,2,...,n \text{ and } |\bar{t}-t'| \geq \Delta t, \tag{B.23}$$

$$\rho'(t') > \Delta t \sup_{0<\bar{t}\leq T, 1\leq j\leq m} |u_{tt}(\bar{t},x_j)| \quad \text{for} |\bar{t}-t| \leq \Delta t. \tag{B.24}$$

Then, $\sup_{0\leq j\leq m+1}|(u(t_i,x_j)-v_j^i| \leq \rho(t_i)$ *for* $i = 0,1,...,n$*, where v_j^i is the solution to B.14.*

B.1.3 Finite Element Methods

Finite elements are a useful tool, especially for elliptic problems. One of the main advantageous of using finite elements is that they simplify the method for complex domains. This is due to the triangulation that is applied to the domain rather than a grid. In order to introduce the finite element method some background around weak derivatives and Sobolev spaces needs to be introduced.

Sobolev Spaces and Weak Solutions

Definition 8. Let $\Omega \subset \mathbb{R}^n$. We define the function space $L^P(\Omega)$ as the space of all functions, f, such that the integral, $\int_\Omega |f|^p$, is finite.

Definition 9. Let u be a function in $L^p(\Omega)$. Then if there exists a function $w \in L^P(\Omega)$ such that,

$$\int_\Omega uv' = -\int_\Omega wv,$$

for all smooth, compactly supported test functions v, then w is the weak derivative of v.

Definition 10. Sobolev Space We define the space $W^{k,p}(\Omega)$ as the space for which the k^{th} weak derivatives are all in $L^P(\Omega)$.

To investigate the finite element method we restrict our work to 2nd order differential operators. Consider a second order differential operator P, such that

$$Pu(x) = f(x), \qquad x \in \Omega, \tag{B.25}$$

$$u(x) = g(x), \qquad x \in \partial\Omega, . \tag{B.26}$$

where $u(x) \in C^2(\Omega)$, $f(x) \in C(\Omega)$ and $\Omega \subset \mathbb{R}$. If u is a solution of equation (B.25), then u is a solution of,

$$\int_\Omega Pu(x)v = \int_\Omega f(x)v, \qquad \forall v \in W^{1,2}(\Omega), \tag{B.27}$$

for $u \in W^{1,2}$ and $f \in L^2(\Omega)$ This form is called the weak formulation of the differential equation. Additionally to the classical solutions there are weak solutions that solve the weak formulation of the differential equation. These solutions exist in the Sobolev space.

Example B.1.2. The Poisson problem is the simplest problem to investigate. Consider the differential equation,

$$- \triangle u = f,$$

with domain Ω, and boundary condition $u = g$. The weak formulation for this is the equation,

$$\int_\Omega - \triangle uv = \int_\Omega Fv, \qquad \forall v \in W_g^{1,2}(\Omega),$$

where $W_g^{1,2}(\Omega)$ are the functions in $W^{1,2}(\Omega)$ that are equal to the function g on the boundary.

The weak formulation increases the solution space for the functions u and thus we may obtain a weak solution to the problem. Applying the divergence theorem this reduces to,

$$\int_\Omega \nabla u \cdot \nabla v - \int_{\partial\Omega} \frac{\partial u}{\partial n} v = \int_\Omega Fv, \qquad \forall v \in W^{1,2}(\Omega)$$

For a more rigorous treatment of the weak solutions to PDE's and of Sobolov spaces we recommend Evans [148]. For an explanation of the finite element method the concepts we have defined are sufficient. To explore the convergence of the schemes then a greater understanding of Sobolev spaces and weak solutions are required.

B.1.4 The Method

The idea behind finite element methods is that we can take this weak formulation, but rather than solve this problem using the infinite function space $W^{1,2}$, we approximate this function space with a discrete function space. We then solve this problem on the

discrete function space numerically by using the basis functions for that function space. One standard way to approach this is to consider a discrete mesh over the domain and use basis functions that are one at each node in the domain and 0 at every other node. There are multiple ways to do this but piecewise linear functions are the simplest way. We consider an example of a simple Poisson problem on a one dimensional domain. We let $W_h^{1,2}$ describe this discrete function space where the maximum distance between two nodes is h. We let u_h denote the solution to,

$$\int_\Omega \nabla u \cdot \nabla v - \int_{\partial\Omega} \frac{\partial u}{\partial n} v = \int_\Omega Fv, \qquad \forall v \in W_h^{1,2}(\Omega).$$

It can be shown that the norm $||u - u_h|| \leq Ch$ where C is some constant. This relationship demonstrates the convergence property of the solutions.

B.1.5 The Poisson Problem

The Poisson problem on the domain $[0,1]$ is,

$$-\triangle u(x) = f(x), \qquad x \in [0,1], \tag{B.28}$$

with $u(0) = u(1) = 0$. This can be written in the weak formulation,

$$\int_\Omega \nabla u \cdot \nabla v = \int_\Omega Fv \qquad \forall v \in W^{1,2}(\Omega)$$

To simplify the notation let $B(u,v\int_\Omega \nabla u \cdot \nabla v$ and $F(v) = \int_\Omega Fv$. We consider the nodes $x_i = i/10$ for $i = 0,1,...,10$. These nodes defines a mesh that splits the interval $[0,1]$ into the intervals $[x_i, x_{i+1}]$ for $i = 0,1,...9$. With this interval mesh we can define a function space,

$$W_h = \{u \in W^{1,2}([0,1]) : u \text{ is linear on } [x_i, x_{i+1}], i = 0,1,...,9\}. \tag{B.29}$$

This function space is discrete and as such can be spanned by the functions,

$$v_i = \begin{cases} \frac{x - x_{i-1}}{x_i - x_{i-1}} & x \in [x_{i-1}, x_i] \\ \frac{x - x_i}{x_{i+1} - x_i} & x \in [x_i, x_{i+1}] \\ 0 \quad \text{elsewhere} \end{cases}, \tag{B.30}$$

for $i = 1,2,...,9$. Rephrasing the problem on the discrete function space we have.

$$B(u,v) = F(v) \quad \forall v \in W_h.$$

Given that the bilinear form, $B(u,v)$ is linear in v and the functional $F(v)$ is also linear, and all functions $v \in W_h$ can be written in the form $v = \sum_i c_i v_i$, then the weak formulation

is equivalent to,

$$B(u, v_i) = F(v_i), \quad \forall i = 1, 2, ..., 9.$$

This equation defines our solution for u. Given that u_h is also contained in the discrete function space W_h, we can write this in terms of the functions v_i such that $u_h = \sum_i c_i v_i$. Thus the solution to the discrete problem is,

$$\sum_j c_j B(v_j, v_i) = F(v_i) \quad \forall i = 1, 2, ..., 9.$$

Since the functions v_i are known, these can be evaluated and the result is the equation,

$$BC = F, \tag{B.31}$$

where $B_{ij} = B(v_i, v_j)$ and $C_j = c_j$ are the coefficients of the basis functions of our solution, and $F_i = F(v_i)$. Thus our differential equation has been reduced to an algebraic equation. Although this is a one dimensional example, this approach can be extended into higher dimensions where we consider a triangulation of the domain. This consists of using triangles to cover the domain in two dimensions and tetrahedron in three dimensions.

B.2 Application of the methods

The theoretical basis for these methods has now been established. Here we consider the specific application of each of these methods that we have used.

B.2.1 Runge-Kutta

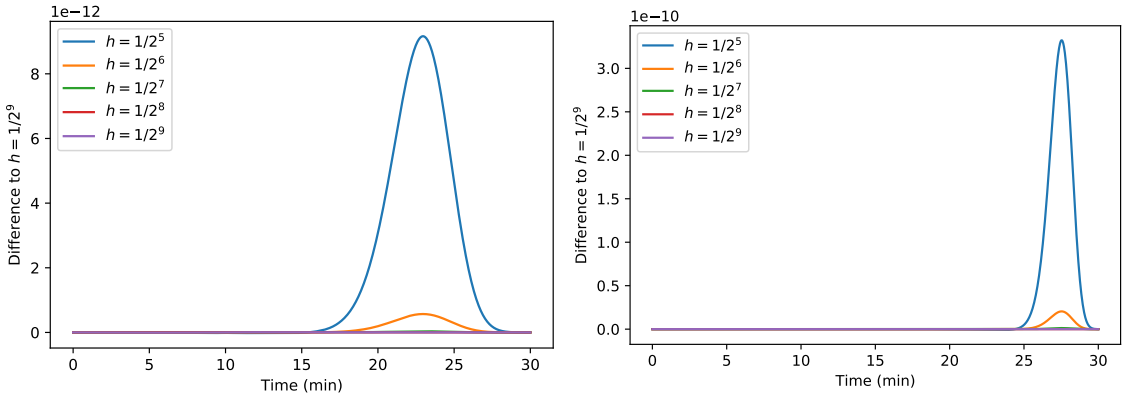
We implement a Runge–Kutta scheme for the ODE equations that we use. Specifically we use the scheme commonly referred to as RK4. This is a 4th order Runge–Kutta scheme with the number of stages $s = 4$. The RK4 method we use has the Butcher Tableau,

0				
1/2	1/2			
1/2	0	1/2		
1	0	0	1	
	1/6	1/3	1/3	1/6

Implementing a RK4 is simple and done using the code, in Listing B.1. This is a basic solver that requires the function input into the solver.

```
1 def RK4(f,IC=0,t0=0,h=0.01,n=1000):
2     y0=np.array(IC);
```

![!ht]



(a) Reaction 1.

(b) Reaction 2.

Figure B.1: Convergence plots for the RK4 numerical scheme for the two reaction model.

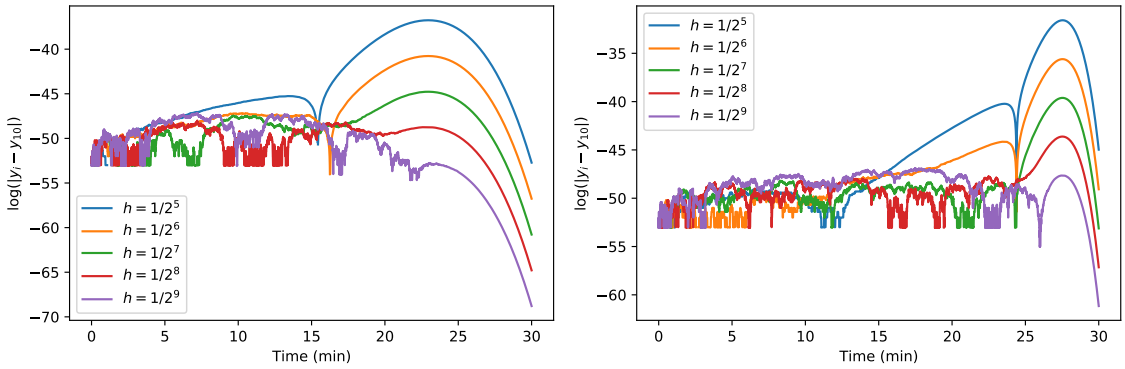
```
    y=np.zeros((y0.size,n+1))
    y[:,0]=y0
    ti=t0
    for i in range(n):
        k1=np.array(f(t=ti,x=y[:,i]))
        k2=np.array(f(t=ti+h/2,x=y[:,i]+k1*h/2))
        k3=np.array(f(t=ti+h/2,x=y[:,i]+k2*h/2))
        k4=np.array(f(t=ti+h,x=y[:,i]+k3*h))
        y[:,i+1]=y[:,i]+1/6*h*(k1+2*k2+2*k3+k4)
        ti=ti+h
    return y
```

Listing B.1: Runge–Kutta solver.

We test this solver out by considering the convergence of our solution. To do this we compare our solution after decreasing the number of step size h and comparing it to the smallest step size we consider which is $h = 2^{-10}$. To do this we introduce some notation to simplify things. Let y^i denote the numerical solution to the equation with step size $h = 2^{-i}$. The convergence plots for the simulated model are displayed in Figure B.1. This figure indicates how quickly the numerical scheme converges. We observe that the discrepancy between the solutions is very low and certainly below the sensitivity of any of the mass scales in the experimental dataset.

We can also consider the difference between two solutions, $y^i - y^{i-1}$. We also choose to plot this difference on the log scale. Using a log scale makes it easier to determine how quickly the numerical scheme converges and whether the order of the numerical scheme is consistent with what we expect. We show this result in Figure B.2, for which we can see the plot has two regions of interest. For both reactions there is a region where it does not appear to converge. In this region, the change in the reactant concentration is very small and subsequently the difference between the solutions is affected significantly by

the computational error of floating point numbers. In the latter part of the integration period, the convergence behaves as we expected with the difference between consecutive solutions differing by 4, which is the order of convergence for the RK4 scheme.



(a) Reaction 1. **(b)** Reaction 2.

Figure B.2: Convergence plots for the RK4 numerical scheme for the two reaction model.

B.2.2 Finite Differences

We use finite differences to solve the non-dimensional stockpile equation. To do this we use a first order forward difference in time and a second order centred difference in space. Consider the equation,

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + q(u). \tag{B.32}$$

We consider the domain $x \in [0,1]$ and $t \in [0,\infty)$. We have an initial condition, $u(0,x) = f(x)$ and boundary condition, $u(t,0) = 0$ and $u(t,1) = a$. This equation describes our self heating problem in a more general form. Consider a step size h and grid size k with the condition $kM = 1$, where M is an integer and is the number of intervals. The we can define our grid across our domain as,

$$x_m = km, \qquad m = 0,1,\ldots,M, \tag{B.33}$$

$$t_n = nk, \qquad n = 0,1,\ldots, \tag{B.34}$$

with $x_0 = 0$ and $x_M = 1$. We let u_m^n be the numerical solution at the point (t_n, x_m). To ensure that our solution satisfied the initial condition we set $u_m^0 = f(x_m)$, and in order to satisfy the boundary conditions we set $x_0 = x_1$ and $u_M = a$. In order to calculate the values u_m^n we use the equation,

$$u_m^{n+1} = u_m^n + \frac{h}{k^2}\left(u_{m-1}^n - 2u_m^n + u_{m+1}^n\right) + q(u_m^n). \tag{B.35}$$

This method is included in the Matlab code in Listing B.2

```
%% one dimensional code using the finite difference stencils.

%Source parameters.
delta=3.639;
epsilon=0.027;
Para_src=[delta,epsilon];

%Parameters for the boundary condition function;
theta0=0.637;
omega=0.3;
Para_BC=[theta0,omega];

%specify the bounds in the x direction.
x0=0;
x1=1;

%determine the number of intervals
Ix=100;

%node numbers
nx=Ix+1;

% boundary conditions
%[x0,x1]
BC=[1,0]; %This specifies the boundary condition type on each node
%0= Dirichlet Condition
%1= Neumann Condition
%2 Is a robin condition
a=[1,1]; % The coefficient for the u term in a robin condition.
%Note: a is only required for a robin condition and any value will
    run if
%it is not a robin condition

%Specify the grid size. Note it doesn't have to be a uniform grid
    in each direction.
hx=(x1-x0)/Ix;

%Specify the initial and final times.
t0=0;
tf=100;

%Specify the time step size;
k=0.000005;

% Specify time and space vectors;
x=x0:hx:x1;
t=t0:k:tf;
```

```
46
47 %Specify the initial condition
48 u0=zeros(nx,1);
49 for i=1:nx
50     u0(i)=initialCondition1(x(i));
51 end
52
53 %initialise the matricies for computation
54 uold=u0;
55 unew=u0;
56
57 %Define some indexes
58 index=0;
59 jndex=0;
60
61 %Determine the necessary finite difference stencils that are
       required.
62 cd2x=[1,-2,1]./hx^2;  % the centred difference stencil
63
64 for T=t
65
66     %denote an indexing parameter to pick out certain time
       intervals
67     index=index+1;
68
69     %add the effect of diffusion in the x direction
70     unew=unew+k*conv(uold,cd2x,'same');
71
72     %add the effect of the source term
73     unew=unew+k*src1(x,T,uold,[delta;epsilon]);
74
75     %Boundary condition along the x0 boundary
76     i=1;
77     switch BC(1)
78         case 0  %dirichlet BC
79             unew(i)=boundarycond1(T,x(i),Para_BC);
80         case 1 %neuman BC
81             unew(i)=(4*unew(i+1)-unew(i+2)-2*hx*boundarycond1(T,x(i
       ),Para_BC))/3;
82         case 2 %Robin Condition
83             unew(i)=(4*unew(i+1)-unew(i+2)-2*hx*boundarycond1(T,x(i
       ),Para_BC))/(3+2*a(1)*hx);
84     end
85
86     %Boundary condition for x1 boundary
87     i=nx;
88     switch BC(2)
```

```
89          case 0  %dirichlet BC
90              unew(i)=boundarycond1(T,x(i),Para_BC);
91          case 1 %neuman BC
92              unew(i)=(4*unew(i-1)-unew(i-2)-2*hx*boundarycond1(T,x(i
    ),Para_BC))/3;
93          case 2 %Robin Condition
94              unew(i)=(4*unew(i-1)-unew(i-2)-2*hx*boundarycond1(T,x(i
    ),Para_BC))/(3+2*a(2)*hx);
95      end
96
97      uold=unew;  %Reset the matrix for the next step
98
99  end
```

Listing B.2: Finite Difference Solver.

Similarly to the Runge–Kutta method we can investigate the convergence of the finite difference method. Now that we have a PDE, we need to address convergence in both the spatial and time dimensions. We use the same approach as we had for the RK4 method where we halve the step size and compare the solutions. Halving the grid size k, doubles the number of intervals. We also maintain nodes of the solutions with a larger grid size which means we can consider convergence pointwise. We consider the number of intervals $M = 50, 100, 200, 400, 800$ and compare these to the solution with $M = 1600$. We consider the error absolute value of the difference between the two solutions, as we had with the RK4 method.

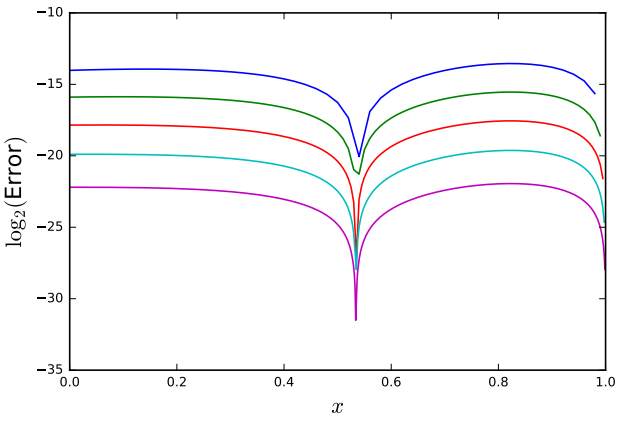


Figure B.3: The $\log_2$ error of the numerical scheme at time $t = 1$.

Figure B.3 indicates that as we halve the grid size, the solutions converge to the solution provided by the most refined grid. Additionally each time we halve the grid size the error is reducing by the same factor 2, with the error profile remaining consistent. The difference of 2 in the log of the error indicates that the numerical scheme is a second order scheme which agrees with our expectation of using the second order finite difference

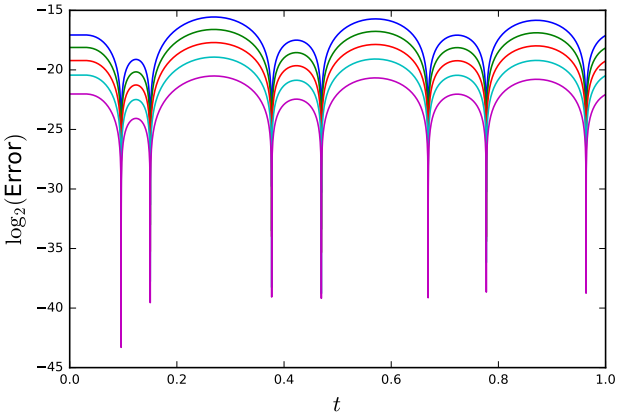


Figure B.4: The $\log_2$ error in time.

approximation for space.

We also check for the convergence in time. As before we halve the step size between each simulation. We compare the temperature at the centre centre of the domain. We compare these solutions to the solution with the smallest step size. Figure B.4 indicates the expected behaviour. We find that the numerical scheme is converging and the difference between the plots is 1, since we have a first order method for the time derivative.

Figures B.3 and B.4 indicate that the numerical schemes are converging in a consistent manner. These convergence plots provide a justification that the numerical scheme is approximating the true solution to the PDE.

B.2.3 Finite Element Method

The finite element methods are more complex, particular from a coding perspective. Because of this we use the FEniCS project [137, 136] to construct the finite element solutions. The FEniCS environment allows the construction of meshes and formulation of the weak problem very simply. Using this environment we consider the function space P, of piece-wise linear functions over the prescribed mesh. This is used for the spatial derivatives but we discretise in time using a finite difference method, similar to a standard backward Euler Method. We use the domain $I = [-1, 1]$, with Dirichlet boundary conditions. In this instance the weak solution solves,

$$\int_I u^n v = h \int_I \nabla u^n \nabla v + \int_I u^{n-1} v + \int_I q(u^{n-1}) v, \qquad \forall v \in P.$$

This forms an iterative scheme over time. One key thing we have implemented here is that we use the function u^{n-1} at the previous time interval in the function $q(u)$. This is due to the non-linear nature of this function. Using this approximation gives a linear equation in order to solve for u. The code used to solve the PDE using this finite element method is in Listing B.3

```
from __future__ import print_function
from fenics import *
import numpy as np
import math

def fenicsSolver(delta=0.9,epsilon=0.027):
  T = 0.3            # final time
  num_steps = 1000     # number of time steps #increase time steps
  dt = T / num_steps # time step size
  #T0=0.637

  # Create mesh and define function space
  M=1024                      #number of intervals
  mesh=IntervalMesh(M,-1,1)
  V = FunctionSpace(mesh, 'P', 2)

  # Define boundary condition
  u_D = Expression('0',degree=2)
  bc = DirichletBC(V, u_D, boundary)

  # Define initial value
  u0=Expression('0',degree=0)
  u_n = interpolate(u0, V)

  # Define variational problem
  u = TrialFunction(V)
  v = TestFunction(V)
  f = Expression('exp(u_n/(1+epsilon*u_n))*delta',
        degree=1, delta=delta,u_n=u_n,epsilon=epsilon)

  F = u*v*dx + dt*dot(grad(u), grad(v))*dx - (u_n+dt*f)*v*dx
  a, L = lhs(F), rhs(F)


  # Time-stepping
  u = Function(V)
  t = 0
  for n in range(num_steps):

    # Update current time
    t += dt
    u_D.t = t
    f.u_n=u_n

    # Compute solution
    solve(a == L, u, bc)

```

```
48      # Update previous solution
49      u_n.assign(u)
50
51  return u
```

Listing B.3: Finite Element solver.

We check the convergence of our numerical scheme using the same method as before. Defining some notation, let u^n denote the solution with 2^n intervals in time and let u_m denote the solution using an interval number 2^m. To test convergence we compare solutions doubling the number of intervals to a solution with interval number 2^9 and compare the solutions at the final time. This tests the convergence in the spatial parameter. Similarly for time we double the number of time intervals, halving the step size, and compare the solution at the point $x = 0$, to the solution with 2^{11} time intervals

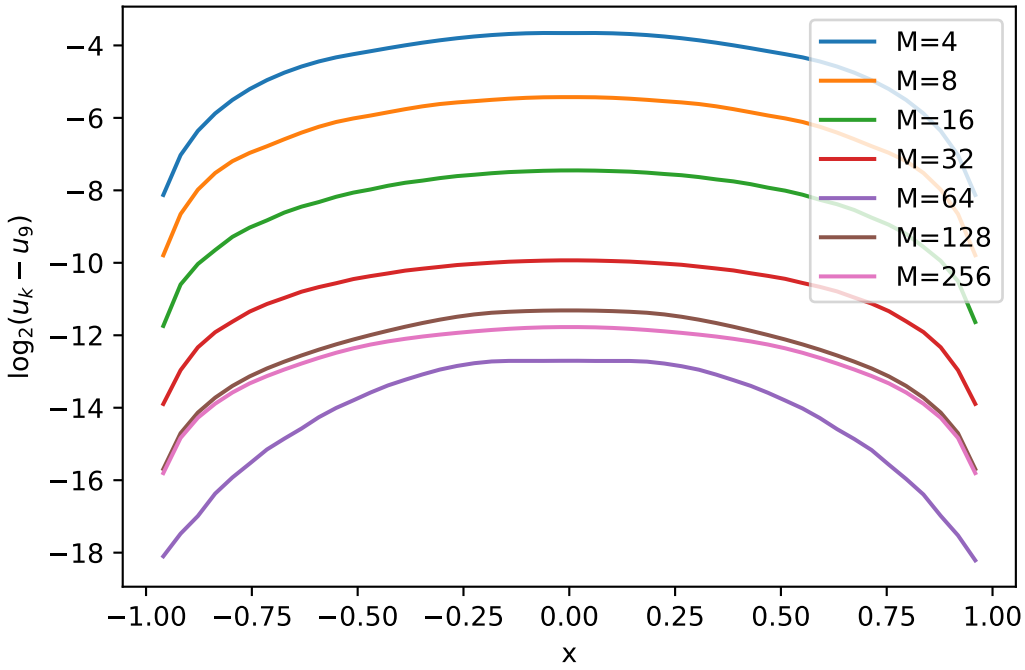
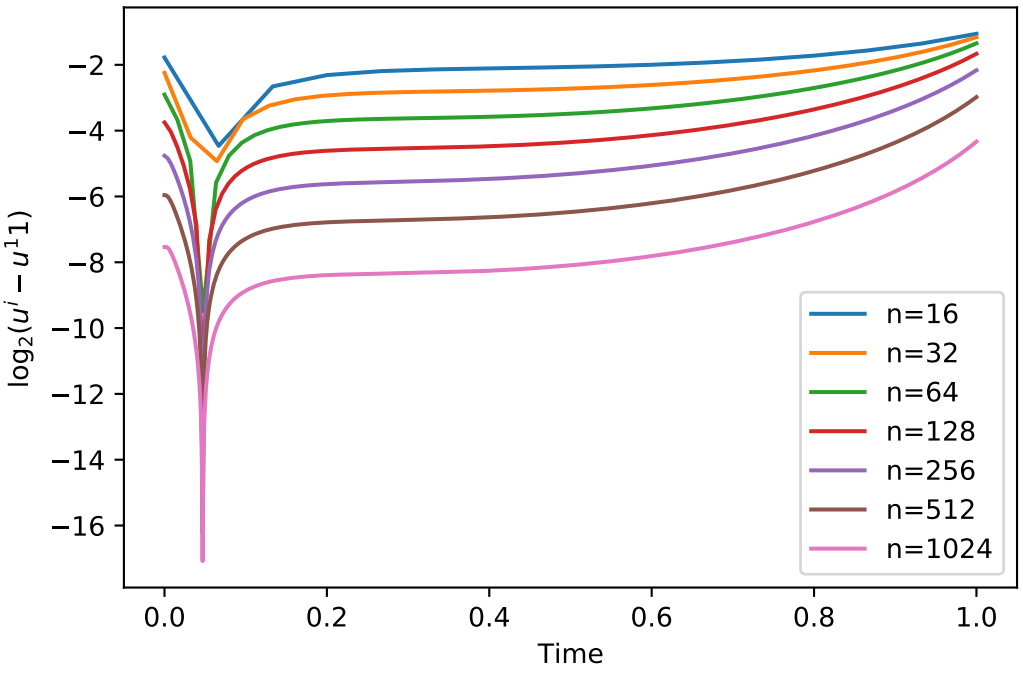


Figure B.5: Convergence of the finite element method in space.

Figure B.5 indicates that we have convergent behaviour in space. There is a slight anomaly when we consider the smaller interval sizes against each other. At these points, as the grid size becomes smaller, then the numerical error due to the rounding of floating point numbers becomes more of an issue. Considering that solving the finite element equation involves solving a linear matrix equation, these rounding errors can be amplified and the effects noticed through these calculations. Given that the convergence in the earlier stages indicate that the scheme is behaving as we expect with order 2, then we can reasonably believe that the numerical schemes are converging. The convergence plots for time in Figure B.6 behave as we expect for a first order finite difference method. The figure indicates we have a first order reaction scheme, validating our numerical scheme.



![ht]

Figure B.6: Convergence of the finite element method in time.

B.3 Summary

The numerical techniques we used in this thesis are well established. The convergence plots for each of the reaction schemes indicate that our numerical solvers are behaving according to the theory. This provides strong justification for their convergence and provides a solid basis for the results that we obtain using these solvers.

Appendix C

Supplementary Information for Bayesian Inference

In this appendix we include some supplementary details and additional plots for the inference carried out in Chapter 4.

C.1 Parameter Values

During Chapter 4 we simulate the experiments with the specified reactions from the experimental data. The parameter values are listed in table C.1.

Parameter	Value
$T(0)$	273 K
$W(0)$	26.6g
$I(0)$	15.0g
α	10 K/min
$w_{c,1}$	-0.376
$w_{c,2}$	-0.074
Q_1	7377
Q_2	1950

Table C.1: Parameter values used in the simulated experiments.

C.2 Distributions

During Chapter 4 we made note of the prior and proposal distributions that were used. Throughout the chapter the priors that we implement for the parameters are,

$$\begin{aligned}
\widetilde{A} &\sim U(0,30), \\
\widetilde{E} &\sim U(2,6), \\
\widetilde{Q} &\sim U(2,6), \\
T_m &\sim U(a,b), \\
\log(\sigma) &\sim N(-1,1),
\end{aligned}$$

unless otherwise specified. The priors that we used for T_m are uniform but the choice of the bounds are dependant upon each application. For the single reaction these values are $a = 600$ and $b = 700$.

In addition to the priors we have a random walk proposal distributions. Given the previous parameters θ, the proposed parameters θ^* are distributed by,

$$\theta^* \sim \mathscr{N}\left(\theta, \frac{\Sigma}{s}\right), \tag{C.1}$$

where Σ is a diagonal covariance matrix that varied between each application. In this equation the parameter s is a scale parameter that changes at fixed iterations. Let N be the length of the Markov chain, then we change the scale at either $N/10$ and $N/3$ or we change the scale at $N/4$ and $N/2$. The decision of when to change is dependant upon how quickly the Markov chain is converging and also the effective sample size (ESS). If the ESS is low, then this indicates the chains need to be longer for longer. This is particularly important since all of the points in the Markov chain that occur before we set the final scale parameter are discarded as burn-in.

C.3 Trace Plots and Summary Statistics

In the thesis we used a Markov chain Monte Carlo (MCMC) algorithm to simulate from the posterior distribution. We assessed the convergence using the ESS and the Split-$\widehat{R}$ diagnostics. Additionally, we can inspect the trace plots of the Markov chains and this yields additional information. We used the trace plots to assess convergence of the MCMC algorithm for the experimental data set but it is also useful to consider these in the simulated cases.

In addition to the histograms of the samples, it is valuable to consider some summary statistics of these samples. In this instance we are most interested in the mean, median,

minimum, maximum and a 90% credible interval (CI). These statistics provide valuable input into these posterior distributions. In this section we present these summary statistics and the trace plots that are not included in Chapter 4.

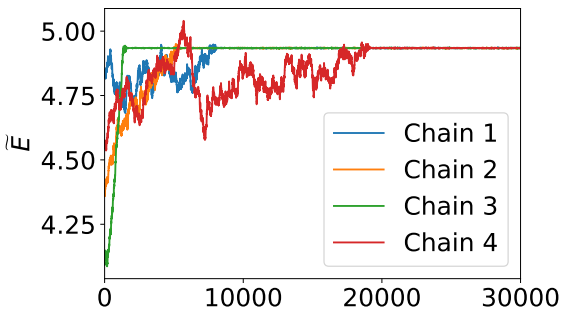
C.3.1 Single Reaction

In addition to the histograms in Chapter 4 we can consider the summary statistics of the sample. We present these in Table C.2. This table indicates that we have successfully obtained the true values within the 90% CIs.

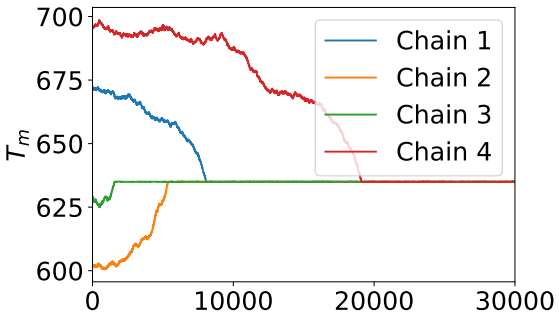
	Mean	Minimum	Median	Maximum	90% CI	True Values
$\widetilde{A}$	6.47	6.43	6.48	6.53	(6.45, 6.50)	6.50
$\widetilde{E}$	4.934	4.931	4.934	4.936	(4.933, 4.935)	4.935
σ	0.046	0.035	0.046	0.074	(0.040, 0.054)	0.05
T_m	634.98	634.83	634.98	635.16	(634.91, 635.04)	635

Table C.2: Summary statistics for the single simulated experiment.

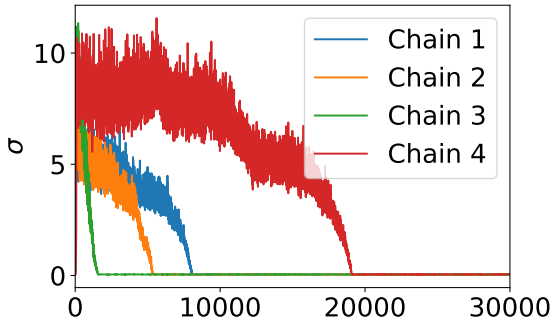
The trace plots for the Markov chains are presented in Figure C.1. These indicate that the Markov chains have converged from different starting points onto the same stationary distribution.



(a) Trace plot for $\widetilde{E}$.

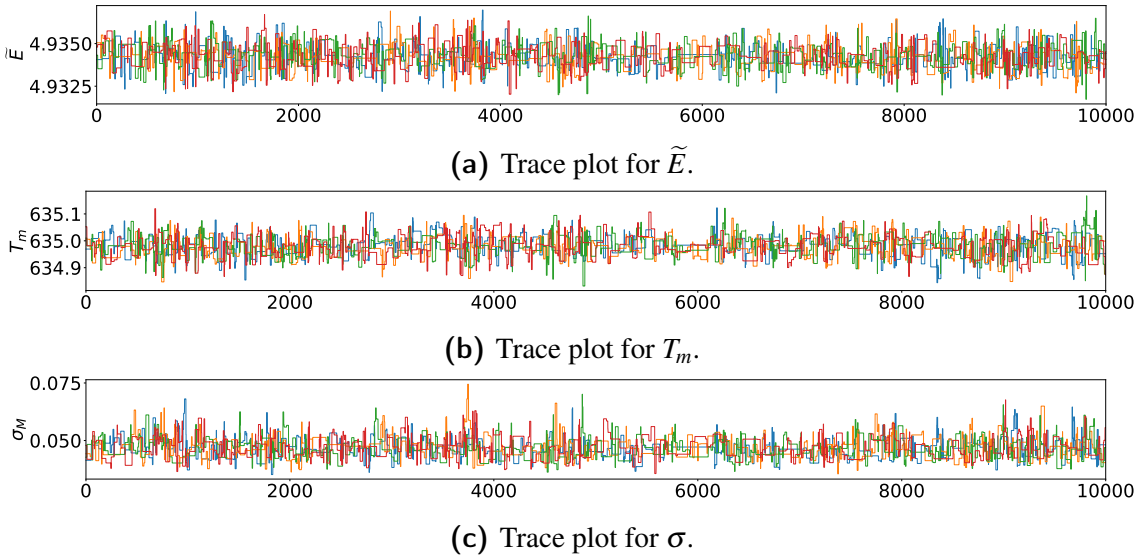


(b) Trace plot for T_m.



(c) Trace plot for σ.

Figure C.1: Trace plots for the simulated single reaction.



(a) Trace plot for $\widetilde{E}$.

(b) Trace plot for T_m.

(c) Trace plot for σ.

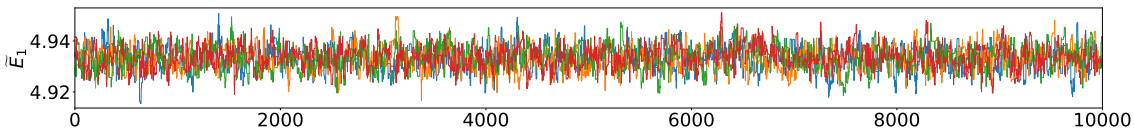
Figure C.2: Trace plots for the simulated single reaction after discarding the burn-in period.

The Markov chains are displayed in Figure C.2, after discarding the burn-in. These trace plots indicate that whilst the chains have converged, they do not sample the space in the most efficient manner and appear to be stuck on certain nodes for an extended period of time. This is typically an indication that a smaller step size is required [8] and this MCMC algorithm used a proposal distribution with a fixed parameter.

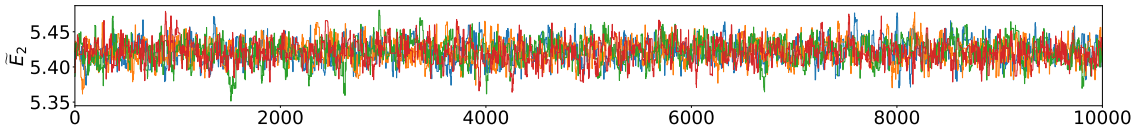
C.3.2 Two reaction Model

For the two reaction we conducted multiple simulations. The trace plots can be used to assist with determining if the Markov chains have converged and hence we present them in this section. Given that we have multiple simulations we simply present the various trace plots for each different iteration without any additional analysis.

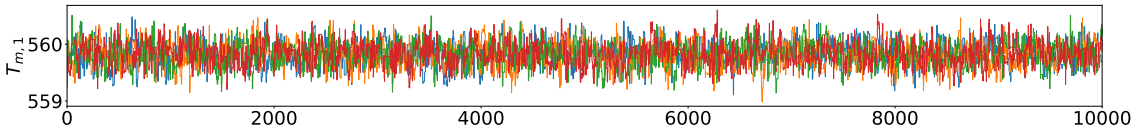
Trace Plots



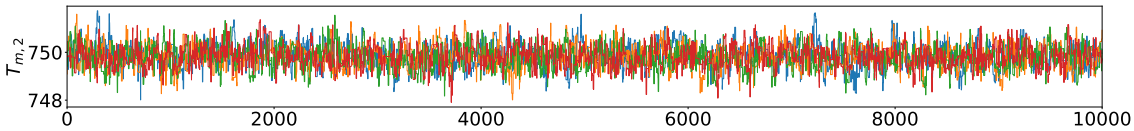
(a) Trace plot of $\widetilde{E_1}$.



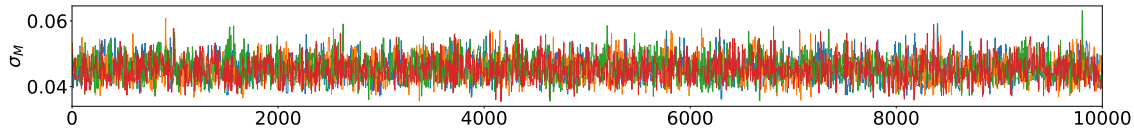
(b) Trace plot of $\widetilde{E_2}$.



(c) Trace plot of $T_{m,1}$.

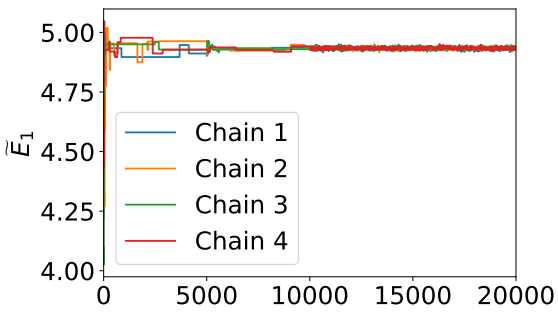


(d) Trace plot of $T_{m,2}$.

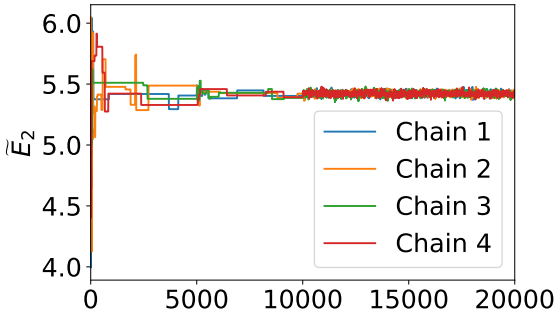


(e) Trace plot of σ.

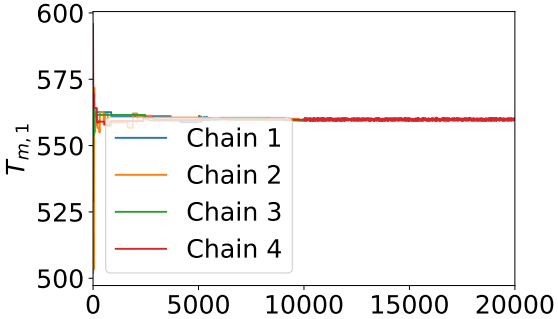
Figure C.3: Trace plots after discarding the burn-in for the two reaction simulated data without DSC for well separated reactions.



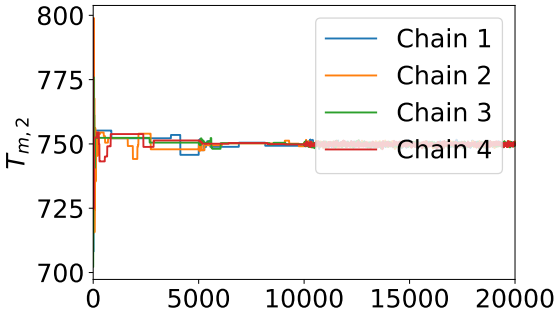
(a) Trace plot of $\widetilde{E_1}$.



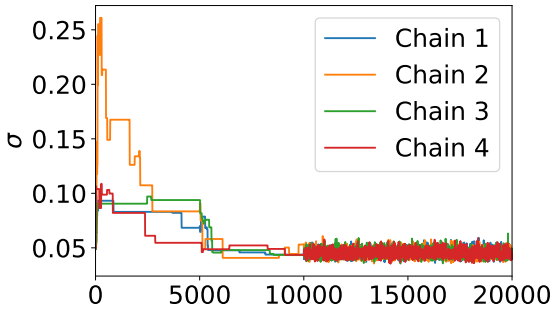
(b) Trace plot of $\widetilde{E_2}$.



(c) Trace plot of $T_{m,1}$.

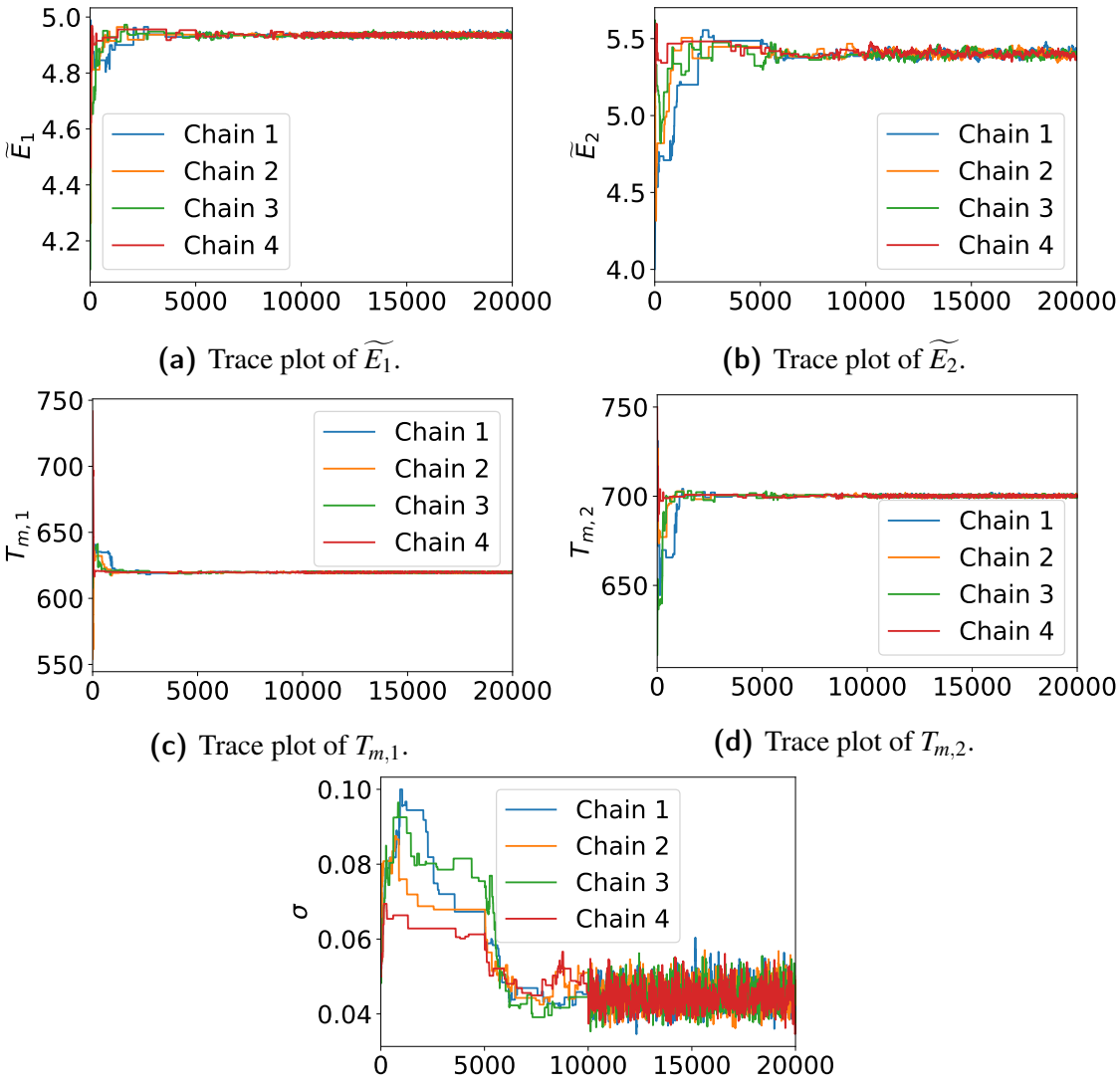


(d) Trace plot of $T_{m,2}$.



(e) Trace plot of σ.

Figure C.4: Trace plots for the two reaction simulated data without DSC and well separated reactions.



(a) Trace plot of $\widetilde{E_1}$.

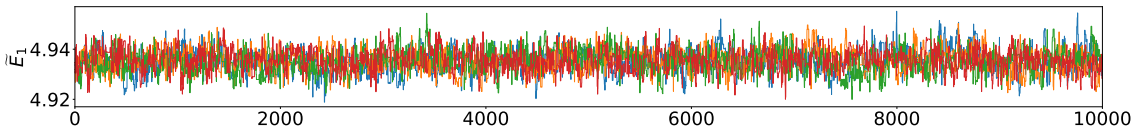
(b) Trace plot of $\widetilde{E_2}$.

(c) Trace plot of $T_{m,1}$.

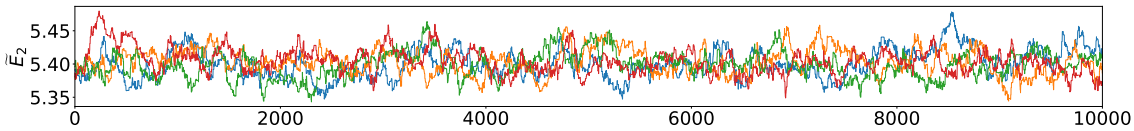
(d) Trace plot of $T_{m,2}$.

(e) Trace plot of σ.

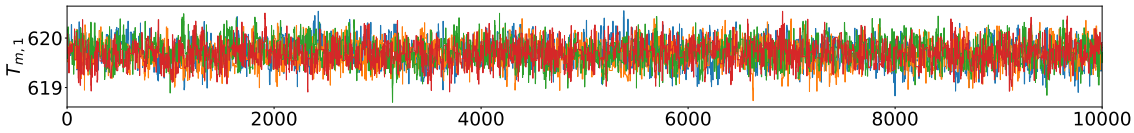
Figure C.5: Trace plots for the two reaction simulated data without DSC and close reactions.



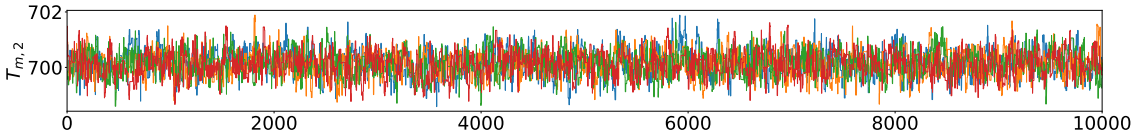
(a) Trace plot of $\widetilde{E_1}$.



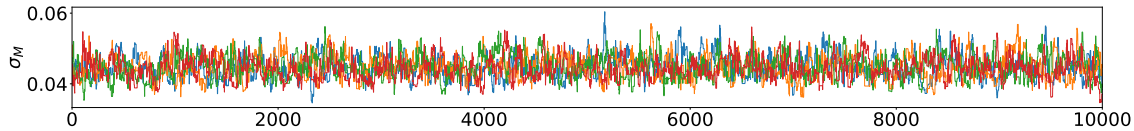
(b) Trace plot of $\widetilde{E_2}$.



(c) Trace plot of $T_{m,1}$.

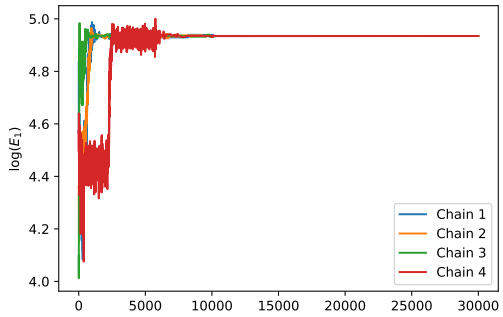


(d) Trace plot of $T_{m,2}$.

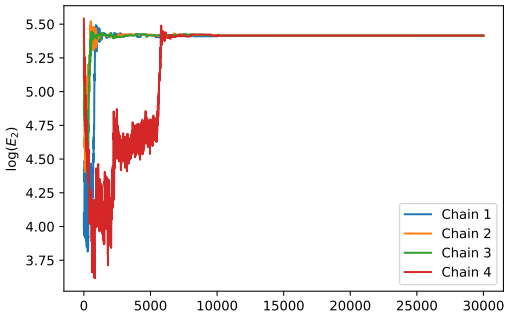


(e) Trace plot of σ.

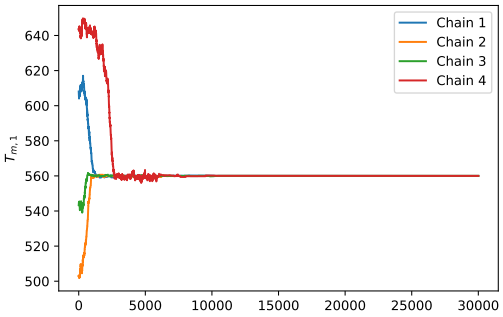
Figure C.6: Trace plots after discarding burn-in for the two reaction simulated data without DSC and close reactions.



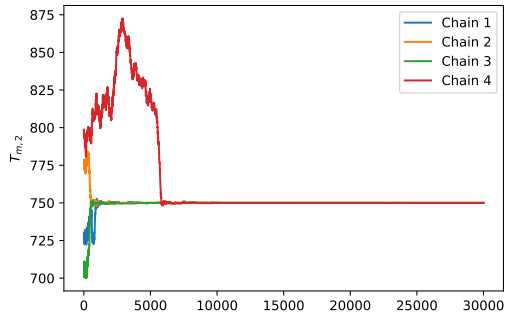
(a) Trace plot of $\widetilde{E_1}$.



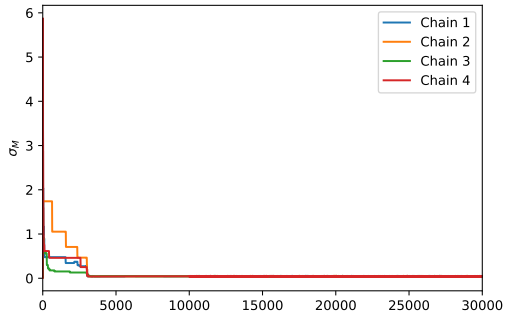
(b) Trace plot of $\widetilde{E_2}$.



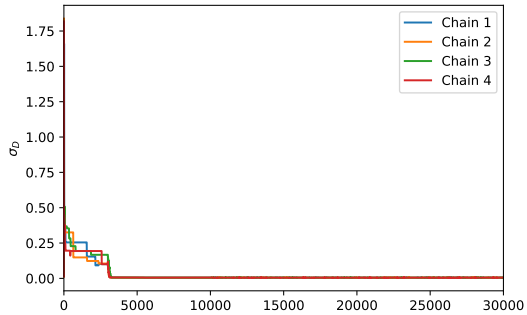
(c) Trace plot of $T_{m,1}$.



(d) Trace plot of $T_{m,2}$.

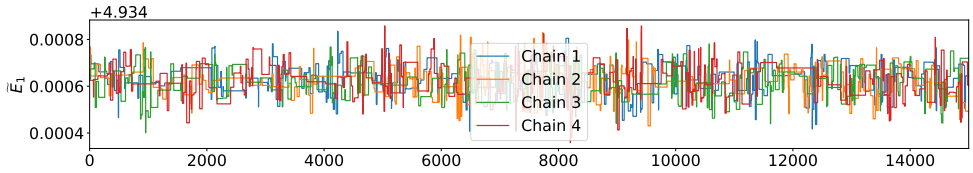


(e) Trace plot of σ_m.

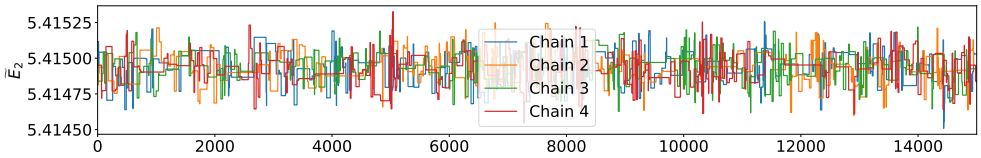


(f) Trace plot of σ_D.

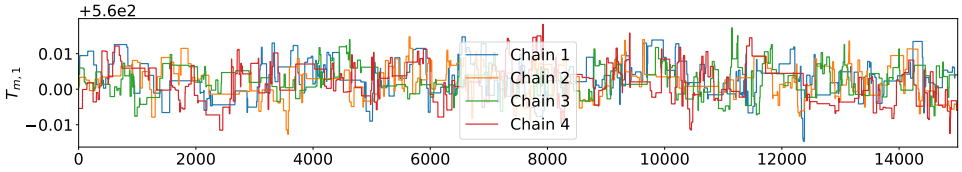
Figure C.7: Trace plots for the two reaction simulated data with DSC and fixed heat parameters.



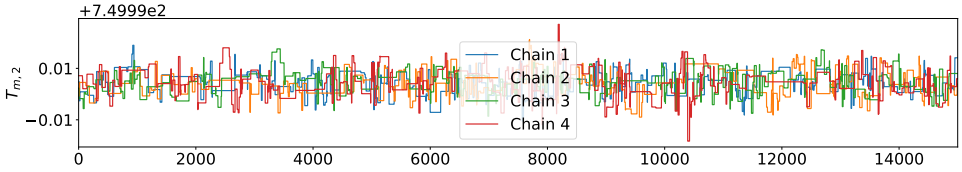
(a) Trace plot of $\widetilde{E_1}$.



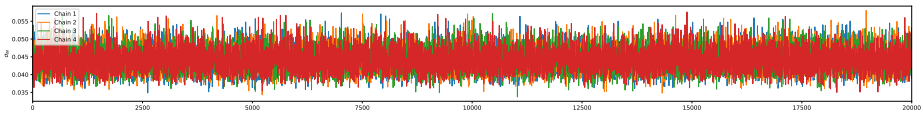
(b) Trace plot of $\widetilde{E_2}$.



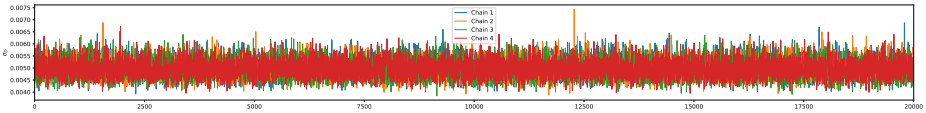
(c) Trace plot of $T_{m,1}$.



(d) Trace plot of $T_{m,2}$.

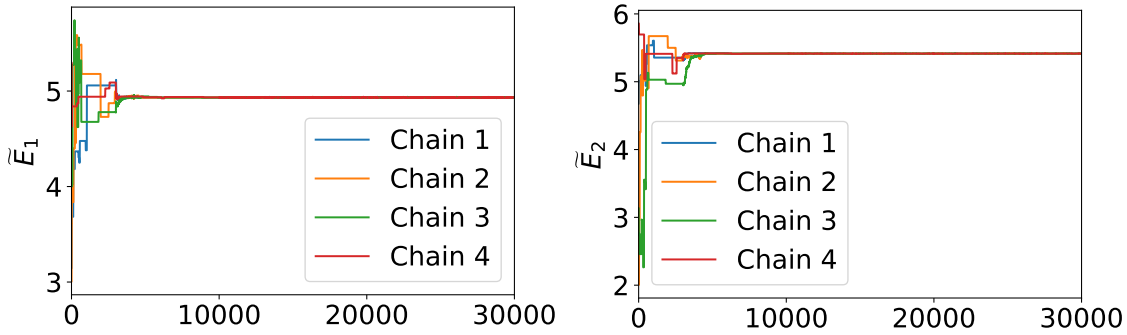


(e) Trace plot of σ_m.



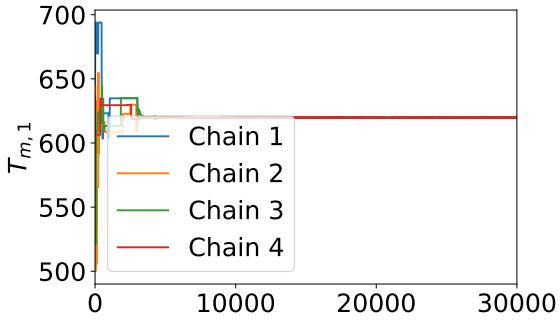
(f) Trace plot of σ_D.

Figure C.8: Trace plots after discarding burn-in for the two reaction simulated data with DSC and fixed heat parameters.

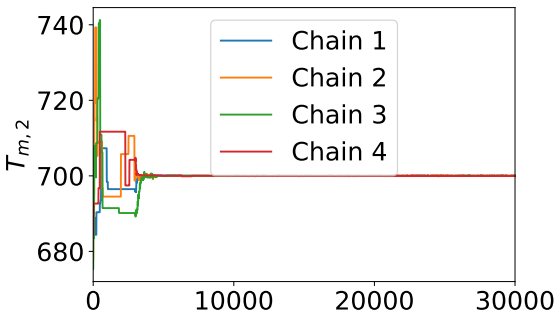


(a) Trace plot of $\widetilde{E_1}$.

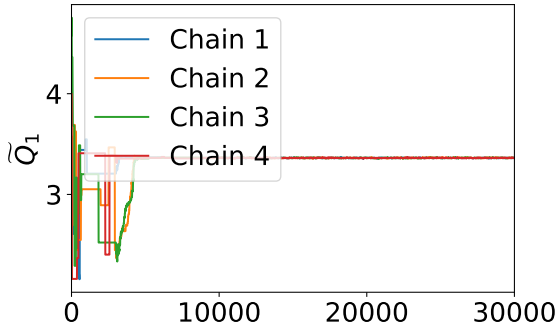
(b) Trace plot of $\widetilde{E_2}$.



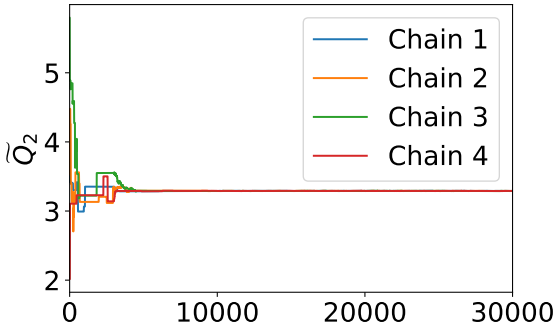
(c) Trace plot of $T_{m,1}$.



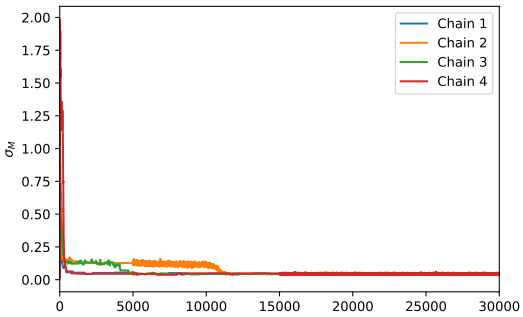
(d) Trace plot of $T_{m,2}$.



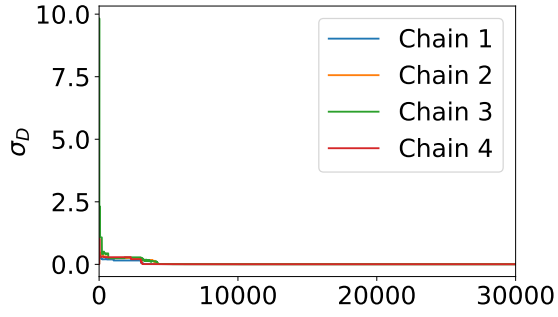
(e) Trace plot of $\widetilde{Q_1}$.



(f) Trace plot of $\widetilde{Q_2}$.

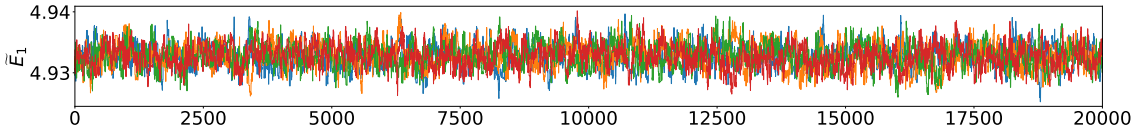


(g) Trace plot of σ_m.

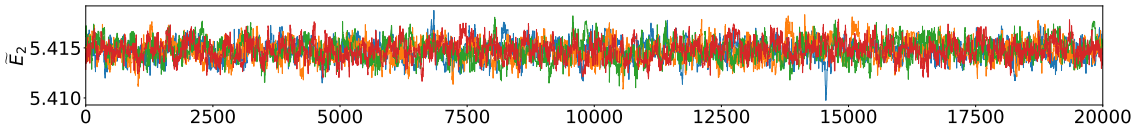


(h) Trace plot of σ_D.

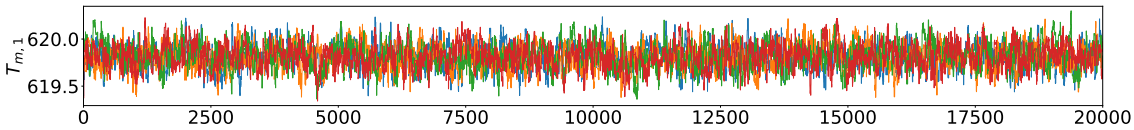
Figure C.9: Trace plots for the to reaction simulated data with DSC and fixed heat parameters.



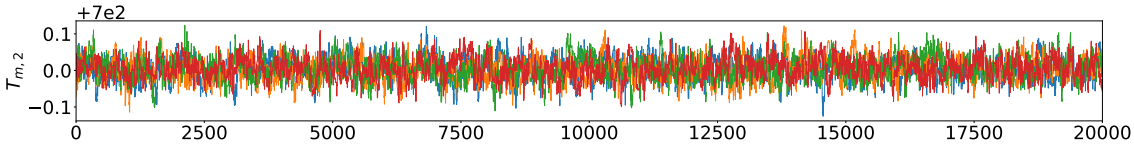
(a) Trace plot of $\widetilde{E_1}$.



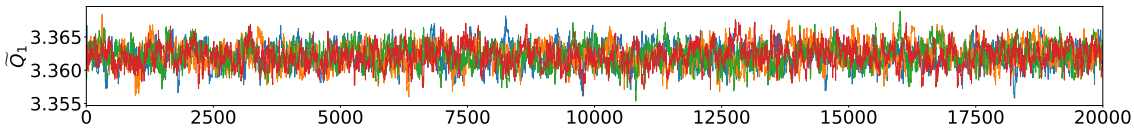
(b) Trace plot of $\widetilde{E_2}$.



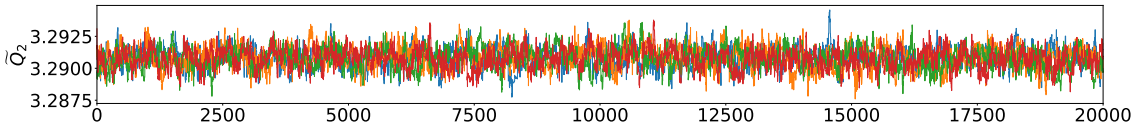
(c) Trace plot of $T_{m,1}$.



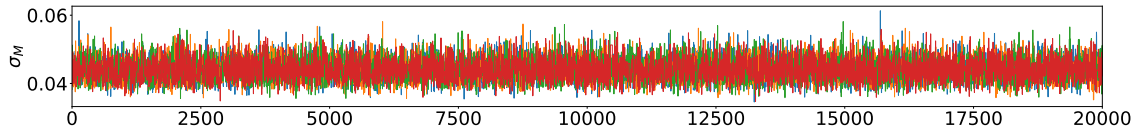
(d) Trace plot of $T_{m,2}$.



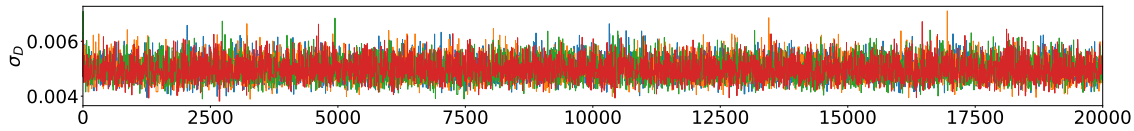
(e) Trace plot of $\widetilde{Q_1}$.



(f) Trace plot of $\widetilde{Q_2}$.



(g) Trace plot of σ_m.



(h) Trace plot of σ_D.

Figure C.10: Trace plots after discarding burn-in for the two reaction simulated data with DSC and fixed heat parameters.

Summary Statistics

	Mean	Minimum	Median	Maximum	90% CI	True Values
$\widetilde{A}_1$	7.52	7.17	7.52	7.87	(7.38, 7.67)	7.54
$\widetilde{A}_2$	18.2	15.3	18.1	20.16	(17.1, 19.3)	17.9
$\widetilde{E}_1$	4.933	4.915	4.933	4.951	(4.925, 4.941)	4.934
$\widetilde{E}_2$	5.42	5.35	5.42	5.48	(5.40, 5.45)	5.41
$T_{m,1}$	559.83	558.99	559.83	560.60	(559.48, 560.17)	560
$T_{m,2}$	749.8	747.9	749.8	751.7	(749.0, 750.6)	750
σ	0.045	0.035	0.045	0.063	(0.040, 0.051)	0.05

Table C.3: Summary statistics for the case without DSC for well separated reactions.

	Mean	Minimum	Median	Maximum	90% CI	True Values
$\widetilde{A}_1$	6.69	6.40	6.69	7.09	(6.55, 6.83)	6.67
$\widetilde{A}_2$	18.6	15.5	18.6	21.8	(17.1, 20.0)	19.2
$\widetilde{E}_1$	4.935	4.918	4.935	4.957	(4.928, 4.943)	4.934
$\widetilde{E}_2$	5.401	5.326	5.401	5.469	(5.367, 5.433)	5.415
$T_{m,1}$	619.7	618.7	619.7	620.8	(619.3, 620.1)	620
$T_{m,2}$	700.1	698.4	700.1	701.7	(699.4, 700.9)	700
σ	0.045	0.036	0.045	0.064	(0.040, 0.051)	0.05

Table C.4: Summary statistics for the case without DSC for close reactions.

	Mean	Minimum	Median	Maximum	90% CI	True Values
$\widetilde{A}_1$	6.65	6.53	6.64	6.76	(6.60, 6.68)	6.68
$\widetilde{A}_2$	19.22	19.11	19.22	19.35	(19.17, 19.27)	19.21
$\widetilde{E}_1$	4.932	4.926	4.933	4.940	(4.930, 4.935)	4.934
$T_{m,1}$	619.87	619.42	619.87	620.32	(619.67, 620.06)	620
$T_{m,2}$	700.01	699.90	700.01	700.12	(699.97, 700.05)	700
$\widetilde{E}_2$	5.415	5.412	5.415	5.418	(5.414, 5.416)	5.415
σ_M	0.044	0.033	0.044	0.058	(0.039, 0.050)	0.050
σ_D	0.0049	0.0038	0.0049	0.0074	(0.0043, 0.0055)	0.005

Table C.5: Summary statistics for the case where the DSC was included with fixed heat paramters.

	Mean	Minimum	Median	Maximum	90% CI	True Values
$\widetilde{A}_1$	6.64	6.51	6.64	6.78	(6.59, 6.70)	6.68
$\widetilde{A}_2$	19.20	18.97	19.19	19.38	(19.12, 19.27)	19.20
$\widetilde{E}_1$	4.932	4.925	4.932	4.940	(4.929, 4.936)	4.934
$\widetilde{E}_2$	5.414	5.401	5.415	5.419	(5.413, 5.416)	5.415
$T_{m,1}$	619.8	619.3	619.8	620.3	(619.6, 620.0)	620
$T_{m,2}$	700.00	699.87	700.00	700.12	(699.95, 700.06)	700
$\widetilde{Q}_1$	3.362	3.355	3.362	3.369	(3.359, 3.365)	3.362
$\widetilde{Q}_2$	3.291	3.287	3.291	3.295	(3.289, 3.292)	3.290
σ_M	0.044	3.45	0.044	0.061	(0.039, 0.049)	0.05
σ_D	0.0050	0.0038	0.0049	0.0071	(0.0044, 0.0056)	0.005

Table C.6: Summary statistics for the case where the DSC is included with inferred heat paramters

Appendix D

MCMC and SMC Algorithms

In this appendix we present the Code used for the Markov Chain Monte Carlo (MCMC) algorithm in Chapter 4 and the Sequential Monte Carlo (SMC) algorithms in Chapter 5. These algorithms are written in python. In this Appendix we present the necessary code for the various algorithms with the complete code for all of the simulations available on Github https://github.com/mtb980/Thesis_Code.

D.1 MCMC algorithm

For the MCMC algorithm we present the code that is used for the experimental data. Each other application of the algorithm follows the same process, but the functions, prior distributions, and proposal distributions follow the same format. Listing D.1 provides a sample of the MCMC code used for the Markov Chain.

```
"""
Created on Sun Apr 11 08:17:20 2021
  Markov Chain Monte Carlo Algorithm to infer
  the parameters from the experimental data.
@author: matt
"""
#Set random seed for reproducibility
np.random.seed(1)

#Import the Data
m0=98.36
data=pd.read_csv('data/TGA_data.csv',header=0,usecols=([0,1,2,3]))
data.columns=['temp','time','DSC','FWC']
data_restrict=data[(data['time']>30) & (data['time']<60)]
data_mass=(data_restrict.values[:,3]*m0/100)
data_DSC=(data_restrict.values[:,2])
data_temp=(data_restrict.values[:,0])

```

```
19
20 def prior(*theta):
21     p1=scist.uniform.logpdf(theta[0],loc=4,scale=2)     #Prior on E1
22     p2=scist.norm.logpdf(theta[1],loc=613,scale=80)     #Prior on
       Tm1
23     p3=scist.uniform.logpdf(theta[2],loc=2,scale=4)     #Prior on Q1
24     p4=scist.uniform.logpdf(theta[3],loc=4,scale=2)     #Prior on E2
25     p5=scist.norm.logpdf(theta[4],loc=700,scale=50)     #Prior on
       Tm2
26     p6=scist.uniform.logpdf(theta[5],loc=2,scale=4)     #Prior on Q2
27     p7=scist.gamma.logpdf(theta[6],1,scale=1000)        #Prior on
       TauM
28     p8=scist.gamma.logpdf(theta[7],1,scale=1000)        #Prior on
       TauD
29     p=p1+p2+p3+p4+p5+p6+p7+p8
30     return p
31
32 def f(j,data):
33     #define sample size
34     sample_size=30000
35
36     #data
37     data_mass=data[0]
38     data_DSC=data[1]
39     N=len(data_mass)
40
41     #Initialise the MCMC algorithm
42     np.random.seed(seed=10*j+1)
43
44     #Initalise parameter lists
45     E1=[0]*sample_size
46     Tm1=[0]*sample_size
47     Q1=[0]*sample_size
48     E2=[0]*sample_size
49     Tm2=[0]*sample_size
50     Q2=[0]*sample_size
51     TauM=[0]*sample_size
52     TauD=[0]*sample_size
53     qs=[0]*sample_size
54
55     #Initalise subsequent parameter
56     A1=[0]*sample_size
57     A2=[0]*sample_size
58     sigmaM=[0]*sample_size
59     sigmaD=[0]*sample_size
60     acceptance=[0]*sample_size
61
```

```
    #define Auxillary parameters
    alpha0=1
    a=alpha0+N/2
    beta0=0.001

    #Initialise Parameter values
    E1[0]=scist.uniform.rvs(loc=4,scale=2)
    Tm1[0]=scist.norm.rvs(loc=613,scale=40)
    Q1[0]=scist.uniform.rvs(loc=2,scale=4)
    E2[0]=scist.uniform.rvs(loc=4,scale=2)
    Tm2[0]=scist.norm.rvs(loc=700,scale=25)
    Q2[0]=scist.uniform.rvs(loc=2,scale=4)
    TauM[0]=scist.gamma.rvs(a,scale=1/beta0)
    TauD[0]=scist.gamma.rvs(a,scale=1/beta0)


    A1[0]=log_Pre_exp(Tm1[0],10**E1[0],alpha)
    A2[0]=log_Pre_exp(Tm2[0],10**E2[0],alpha)
    sigmaM[0]=1/np.sqrt(TauM[0])
    sigmaD[0]=1/np.sqrt(TauD[0])

    #Solve the ODE equation
    sol=np.array(TGA(math.pow(10,A1[0]),pow(10,A2[0]),math.pow(10,
    E1[0]),
                     pow(10,E2[0]),10**Q1[0],10**Q2[0],alpha))


    #Determine Residuals
    eM=sol[0]-data_mass
    eD=sol[1]-data_DSC

    #calculate betas
    betaMi=beta0+np.linalg.norm(eM)**2/2
    betaDi=beta0+np.linalg.norm(eD)**2/2

    #Evaluate the Liklihood
    p1_M=sum(scist.norm.logpdf(eM,scale=sigmaM[0]))
    p1_P=prior(E1[0],Tm1[0],Q1[0],E2[0],Tm2[0],Q2[0],TauM[0],TauD
    [0])
    p1_D=sum(scist.norm.logpdf(eD,scale=sigmaD[0]))
    p1=p1_M+p1_D+p1_P
    #initalise some miscellaneous parameters
    acceptance[0]=0
    counter=1
    scale=6
    pf=1*scale
    pf2=1*scale
```

```
    for i in range(1,sample_size):

        if i>sample_size/3:
            pf=50*scale
            pf2=25*scale

        elif i>sample_size/10:
            pf=4*scale
            pf2=4*scale

        #Propose new values for the parameters from random walks
        Tm1t=scist.norm.rvs(loc=Tm1[i-1],scale=12/pf2)
        E1t=scist.norm.rvs(loc=E1[i-1],scale=0.2/pf2)
        Q1t=scist.norm.rvs(loc=Q1[i-1],scale=0.3/pf2)
        Tm2t=scist.norm.rvs(loc=Tm2[i-1],scale=12/pf)
        E2t=scist.norm.rvs(loc=E2[i-1],scale=0.2/pf2)
        Q2t=scist.norm.rvs(loc=Q2[i-1],scale=0.3/pf)


        #Propose new Tau from proposal distribution
        TauMt=scist.gamma.rvs(a,scale=1/betaMi)
        TauDt=scist.gamma.rvs(a,scale=1/betaDi)

        #Calculate additional parameters
        A1t=log_Pre_exp(Tm1t,10**E1t, alpha)
        A2t=log_Pre_exp(Tm2t,10**E2t, alpha)
        sigmaMt=1/np.sqrt(TauMt)
        sigmaDt=1/np.sqrt(TauDt)

        #Solve the Differential Equation
        sol=TGA(10**A1t,10**A2t, 10**E1t,10**E2t,10**Q1t,10**Q2t,
    alpha)

        #Calculate Residuals
        eM=sol[0]-data_mass
        eD=sol[1]-data_DSC

        #Calculate new Betas
        betaMt=beta0+np.linalg.norm(eM)**2/2
        betaDt=beta0+np.linalg.norm(eD)**2/2

        #Evaluate the Log-Liklihood
        p2_M=sum(scist.norm.logpdf(eM,scale=sigmaMt))
        p2_P=prior(E1t,Tm1t,Q1t,E2t,Tm2t,Q2t,TauMt,TauDt)
        p2_D=sum(scist.norm.logpdf(eD,scale=sigmaDt))
        p2=p2_M+p2_D+p2_P

```

```
153         #Evaluate the ratio of proposals
154         q_M=scist.gamma.logpdf(TauM[i-1],a,scale=1/betaMt)-scist.
    gamma.logpdf(TauMt,a,scale=1/betaMi))
155         q_D=scist.gamma.logpdf(TauD[i-1],a,scale=1/betaDt)-scist.
    gamma.logpdf(TauDt,a,scale=1/betaDi)
156         q=q_M+q_D
157
158         #Evaluate the acceptance probability
159         p=p2-p1+q
160         #Determine the acceptance parameter
161         beta=np.log(scist.uniform.rvs(loc=0,scale=1))
162         if p>beta:
163             #Accept the new Values
164             E1[i]=E1t
165             Tm1[i]=Tm1t
166             Q1[i]=Q1t
167             E2[i]=E2t
168             Tm2[i]=Tm2t
169             Q2[i]=Q2t
170             TauM[i]=TauMt
171             TauD[i]=TauDt
172
173             A1[i]=A1t
174             A2[i]=A2t
175             sigmaM[i]=sigmaMt
176             sigmaD[i]=sigmaDt
177
178             betaMi=betaMt
179             betaDi=betaDt
180
181             p1=p2
182             counter+=1
183         else:
184             #Set to previous values
185             E1[i]=E1[i-1]
186             Tm1[i]=Tm1[i-1]
187             Q1[i]=Q1[i-1]
188             E2[i]=E2[i-1]
189             Tm2[i]=Tm2[i-1]
190             Q2[i]=Q2[i-1]
191             TauM[i]=TauM[i-1]
192             TauD[i]=TauD[i-1]
193
194             A1[i]=A1[i-1]
195             A2[i]=A2[i-1]
196             sigmaM[i]=sigmaM[i-1]
197             sigmaD[i]=sigmaD[i-1]
```

```
198
199         #Set Miscellaneous Parameters
200         acceptance[i]=(counter/(i+2))
201         qs[i]=p
202     col_names=['A1','E1','sigmaM','acceptance','Tm1',
203                'A2','E2','Tm2','sigmaD','Q1','Q2','ps']
204     data=np.transpose(np.array([A1,E1,sigmaM,acceptance,Tm1,
205                                 A2,E2,Tm2,sigmaD,Q1,Q2,qs]))
206     return(pd.DataFrame(data,columns=col_names))
```

Listing D.1: MCMC algorithm for experimental data.

D.2 SMC Algorithms

In Chapter 5 we introduced a SMC algorithm that we used to infer data using multiple sets of experiments. To do this we used the python code in Listing D.2.

```
1  """
2  Compacted Code for the SMC algorithm
3  """
4
5  def evaluateWeights(x,data=np.zeros(60),alpha=10):
6      A1=x.A1
7      E1=x.E1
8      Q1=x.Q1
9      sigmaM=x.sigmaM
10     A2=x.A2
11     E2=x.E2
12     Q2=x.Q2
13     sigmaD=x.sigmaD
14
15     sol=np.array(TGA(pow(10,A1),10**A2,pow(10,E1),10**E2,10**Q1
       ,10**Q2,alpha))
16     eM=sol[0]-data[0]
17     eD=sol[1]-data[1]
18     wM=sum(scist.norm.logpdf(eM,scale=sigmaM))
19     wD=sum(scist.norm.logpdf(eD,scale=sigmaD))
20     return wM+wD
21
22
23 def MCMC_SMC(theta0,var,likelihood,prior,data,alphas,steps=10):
24     theta=np.array(theta0)
25     p=likelihood(theta,data,alphas)
26     counter=0
27     for i in range(steps):
28         thetat=scist.multivariate_normal.rvs(mean=theta,cov=var)
```

```
        pt=likelihood(thetat,data,alphas)
        pc=pt-p
        beta=np.log(scist.uniform.rvs(loc=0,scale=1))
        if beta<pc:
            theta=thetat
            p=pt
            counter+=1
    A1=log_Pre_exp(theta[1],10**theta[0],10)
    A2=log_Pre_exp(theta[5],10**theta[4],10)
    theta=np.append(theta,[A1,A2,counter/steps])
    return theta


num_cores=50

#Import sample
sample=processData('SIM_Q',Burnin=20000)

#Datas
#The data from experiment i is contained in EXPi
alphas=[10,10,5,20]
EXPs=[EXP1,EXP2,EXP3,EXP4]

N=40000      #Sample Size
sample['weights']=np.ones(N)
#Create a unique sample to reduce computation requirements
unique_sample =sample.drop_duplicates()

#SMC Resampling Process
for i in range(1,4):
    #Select experiment to weight
    data=EXPs[i]
    #Calculate Weights
    weights = np.array(Parallel(n_jobs=num_cores)(delayed(
   evaluateWeights)(x,data=data,alpha=alphas[i]) for x in
   unique_sample.itertuples()))
    unique_sample['weights']=weights+unique_sample['weights']
    sample=merge_weights(sample, unique_sample)
    allweights=sample['weights'].to_numpy()
    #Compute ESS
    ESS=np.exp(2*scisp.logsumexp(allweights)-scisp.logsumexp(2*
   allweights))
    if ESS<100:
        #Resample according to the weights
        probs=logProb(weighted_sample['weights'].to_numpy())
        resampled=pd.DataFrame(Parallel(n_jobs=num_cores)(delayed(
   resample)(weighted_sample,probs) for j in range(N)))
```

```
        #Initiate MCMC resampling
        Thetadf=resampled[['E1','Tm1','Q1','sigmaM','E2','Tm2','Q2'
    ,'sigmaD']]
        variances=Thetadf.cov().to_numpy()
        sample=pd.DataFrame(Parallel(n_jobs=num_cores)(delayed(
    MCMC_SMC)(theta, var=variances,likelihood=liklihood,prior=prior,
     data=EXPs[:i],alphas=alphas[:i]) for theta in resampled.
    itertuples(index=False)))
        #reset weights
        sample['weights']=np.ones(N)
```

Listing D.2: SMC Algorithm.

In addition to the code for the SMC method in Listing D.2, we change our approach when considering the experimental data. One way we did this was to consider an SMC algorithm that used the distributions from the MCMC algorithms to define a new prior. With this new prior we implemented importance sampling (IS) to sample from the posterior distribution after sampling from our prior. The python code to do this is presented in Listing D.3.

```
#Code that loads the output from the MCMC algorithm and conducts IS
    .

#Define the weight function
def calculatWeights(theta,data=np.zeros(60),alpha=10):
    sol1=np.array(TGA1(theta,alpha))
    sol2=np.array(TGA2(theta,alpha))
    e1=sol-data[0]
    e2=sol2-data[1]
    w1=sum(scist.norm.logpdf(e1,scale=sigma))
    w2=sum(scist.norm.logpdf(e2,scale=sigma))
    return w1+w2

#import TGA data
data1=loadData('EXP1')
data2=loadData('EXP2')
#Load Samples
sample1=loadSample('EXP1')
sample2=loadSample('EXP2')
sample=pd.concat([sample1,sample2])
#Set mean and covariance for Prior Distribution
var=sample.cov().to_numpy()
mean=sample.mean().to_numpy()

#Sample from Prior Distribution
sample_size=100000
theta=pd.DataFrame(scist.multivariate_normal.rvs(mean,var,size=
    sample_size))
```

```
27
28 #Calculate weights
29 weights = Parallel(n_jobs=num_cores)(delayed(calcuateWeights)(x,
       data=[data1,data2],alpha=10) for x in theta.itertuples())
30
31 #resample according to the weights.
32 resample= Parallel(n_jobs=num_cores)(delayed(resample)(theta,
       weights) for i in range(sample_size))
```

Listing D.3: IS algorithm.

The code that has been provided in this appendix is minimal and focuses on the translation of the general algorithms into written code. The full extent of code is provided on Github https://github.com/mtb980/Thesis_Code.