MATERIALS DESIGN UNDER BAYESIAN UNCERTAINTY QUANTIFICATION

A Dissertation

by

PEJMAN HONARMANDI

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Raymundo Arroyave
Committee Members,	Dimitris C. Lagoudas
	Ibrahim Karaman
	Edward R. Dougherty
Head of Department,	Ibrahim Karaman

August 2019

Major Subject: Materials Science and Engineering

ABSTRACT

Uncertainty quantification and its propagation across multi-scale model/experiment chains are key elements of decision-based materials design in the framework of Integrated Computational Materials Engineering. In this context, understanding the sources of uncertainty and their quantification can provide a confidence for the applicability of models for decision making in materials design, which is generally overlooked in the field of materials science.

Based on the above mentioned motivation, different case studies are considered in this work to indicate how Bayesian inverse uncertainty quantification and forward uncertainty propagation approaches operate in various applications and procedure conditions. In this dissertation, inverse uncertainty quantification of model parameters is performed through a Markov Chain Monte Carlo approach; and all propagation of uncertainties from model parameters to model responses are accomplished through the first order second moment approach and/or the forward model analysis of parameters sampled from the posterior probability distribution after the parameter probabilistic calibrations. Moreover, different information fusion approaches are proposed here to smartly combine the probabilistic information obtained from different sources of information for more precise probabilistic predictions of physical systems behaviors.

This dissertation starts with the importance of uncertainty quantification for product design and engineering. This is followed by some fundamentals about different sources of uncertainties, different statistical views and approaches for uncertainty quantification and propagation in computational modeling, and the previous work in literature for uncertainty quantification and propagation in materials science problems. Then, uncertainties are evaluated in the case of plastic flow behavior modeling of transformation induced plasticity steels using two different procedures of data training, including sequential training with each experimental data as independent evidence and simultaneous training of all data together as overall evidence. A multi-objective probabilistic calibration of an Ni-Ti precipitation model in MatCalc© are also performed against all experimental data simultaneously to quantify the uncertainties of resulting micro-structural features from the

model. It should be noted that an empirical relationship for matrix/precipitate inter-facial energy in terms of aging temperature and nominal composition have also been introduced according to the values of inter-facial energy obtained from the model calibration with each given experimental data individually. However, large discrepancies and uncertainties obtained for model results are important reasons to apply co-kriging surrogate modeling for more precise prediction of precipitation behavior and its uncertainty based on the establishment of a linear correlation between the experimental responses and the fitted surrogate model over the results of the precipitation model. In addition, a constrained probabilistic calibration is carried out for a thermo-mechanical model using a distance based comparison metric of transformation strain-temperature curves. In this case, a design of experiment followed by a variance-based sensitivity analysis are also performed to identify the most influential model parameters before the calibration. Uncertainty quantification in the calculation of Hf-Si binary phase diagram are also discussed in this dissertation. In this case, Bayesian model averaging and an error correlation-based model fusion are also applied to combine all the results obtained from randomly generated models together to make the calculation of phase diagrams more objective rather than being subjective to the expert opinions for the model selection. In the end, the major efforts for uncertainty quantification of thermodynamic properties and phase diagrams are reviewed.

For future work, the probabilistic calibration of the influential parameters in expensive models (such as phase field models) can be performed through the efficient global optimization approach. In these cases, the uncertainty propagation of parameters to the model responses through an efficient method is also very important. Moreover, the proposed information fusion approaches can be applied to combine the model and experimental results for the efficient optimization cases (such as efficient global optimization or knowledge gradient) since the fused responses can be more beneficial than just physical models/simulations in solving the inverse problems in materials design under the Integrated Computational Materials Engineering framework.

DEDICATION

To my lovely wife, my family, my in-laws, and my friends.

ACKNOWLEDGMENTS

I would like to express my special thanks to my advisor, Prof. Raymundo Arroyave for all of his invaluable guidance and support during my doctoral program. I would like to thank my dissertation committee members for their constructive comments and feedback that have improved the quality of my research and the content of this dissertation. I am very thankful to have my fellow graduate students and collaborators, especially Dr. Thien Duong, Dr. Luke Johnoson, Dr. Anjana Talapatra, Dr. Sean Gibbons, Dr. Alex Solomou, Dr. Seyede Fatemeh Ghoreishi, and Vahid Attari, for their insightful discussions and help in my research. I would also like to thank the National Science Foundation (NSF)–United States [grant nos. CMMI-1534534, CMMI-1663130, CMMI-0900187, CMMI-095398, DGE-1545403] for their financial support in my research projects.

In addition, I would like to thank my family and in-laws for their financial and emotional supports during my Ph.D. program. I am also very grateful to have the OASIS friends who helped me get settled in America, improved my English and also provided a cheerful environment that was extremely helpful to better handle hardships in life during my graduate study.

More importantly, I am very blessed to have a supportive, thoughtful, understanding and calming wife, Rebecca Honarmandi, who has always been by me over the last three years, and definitely without her help getting my Ph.D. degree would be much more difficult. I am very thankful to have her in my life.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professors Raymundo Arroyave (advisor) and Ibrahim Karaman of the Department of Materials Science and Engineering, Professor Dimitris Lagoudas of the Department of Aerospace Engineering, and Professor Edward Dougherty of the Department of Electrical and Computer Engineering.

The codes for the precipitation model in Chapter 3, the thermo-mechanical model in Chapter 4, CALPHAD models in Chapter 5, and the CMF approach in Chapter 5 were provided by Doctor Luke Johnson from Professor Arroyave's group of the Department of Materials Science and Engineering, Doctor Alex Solomou from Professor Lagoudas' group of the Department of Aerospace Engineering, Doctor Thien Duong from Professor Arroyave's group of the Department of Materials Science and Engineering, and Doctor Seyede Fatemeh Ghoreishi from Professor Allaire's group of the Department of Mechanical Engineering. Moreover, Sections 6.1, 6.2, 6.4, 6.6, and 6.7 in this dissertation were written with the contributions of Doctor Noah Paulson and Professors Raymundo Arroyave and Marius Stan.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study was supported by the National Science Foundation (NSF)–United States [grant nos. CMMI-1534534, CMMI-1663130, CMMI-0900187, CMMI-095398, DGE-1545403].

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

uncertainty quantification (UQ) is all about the determination of bounds on the behavioral predictions of the system of interest, which can provide confidence for decision making in the context of robust or reliability-based design. UQ in modeling play and important role in materials design under the framework of Integrated Computational Materials Engineering (ICME) where computational models/simulations and experiments that are applied to close the process-structure-property-performance loop are systematically combined together in order to obtain the alloy and processing specifications required for the desirable performance in a system. Generally, UQ incorporates the detection of uncertainty sources and the development of corresponding appropriate mathematical approaches to calculate the error bounds for any quantity of interest in models [1, 2].

Traditionally, predictions in computational modeling in the field of materials science and engineering are deterministic, which result from the application of the deterministic approaches, such as least squares, for the parameter calibrations. In these approaches, a single estimate is obtained for each model parameter based on the available data for the system of interest. However, in practice, different combinations of the parameter values may result in similar values of the model predictions, which can have the chances to be the optimum parameter values in the probabilistic point of view. Since just one of these combinations are introduced in the deterministic approaches and the rest are ignored, the corresponding predictions cannot provide useful information for robust or reliability-based design due to the absence of uncertainty bounds on the model predictions. However, the probabilistic approaches can find a group of possible parameter combinations in the form of probability distributions which can be translated into the probability distributions for the model predictions through the propagation of uncertainty. The probabilistic evaluation of models can provide some useful information for robust materials design where the predictions should be independent of the parameter/input variable variations in the system. Therefore, the values of the design parameters/input variables are selected such that their noises have very small influences on the system responses (predictions). In addition, the importance of UQ for reliable designs can be comprehended through the schematic illustrations in Figure 1.1. As can be observed in this figure, the reliable design region (green region) in the total design input/output space is determined by finding the furthest region from the boundaries for the allowable design regions (red line) or from the uncertainty bounds around these boundaries in the deterministic or probabilistic point of view, respectively. For instance, the allowable design region can be defined as a region in the multi-dimensional space of temperature-time-heating and cooling rate for the heat treatment of a given alloy that allows or prevents the formation of specific micro-structural phases. In probabilistic view, the reliable design region is not necessarily in the center of the allowable design region as it is in the deterministic way, since the confidence intervals can change across the boundaries, as shown in Figure 1.1(b). Therefore, the region proposed by the deterministic view may intersect the uncertainty bounds around the boundaries which implies the possibility of having the design outside of the allowable region.

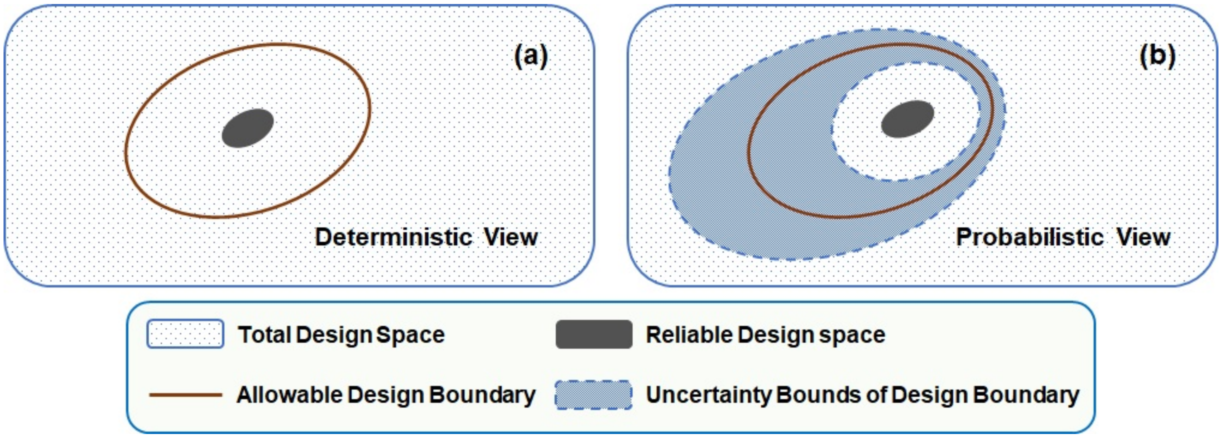


Figure 1.1: The schematic illustrations of reliable designs in deterministic and probabilistic point of view

The importance of UQ can be discussed in the context of safety factor in design which is one of the major design factors to prevent the failure of the system. In deterministic approaches, the safety factors are usually determined based on the degree of concern about the detrimental effects of the system failure; however, very high values are usually selected for this factor that might not be always necessary in design. On the other hand, the generation of confidence bounds in probabilistic approaches can provide better estimations of the safety factors that can generally lead to cost reduction in design. For more clarification, the deterministic and probabilistic approaches to determine the safety factor for a load-bearing material/structure have been shown in Figure 1.2. As can be observed in Figure 1.2(a), safety factors that have been standardized based on the degree of concern in design are considered for both calculated load and strength (resistance) in the deterministic methods in order to find final design load and resistance. If the final design resistance is larger than the final design load, the safety requirements in design are fulfilled; otherwise, the strength of the material/structure must increase further. However, the considerations of load and resistance as the random variables in the form of the probability distributions can be very useful for better estimation of the safety factor. In this probabilistic scheme, less intersection area (blue hatched area in Figure 1.2(b)) is equivalent to safer design since the probability of failure, i.e. resistance < load, reduces. Therefore, the safety factor can be determined by defining a maximum allowable intersection area based on the degree of concern in design. As a result, the probabilistic approaches can provide more precise and less conservative safety factor compared to the deterministic approaches which can reduce the design cost.

Before proceeding our discussion of UQ in computational modelling of materials, it is also important to know what are the sources of uncertainties. The sources of uncertainty are generally classified in different ways; however, there are two broad categories for uncertainty based on the capability for uncertainty reduction, known as aleatoric and epistemic uncertainty [1]. Aleatoric uncertainty refers to the inherent random error in a system that results in a frequency (probability) distribution for the response of the system. This type of uncertainty is measurable, but not reducible by the current available technologies. on the other hand, epistemic uncertainty describes the controllable and reducible errors in the system that result from the lack and/or inaccuracy of information. This discussion about aleatoric and epistemic uncertainty can become more clear in their analogy to accuracy and precision in target shooting. As shown in Figure 1.3, the deviations of the shots from the target center and the amounts of their dispersion indicate aleatoric and

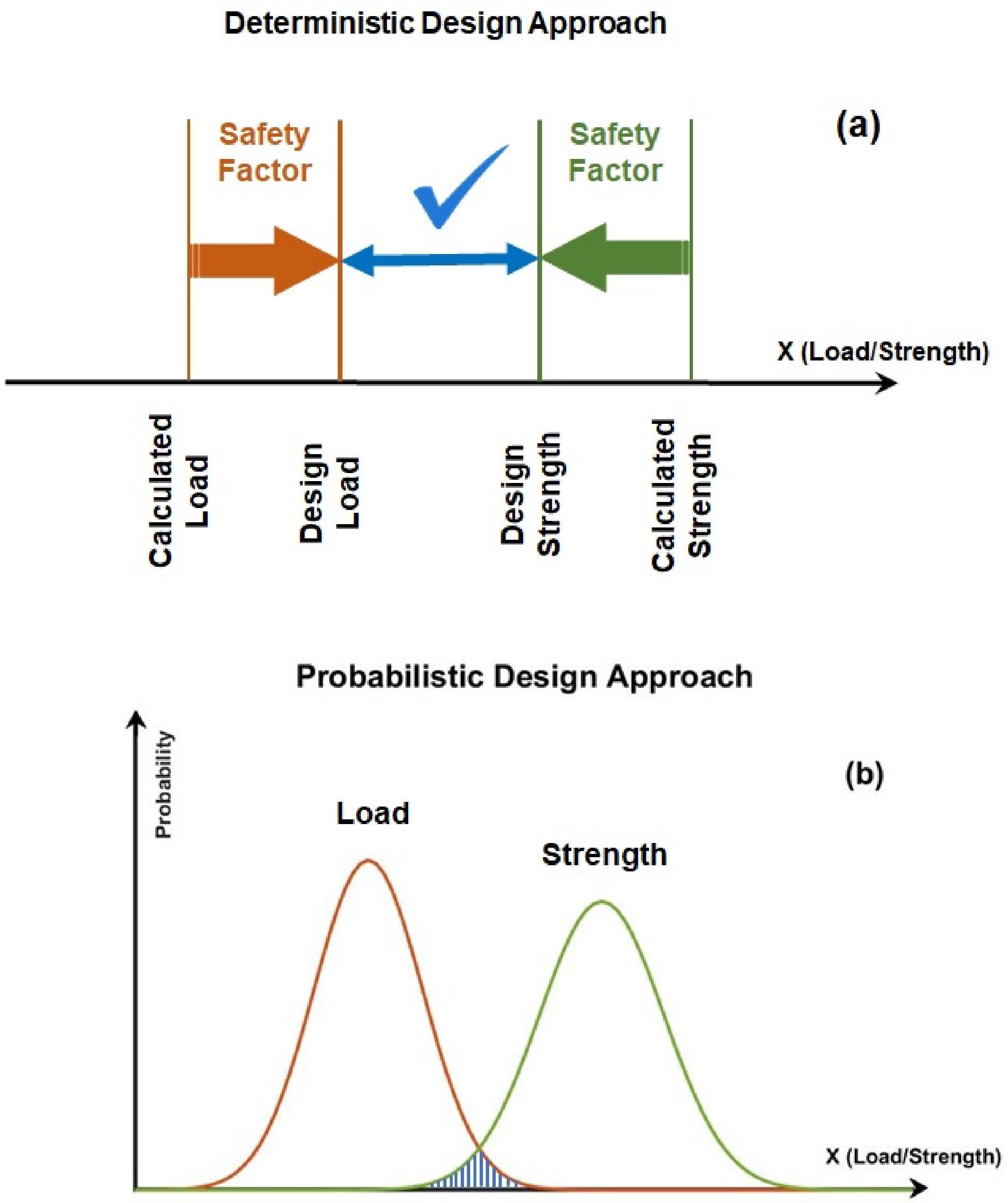


Figure 1.2: The comparison between the deterministic and probabilistic approaches in the determination of the safety factor in materials or structural design

epistemic uncertainty, respectively. It can be observed that these two sources of uncertainty can have different contributions (small or large) to the overall uncertainty. It should also be noted that the epistemic uncertainties shown in the figure can be identified and reduced by aiming the point symmetry of the center of the shots on target rather than the target center; in contrast, aleatoric uncertainties are measurable, but not reducible (see Section 6.3 for more explanations about aleatoric and epistemic sources of uncertainty).

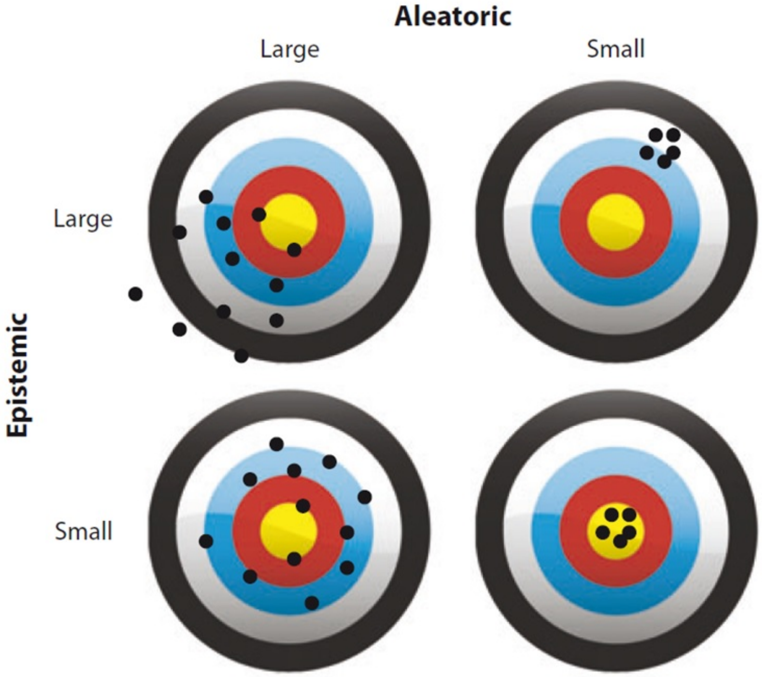


Figure 1.3: The analogy between precision/accuracy and aleatoric/epistemic uncertainty. Reprinted with permission from [2].

It is worth noting that the sources of uncertainty in computational modeling can be classified in more details as: 1) natural uncertainty (NU) that is again the the inherent random error in the system, 2) model parameter uncertainty (MPU) that is the error corresponding to the incomplete or inaccurate knowledge about the model parameters, 3) propagated uncertainty (PU) that refers to the error propagating in a chain of multi-scale models, and 4) model structure uncertainty (MSU) that results from the lack of knowledge about the underlying physics of the system, simplifications, inaccurate assumptions, and/or numerical errors [3].

The main goal of UQ in the computational modeling is to determine the uncertainties associated with models and their predictions. In this regard, the quantification process of the overall uncertainties reflected on model parameters and predictions are known as inverse uncertainty quantification and forward uncertainty propagation that will be referred as UQ and UP in this dissertation, respectively. The inverse UQ of the model parameters given data can be evaluated in two different statistical perspectives, i.e. Frequentist and Bayesian, which are explained in detail in Section 6.3.2. However, it can be briefly stated that the Frequentist view defines the probability as the rate of occurrence (frequency) of an event during a long-term observation or numerous replications of a particular measurement, while the Bayesian prospective describes the probability based on the degree of belief about the occurrence of an event which can be determined through the combination of the prior knowledge and any new data about the event. Based on these definitions, data is considered a random variable while the parameters are fixed (with noises) in the Frequentist methods. An opposite view is taken into account for the Bayesian approaches, i.e. fixed data coupled with parameters as random variables. The consideration of prior knowledge in the context of Bayes' theorem differentiates the Bayesian statistical view from its Frequentist counterpart. The presence of prior makes the Bayesian inference subjective that is the main criticism from the Frequentist point of view since wrong prior knowledge can be misleading in statistical inference. However, The true definition of prior probability distribution in Bayesian approaches can be beneficial in statistical inference. The focus of this dissertation is mostly on the Bayesian UQ in materials computational science. Among Bayesian approaches, Markov Chain Monte Carlo (MCMC) algorithms are the powerful techniques that are increasingly applied in UQ of different science and engineering problems nowadays due to the advent of faster computing capabilities [4]. In contrast to conventional numerical and analytical methods, MCMC sampling techniques can be more simple and robust in solving high-dimensional intractable integrals [5,6]. MCMC techniques also have the ability to sample from very complicated multi-variate distributions which is very difficult to do using the traditional sampling approaches, such as rejection or inversion sampling [6]. There are two commonly used MCMC approaches known as Gibbs sampling and Metropolis-Hastings whose similarities and differences are discussed in Sections 3.2 and 6.3.2.2. In this dissertation, a Metropolis-Hastings algorithm has been used for the probabilistic calibrations of a plastic flow behavior model for transformation induced plasticity (TRIP) steels, an Ni-Ti precipitation model in MatCalc©, a thermo-mechanical model for the strain-temperature hysteresis loop in Ni-Ti shape memory alloys (SMAs) and CALPHAD models against their corresponding experimental or calculated data. The details of the applied method has been explained thoroughly in Sections 2.3.3, 3.3.2.1, 4.5.1 and 5.4.1. UP from model parameters to outcomes or across a chain of multi-level models in the ICME framework is also an important task for UQ of model predictions, which can generally be performed through analytical or numerical approaches, which are discussed in Section 6.3.1. Briefly, analytical approaches take advantage of surrogate models to propagate the uncertainties, such as first order second moment (FOSM) and second order second moment (SOSM) of the approximated model through Taylor expansion (see Section 4.5.2), Polynomial Chaos Expansion (PCE) and Gaussian process (GP) regression (see Section 3.3.3). On the other hand, numerical approaches, e.g. Monte Carlo simulations, propagate the uncertainties through the model predictions at discrete sampled points in the parameter space. Generally, analytical approaches are cheaper and faster but not as precise as numerical approaches, which are often applied in the case of expensive models.

It is also important to state that the application of MCMC sampling approaches in inverse UQ is usually very costly, especially when a high-dimensional parameter space exists. Therefore, a systematic reduction in the number of the model parameters can considerably increase the efficiency of the MCMC sampling approaches. For this reason, a parameter sensitivity analysis (SA) can be performed before MCMC sampling in order to find the most influential model parameters whose variations have the highest impacts on the total variations of the model predictions [7]. The most common used approach for the process of finding the most influential model parameters is known as Design of experiments (DOE) that includes a description of experiments (e.g. complete/fractional factorial design (CFD/FFD)) followed by an analysis of the experiments (e.g., analysis Of variance (ANOVA)). CFD/FFD which was pioneered by Fisher [8] (see Section 4.4.1 for more detailed explanation) is used to consider the model predictions associated with all/fractional combinations of the levels assigned for the parameters in the subsequent ANOVA hypothesis testing for sensitivity analysis. ANOVA is a powerful variance based sensitivity analysis (VBSA) with no limitations in the number of parameters/factors, which performs the sensitivity testings based on the contributions of the individual parameters and their interactions to the total uncertainties of the model predictions (see Section 4.4.2 and Appendix A for more detailed explanation).

Another important discussion in this dissertation is how to fuse different sources of information (models) together to not lose any useful information about the system in the case that several competing models or models with different fidelities and costs are available to describe the same physical phenomenon. Generally, it is difficult and sometimes expensive to decide about which model should be applied in our design problems such that a reasonable balance between precision and cost is obtained. However, if the probabilistic information is appropriately gained from these models (considering the design budget) and smartly combined together, the fused model can probably provide more robust or accurate information for decision making in materials design compared to just using one of the models. The reason is the fact that all the physically reasonable models potentially have useful information to some extent which can contribute to better behavioral predictions of the given system [9–11]. It should be noted that the probabilistic information can be gained through various probabilistic calibration approaches, such as the mentioned MCMC methods or kriging/Gaussian process (GP) surrogate modeling in the presence of expensive physical models/simulations. Among information fusion approaches, Bayesian model averaging (BMA), co-kriging and error correlation based approaches are the focus of this dissertation which are briefly discussed here. Co-kriging approach was proposed by Kennedy and O'Hagan [12] in 2000, which is an extension of Gaussian Process surrogate modeling where a stochastic process is fitted to the existing data points based on a distance-based correlation between the errors associated with these data points. In this information fusion approach, any two consecutive models in fidelity order are correlated through a linear relationship. In this context, the response of the model with higher fidelity is estimated through the multiplication of a constant linear coefficient by the response of a GP fitted over the data points of the model with lower fidelity, plus a constant term (the error or discrepancy terms) that is defined by another GP surrogate model. It should be noted that the linear coefficient is optimized with the other hyper-parameters during the co-kriging process [3, 13]. In the end, the statistical combination of the two GP models results in a co-kriging model over the data points of the model with higher fidelity. In this combination, a co-variance matrix is constructed from the correlation between the errors of the low- and high-fidelity data points and their interactions which leads to a more precise surrogate model with less uncertainties compared to a GP surrogate model [12, 13]. More detailed explanations can be found in Section 3.3.3 about kriging and co-kriging surrogate modeling. BMA is a weighted average approach based on Bayes' factor proposed by Jeffreys [14] that shows the evidence in favor of each given hypothesis or model. In this approach, the weight of each model is the posterior of that model given the existing data, i.e., $P(M|D)$, which can be calculated using the Bayes' factor (see Section 5.4.2 for more details). Besides building a new model by taking the weighted average of all given models, BMA can provide larger confidence bounds around the average predictions which can be useful for the safety-critical applications where a highly conservative robust design is required. However, the models in BMA are assumed to be statistically independent from each other that is not always the case. Therefore, an information fusion approach that considers some sorts of correlations between models can be beneficial in the presence of correlated models. In this regard, Allaire and Willcox [10] developed an error correlation based model fusion (CMF) that constructs a fused model using the correlation between the discrepancies (errors) of models in the form of a co-variance matrix (see Section 5.4.3). The error correlations in this information fusion approach yields a fused model with less uncertainties (more precision) compared to any applied individual model.

In spite of the importance of UQ/UP in materials design under the ICME framework, there are very few UQ works in materials science and engineering so far which have been reviewed by Chernatynskiy et al. [2] in two different scales, i.e., UQ in electronic/atomic-level and multi-level simulations. These efforts are briefly described in Section 5.2. Few existing UQ works in literature specifically about thermo-mechnical modeling of SMAs are also mentioned in Section 4.2. In addition, the important efforts associated with UQ in CALPHAD are thoroughly reviewed in Chapter 6 and Section 5.2.

In this dissertation, a complete systematic analysis of UQ/UP is demonstrated using various computational models in materials science and engineering. Different information fusion approaches are also introduced to the community, which is very useful in most of the computational materials problems since there are usually different models/simulations with different fidelities and costs that try to explain the same physical phenomenon.

2. USING BAYESIAN FRAMEWORK TO CALIBRATE A PHYSICALLY BASED MODEL DESCRIBING STRAIN-STRESS BEHAVIOR OF TRIP STEELS [*]

2.1 Overview

Often times, the development of physical models of materials behavior is hindered not only by the incompleteness of the theoretical approach, but also by uncertainties that arise from limitations in the experimental observations used to validate and calibrate these models. In this work, we present a Bayesian framework for both the calibration of physical models as well as the quantification of likely uncertainty in experimental observations and apply it to a model for the plastic response of multi-phase Transformation Induced Plasticity (TRIP) steels. The model is based on a formulation of irreversible thermodynamics of plastic deformation and accounts for the presence of multiple phases through homogenization theories based on the iso-work approximation. Bayesian calibration through Metropolis-Hastings Markov Chain Monte Carlo has been used to calibrate a subset of the model parameters against experimental data sequentially and simultaneously. The calibrated parameters obtained from sequential training were in turn used to assess the uncertainty in a subset of experimental data-i.e. phase volume fractions-used as input to the models themselves. The viability of the calibration approach has also been examined using synthetic data generated from simultaneous calibrated model.

2.2 Introduction

The requirements for better manufacturability, higher safety, and lower fuel consumption in the automotive industry have led to the development of advanced materials, such as TRIP steels [15–18]. Generally, TRIP steels are considered as composite steels that contain different phases: ferrite as the matrix with bainite, retained austenite, and/or martensite inclusions [19,20]. The presence of retained austenite in the microstructure of these low alloy steels can give an enhanced combination

---

[*] Reprinted with permission from "Using Bayesian framework to calibrate a physically based model describing strain-stress behavior of TRIP steels" by Pejman Honarmandi, Raymundo Arroyave, 2017. *Computational Materials Science*, 129:66-81, Copyright 2017 by Elsevier B.V. or its licensors or contributors

of strength and fracture toughness. This results from high strain hardening that occurs during plastic deformation by strain-induced martensitic transformation (SIM) [21–24]. SIM behavior is directly related to the stability of the retained austenite and mainly influenced by its volume fraction, grain size, and carbon content [20].

Either to rationalize observations or to assist in the design of novel TRIP steel alloying and thermo-mechanical processing protocols it is necessary to develop physically motivated models that account for the effects of phase and chemical constitution on plastic deformation [25]. However, to date, work to address this challenge is rather limited [26] due to the modeling complexities that arise when considering the contribution of different phases to the overall mechanical response of these complex microstructures, in addition to phenomena associated with SIM.

In this regard, Bouquerel et al. [27] proposed a physical micromechanical method based on Mecking-Kocks theory for plastic flow behavior of each individual microstructural phase, and combined them using a Gladman-type mixture law. Moreover, they used a modified Olsen-Cohen approach to model SIM [27]. Perlade et al. [28] also suggested another physically-based model that considers the effects of morphology and chemical composition of each phase on stress-strain behavior, and subsequently homogenized the flow behavior of all the phases using the iso-work hypothesis. A generalized Olsen-Cohen was also applied to treat SIM in that work [28]. In some other related studies, dual or multi-phase microstructures with phase volume fraction evolutions have been assumed in the framework of constitutive models [26] or mean field approaches [25, 29] in order to predict stress-strain partitioning within microstructural phases. In addition, there is a lot of models for describing the mechanical behavior of structural materials undergoing SIM, pioneered by Zackay et al. [30], Olsen-Cohen [31], and Olson and Azrin [32]. Among recent studies, Tomita's constitutive equation in terms of applied stress, strain rate and temperature [33], Thibaud's phenomenological approach [34], and Han's kinetic model [18] based on martensitic variant selection through the Kurdjumov-Sachs orientation relationship are good examples of SIM modeling investigations.

The examples of physically motivated mechanical models provided above, as well as many others not discussed in this work for lack of space, are always subject to limitations inherent in the theoretical and simulation frameworks used, as well as in the uncertainty associated with the experimental observations used to parameterized and/or validate such models. Models are naturally limited in that it is usually the case that not all relevant physical phenomena are accounted for in a given formulation. Further limitations arise from the need to compromise between the need for facile implementation and execution of such models versus the need to accurately represent the microstructural features ultimately responsible for the mechanical behavior of the systems under study-i.e., the tension between analytical vs microstructure-sensitive approaches. Further limitations arise from limitations in the experimental observations themselves, as it is often the case that a considerable fraction of the inputs to the models are in turn determined by characterization techniques that are either limited in their accuracy/resolution or that are costly in resources.

Most approaches to model calibration (parameterization)/validation either do not account for uncertainties in the models and/or experiments and treat the exercise as a purely deterministic one. While such an approach provides valuable insight into the qualitative response of a physically-based model, the neglect of uncertainty quantification limits the power of the models as it is not possible to establish confidence bounds to the estimated response of the system. Furthermore, neglect of UQ makes it harder to determine in a quantitative manner the value derived from specific experimental observations either based on whether such observations can be used to reduce the model and/or experimental uncertainty. Additionally, neglect of UQ makes it difficult to deploy the models in materials design/optimization frameworks that account for uncertainty-e.g. robust design. Finally, we would like to note that UQ can be used to estimate values for physical features/parameters of a system that are difficult to determine experimentally.

In the context of UQ, the relevant uncertainties can result from different sources: (1) inherent uncertainty of a physical system, (2) model parameter uncertainty due to lack of knowledge/data about parameters, (3) propagated uncertainty in the case of a sequence of models, and (4) model structure uncertainty owing to the lack of physics in the models. Among these uncertainties, the quantification of the model uncertainty is usually the hardest one because of limited knowledge about the physical systems and inaccurate experimental data. This uncertainty quantification includes the model calibrations with data [3], which is often disregarded or performed superficially in most materials modeling.

It has been found [35] that the Bayesian approach based on Markov Chain Monte Carlo (MCMC) techniques is highly robust compared to other methods for parameter estimation. In fact, MCMC methods resolve the issue associated with the parameter calibration of complex non-linear models, which usually fail using traditional statistical methods [36]. In the case of multi-level models, Browne and Draper [37] indicated that the MCMC-based Bayesian approaches can provide better calibration compared to different likelihood-based methods, i.e. maximum likelihood (ML), marginal quasi-likelihood (MQL), penalized quasi-likelihood (PQL), and restricted maximum likelihood (REML).

In the current computational work, Rivera [38] and Haidemenopoulos [39] models have been applied to simulate the plastic flow behavior of multiple phases and SIM within dispersed retained austenite, respectively, to account for the mechanical deformation of multi-phase TRIP steel. In the first model, the dislocation density evolution for any individual phase is determined using a thermo-statistical theory of plasticity, and overall plastic flow behavior has subsequently been estimated through an iso-work approximation. This approximation realizes the coupling among multiple phases by assuming that the work of deformation is equally partitioned among all phases in the microstructure at any strain interval. The model that accounts for SIM in turn can be considered as a modified Olson-Cohen model for the distribution of nucleation site potency, which predicts the SIM behavior in terms of austenite particle size, austenite chemical composition, and stress triaxiality. These models are applied as case studies for the introduction of Bayesian parameter calibration based on Metropolis-Hastings MCMC algorithm, which is our main direction in this paper.

2.3 Models and Methodology

2.3.1 Microstructural evolution during plastic deformation

In this model, shear stress for each microstructural phase (i) is defined in terms of the contribution of the following mechanisms [40]:

$$\tau_i = \tau_0 + \tau_s + \tau_b + \sqrt{\tau_{in}^2 + \tau_p^2} \tag{2.1}$$

where τ_0, τ_s, τ_b, τ_{in}, and τ_p are Peierls force, solid solution strengthening, long-range back stress, dislocation strengthening, and precipitation strengthening contributions within phase i, respectively. Here, τ_0 and τ_s are dependent on chemical compositions of each phase through the existing empirical equations [41–43]. The following relationships have also been proposed for τ_b and τ_p [40]:

$$\tau_b = \frac{\mu b}{D} n = \frac{\mu b}{D} n_{max} \Big(1 - exp(-\frac{\lambda}{b n_{max}} \gamma)\Big) \tag{2.2}$$

$$\tau_p \cong 0.6 \mu b \frac{\sqrt{f_\nu}}{R} \tag{2.3}$$

In which l is the shear modulus, b is the Burger's vector magnitude, D is the grain size, n is the evolution of piled-up dislocations at the grain boundaries in terms of shear strain (γ), n^* is the maximum possible number of piled-up dislocations, λ is the average spacing between slip planes at the grain boundaries, f_ν is the precipitate, volume fraction, and R is the mean radius of precipitates.

τ_{in} is defined by Rivera's model [38] based on dislocation density evolution of individual phases during plastic deformation which is described as the difference between the dislocation generation and annihilation rates according to Kocks and Mecking theory [44]:

$$\frac{d\rho(\gamma)}{d\gamma} = \frac{d\rho^+(\gamma)}{d\gamma} - \frac{d\rho^-(\gamma)}{d\gamma} \tag{2.4}$$

The dislocation storage rate is proportional to the mean free path for dislocations which incorporates the effects of dislocation-dislocation (Λ_{pure}^{-1}) and solute atom-dislocation interactions, including substitutional (Λ_{subs}^{-1}) and interstitial solute atoms (Λ_{inter}^{-1}) effects:

$$\frac{d\rho^{+}(\gamma)}{d\gamma} = \frac{1}{b}\left(\frac{1}{\Lambda_{pure}} + \frac{1}{\Lambda_{subs}} + \frac{1}{\Lambda_{inter}}\right) \tag{2.5}$$

$\Lambda_{pure}^{-1} = K_1\sqrt{\rho}$, $\Lambda_{subs}^{-1} = K_{ss}\sqrt{\rho}$, and $\Lambda_{inter}^{-1} = x_c^{1/3}\Lambda_{pure}^{-1}$. μ_0 is the shear modulus at 0 K, x_i is the atomic fraction of element i, and ΔG_{sys} is the free energy of the applied alloy [44].

The dislocation annihilation rate due to dynamic recovery is also expressed as follows:

$$\frac{d\rho^{-}(\gamma)}{d\gamma} = f_{DRV}\rho(\gamma) \tag{2.6}$$

where f_{DRV} is the dynamic recovery coefficient. This can be determined by considering the dislocation annihilation energy barrier ΔG, which contains the formation energy associated with the annihilating dislocation segment with the length of l, the migration energy of this segment during slip, the contribution of vacancy energy for the dislocation annihilation (especially at high temperatures), the energy term associated with dislocation-substitutional solute atom interaction, and the contribution of statistical entropy that takes account of the number of possible dislocation migration paths [38, 44], as follows:

$$\begin{aligned}\Delta G(l) &= \frac{b}{l}(U_{form} + U_{mig} + U_{vac} + U_{ss} - T\Delta S)\\ &= \frac{b}{l}\Bigg[\frac{1}{2}\mu b^2 l + \sigma_Y A_{act} b^2 l + \frac{V_{sys}}{b^3}\delta(T)(E_f + k_B T ln(c_m))\\ &+ V_{sys}\frac{1}{b^3}\sum_i x_i \Delta G_{sys} - k_B T ln(\Omega)\Bigg]\end{aligned} \tag{2.7}$$

where σ_Y , A_{act} , $\delta(T)$, and E_f are the yield stress, the activation area associated with dislocation cross slip, a continuum piece-wise function in terms of temperature, and the vacancy formation energy, respectively. $V_{sys} = bl^{*}l$ is the annihilating system volume per dislocation, which is proportional to the length of dislocation distortion field over the slip plane, l^{*}. This parameter is approximated by $12.5b$. $c_m = exp(\frac{-\Delta S_{form}}{R})exp(\frac{-E_f}{RT_m})$ is the vacancy concentration at melting temperature, where ΔS_{form} is the vacancy formation entropy. With regards to the entropy term, dislocation slip and climb are considered to calculate the total number of micro-states in the microstructure. In pure materials, $\Omega_{pure} = \Omega_{dis} + \Omega_{\nu-d} = \left(\frac{C}{\langle\nu\rangle} + \frac{C_{int}}{\langle\nu\rangle}\right)^{\mathcal{N}} = \left(\frac{\dot{\varepsilon}_0}{\dot{\varepsilon}} + \frac{\vartheta}{\dot{\varepsilon}}\right)^{\mathcal{N}}$ in which C, C_{int}, $\langle\nu\rangle$, $\dot{\varepsilon}$, and $\dot{\varepsilon}_0$ are the sound speed in the material, the effect of vacancies on dislocation annihilation velocity, dislocation average velocity during annihilation process, axial strain rate, and the maximum possible strain rate in the material, respectively. It should be noted that $C_{int} = \frac{\vartheta}{b\rho_Y}$ and $\dot{\varepsilon}_0 = Cb\rho_Y$ where $\rho_Y = \left(\frac{\sigma_Y}{0.9\mu b}\right)^2$ is the dislocation density corresponding to the yield point. Moreover, $\mathcal{N}$ is associated with the impingement effect resulting from the overlap of contiguous dislocations strain fields, and $\vartheta = \vartheta_D exp(-\frac{E_m}{RT})$ is the frequency for vacancy migration in which $\vartheta_D = 10^{13}s^{-1}$ is Debye frequency and E_m is the vacancy migration energy. Considering the effect of solid solution solute atoms, the above-mentioned equation for the total number of micro-states can be extended to $\Omega = \Omega_{pure}.\Omega_{subs}.\Omega_{inter}$. It is found that $\Omega_{subs} = (\Omega_{pure})^{N_{subs}} = \left(\frac{\dot{\varepsilon}_0+\vartheta}{\dot{\varepsilon}}\right)^{N_{subs}}$ and $\Omega_{subs} = (\Omega_{pure})^{x_C^{1/3}(1+N_{subs})} = \left(\frac{\dot{\varepsilon}_0+\vartheta}{\dot{\varepsilon}}\right)^{x_C^{1/3}(1+N_{subs})}$, where $N_{subs} = \sum_i x_i$. Therefore, $\Omega = \left(\frac{\dot{\varepsilon}_0+\vartheta}{\dot{\varepsilon}}\right)^{\mathcal{N}(1+x_C^{1/3})(1+N_{subs})}$ [38].

Dislocation average velocity during annihilation process can also be expressed in Arrhenius form in terms of the annihilation barrier as follows:

$$\left(\frac{\langle\nu\rangle}{C + C_{int}}\right)^{\mathcal{N}(1+x_C^{1/3})} = \left(\frac{\dot{\varepsilon}}{\dot{\varepsilon}_0 + \vartheta}\right)^{\mathcal{N}(1+x_C^{1/3})} = exp\left(-\frac{\langle\Delta G\rangle}{k_B T}\right) \tag{2.8}$$

The average length of annihilating dislocation segments that is obtained by substituting Equation 2.7 in Equation 2.8 is applied to determine the dynamic recovery coefficient:

$$\begin{aligned} f_{DRV} &= \frac{N_A}{w_a}\rho_a V_{sys} = \frac{N_A}{w_a}\rho_a b l^* l \\ &= \frac{\frac{N_A}{w_a}\rho_a b^2 l^*(1 + N_{subs})(1 + x_C^{1/3})k_B T ln\left(\frac{\dot{\varepsilon}_0+\vartheta}{\dot{\varepsilon}}\right)^{\mathcal{N}}}{\frac{1}{2}\mu b^3 + \sigma_Y A_{act} b^3 + \frac{l^*}{b}\delta(T)(E_f + k_B T ln(c_m)) + \frac{l^*}{b}\sum_i x_i \Delta G_{sys} - (1 + x_C^{1/3})k_B T ln\left(\frac{\dot{\varepsilon}_0+\vartheta}{\dot{\varepsilon}}\right)^{\mathcal{N}}} \end{aligned} \tag{2.9}$$

N_A, w_a, and ρ_a are the Avogadro's number, the atomic weight of alloy, and the alloy density, respectively [38].

A formulation in the form of Kocks and Mecking equation is employed to predict the dislocation density evolution of microstructural phases during plastic deformation:

$$\frac{d\rho_i(\gamma_i)}{d\gamma_i} = A(\frac{K}{b})\sqrt{\rho_i(\gamma_i)} - Bf_{DRV}\rho_i(\gamma_i) \tag{2.10}$$

$K = (1 + x_c^{1/3})K_1 + K_{ss}$. A and B are structure dependent constants that are considered as 1 for FCC phase (austenite), while they can be calculated for BCC (ferrite/bainite) and BCT phase (martensite) through the ratio of the number of slip system times coordination number in these structures to their counterpart in FCC phase [45], as shown in Table 2.1.

Table 2.1: Phase dependent coefficients in Kocks and Mecking formulation

Phase	Structure	Slip System	Number of Slip Systems	Coordination Number	A	B
Austenite	FCC	$\{111\}\langle\bar{1}10\rangle$	12	12	1	1
Ferrite/Bainite	BCC	$\{110\}\langle111\rangle$	12	8	$\frac{3}{8}$	$\frac{8}{3}$
		$\{112\}\langle111\rangle$	12			
		$\{123\}\langle111\rangle$	24			
Martensite	BCT	$\{101\}\langle\bar{1}01\rangle$	2	8	$\frac{9}{2}$	$\frac{2}{9}$
		$\{\bar{1}01\}\langle101\rangle$	2			

The dislocation density in each individual phase can be obtained as a function of shear strain by solving Equation 2.10. According to Taylor relation, shear stress during plastic deformation can also be obtained in terms of shear strain for each microstructural phase as follows:

$$\tau_i = \alpha M\mu b\sqrt{\rho_i(\gamma_i)} \tag{2.11}$$

where a is a constant and M is the Taylor factor [46].

To obtain the stress-strain curve, the iso-work approximation has been taken into account to homogenize the flow behavior in the entire multi-phase structure, which is achieved by assuming that every microstructural phase has the same energy dissipation during deformation [40]:

$$\tau_i d\gamma_i = \tau_{i+1} d\gamma_{i+1} \tag{2.12}$$

where i and $i+1$ correspond to different microstructural phases.

The values of the model parameters in the literature have been listed for different microstructural phases in Table 2.2.

Table 2.2: Model parameters in different microstructural phases. Adapted from [40, 45, 47].

Parameter	Phase			
	Ferrite	Austenite	Bainite	Martensite
$\mu (MPa)$	75,200			
$\mu_0 (MPa)$	85,335			
$b(m)$	2.5E-10			
$D(m)$	20E-6	15E-6	Artificial neural network [48]	0.13E-6
n^*	4	20	4	4
$\lambda(m)$	1.5E-7	1.5E-7	1.5E-8	5.0E-8
$R(m)$	4.5E-9	50.0E-9	1.5E-8	5.0E-8
α	0.25			
M	3.06			
$\tau_0 (MPa)$	18.0			
$\rho_0 (m^{-2})$	1E-13	1E-13	$(0.7 + 3.5w_c) \times 10^{15}$	$10^{9.284 + \frac{6880}{T_{BIT} - 298} - \frac{178060}{(T_{BIT} - 298)^2}}$ [49]
$\mathcal{N}$	2	1	2	2
$\lambda_{ss} (m)$	16,000			
$\delta (T)$	$\frac{T - T_0}{T_f - T_0}$			
$\chi (\frac{J}{m^2})$	0.077			
$l^* (m)$	31.25E-10			
$E_f (eV)$	1.60			
$E_m (eV)$	1.28			
$\Delta S_{form} (\frac{J}{mol \quad K})$	-12.47			
ν	0.27			
$\Delta G_{sys} (\frac{kJ}{mol})$	$16 - 0.45T - 1.4 \times 10^{-5}$			
$\sigma_Y (MPa)$	475	720	475	2,000

2.3.2 Strain-induced martensitic transformation model

The martensitic transformation response in TRIP steels during deformation is described by Haidemenopoulos model [39] based on the Olson-Cohen theory for heterogeneous martensitic nucleation. In this model, a composite microstructure of dispersed retained austenite particles in a ferritic matrix is assumed in which martensitic nucleus is formed in austenite phase owing to the dissociation of existing microstructural defects. Generally, there are two types of mechanically-induced martensitic nucleation, including stress-assisted nucleation on traditional pre-existing sites (or defects) and strain-induced nucleation on sites produced during plastic deformation. These sites require a sufficient potency for instantaneous nucleation of martensitic phase. The potency of nucleation sites are determined by the number of crystal planes n_{crys} in the fault formed from their dissociation. Therefore, the sufficient potency corresponding to a critical number of crystal planes n^*_{crys} can be calculated when the formation energy of the n_{crys}-plane fault $\gamma_f(n_{crys})$ becomes zero:

$$\gamma_f(n_{crys}) = n_{crys}\rho(\Delta G_{ch} + E_{str} + W_f) + 2\gamma_s \tag{2.13}$$

$$\gamma_f(n_{crys}) = 0 \rightarrow n^*_{crys} = -\frac{2\gamma_s}{\rho(\Delta G_{ch} + E_{str} + W_f)} \tag{2.14}$$

where ρ, γ_s, ΔG_{ch}, E_{str} , and W_f are the atom density of the fault plane, the interfacial energy between fault and matrix, the chemical driving force of martensitic nucleation, the elastic energy at fault/matrix interface, and the frictional work resulting from the movement of fault/matrix interface, respectively.

According to the above-mentioned nucleation mechanisms, the total number of nucleation sites with sufficient potency per unit austenite volume N_ν is the summation of operational stress-assisted nucleation sites N^σ_ν and operational strain-induced nucleation sites N^ε_ν, which are defined as follows:

$$N_\nu(\varepsilon) = N^\sigma_\nu + N^\varepsilon_\nu(\varepsilon) \tag{2.15}$$

$$N_\nu^\sigma = N_\nu^{\sigma 0} exp(-a_\sigma n^*) \tag{2.16}$$

$$N_\nu^\varepsilon(\varepsilon) = N_\nu^{\varepsilon 0}(\varepsilon) exp(-a_\varepsilon n^*) = N\big(1 - exp(-k\varepsilon^m)\big) exp(-a_\varepsilon n^*) \tag{2.17}$$

where k and m are constants, N is the maximum possible number of strain-induced sites, a_σ and a_ε are the constants in the above exponential functions called shape factors of stress-modified and strain-modified potency distribution, respectively. It is assumed that the nucleation sites formed during plastic deformation possess higher potency compared to the pre-existing nucleation sites, which corresponds to a lower value of a_ε.

Accordingly, the evolution of martensite volume fraction (f) in terms of plastic strain can be written as:

$$f(\varepsilon) = 1 - exp\big(-\nu_p N_\nu(\varepsilon)\big) \tag{2.18}$$

It is worth noting that the effect of austenite particle size is also taken into account in this model. In Table 2.3, the parameter values for the applied SIM model have been indicated [39].

2.3.3 Metropolis Hastings algorithm

Metropolis-Hastings algorithm is one of the MCMC techniques to randomly sample from a multi-dimensional probability distribution. It is usually applied to obtain the statistical characteristics in the framework of Bayes' theorem for a given set of variables. Since solving the intractable integrals associated with the application of Bayesian framework for model calibration through analytical or numerical approaches is very hard in the high-dimensional cases, MCMC sampling from a joint posterior probability distribution is considered as a good solution [6].

In this approach, experimental data and prior knowledge including the initial values of parameters and their range obtained from literature and/or prior deterministic optimizations are fed to the Matlab MCMC toolbox [50]. In this work, least square optimization techniques have been em-

Table 2.3: SIM Model parameters. Adapted from [39].

Parameter	Value
$v_p(m^3)$	4.18E-18
$\gamma_s(\frac{J}{m^2})$	0.15
$\Delta G_{ch}(\frac{J}{mol})$	-3,285
$E_{str}(\frac{J}{mol})$	500
$W_f(\frac{J}{mol})$	$1.893E3 \times X_{Mn}^{2/3} + 1.310E4 \times X_C^{2/3}$
$\rho(\frac{mol}{m^2})$	3E-5
$N_{\nu}^{\sigma 0}(m^-3)$	1.5E17
$N(m^-3)$	1.9E19
a_σ	0.1
a_ε	0.03
k	46
m	3.45

ployed to suggest the beginning point in the parameter space (θ_0) in order to improve the MCMC efficiency for parameter convergence to their plausible mean values, which results in a lower computational cost. Non-informative uniform probability density functions (PDFs) have been also considered as parameter prior distributions since no statistical information has been found for the given model parameters. MCMC toolbox uses random walk in multi-dimensional parameter space to generate n parameter samples. A new sample or candidate has been drawn from an arbitrary proposal distribution q. In our research, an adaptive Gaussian distribution is considered as the proposal density which is centered at the initial or the last accepted parameter with an adaptive variance-covariance. In each generation step, the adaption takes place using the variance-covariance of the previously generated samples based on Haario et al. work [51]. Acceptance/rejection of new sampled candidate is performed by calculating the Metropolis-Hastings ratio:

$$MH = \frac{\pi(\theta^{cand})}{\pi(\theta^i)} \frac{q(\theta^i|\theta^{cand})}{q(\theta^{cand}|\theta^i)} = \frac{P(\theta^{cand})P(D|\theta^{cand})q(\theta^i|\theta^{cand})}{P(\theta^i)P(D|\theta^i)q(\theta^{cand}|\theta^i)} \tag{2.19}$$

where $\frac{\pi(\theta^{cand})}{\pi(\theta^i)}$ is the Metropolis ratio that is the joint probability distribution for the new candidate over the previous one, and $\frac{q(\theta^i|\theta^{cand})}{q(\theta^{cand}|\theta^i)}$ is the ratio of proposal density in two different directions, i.e., jump from θ^i to θ^{cand} and vice versa. The new candidate can be accepted with the minimum probability between the above ratio and 1. It should be noted that the previous parameter vector is used as the new sample in the case of rejection. At the end of this process, n samples of parameter vectors $\left\{\theta^{(0)}, ..., \theta^{(n)}\right\}$ are produced, which show the posterior PDFs of given parameters. It is worth noting that the likelihood function is defined as the difference of the model result at θ and corresponding data values ($Modelfun(\theta) - D$). In this work, the likelihood function is considered as a Gaussian function that is centered at data; its variance (error of data) is regarded as a hyper-parameter and updated in the framework of Bayes' theorem during MCMC sampling. More details can be found in Gelman et al. [52]. Flow-chart of this Approach has been shown in Figure 2.1.

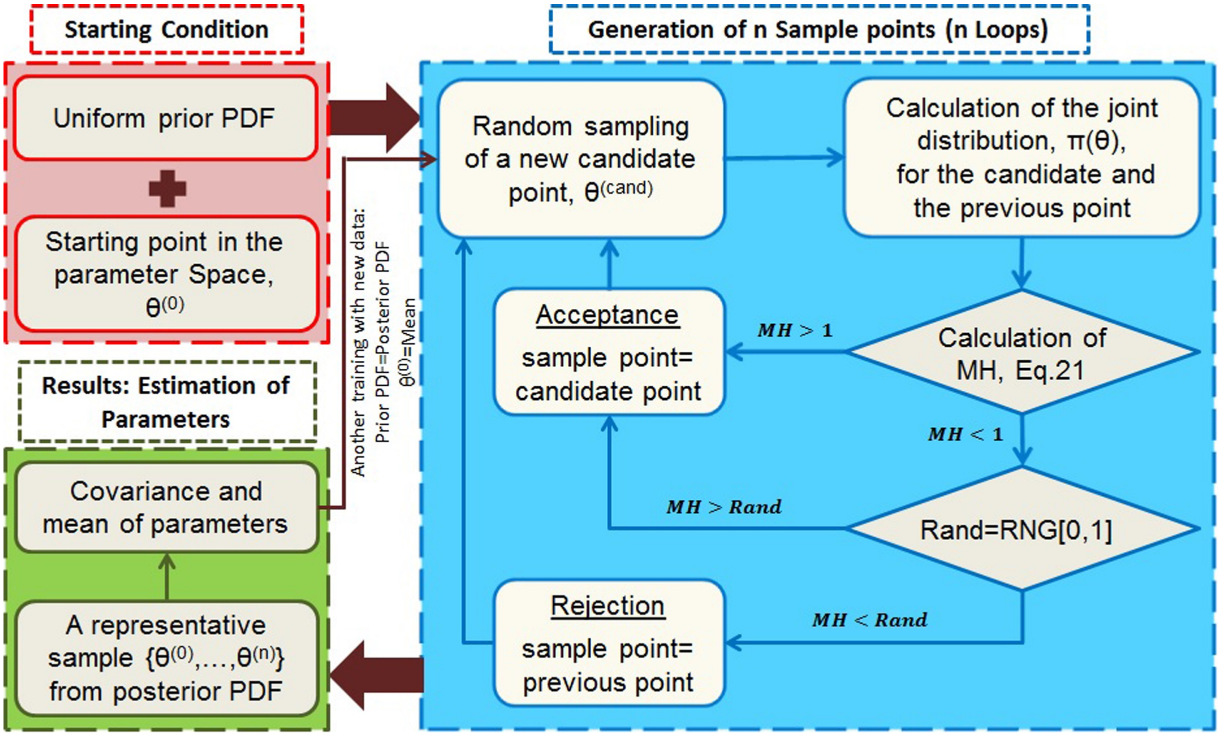


Figure 2.1: Flow-chart of MCMC Metropolis-Hastings method (RNG: Random Number Generation)

2.3.4 Propagation of uncertainty

The mean and covariance of generated samples in MCMC toolbox are used to calculate the uncertainty of the model result, which is the uncertainty of deformation stress in terms of plastic strain in our work. This can be performed through propagation of uncertainty as follows:

$$\sigma_S^2 \approx \sum_{i=1}^{N} |\frac{\partial M}{\partial a_i}|^2 \sigma_i^2 + \sum_{i=1}^{N}\sum_{j\neq i}^{N} \frac{\partial M}{\partial a_i}\frac{\partial M}{\partial a_j}\sigma_{ij} \tag{2.20}$$

where M, $a_{i/j}$, σ_i^2, and σ_{ij} are the model function, applied model parameters, parameter variances that is the diagonal elements of variance-covariance matrix, and the covariance of any two parameters that is the off-diagonal elements of variance-covariance matrix, respectively.

The partial derivatives with respect to the parameters have numerically been calculated using the following method ($\Delta a_i = 10^{-5}$) [53]:

$$\frac{\partial M(a_1, a_2, ..., a_i, ...)}{\partial \theta_i} \approx \frac{M(a_1, a_2, ..., a_i + \frac{\Delta a_i}{2}, ...) - M(a_1, a_2, ..., a_i - \frac{\Delta a_i}{2}, ...)}{\Delta a_i} \tag{2.21}$$

2.3.5 Applied experimental data

Experimental data reported by Jacques et al. [54] for stress-strain behavior in multi-phase TRIP and dual phase steels has been taken into account for the calibration of the models' parameters. Nominal chemical composition, volume fraction of phases, carbon weight percentage of phases, bainitic isothermal transformation (BIT) temperature, tensile deformation temperature (RT) and strain rate (6.67E-4 s^{-1}) of these experiments are used as inputs for our models, which are listed in Table 2.4. It should be noted that inter-critical annealing (IA) temperature and time besides BIT time are not directly among the model inputs, but their effects will be considered through phase volume fractions and carbon contents. For instance, in the case of BIT temperature at 370 °C, prolonged BIT time results in the formation of dual phase ferrite-bainite microstructure, whereas short BIT time of 0.5, 1, and 3 min for TRIP 4, 5, and 6 causes the formation of multi-phase microstructures, which have austenite and/or martensite besides bainite and ferrite.

Table 2.4: Applied Experimental information as model inputs. Adapted from [54].

Alloy	Chemical Composition (wt%)						Volume Fraction (%)				Carbon Content (wt%)				BIT (°C)
	C	Si	Mn	Al	Ni	Cr	Fer	Aus	Bai	Mar	Fer	Aus	Bai	Mar	
TRIP 1	0.13	1.5	1.42	0.027	0.02	0.013	75	8	15	0	0.014	0.93	0.3	0	410
TRIP 2							60	7.9	30	0	0.01	0.85	0.2	0	360
TRIP 3							60	7.8	30	0	0.01	0.97	0.16	0	410
TRIP 4	0.16	0.38	1.3	0.03	0.027	0.019	75	8.8	9	7	0.02	0.61	0.3	0.97	370
TRIP 5							75	8.1	12	5	0.02	0.68	0.3	1.1	
TRIP 6							75	6.1	19	0	0.02	0.73	0.53	0	
DP							75	0	25	0	0.02	0	0.58	0	

*Fer: ferrite, Bai: bainite, Aus: retained austenite, Mar: martensite, and DP: Dual Phase.

2.4 Results and Discussion

2.4.1 Forward analysis of the models

Using the parameter values in Tables 2.2 and 2.3, stress-strain curves have been estimated in Figure 2.2 based on different conditions mentioned in Table 2.4. These plots show that there are significant discrepancies between model results and experimental data, except for dual phase case. Regardless of inherent uncertainty, these disagreements can be due to three different reasons as mentioned before: (1) the uncertainty in experiments corresponding to the error of tensile test and microstructural characterization, (2) the uncertainty in the models due to their lack of physics, and (3) uncalibrated model parameters.

In order to evaluate the uncertainty of the models, the calibration of the model parameters is a necessary task. Since MCMC approach is an expensive method for model calibration, the most sensitive model parameters are selected using model forward analysis. This has been performed by changing the parameters one by one to identify the sensitivity of model parameters, some of which are shown in Figure 2.3 for TRIP 4 as examples. In Figures 2.3(a) and (b), high sensitivity of Taylor constant (alpha) and Peierls stress τ_0 (Stress-Fe) is noticeable. It is worth noting that

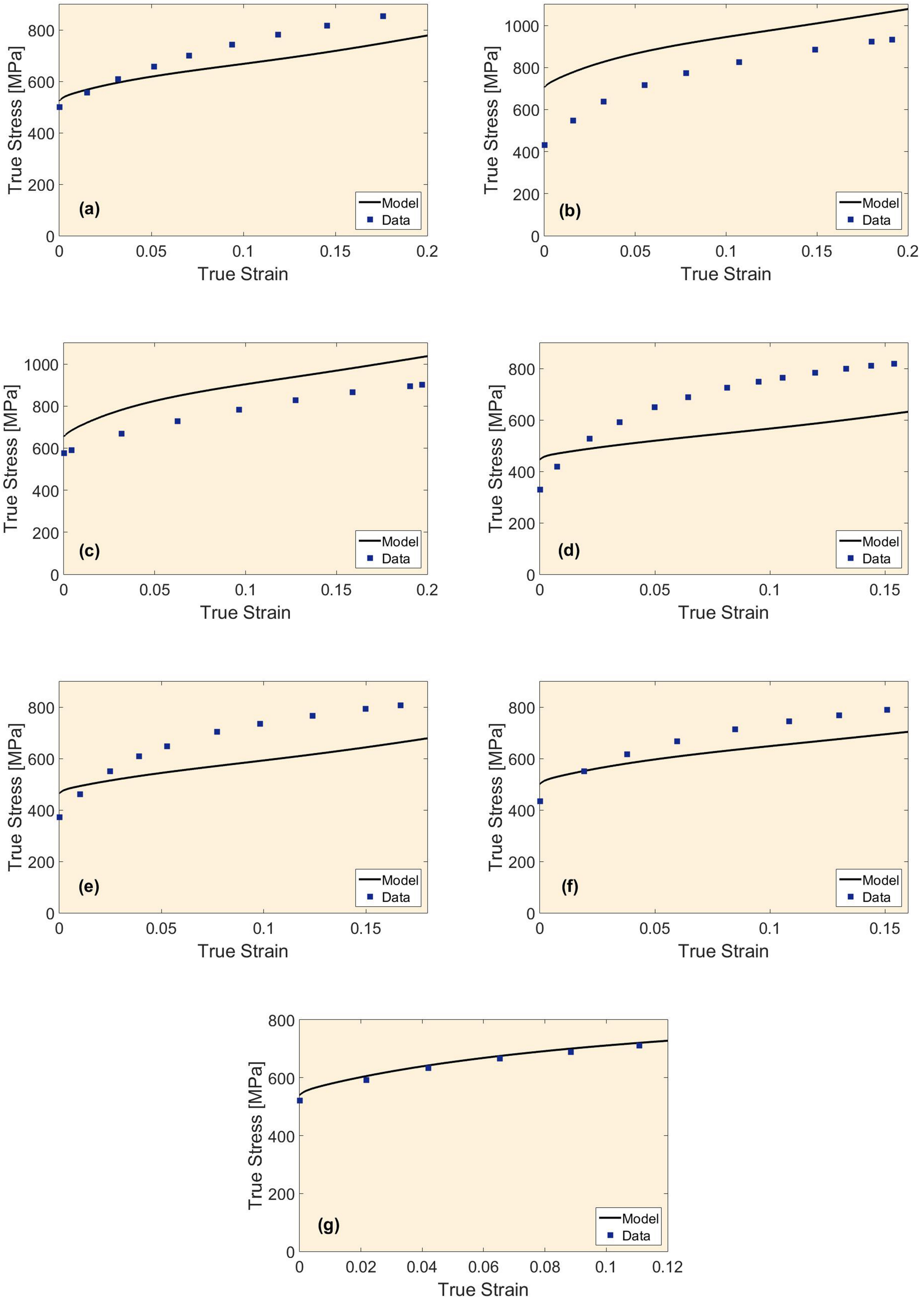


Figure 2.2: (a-g) comparison of experimental data with model results for plastic flow behavior in TRIP (1-6) and DP

these two parameters mainly affect the initiation of the model response and the curvature remains unchanged. However, the constants in SIM model, i.e., m-SIM and K-SIM, and n^* in Equation 2.2 for bainite, ferrite and martensite (n-max(BFM)) change the curvature of the model results without any variation at initial point, as shown in Figures 2.3(c-e). In other words, these parameters influence the strain hardening response of the model which is much more significant in the case of m-SIM. The effect of the variation of n^* in austenite (n-max(A)) has also been brought in Figure 2.3(f) in order to show one of the insensitive model parameters that has been disregarded for calibrations.

After the forward analysis, the eight most sensitive model parameters have been chosen for calibration, which include four parameters from each model, i.e., microstructural evolution model and SIM model.

2.4.2 Sequential and simultaneous calibration of model parameters

The model has been sequentially calibrated for the selected parameters with three different experimental data, including TRIP 2, 1, and 5 in turn. As mentioned in Section 2.3, the parameter posteriors obtained from previous calibration are considered as the parameter prior for next calibration in this approach. The three experimental data are applied all at once during the simultaneous calibration as well. Then, the sequentially/simultaneously calibrated parameters are used to obtain strain-stress curves under other experimental conditions in order to test the validity of the model. In each one of the sequential or simultaneous calibrations, 10,000 samples are generated. The mean and covariance of the samples can subsequently be introduced as calibrated parameters and their uncertainties, as listed in Table 2.5. Histograms of parameter posterior PDFs have also been plotted for both calibration methods in Figures 2.4 and 2.5. In the case of sequential calibration (Figure 2.4), m-SIM, alpha, n-max, and shear modulus (μ) contain a clear peak in their posterior distribution unlike a_σ (a-stress), a_ε (a-strain), K-SIM, and Stress-Fe whose histograms show multiple peaks suggesting equi-finality problem, which is attributed to the presence of too many parameters for calibration (over-parameterization) and/or the lack of data to constrain the parameter values [55]. Simultaneous calibration also results in similar trends for parameter poste-

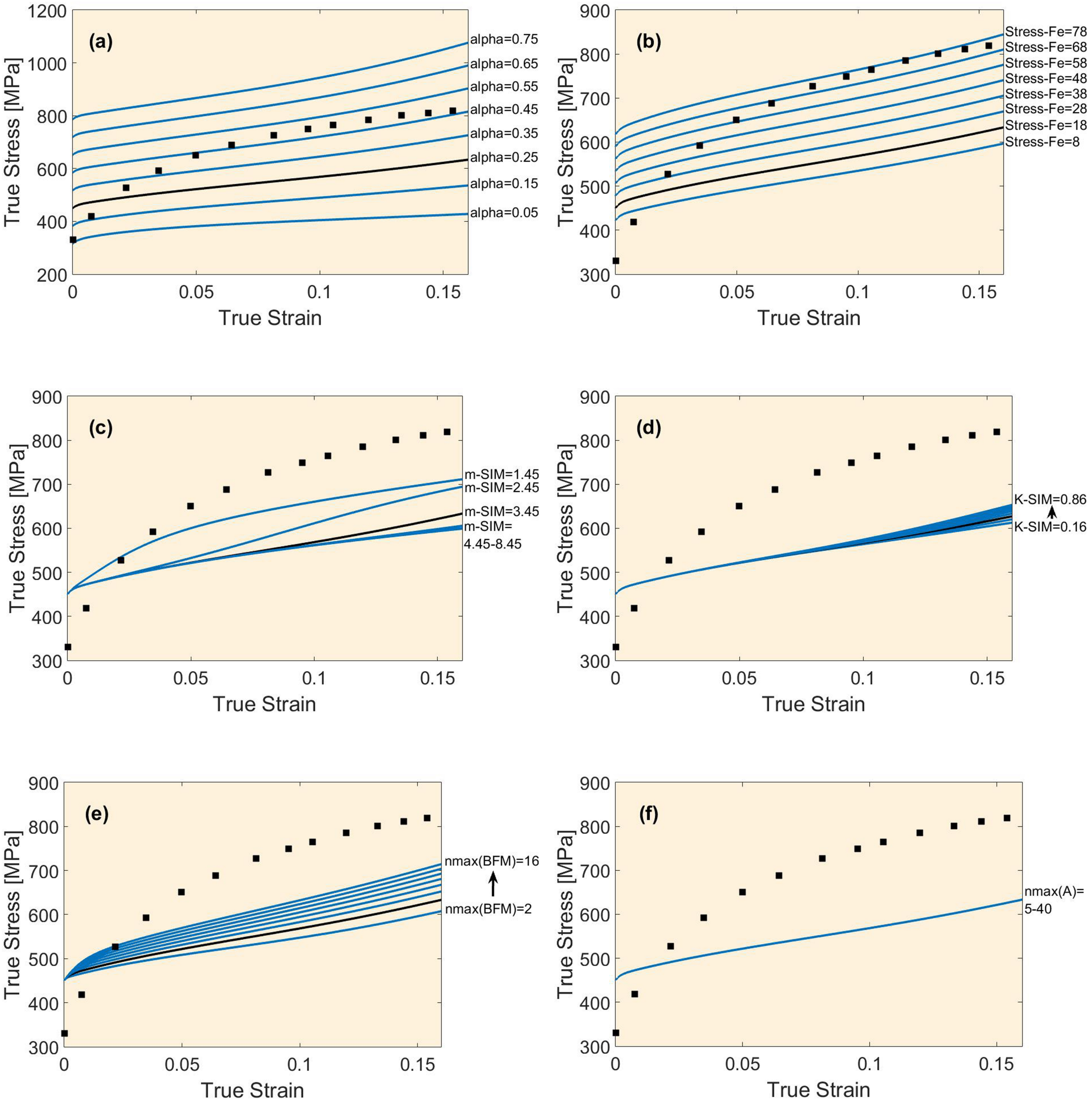


Figure 2.3: Model forward analysis based on data corresponding to TRIP 4 for evaluation of the sensitivity of model parameters

rior distributions as shown in Figure 2.5; however, the posterior PDFs are more defined for K-SIM and Peierls stress, and the distribution peak of alpha, n-max and shear modulus shift to lower values compared to their counterparts in sequential calibration. The latter is also confirmed by lower mean values for alpha, n-max and shear modulus in Table 2.5. It should be noted that the significant narrow peak in shear modulus posterior distribution is attributed to the acceptance of lots of candidates between 75,200 and 75,300 MPa in this case.

Table 2.5: Applied parameter values and their uncertainties after sequential and simultaneous calibrations compared to initial values before calibrations

Parameters	4 SIMT model parameters				4 strain-stress model parameters			
	a_σ	a_ε	k	m	α	$n-max$	μ	σ_p
Before Calibration	0.1	0.03	46	3.45	0.25	4	75,200	18
After Sequential Calibration	0.28 ± 0.125	0.056 ± 0.026	50.8 ± 16.54	1.77 ± 0.51	0.256 ± 0.039	6.31 ± 1.52	76,600 ± 2612	20.4 ± 5.85
After Simultaneous Calibration	0.27 ± 0.13	0.051 ± 0.025	49.55 ± 15.2	1.92 ± 0.74	0.176 ± 0.036	4.55 ± 1.80	75,155 ± 1502	23.41 ± 5.20

The parameter posterior distributions can also reflect the correlation between parameters that are missed in deterministic approaches for parameter calibrations such as least squares. Two dimensional graphs can demonstrate the correlation of pair parameters (their joint probability distributions), some of which is shown in Figures 2.6 and 2.7 for sequential and simultaneous calibration, respectively. The Pearson correlation coefficient is applied to determine the linear correlation between each two model parameters using the variance-covariance matrix obtained from MCMC sampling as follows [56]:

$$\rho_{x,y} = Corr(x,y) = \frac{cov(x,y)}{\sigma_x \sigma_y} \tag{2.22}$$

This can be between -1 and 1. The maximum and minimum values correspond to perfect correlations and zero means no correlation between parameters. Positive or negative signs indicate increasing or decreasing change of one parameter by the variation of the other one. It is worth noting that the slope of the relationships in the graphs is nothing to do with Pearson correlation.

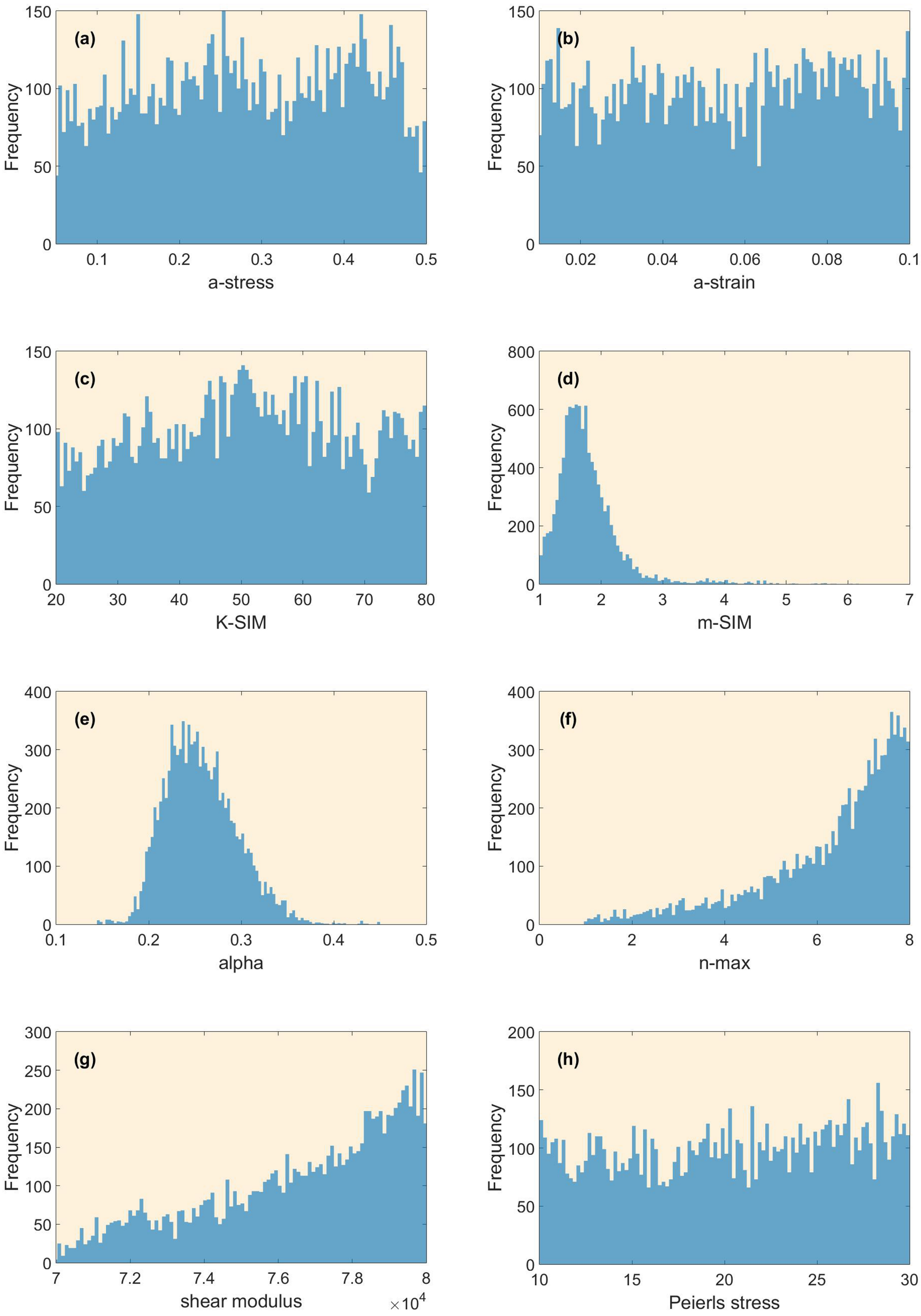


Figure 2.4: Histograms of posterior PDFs of applied parameters after sequential calibration

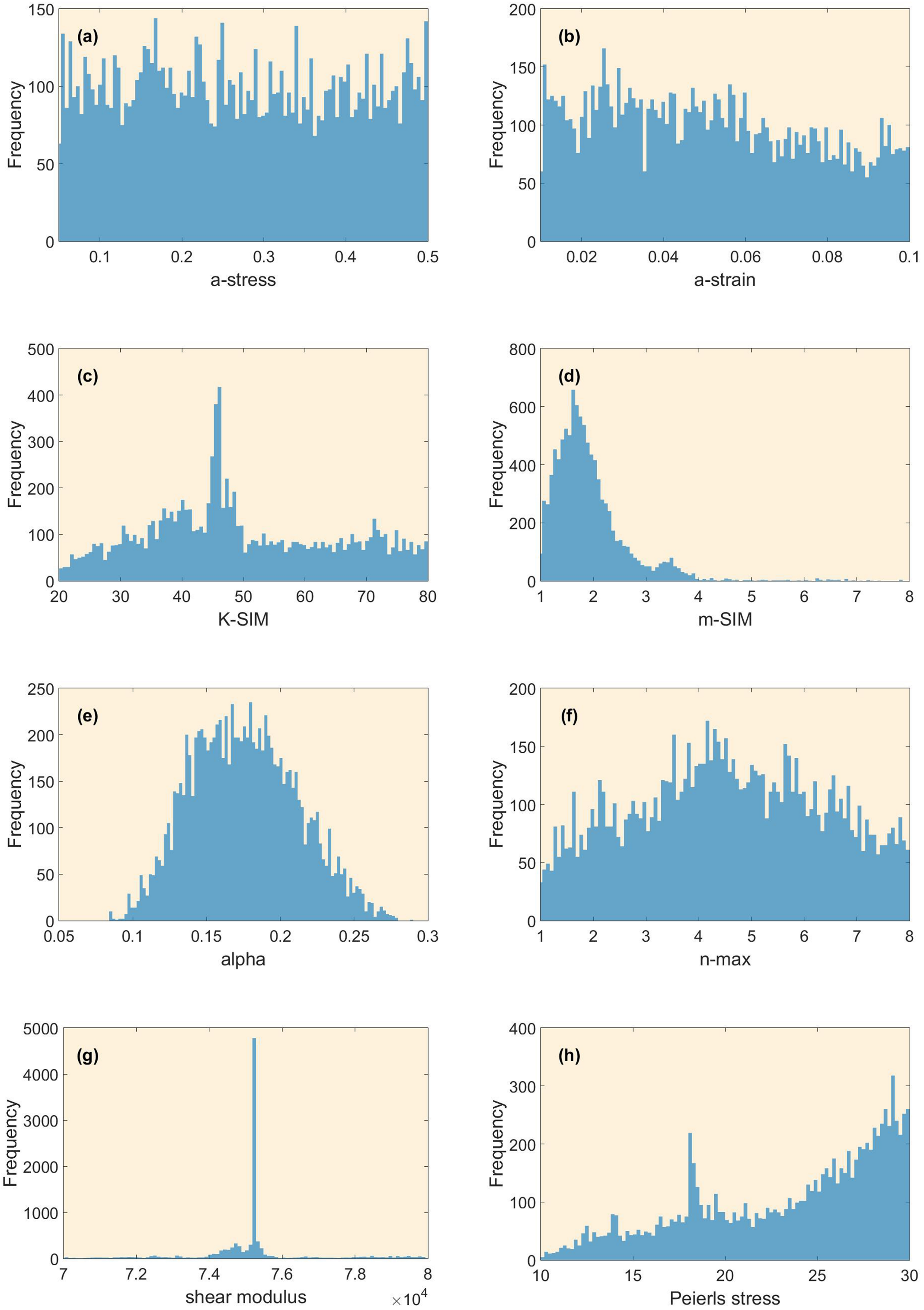


Figure 2.5: Histograms of posterior PDFs of applied parameters after simultaneous calibration

In fact, the correlation depends on how linear the relationship is. The correlation coefficients are calculated for each two model parameters in both of the calibration cases, as observed in Tables 2.6 and 2.7.

In both methods of calibration, there are no/low correlation coefficients between applied parameters, except for alpha, that show some correlations with m-SIM, Peierls stress, and n-max; although, some correlations are also observable between alpha and shear modulus as well as m-SIM and K-SIM in the case of sequential calibration. These can be also evaluated through the correlation graphs in Figures 2.6 and 2.7. In these graphs, red regions indicate the convergence regions of the generated MCMC samples that include the plausible optimal values for the model parameters. The density of samples in 2D parameter spaces decreases from red to blue regions, so the dispersion of points and their linearity in each case can be visually analyzed and verified with the correlation coefficients obtained by Equation 2.22. In fact, more scattering of sample points and less linearity (corresponding to a wider ellipsoid and more circular shape) lead to lower correlation coefficient. Another important feature of these graphs is that the equifinality problem is easily distinguishable. For instance, several density peaks are noticeable in Figures 2.6(a) and (d). However, this feature cannot be seen in Figures 2.7(a) and (d) anymore, which suggests less equifinality problem and more defined posterior PDFs in the case of simultaneous calibration.

Based on the information in Table 2.5 about the calibrated parameter values and their uncertainties and considering the propagation of uncertainty, stress-strain curves and their uncertainty bands after sequential and simultaneous calibration have been demonstrated for all experimental conditions in Figures 2.8 and 2.9. The black lines correspond to the plausible mean parameter values obtained after calibrations, and the blue and (blue + green) shaded areas are related to 68% and 95% Bayesian confidence intervals of the model predictions, respectively, i.e., $\sigma(\varepsilon) \pm \sigma_s(\varepsilon)$ and $\sigma(\varepsilon) \pm 2 \times \sigma_s(\varepsilon)$ in which rs can be determined by Equation 2.20 at any strain. According to these plots, it can be mentioned that the model gives fairly close results to experimental data in the cases of TRIP 1, 5, 6, and 7 after sequential calibration, and 1, 3, 6, and 7 after simultaneous calibration. The outcomes of the model imply a requirement for improvement of the model; al-

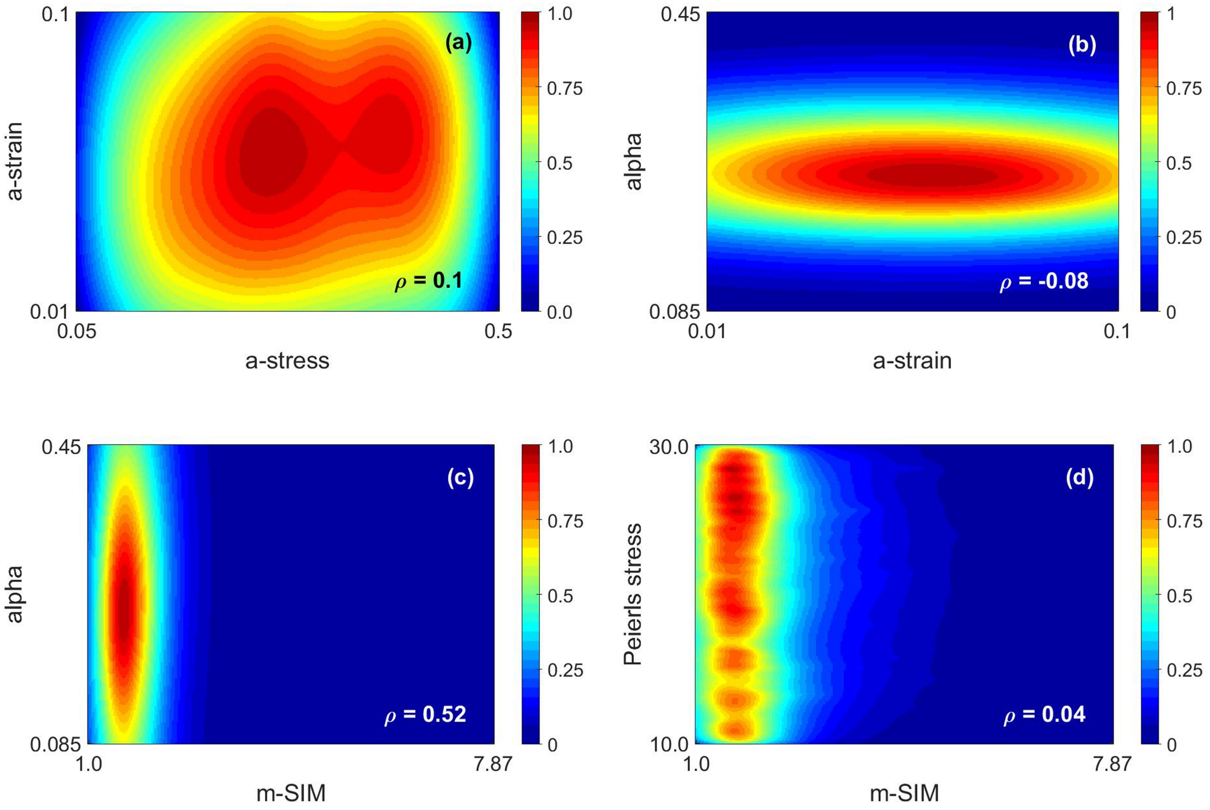


Figure 2.6: 2D correlation graphs with normalized color bars for some of pair model parameters in the case of sequential calibration

Table 2.6: Pearson correlation coefficients between model parameters in the case of sequential calibration

	a_σ	a_ε	k	m	α	$n-max$	$\mu(MPa)$	$\sigma_p(MPa)$
a_σ	1	0.11	-0.04	-0.06	-0.11	-0.03	0.01	0.14
a_ε	0.11	1	0.035	-0.14	-0.08	0.08	0.095	0.01
k	-0.04	0.035	1	0.23	0.03	-0.055	0.07	-0.03
m	-0.06	-0.14	0.23	1	0.52	-0.12	-0.15	0.04
α	-0.11	-0.08	0.03	0.52	1	-0.39	-0.40	-0.57
$n-max$	-0.03	0.08	-0.055	-0.12	-0.39	1	0.11	0.08
$\mu(MPa)$	0.01	0.095	0.07	-0.15	-0.40	0.11	1	0.01
$\sigma_p(MPa)$	0.14	0.01	-0.03	0.04	-0.57	0.08	0.01	1

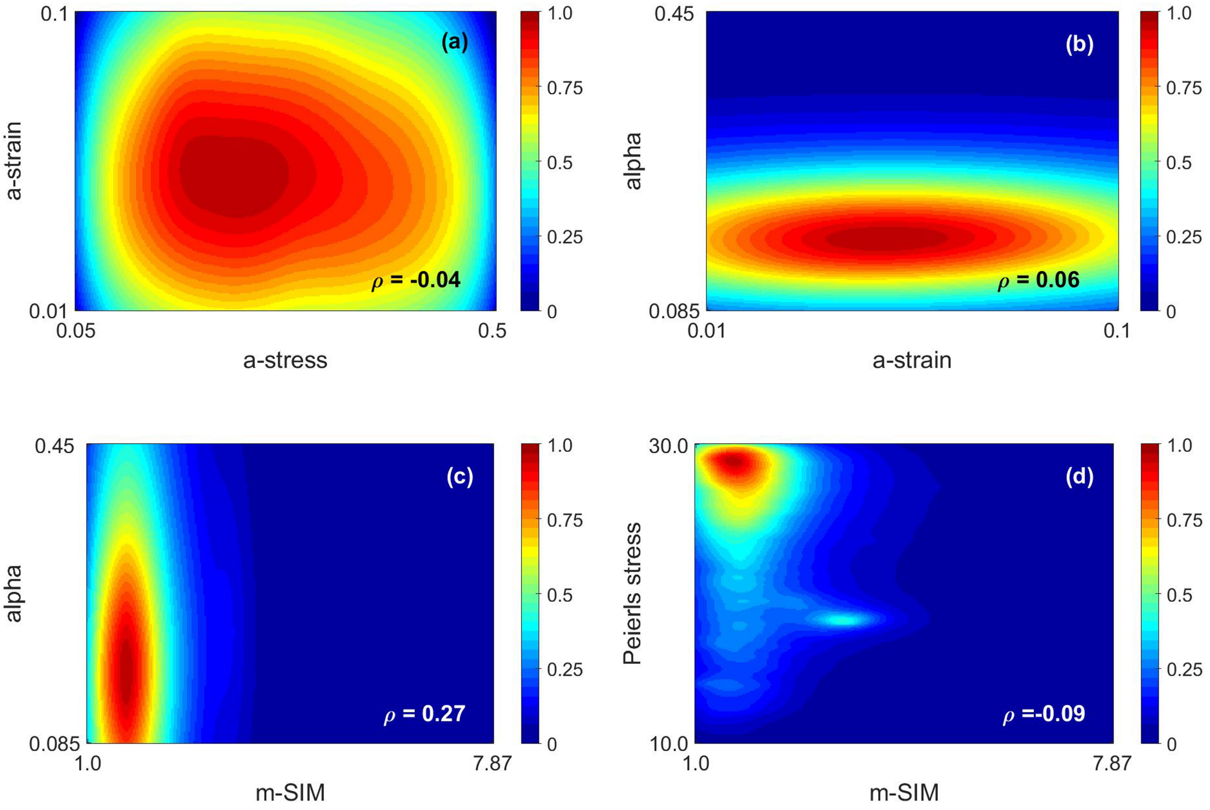


Figure 2.7: 2D correlation graphs with normalized color bars for some of pair model parameters in the case of simultaneous calibration

Table 2.7: Pearson correlation coefficients between model parameters in the case of simultaneous calibration

	a_σ	a_ε	k	m	α	$n-max$	$\mu(MPa)$	$\sigma_p(MPa)$
a_σ	1	-0.04	0.002	-0.115	-0.05	0.025	0.075	-0.02
a_ε	-0.04	1	0.09	-0.02	0.06	-0.006	0.04	-0.06
k	0.002	0.09	1	0.05	-0.015	0.025	0.09	0.0002
m	-0.115	-0.02	0.05	1	0.27	0.09	-0.001	-0.09
α	-0.05	0.06	-0.015	0.27	1	-0.685	-0.13	-0.495
$n-max$	0.025	-0.006	0.025	0.09	-0.685	1	-0.05	-0.003
$\mu(MPa)$	0.075	0.04	0.09	-0.001	-0.13	-0.05	1	0.085
$\sigma_p(MPa)$	-0.02	-0.06	0.0002	-0.09	-0.495	-0.003	0.085	1

though the discrepancies between model and experiments cannot only be attributed to the model deficiency and the effect of experimental errors should also be considered. Therefore, the volume fractions of microstructural phases are estimated in Section 2.4.3 due to high uncertainty in their experimental measurement.

2.4.3 Estimation of phase volume fractions using sequentially calibrated parameters

Generally, the determination of phase volume fractions in steels using image analysis is a hard task and coupled with a lot of uncertainties, especially for ferrite, bainite, and martensite which possess similar crystal structures but different dislocation densities. Therefore, the disagreement between model results and experimental data can be partly related to the experimental uncertainty of phase volume fractions. Assuming the uncertainty of the model is zero and the model is physically correct and complete, the sequentially calibrated parameters have selected for the model to estimate the phase volume fractions by our MCMC approach. Similar to the calibrated parameters in Section 2.4.2, these values and their uncertainties (listed in Table 2.8) can be applied to plot stress-strain curves and their uncertainty bands, as shown in Figure 2.10. It can be observed that the correction of phase volume fractions either results in a very good agreement between the model prediction and corresponding experimental data or at least brings the experimental data to 95% Bayesian confidence intervals of the model predictions. These estimations can similarly be performed using simultaneously calibrated parameters.

2.4.4 Accuracy evaluation of the MCMC calibration using generated synthetic data

Regarding the disagreement between model and experimental results, some can state that the uncertainty may result from the inaccuracy of the calibration approach. In order to prove that our approach is reliable for model calibration, the model parameters are simultaneously recalibrated using three synthetic datasets for TRIP 1, 2, and 3, which are generated from the model with simultaneously calibrated parameters in Section 2.8 and experimental conditions in Table 2.4. In fact, the reproducibility of the calibration method is examined by synthetic datasets. As observed in Figure 2.11(a), the model with initial parameter values in Tables 2.2 and 2.3 leads to the stress-

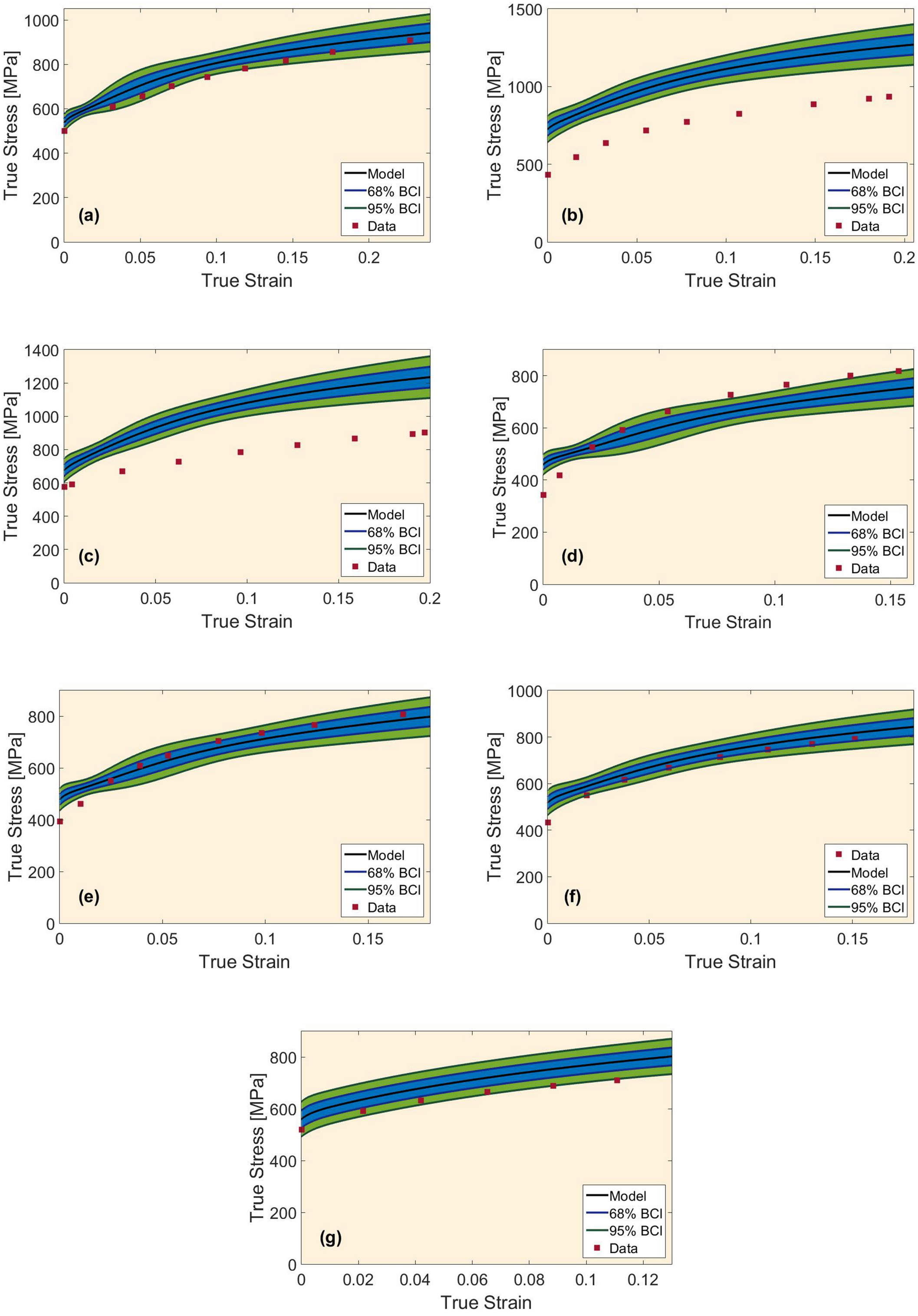


Figure 2.8: Model prediction of strain-stress curves and their uncertainty bands after sequential calibration in different experimental conditions: (a-g) TRIP (1-6) and DP, BCI: Bayesian Confidence Interval

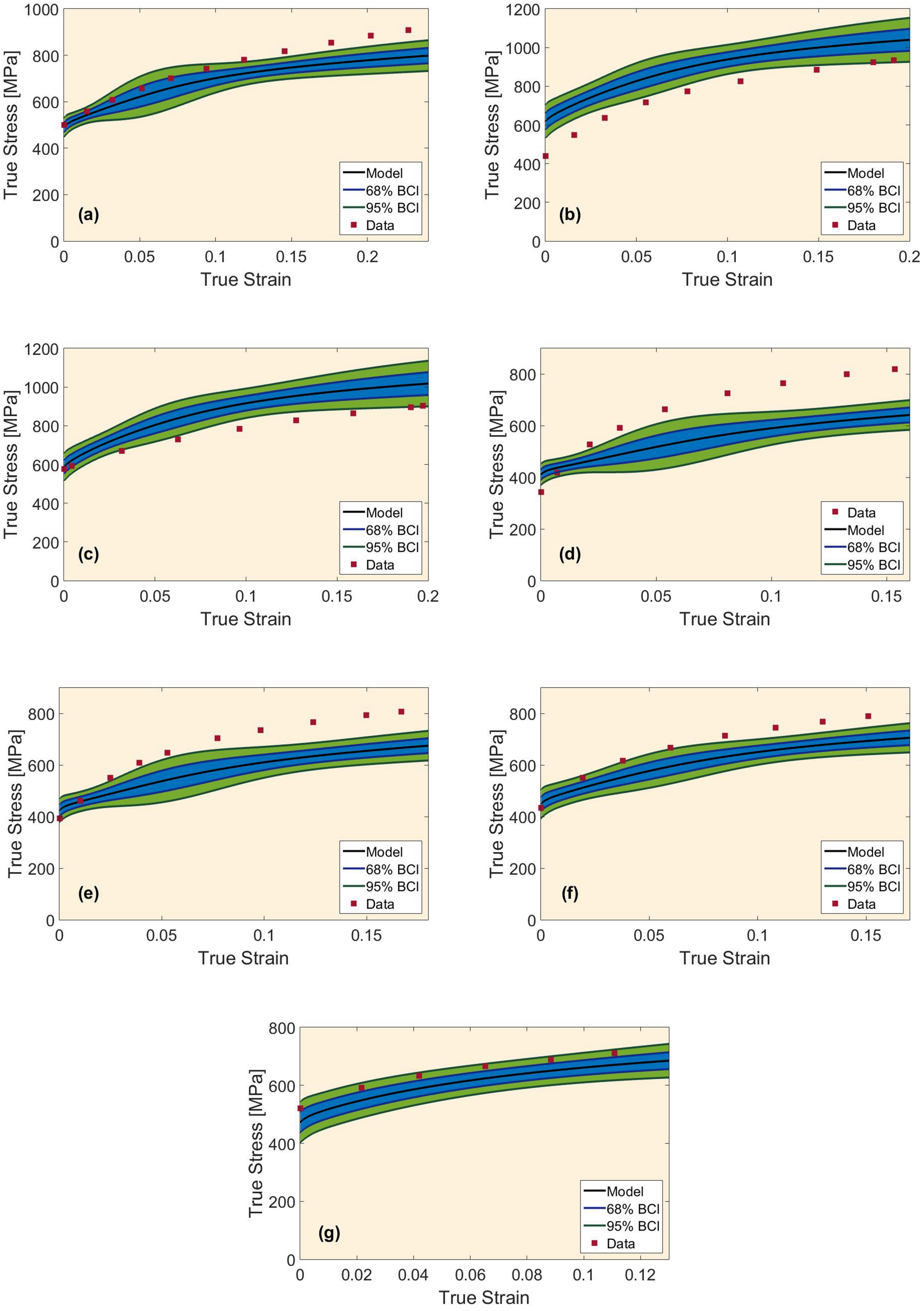


Figure 2.9: Model prediction of strain-stress curves and their uncertainty bands after simultaneous calibration in different experimental conditions: (a-g) TRIP (1-6) and DP, BCI: Bayesian Confidence Interval

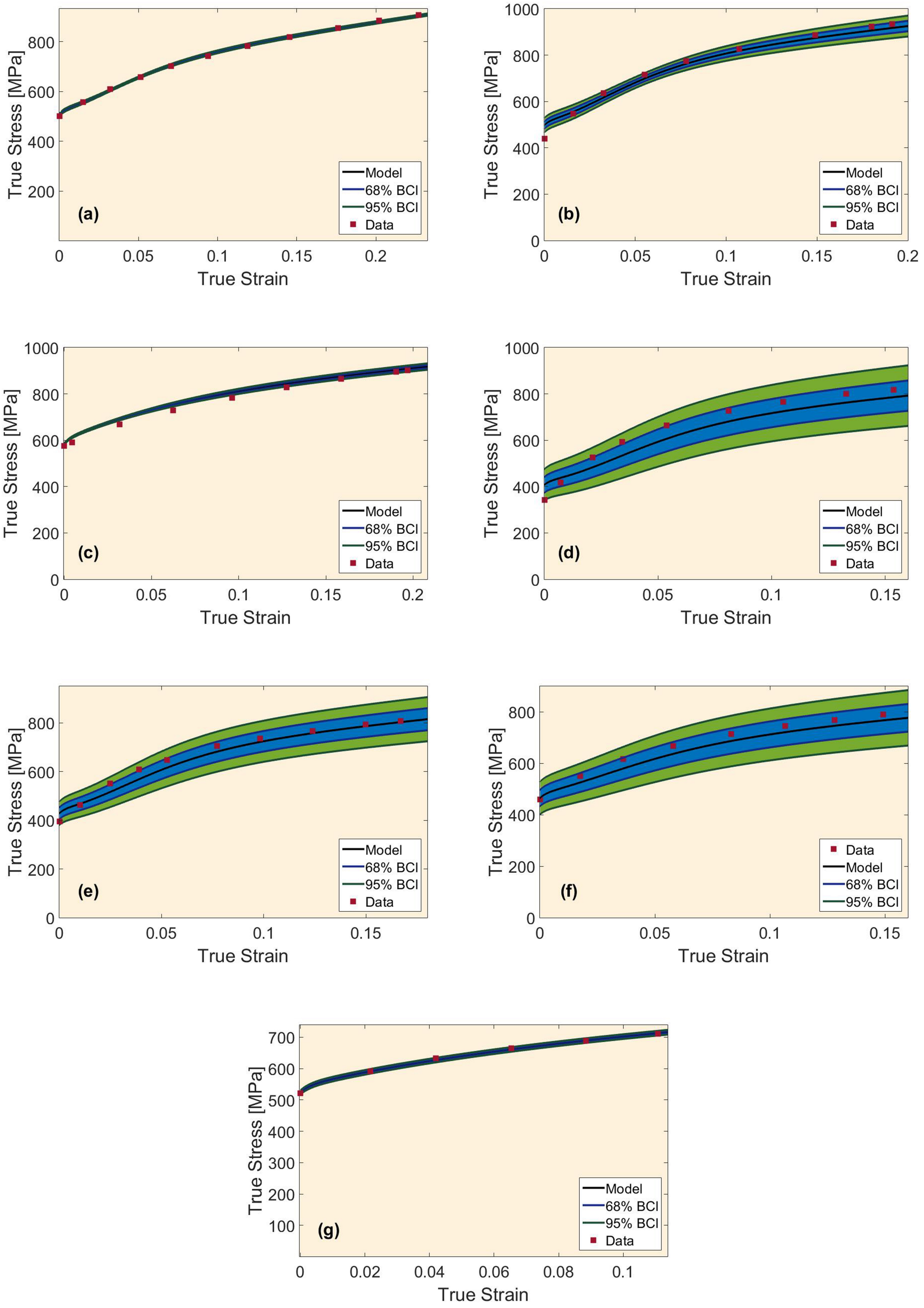


Figure 2.10: Model prediction of strain-stress curves and their uncertainty bands after the estimation of phase volume fractions using sequentially calibrated parameters in different experimental conditions: (a-g) TRIP (1-6) and DP, BCI: Bayesian Confidence Interval

Table 2.8: Applied parameter values and their uncertainties after sequential and simultaneous calibrations compared to initial values before calibrations

Alloy	Volume Fraction (%)			
	Ferrite	Bainite	Austenite	Martensite
TRIP 1	74.8±0.62	3.6±0.63	10.5±0.54	11.1±1.51
TRIP 2	76.6±2.24	6.5±1.06	16.9±1.82	0.0
TRIP 3	73.8±0.26	18.3±0.35	1.2±0.44	6.7±0.58
TRIP 4	75.0±5.43	4.6±2.18	20.4±2.98	0.0
TRIP 5	75.6±5.56	5.8±2.22	18.6±2.62	0.0
TRIP 6	75.9±4.06	11.3±1.97	10.9±1.78	1.9±0.58
DP	80.0±1.96	20.0±0.28	0.0	0.0

strain results being very far from the synthetic datasets. However, excellent fits between model results and the synthetic data with very low uncertainties in Figure 2.11(b) show that simultaneous recalibration delivers the same parameters used for dataset generations. This confirms a high reproducibility of the introduced MCMC approach. The same scenario is believed to happen in the case of considering sequentially calibrated parameters to generate datasets. This means any discrepancy between model and experimental results can only be attributed to the uncertainty of model and/or experiments, not the applied calibration approach.

2.5 Summary and Conclusion

In this work, plastic flow behavior of TRIP steels is predicted through a novel physical-based model that considers different hardening mechanisms, especially dislocation strengthening grains based on RiveraâĂŹs work, which are considered for each microstructural phase and then homogenized for the entire microstructure by an iso-work approximation. Strain-induced martensitic transformation is also incorporated in the model through Haidemenopoulos approach.

The major point of current work is the introduction of Bayesian calibration based on Metropolis-Hastings MCMC using the abovementioned models as case studies. After finding eight sensitive parameters through forward analysis of models, two different techniques, i.e., sequential and simul-

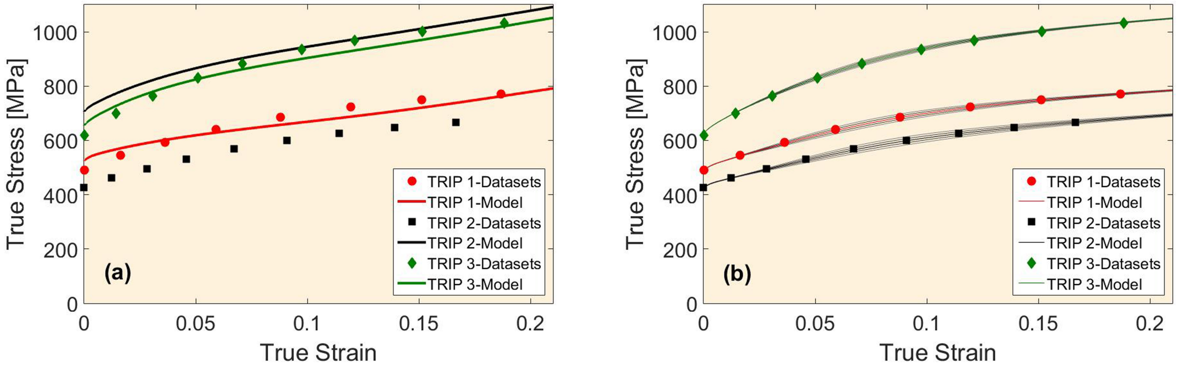


Figure 2.11: Model prediction of strain-stress curves (a) before and (b) after simultaneous recalibration with synthetic datasets generated using simultaneously calibrated parameters and experimental conditions corresponding to TRIP 1, 2, and 3

taneous experimental training have been applied to calibrate these parameters. Obtained posterior PDFs and 2D correlation graphs from both calibration methods show a clear peak for m-SIM, alpha, n-max, and shear modulus although the distribution peak of alpha, n-max and shear modulus are lower in the case of simultaneous calibration. For the other four parameters, sequential calibration leads to multiple peak or equifinality problems, which is less distinguishable in the results of simultaneous calibration. It should be noted that simultaneous calibration produces a more defined shape of parameter posterior distribution compared to sequential calibration. In both calibration approaches, some of the stress-strain curves resulting from the models with calibrated parameters show disagreements with experimental data; however, the results are better than ones obtained from uncalibrated models overall.

Considering high experimental uncertainties in determination of phase volume fractions and assuming a high accuracy in physics of the models, Bayesian estimation of volume fractions using the calibrated model causes the stress-strain predictions to either fit or be close to experimental results in a way that all data points at least fall in 95% Bayesian confidence intervals. In order to prove that the uncertainties result from the models and/or experiments and not from our calibration approaches, the models have simultaneously been recalibrated with synthetic datasets generated from the models with simultaneously calibrated parameters. Excellent agreement between model results and synthetic datasets after recalibration imply a high reproducibility of the applied MCMC method.

To close the present discussion, we would like to note that while the present work is rather specific in the system under study (TRIP steels) and the selected model (irreversible-thermodynamics of dislocation evolution + iso-work approximation + SIM), the framework presented here is highly generalizable to many other computational materials science problems that require rigorous uncertainty quantification (UQ). In fact, we believe that current emphasis on the use of computational approaches to accelerate the development/design of materials require UQ in order to provide a measure of robustness to the results of materials simulations. Moreover, UQ can also be used to assess the uncertainty in experimental observations and to determine parameters difficult to observe experimentally. While Bayesian methods have been amply used in other fields, they—with some notable exceptions—have not been widely used in computational materials science but their increased use is warranted by the needs of the field.

3. BAYESIAN PROBABILISTIC CALIBRATION OF NI-TI PRECIPITATION MODEL IN MATCALC SOFTWARE USED IN AN INFORMATION FUSION APPROACH FOR MORE PRECISE PREDICTION OF PRECIPITATION BEHAVIOR

3.1 Overview

Ni-Ti alloys are the most popular shape memory alloys in different industrial applications due to their capability for precipitation of second phases. Therefore, precipitation modelling and associated model calibration coupled with the uncertainty estimations seem to be crucial. In this work, the calibration and uncertainty quantification of influential model parameters have been performed using a Bayesian approach based on Markov Chain Monte Carlo-Metropolis Hastings algorithm. Then, the resulting parameters' uncertainties have been propagated to the uncertainty of model results. An empirical relationship for matrix/precipitate interfacial energy in terms of aging temperature and nominal composition has been introduced according to the values of interfacial energy obtained from the model calibration with each given experimental data individually. Using this equation for interfacial energy in the model, all experimental data has been used together to calibrate the other four model parameters. The results of the probabilistic calibration showed some discrepancies between the model responses and the experimental data that are the important reasons to apply co-kriging surrogate modelling for more precise prediction of experimental precipitation behavior in this system through an error correlation-based fusion of the model results and the experimental data.

3.2 Introduction

Over last decades, Ni-Ti shape memory alloys (SMAs) have been in the center of attention for different engineering and medical applications due to their excellent combination of mechanical properties and shape memory behaviors [57, 58]. Among the SMAs, Ni-Ti alloys are more attractive due to the tailorability of transformation temperature (M_S) resulting from a change in matrix composition through precipitation of second phases [58, 59]. In these alloys, different thermo-mechanical processings can result in the formation of various types of (meta)stable precipitates, among which coherent $\mathrm{Ni_4Ti_3}$ phase has been studied more because of its significant effects on transformation temperature, phase constitution, and cyclic fatigue resistance [60]. Despite the recent experimental studies of the $\mathrm{Ni_4Ti_3}$ precipitates influence on multi-stage martensitic transformation path [58, 61, 62] and shape memory behavior (temperature hysteresis) [63, 64], there are few studies with regards to modelling of precipitation hardening for Ni-Ti alloys. In this regard, there are some phase field studies, including the determination of the precipitate equilibrium shape and its surrounding concentration in addition to the effects of stress field on the subsequent martensitic transformation response [65], the homogeneity assessment of precipitate distribution across the microstructure during stress-free and stress-assisted aging heat treatment [66], and the investigation of the precipitate morphology and growth during stress-assisted aging heat treatment in a porous Ni-Ti system [67]. Another important work with the purpose of connecting chemistry and processing to microstructure under Integrated Computational Materials Engineering (ICME) framework is the evaluation of $\mathrm{Ni_4Ti_3}$ precipitation kinetics through the Kampmann-Wagner Numerical (KWN) [68] model in MatCalc© package [60].

Generally, model calibration and associate uncertainty quantification (UQ) is one of the most required tasks for a robust and/or reliability based design, which is usually ignored in materials modeling. Among the few UQ works in computational materials science, our previous studies on the uncertainty analyses of mechanical properties in TRIP steels [69] and UQ in calculation of phase diagrams (CALPHAD) model [70] are examples of thorough probabilistic assessments of the models' parameters and their predicting responses. It is worth noting that the modeling uncertainties can result from different sources: 1) Natural uncertainty (NU), 2) Model parameter uncertainty (MPU), 3) Propagated uncertainty (PU), and 4) Model structure uncertainty (MSU) [3]. Simplification assumptions, numerical approximations, and lack of knowledge in physical models result in MSU where the uncertainty quantification is more complex and challenging compared to the other types of uncertainties. In this regard, it should be noted that model calibration is the first task to do before the quantification of MSU.

In the context of Bayesian statistics, the estimation of parameters' values and their uncertainties incorporate multi-dimensional intractable integrals with respect to parameter vector (θ), which is hard or sometimes impossible to solve via the conventional numerical approaches. Markov Chain Monte Carlo (MCMC) techniques have already been introduced as an efficient sampling tool to approximate these integrations [5]. It should be stated that MCMCs are considered as highly robust non-linear techniques for the calibration of model parameters. They can provide qualitative information about the correlation between the parameters, the shape of parameter posterior probability density functions (PDFs) in addition to the quantitative information about plausible optimal values and uncertainties of the parameters [35]. In the case of multi-level models, it has been shown [37] that MCMC approaches can provide better data fitting compared to other likelihood-based methods in the model probabilistic calibrations [37]. Generally, Metropolis-Hasting (MH) algorithm and Gibbs sampling are two common MCMC approaches for Bayesian calibration of models. Although 100% acceptance rate makes Gibbs sampling faster than MH algorithm where samples may also be rejected, Gibbs sampling shows restrictions in some cases: 1) when the conditional distributions, which are used for sampling, are difficult to be obtained from the joint distributions, 2) when inversion sampling from the conditional distribution is either very hard or sometimes impossible due to its uncommon distribution shape, and/or 3) in the case that sampling from the conditional distribution may get stuck in a low probability region of the posterior density for a long time that slows down the process of the parameter convergence towards the center of this density. Under these conditions, MH algorithm works more efficient [6].

After the model calibration, MSU can be quantified by assuming that the available experimental data are true. Kennedy and O'Hagan work [12] was one of the pioneers in this regard, where co-kriging surrogate modelling based on a kriging/gaussian process (GP)-based technique has been introduced. In this approach, a linear relationship is assumed between the high-fidelity experimental responses and the fitted GP surrogate model over the low-fidelity model results. Here, the linear regression coefficient is optimized during the process and the regression constant term (the error terms) is estimated through another GP model [3, 13]. The discrepancy between the obtained response for experimental behavior and the model results can be recognizes as MSU.

Continuing the work in our group [60], the calibration and uncertainty quantification of the precipitation model parameters in MatCalc© have been carried out in the present work by a rigorous Bayesian approach in the context of Markov Chain Monte Carlo (MCMC). This model is then used as a low-fidelity model in Kennedy and O'Hagan co-kriging approach to cover the discrepancy between model and data (i.e., MSU) for more precise prediction of precipitation behavior in this system. This approach is very helpful in order to reduce the uncertainties in design under ICME framework since the precipitation model is one of the important puzzle pieces in the process-structures-properties-performance that connects processing to microstructure.

3.3 Model Description and Methodology

3.3.1 Precipitation model

A precipitation model implemented in MatCalc© version 5.61 has been used in this work. This model includes thermodynamic and kinetics assessment of $\mathrm{Ni_4Ti_3}$ precipitation in binary Ni-Ti system. In this numerical model, all the calculations are performed in a time loop. In each time iteration, the equilibrium calculations are carried out first in the thermodynamics module for the evaluation of the phases' stability. These calculations are based on the minimization of total Gibbs free energy of the system as follows:

$$\Delta G = \Sigma_i f^i . G^i = Minimum \tag{3.1}$$

where i and f^i are the index and volume fractions of each individual phase in the system, and G and G^i are the molar Gibbs free energy of the system and each one of the phases, respectively [71]. Here, G^i is calculated using the sub-lattice model, offered by Hillert and Staffansson [72], which applies the thermodynamic parameters obtained from available commercial database for Ni-Ti system. In the thermodynamics module, the equilibrium composition of phases and thermodynamic driving force for nucleation are also determined for subsequent kinetics calculations. The calculations continue for precipitation evolution if the driving force for precipitation is positive; otherwise, the calculations stop and go to the next time iteration [71].

In the case that precipitation is thermodynamically favorable, the precipitation driving force obtained from the thermodynamics module besides the information related to the diffusion coefficients of the diffusing elements in the system obtained from the mobility database in the software package DICTRA are transferred to the MatCalc© kinetics module. In this module, the calculations corresponding to nucleation, growth, and coarsening of the second phase (i.e., $\mathrm{Ni_4Ti_3}$) are concurrently carried out in each time step, based on the KWN model in the framework of classical nucleation and growth theory. Generally, the precipitation evolution occurs in a way that the total Gibbs free energy dissipation is maximum [73], which is equivalent to the maximization of total entropy of the system [74]. In MatCalc© KWN model, precipitates are assumed to nucleate spherically. The critical nucleation size (r^*), the required critical work to overcome nucleation energy barrier (ΔG^*), and the nucleation rate (J) are determined through the consideration of total Gibbs free energy change in the system due to the competition between the effects of volumetric Gibbs free energy (ΔG_v), free interfacial energy $(\gamma_{\alpha/\beta})$, and elastic strain energy in the case of coherent precipitation (ΔG_ε), as follows:

$$r^* = \frac{2\gamma_{\alpha/\beta}}{\Delta G_v - \Delta G_\varepsilon} \tag{3.2}$$

$$\Delta G^* = \frac{16\pi\gamma^3_{\alpha/\beta}}{3(\Delta G_v - \Delta G_\varepsilon)^2} \tag{3.3}$$

$$J = \frac{dN}{dt} = N_0 Z B^* exp(-\frac{\Delta G^*}{kT}) exp(-\frac{\tau}{t}) \tag{3.4}$$

where N_0, Z, B^*, and τ are the number of nucleation sites per unit volume, the Zeldovich factor (≈ 0.05), the join rate of solute atoms from matrix to nucleus, and the nucleation incubation time, respectively. In the case of heterogeneous nucleation, Equations 3.3 and 3.4 are modified by a function in terms of θ, the wetting angle between the second phase nucleus and its heterogeneous nucleation site [75].

In KWN model implemented in MatCalc©, precipitate growth rate (v) is governed by the following relationship in the case of spherical diffusion fields:

$$v = \frac{dr}{dt} = (\frac{\bar{c} - c_{\alpha/\beta}}{c_\beta - c_{\alpha/\beta}})\frac{D_i}{r} \tag{3.5}$$

where D_i is the diffusion coefficient of the diffusing element and r is the particle size. In addition, $\bar{c}$, $c_{\alpha/\beta}$, and c_β are the concentrations of solute atom in the matrix, at matrix/precipitate interface, and in precipitate, respectively. It should be noted that $\bar{c}$ and $c_{\alpha/\beta}$ are time dependent while c_β is assumed constant during precipitate growth. Generally, the concentration at the interface $c_{\alpha/\beta}$ is different than the equilibrium solute concentration in matrix (c_e) due to the effect of the matrix/precipitate interface curvature known as Gibbs-Thomson effect. Therefore, the change in $c_{\alpha/\beta}$ results from the variation of the interface curvature as the precipitates grow further. For a spherical growing particle, this change is expressed in terms of the particle radius (r) as follows:

$$c_{\alpha/\beta} = c_e exp(\frac{2\gamma_{\alpha/\beta} V_\beta}{rRT}) \tag{3.6}$$

Here, V_β is the molar volume of the second phase [76]. In the case of non-spherical precipitate growth, a symmetric cylinder has been considered with an equivalent spherical volume but different interfacial area. In this case, Equations 3.5 and 3.6 are modified using shape parameters that are defined as a function of the cylindrical aspect ratio, $H = \frac{h(height)}{d(diameter)}$.

The average matrix composition $(\bar{c})$ can also be obtained during precipitate growth by solving the mass conservation equation [75, 77], as shown in the following:

$$\bar{c} = \frac{V_{domain} c_0 - \sum_{i=1}^{N} \frac{4}{3}\pi \bar{r}_i^3 N_i c_\beta}{V_{domain} - \sum_{i=1}^{N} \frac{4}{3}\pi \bar{r}_i^3 N_i} \tag{3.7}$$

where V_{domain}, c_0, $\bar{r}_i$, and N_i are the initial volume of the matrix, the nominal composition of solute atoms in the matrix, the average radius of the equivalent sphere corresponding to i^{th} precipitate size class and the total number of particles in i^{th} size class, respectively.

In MatCalc©, the number of formed nuclei and the radius change in the equivalent spherical particles are calculated in each time step up to the end of aging time. These changes in the number and the size of particles appear in the form of changes in the particle number of each size class/size interval, $[R_j, R_{j+1}]$. The number and size of particles may also alter by the occurrence of coarsening. Since R^* increases during precipitation process owing to the solute concentration decline in the matrix, the radius of the smallest particles may become smaller than R^* after a while. At this moment, these thermodynamically unstable particles start to dissolve that result in the coarsening of the remaining precipitates [71].

These calculations lead to a histogram for particle size which can be obtained after each time step and eventually at the end of the aging process. These can be utilized for the determination of the precipitation kinetics evolution during heat treatment and the final results for the average solute concentration in the matrix, volume fraction of precipitates, average precipitate size, total number of precipitates, etc.

3.3.2 Uncertainty quantification of model parameters and outcomes

Calibration and uncertainty quantification (UQ) of model parameters have always been an important concern for computational modeling in the context of robust or reliability-based designs, which can provide confidence in product design. Deterministic approaches (such as least squares fitting) are commonly applied for model calibrations, but no uncertainties are quantified in these approaches. Generally, the significance of UQ results from common high uncertainties in the experimental measurements of physical parameters and/or model outcomes in addition to the incomplete physics or the assumptions in models. MCMC techniques in the framework of Bayesian statistics are strong tools for UQ that incorporate all the above mentioned uncertainties into the parameter prior PDFs and likelihood function. During MCMC calibrations against experimental data, prior PDFs for model parameters (prior knowledge) are updated into their posterior PDFs in order to find the plausible optimum values for the parameters and reduce their uncertainties. In robust design where the system responses should be insensitive to the variation of the design variables/parameters, it is very important to find the uncertainties of model responses resulting from the parameters uncertainties. For this reason, the uncertainties are propagated from the parameters to the model responses through a model forward analysis of MCMC parameter samples in this work.

3.3.2.1 Markov Chain Monte Carlo: Metropolis-Hastings algorithm

In the case of multi-dimensional calibration problems, there are multi-dimensional intractable integrals that should be solved to determine the statistical characteristics of model parameters such as parameters' mean (optimum) values and their variance-covariance matrix. Monte Carlo (MC) method can be a plausible way for the estimation of the above-mentioned integrals which are numerically difficult or sometimes impossible to be calculated. However, MC requires a large number of samples from parameters posterior PDF for the integral estimations. MH is one of the efficient approaches for random sampling from a proposal (arbitrary) PDF for the parameters' posterior which continues until the convergence of Markov chain to a stationary PDF [78].

In this work, MCMC-MH toolbox in Matlab has been applied to calibrate the most influential parameters of MatCalc© precipitation model for $Ni-Ti$ system. The first step is the introduction of parameters' prior knowledge to the toolbox, including their initial values ($\theta^{(0)}$), probability distributions, and value ranges. Here, the lack of prior knowledge about the precipitation model parameters resulted in the consideration of an uniform (non-informative) prior PDF for the model parameters. In this toolbox, the parameters (candidates) sampled from a proposal PDF are accepted or rejected based on the calculated value for Metropolis-Hastings ratio in each iteration:

$$MH = \frac{P(\theta^{cand})P(D|\theta^{cand})}{P(\theta^{i})P(D|\theta^{i})} \frac{q(\theta^{i}|\theta^{cand})}{q(\theta^{cand}|\theta^{i})} \tag{3.8}$$

where θ is the parameters' vector, and D is the experimental data vector which includes matrix Ni content ($Ni\%$), precipitate volume fraction (V_f), and average precipitate size (d), explained in details in section 3.3.4. Therefore, a multi-objective calibration is carried out in this work to reduce the risk of equi-finality where there are more than one response for the parameter values after calibration. For instance, multiple combinations of the parameter values may result in the same response for precipitate volume fraction if it is considered as the only objective in this problem.

In Equation 3.8, the first fraction is Metropolis ratio that is the product of prior $(P(\theta))$ and likelihood $(P(D|\theta))$ for new candidate (θ^{cand}) over its counterpart for the last accepted parameters (θ^{i}), and the second fraction is Hastings ratio that compares the probability of a jump from θ^{i} to θ^{cand} with the probability of the reverse jump. Generally, the new candidate is accepted in the case that MH is greater than 1; however, the new candidate can still be accepted with a probability of $MH \times 100$ percent in the case of a MH value less than 1. If rejection happens, the previous accepted θ is repeated again as the new sample. At the end, there will be n samples of parameters $\left\{\theta^{(0)}, ..., \theta^{(n)}\right\}$ which represents a multivariate posterior PDF for parameters. The mean values of these n samples and the square root of diagonal terms in their corresponding variance-covariance matrix can be introduced as the plausible calibrated values and uncertainties of model parameters, respectively.

It is worth noting that likelihood function in this work is a Gaussian distribution centered at D. Here, the likelihood variance is considered as an unknown hyper-parameter and updated against the experimental data, similar to other parameters in the model. In addition, the proposal PDF is defined as an adaptive multivariate Gaussian distribution where its mean is set at the initial or the last accepted θ and its variance-covariance matrix is adapted during sampling in terms of the variance-covariance matrix of all the previous generated samples, as proposed by Haario et al [51].

3.3.2.2 Uncertainty propagation

The propagation of the parameters' uncertainties obtained from MCMC sampling to the responses of the model is an important task in the context of materials design. This can be achieved for any non-linear function using a forward model analysis of the MCMC samples after parameters' convergence. In this approach, the total or a part of samples after MCMC convergence are used in a forward analysis scheme to obtain the corresponding samples for model responses. Then, 95$\%$ Bayesian confidence intervals can be determined through the removal of 2.5$\%$ of samples from the upper and lower limits of the resulting samples for the model responses.

3.3.3 Gaussian process-based surrogate modelling

In the recent years, the development of surrogate models have been the center of attention due to their vital applications for the efficient optimization of expensive models. In fact, the surrogate models provide an approximation for expensive models throughout the entire design space [79]. Generally, surrogate models can be built through two major fitting approaches, ordinary regression analyses and stochastic process (e.g., Bayesian non-parametric GP [80]). The main issues with the regression approaches are the assumption of a known form of regressor to the model function and the consideration of independent errors between points in the design space. On the other hand, the stochastic GP (or kriging in the context of geology literature) emphasizes the correlation between errors instead of the indication of regressor formulation, so that the regressor can reduce to just a constant term [81], as follows:

$$y(X^{(i)}) = \mu + \varepsilon(X^{(i)}) \quad , \quad i = (1, ..., n) \tag{3.9}$$

where $\{X^{(i)} = (x_1^{(i)}, ..., x_k^{(i)})$,y$^{(i)} = y(X^{(i)})\}$ are n sampled data points and their corresponding model function responses. Here, A free noise condition with zero uncertainty is assumed for data points. μ is the average of the stochastic model and $\varepsilon(X^{(i)})$ is the error term defined as a normal distribution with mean zero and the variance σ^2, i.e., $\mathcal{N}(0, \sigma^2)$. In this fitting approach, the correlation between the errors of data points are obtained by a weighted distance function expressed as:

$$d(X^{(i)}, X^{(j)}) = \sum_{l=1}^{k} \theta_l |x_l^{(i)} - x_l^{(j)}|^{p_l} \tag{3.10}$$

$$corr\Big[\varepsilon(X^{(i)}), \varepsilon(X^{(j)})\Big] = exp\Big[-d(X^{(i)}, X^{(j)})\Big] \tag{3.11}$$

where θ_l and p_l indicate the importance or sensitivity of variable l and the smoothness of the correlation function along the coordinate direction l, respectively. It is worth noting that less distance between points means a correlation closer to 1. These correlation values between the errors of different data points can produce a $n \times n$ correlation matrix (R) [81].

In our work, a radial basis function is considered for the correlation function, where $p_l = 2$ in Equation 3.10. However, the optimal values of the other hyper-parameters $\hat{\mu}$, $\hat{\sigma}^2$ and $\hat{\theta}$s are obtained via a Maximum Likelihood Estimation (MLS). These hyper-parameters and R are used to derive the mean value and the mean squared error of the surrogate model at any given point in the design space [81,82], as the following equations:

$$\hat{y}(X) = \hat{\mu} + r(X)^T R^{-1}(\mathbf{y} - I\hat{\mu}) \tag{3.12}$$

$$S^2(X) = \hat{\sigma}^2 \left[1 - r^T R^{-1} r + \frac{(1 - I^T R^{-1} r)^2}{I^T R^{-1} I} \right] \tag{3.13}$$

Here, $r(X)$ is the correlation vector between the errors at the original data points and at any given spatial point, i.e., $r(X) = \left[corr \left[\varepsilon(X^{(i)}, \varepsilon(X)) \right] \right]_{n \times 1}$. I is also a $n \times 1$ identity vector. It should be noted that "T" and "-1" denote the transpose and inverse matrix operations.

In the case that there are multi-fidelity simulations and/or experiments with different levels of accuracy/cost for a real system, co-kriging surrogate modelling is employed to better infer the behavior of higher fidelity/more expensive models with lower uncertainties through the establishment of an error correlation matrix between their data points and the lower fidelity/cheaper data points [12]. This is a very cost-efficient way to predict the behavior of the expensive models with less evaluations. Generally, co-kriging approach is an extension of kriging (GP) surrogate modelling [83], which was pioneered by Kennedy and O'Hagan [12]. In the current work, the precipitation model in MatCalc© and experimental behavior (observations) are considered as the low-fidelity/cheap and high-fidelity/expensive models, respectively. In such cases, a linear relationship is defined between the observation (Γ) and the low-fidelity model result (M) at any given spatial design point based on Kennedy and O'Hagan approach, as follows:

$$\Gamma(X) = \Psi(X) + e(X) = \rho M(X, \theta) + \delta(X) + e(X) \tag{3.14}$$

where Ψ is the real response of the system, ρ is a constant regression parameter, δ is a model discrepancy function that is independent of the low-fidelity model response, and e is the observation error [13]. In this context, two Gaussian Processes are constructed over the low fidelity model $M(X_i, \theta)$ responses and the error terms $\delta(X_i) + e(X_i)$ and statistically combined together in order to predict the behavior of the high-fidelity model $\Gamma(X)$ within an uncertainty band.

Considering $\{X_c = (x_c^1, ..., x_c^{n_c}), \mathrm{y}_c = (y_c^1, ..., y_c^{n_c})\}$ and $\{X_e = (x_e^1, ..., x_e^{n_e}), \mathrm{y}_e = (y_e^1, ..., y_e^{n_e})\}$ data points for low- and high-fidelity models, Y_c and Y_e are two mentioned GP surrogate models which are built over $\{X_c, \mathrm{y}_c\}$ and $\{X_e, \mathrm{y}_d = \mathrm{y}_e - \rho M(X_e, \theta)\}$, according to Equations 3.12 and 3.13. The resulting co-kriging surrogate model is expressed as:

$$Y_e(X) = K\alpha + \phi(X)^T \Phi^{-1}(\mathbf{y} - F\alpha) \tag{3.15}$$

$$S^2(X) = \hat{\rho}^2 \hat{\sigma}_c^2 + \hat{\sigma}_d^2 - \phi^T \Phi^{-1} \phi \tag{3.16}$$

where,

$$K = \begin{bmatrix} \hat{\rho} & 1 \end{bmatrix} \;,\; \alpha = \begin{bmatrix} \hat{\mu}_c \\ \hat{\mu}_d \end{bmatrix} \;,\; F = \begin{bmatrix} \mathrm{I} & 0 \\ \hat{\rho}.\mathrm{I} & \mathrm{I} \end{bmatrix} \;,\; \mathbf{y} = \begin{bmatrix} \mathbf{y}_c \\ \mathbf{y}_e \end{bmatrix} \tag{3.17}$$

$$\phi(X) = \begin{bmatrix} \hat{\rho}\hat{\sigma}_c^2 r_c(X) \\ \hat{\rho}^2 \hat{\sigma}_c^2 r_c(X, X_e) + \hat{\sigma}_d^2 r_d(X) \end{bmatrix} \tag{3.18}$$

$$\Phi(X) = \begin{bmatrix} \hat{\sigma}_c^2 R_c & \hat{\rho}\hat{\sigma}_c^2 R_c(X_c, X_e) \\ \hat{\rho}\hat{\sigma}_c^2 R_c(X_e, X_c) & \hat{\rho}^2 \hat{\sigma}_c^2 R_c(X_e, X_e) + \hat{\sigma}_d^2 R_d \end{bmatrix} \tag{3.19}$$

Again, it is worth noting that the hyper-parameters of both the above mentioned GPs (kriging models), including $\hat{\mu}_c$, $\hat{\mu}_d$, $\hat{\sigma}_c^2$, $\hat{\sigma}_d^2$, $\hat{\theta}_c$ s, and $\hat{\theta}_d$ s, in addition to the correlation linear scale $\hat{\rho}$ are estimated through MLEs approach [79, 84]. To perform the co-kriging surrogate modelling, ooDace package in Matlab has been utilized in this work [85].

In Figure 3.1, an 1-D mathematical example is shown for the sake of comparing the ordinary kriging and co-krigigng results over the high-fidelity simulation data points with the true behavior of this expensive simulation. As observed in this figure, the co-kriging surrogate model provides much more precise results using the existing features in the low-fidelity model [84].

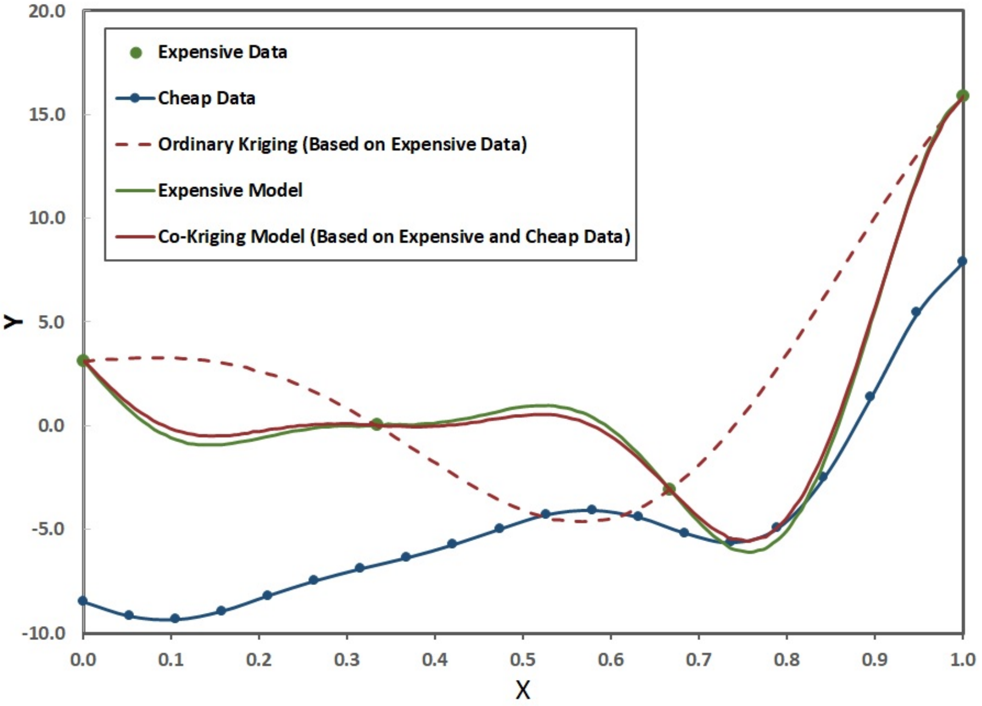


Figure 3.1: Ordinary kriging versus co-kriging surrogate responses over the high-fidelity simulation data points in a 1-D mathematical example. Adapted from [84, 85].

3.3.4 Experimental data

The experiments performed by Panchenko et al [64] have been utilized in this work to calibrate the precipitate model and to build the subsequent co-kriging surrogate model. These experiments are chosen due to the use of single crystal materials where the absence of grain boundaries results in homogeneous distribution of precipitates in the micro-structures after aging process. Therefore, the complexities associated with the heterogeneous nucleation at grain boundaries are avoided in experimental characterization and precipitation modelling. These experiments were carried out at 9 combinations of 4 initial nominal compositions (50.7, 50.8, 51 and 51.5 at% Ni), 3 aging temperatures (400, 500 and 550 °C), and 2 different aging times (1 and 1.5 hour(s)), as shown in Table 3.1. In this work [64], an transmission electron microscope (TEM) was applied for the micro-structural characterization which includes the precipitate inter-distance $(\bar{\lambda})$ and geometry (i.e., average size $(\bar{d})$ and thickness $(\bar{t})$). These measurements were used to calculate the average Ni concentration in the aged matrix $(\bar{c}^{'}_{Ni}\%)$ [64] and volume fraction of precipitates $(V_f\%)$ [58] by the following equations:

$$\bar{c}^{'}_{Ni} = \frac{\bar{c}_{Ni} - AV_f}{1 - V_f} \tag{3.20}$$

$$V_f = \frac{\bar{t}}{\bar{\lambda} + \bar{t}} \tag{3.21}$$

where $\bar{c}_{Ni}$ is the initial nominal Ni content and $A \approx 56.0$ is the approximate Ni atomic percent in $\mathrm{Ni_4Ti_3}$ phase. It should be noted that the uncertainties of all experimental data are reported in [64], except for the uncertainties of matrix Ni contents. Therefore, Equation 3.20 has been used to propagate uncertainty from volume fraction to Ni content according to the approach mentioned in Section 3.3.2.2. The results have been shown in Table 3.1.

3.4 Results and Discussion

3.4.1 Prior knowledge about the model parameters

The influential parameters for calibration have been recognized using the forward analysis of the MatCalc© model, which are matrix/precipitate interfacial energy (γ), diffusion correction (χ), nucleation site density (N_0), nucleation constant (ξ), and precipitate aspect ratio (H). High coherency between NiTi matrix and $\mathrm{Ni_4Ti_3}$ precipitates results in a low value of the interfacial energy. This parameter is one of the most influential parameters in the model since the nucleation rate al-

Table 3.1: Experimental Results and their uncertainties. Adapted from [64].

	Experimental Conditions			Experimental Results			
#	$\bar{c}_{Ni}(\%)$	$T(^oC)$	$t(s)$	$\bar{c}'_{Ni}(\%)$	$\bar{l}(nm)$	$V_f(\%)$	$\bar{d}(nm)$
1	50.7	400	5400	50.50±0.029	100±10	3.6±0.5	35±5
2	50.7	500	3600	50.49±0.029	220±20	3.8±0.5	100±10
3	50.7	550	5400	50.48±0.029	550±20	4±0.5	400±20
4	50.8	550	5400	50.53±0.029	510±20	5±0.5	360±20
5	51	400	3600	50.48±0.030	35±5	9.4±0.5	25±5
6	51	400	5400	50.49±0.030	45±5	9.3±0.5	35±5
7	51	500	3600	50.50±0.030	100±10	9.1±0.5	100±10
8	51	550	5400	50.51±0.030	400±20	9±0.5	400±20
9	51.5	550	5400	50.64±0.032	400±20	16±0.5	400±20

ters exponentially with the cube of the interfacial energy. Diffusion correction is an unknown value to modify Ni diffusivity for the real conditions. For example, there are probably some quenched-in excess vacancies in the micro-structure that might show a significant enhancement in the diffusivity. Moreover, the number of potential nucleation sites are calculated based on the micro-structural pictures captured in Panchenko's work [64]. In this regard, the inter-distance between precipitates can determine the number of nucleation sites per unit volume as $(\frac{1}{\lambda})^3$ with the assumption of a homogeneous precipitate distribution. The calculated value is not precise and should be modified due to the likely presence of some small undetectable particles or some inactivated nucleation sites. As a result, the parameter range has been extended to the larger values. Nucleation constant is a parameter in the nucleation rate equation of the MatCalc© model that is considered as a fine-tuning parameter for the modification of the nucleation rate based on the experimental data. Regarding the aspect ratio parameter, it is worth noting that the formed $\mathrm{Ni_4Ti_3}$ precipitates resemble a disk-like shape, thus its value should be less than 1 for the cylindrical precipitate assumption implemented in the model. The initial values and the range of the model parameters applied for the model calibration are listed in Table 3.2. In addition, a non-informative (uniform) PDF has been suggested for the prior distribution of the model parameters in this work due to the lack of statistical information about the mentioned model parameters.

Table 3.2: Initial value and range of the applied model parameters

	$\gamma(J/m^2)$	χ	$N_0(1/m^3)(\times 10^{18})$	ξ	H
$\theta^{(0)}$	0.0505 [60]	1	1-20000 [64]	1	0.24 [60]
$[\theta_{min}\text{-}\theta_{max}]$	0.01-0.1	0.01-10.0	1-100000	0.01-10.0	0.01-1

3.4.2 Bayesian calibration of the precipitation model

For the model calibration, a 1×5 parameter vector based on Table 3.2 is considered as the initial point $(\theta^{(0)}_{1\times 5})$ in the parameter space, and then new parameter vectors are randomly sampled from a posterior proposal distribution and selected using the defined criterion based on MH ratio in Section 3.3.2.1. First, all the above mentioned model parameters have been calibrated with each experimental data in Table 3.1 individually to find the plausible optimum value for the interfacial energy in each experimental case. These calibrations have been performed as a result of author's intuitions about the variation of interfacial energy with aging temperature and nominal composition. It should be noted that the other parameters are also taken into account in the calibrations because of probable correlations between interfacial energy and other applied parameters that can affect the plausible optimal value of interfacial energy in each calibration case. As a matter of illustration, the results of one of these calibrations is discussed in details as follows.

All the calibrations have been performed for 20000 generations. Here, calibration against experiment 1 from Table 3.1 has been selected as an example to show how the calibration approach works. After MCMC generations, the posterior PDFs of the parameters can be evaluated through their joint and marginal (posterior) frequency distributions. In Figure 3.2(a-e), the accepted parameters during MCMC can be observed which clearly differentiate the convergence region from the "burn-in period" using the red dotted lines. In this regard, it should be noted that the MCMC convergence is identified as fluctuations around a constant mean value rather than just a straight horizontal line with a constant value. Moreover, the burn-in period is some of the generated parameter vectors at the beginning of MCMC sampling (before the convergence region), which should be removed for the further evaluations of the model parameters [6]. Figure 3.3(a-l) demonstrates the joint frequency distributions for each two model parameters. In these plots, colors are representing different density of MCMC sampled points in the parameter space that increases from blue to red color spectrum. Therefore, the red regions can indicate the convergence regions for the model parameters. In addition, the linear correlation between parameters can be qualitatively evaluated by the linearity of the ellipsoidal features in the plots. In this context, higher ratio of major to minor axis (the longest over the shortest diameter) corresponds to higher linear correlation between the given pair of parameters; on the other hand, being closer to a circular shape where the diameters are less different means lower linear correlation. In the case that one of the axis is the density of nucleation sites (Figure 3.3(b), 3.3-e, 3.3-h, and 3.3-k), two dominant sharp regions can be observed in the sampling space. It seems the smaller one which is around the initial value for this parameter $(10^{21}\frac{1}{m^3})$ corresponds to the the burn-in period. This can be also distinguished as two horizontal lines without any considerable fluctuations in Figure 3.2(c) which are separated by a small transition region from about 7000^{th} to 8000^{th} generated samples. In addition, the projections of 3D joint distributions of the color plots in Figure 3.3 (where z axes is the frequency or number of samples) on xz and yz planes result in the marginal frequency distributions for each given parameter.

The marginal frequency distributions can also be obtained directly from the MCMC samples for each model parameter, as shown in Figure 3.4(a-e). The discussion about two sampling regions is more clear in Figure 3.3, not only for the density of nucleation sites but also for interfacial energy which is not detectable through the MCMC accepted sample and joint distribution plots in Figure 3.2(a) and 3.3, respectively. As observed in these figures, it seems MCMC sampling has been trapped around the initial values of these two parameters for a while as indicated by red arrows, and then suddenly moves toward their optimal value ranges as shown with blue arrows. It should be

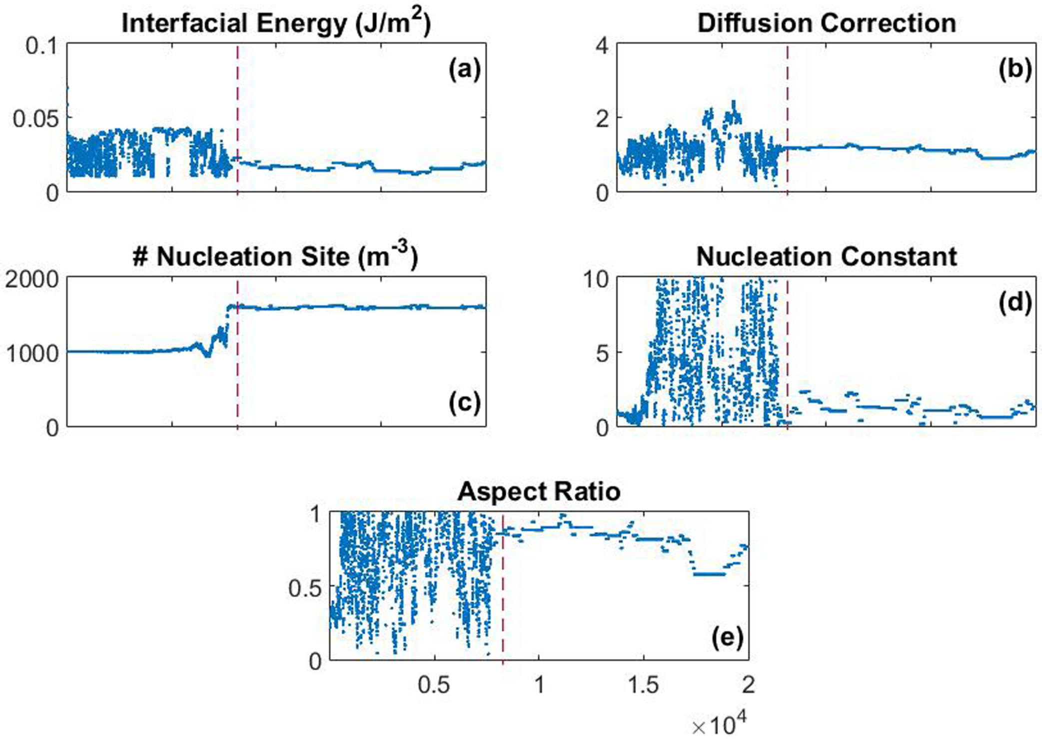


Figure 3.2: MCMC accepted parameters during calibration (the left side of the dashed lines indicate the burn-in period in the parameter space)

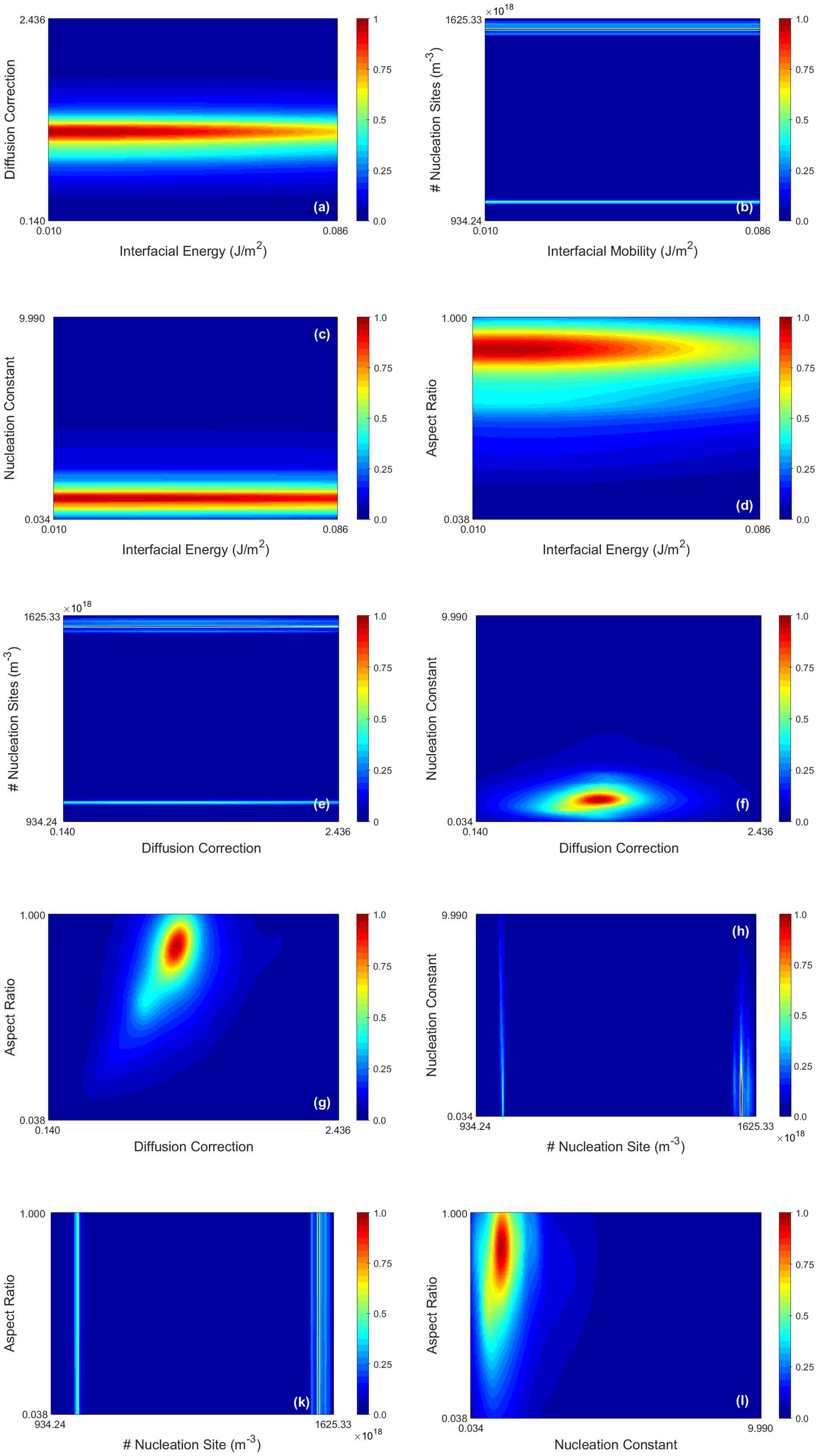


Figure 3.3: 2D Joint frequency distribution of each two model parameters with normalized color bars

noted that the second small peak for aspect ratio (Figure 3.4(e)) results from random noises during the convergence of this parameter and nothing to do with burn-in period in this case. This can be also observed in Figure 3.2(e) where a sudden drop in the sample values occurs close to the end of process, i.e., between 17000^{th} and 19000^{th} generated samples. This MCMC acceptance trend implies that aspect ratio probably has lower sensitivity compared to the other model parameters since its alteration shows less effects on the model results. Continuing sampling probably fixes this issue by generating more samples around the dominant peak that might annihilate the small peak.

Cumulative mean distribution of parameters are usually used to find the number of MCMC parameter generations that belongs to the burn-in period [6]. The cumulative mean plots in Figure 3.5(a-e) demonstrate a transition from noisy trends at the beginning to smooth changes towards constant values after 8000 parameter generations. Therefore, the first 8000 parameter generations are considered as the burn-in period and discarded from the total samples before further analysis of parameters. This period can also be determined through the trace plots in Figure 3.2 in this specific calibration case. Here, the small fluctuations around the mean value after parameters' convergence result in a distinct transition between the burn-in periods and the convergence regions, which is not always the case.

After the removal of the burn-in period, all the samples outside of the optimal value ranges of the parameters disappeared in their posterior frequency distributions, according to Figure 3.6. In other words, no second peak is observed in the case of interfacial energy and the density of nucleation sites. At this point, the remaining 12000 samples are used to calculate the mean values and variance-covariance matrix of the model parameters. The mean values and the square root of the matrix' diagonal elements can be introduces as the parameters' plausible optimal values and their uncertainties, respectively.

This calibration process have been performed exactly the same for the other individual experimental data as well. The results for calibrated parameters and their uncertainties have been listed in Table 3.3 for each experimental case. Using these calibrated parameters, the three model

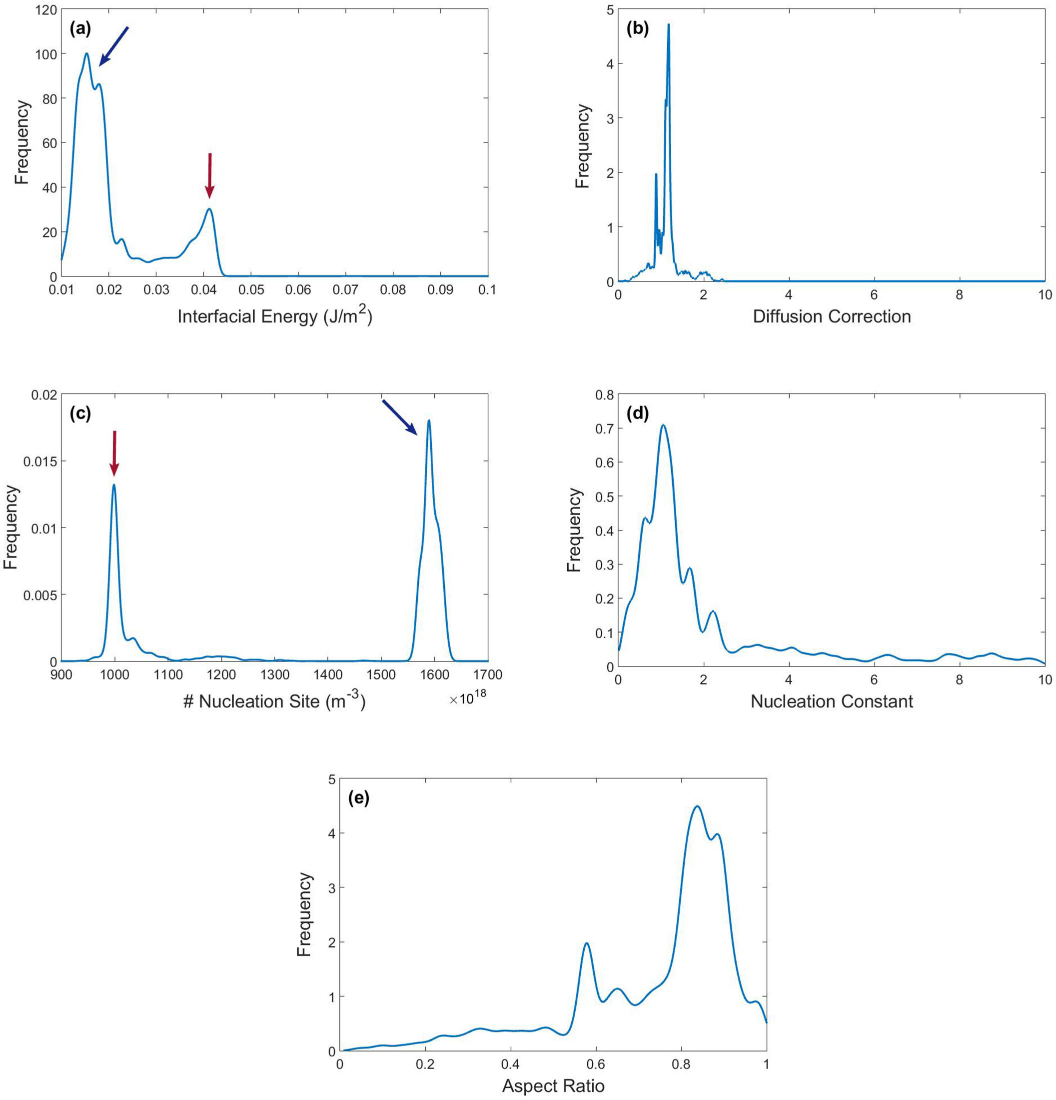


Figure 3.4: Marginal posterior frequency distributions of the model parameters after MCMC sampling

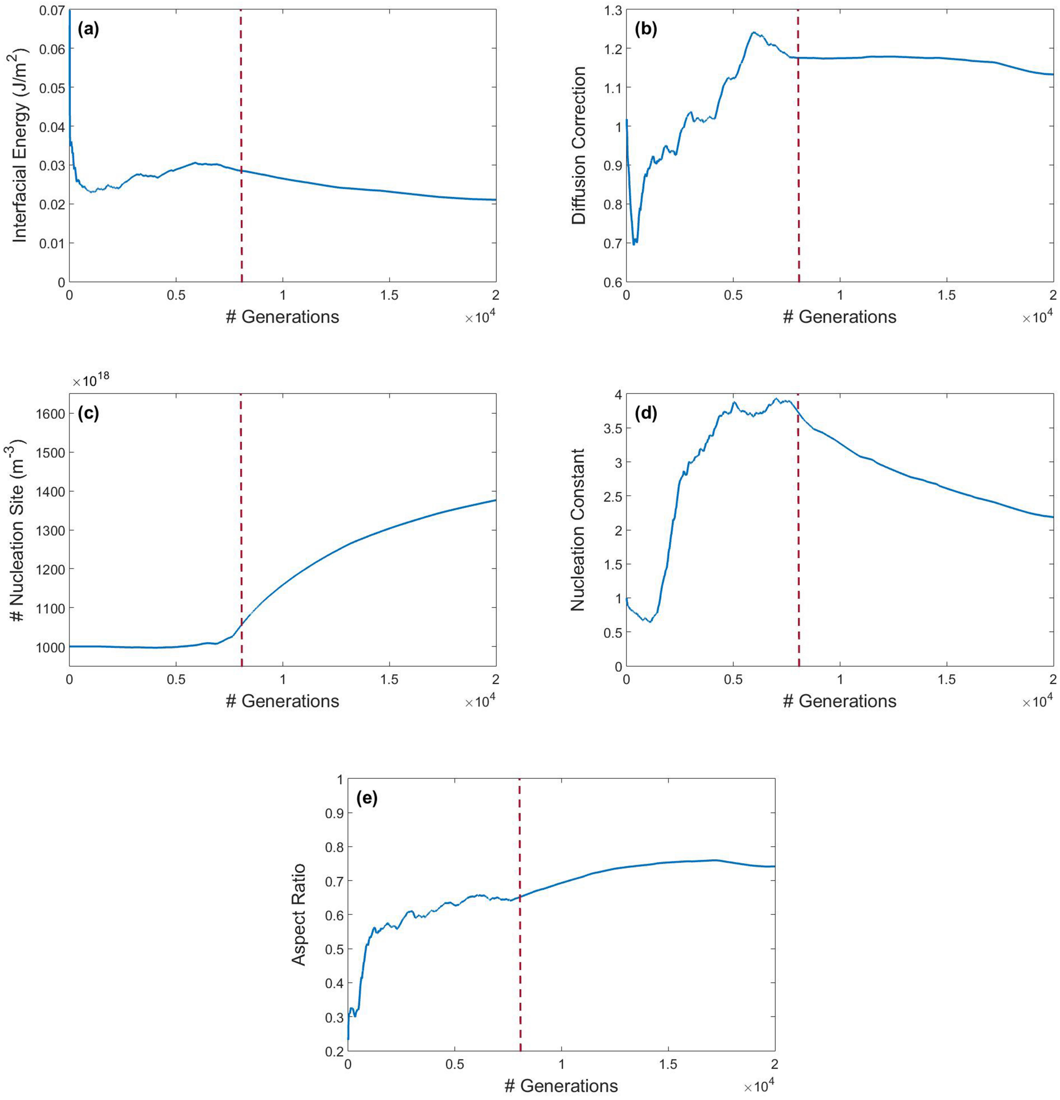


Figure 3.5: Cumulative mean distributions of the model parameters after MCMC sampling

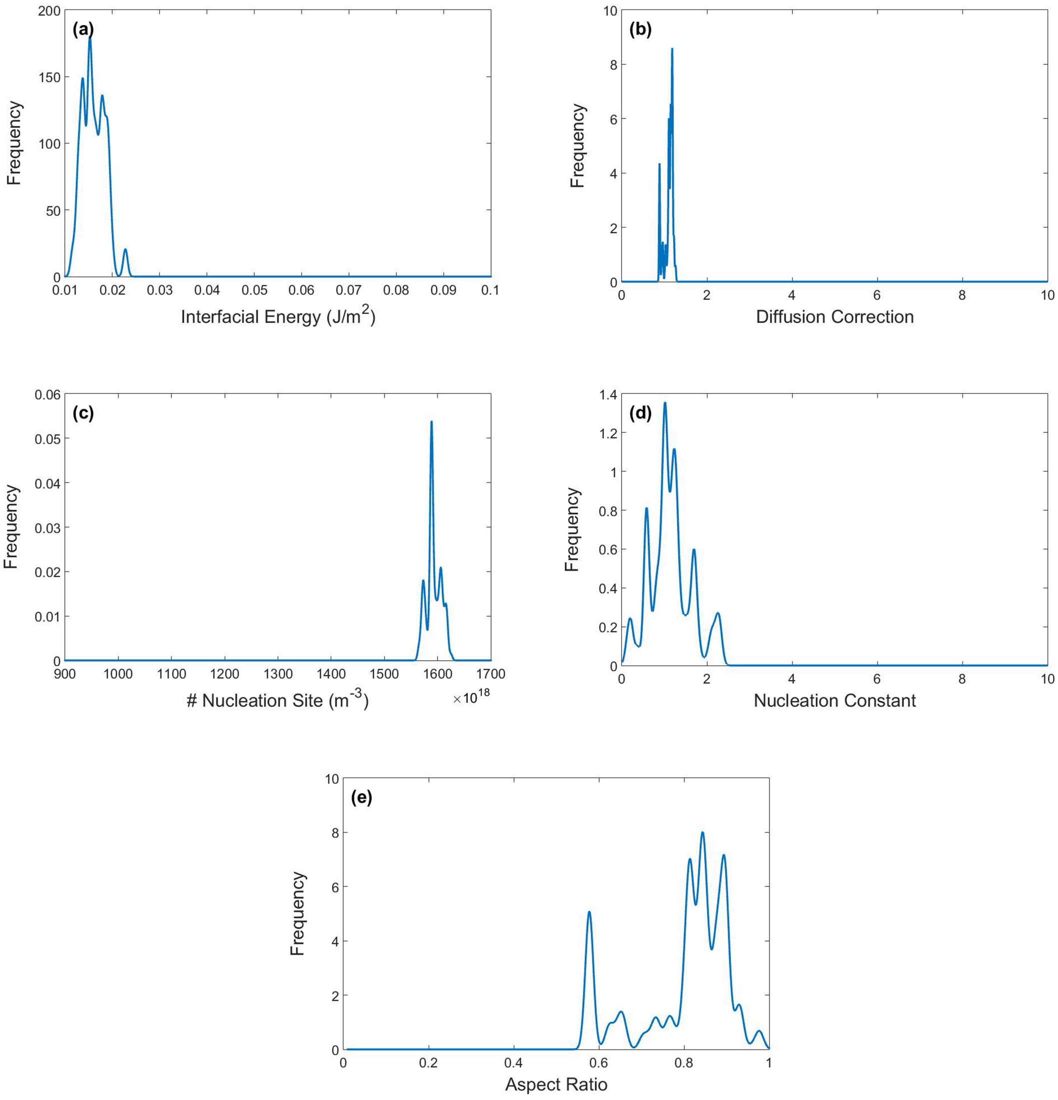


Figure 3.6: Marginal posterior frequency distributions of the model parameters after the removal of the burn-in period from MCMC generated samples

outputs have been calculated and compared with their corresponding experimental data in each case. Moreover, the parameters' uncertainties are also propagated to the model results using the propagation of uncertainty method introduced in Section 3.3.2.2. As can be observed in Table 3.4, there are very good agreements between the mean values of the model results and the experimental data in each calibration case, or at least the 95% Bayesian confidence intervals of the data show noticeable overlaps with the 95% Bayesian confidence intervals of the model results, where 95% Bayesian confidence intervals are defined as the mean values of data/model results plus/minus two times their corresponding standard deviations ($D/M \pm 2\sigma_{D/M}$).

Table 3.3: MCMC Calibrated values and uncertainties of the model parameters obtained for each experimental condition

	Experimental Condition			Model Parameter				
#	$\bar{c}_{Ni}(\%)$	$T(^oC)$	$t(s)$	$\gamma(J/m^2)$	χ	$N_0(1/m^3)(\times 10^{18})$	ξ	H
1	50.7	400	5400	0.016±0.0024	1.1±0.104	1592.87±13.815	1.15±0.491	0.80 ±0.107
2	50.7	500	3600	0.035±0.0013	0.37±0.095	87.27±2.763	5.17±1.768	0.65±0.210
3	50.7	550	5400	0.032±0.0009	0.95±0.215	8.75±3.368	3.50±2.385	0.71±0.228
4	50.8	550	5400	0.019±0.0016	0.29±0.014	1.75±0.172	8.42±0.528	0.35±0.019
5	51	400	3600	0.041±0.0017	1.28±0.400	19999.88±0.182	1.13±0.562	0.63±0.248
6	51	400	5400	0.045±0.0011	1.68±0.419	10456.18±638.232	4.54±2.728	0.68±0.239
7	51	500	3600	0.039±0.0003	0.36±0.042	1001.10±0.056	2.99±0.791	0.80±0.138
8	51	550	5400	0.035±0.0002	0.80±0.063	19.18±0.381	1.85±0.375	0.82±0.101
9	51.5	550	5400	0.023±0.0035	0.54±0.524	4.26±3.605	4.29±1.258	0.87±0.156

As mentioned at the beginning of this section, the goal for the model calibration against each experimental data is to find any possible relationship for interfacial energy in terms of nominal chemical composition and aging temperature, which is very complex and expensive to be determined experimentally. For this reason, the calibrated interfacial energy values have been plotted in terms of aging temperature, T(°C), and nominal Ni content, Ni(at%), and then a polynomial surface has been fitted over the data points, as shown in Figure 3.7. The corresponding fitting equation is:

Table 3.4: Comparison of model results and experimental data obtained after MCMC calibration against each experimental data individually

Model Input				Model Output			Experiment		
#	$\bar{c}_{Ni}(\%)$	$T(^oC)$	$t(s)$	$\bar{c}'_{Ni}(\%)$	$V_f(\%)$	$\bar{d}(nm)$	$\bar{c}'_{Ni}(\%)$	$V_f(\%)$	$\bar{d}(nm)$
1	50.7	400	5400	50.50±0.000	3.6±0.0	35.0±0.0	50.50±0.029	3.6±0.5	35.0±5.0
2	50.7	500	3600	50.50±0.023	3.4±0.5	99.8±2.3	50.49±0.029	3.8±0.5	100.0 Âś 10.0
3	50.7	550	5400	50.50±0.036	3.1±0.7	399.9±3.8	50.48±0.029	4.0±0.5	400.0±20.0
4	50.8	550	5400	50.50±0.016	4.3±0.4	359.9±1.8	50.53±0.029	5.0±0.5	360.0±20.0
5	51	400	3600	50.50±0.035	7.6±0.6	25.1±2.3	50.48±0.030	9.4±0.5	25.0±5.0
6	51	400	5400	50.50±0.018	7.7±0.5	34.8±1.8	50.49±0.030	9.3±0.5	35.0±5.0
7	51	500	3600	50.50±0.010	7.9±0.4	99.8±1.3	50.50±0.030	9.1±0.5	100.0±10.0
8	51	550	5400	50.50±0.031	7.7±0.6	400.2±2.8	50.51±0.030	9.0±0.5	400.0±20.0
9	51.5	550	5400	50.61±0.101	13.6±1.4	399.2±12.0	50.64±0.032	16.0±0.5	400.0±20.0

$$\begin{aligned}\gamma_{\alpha/\beta} &= 8.051532780404465 + 0.026548307462029.T + 22.565282846398304.Ni \\ &- 11.341568269633328.Ni^2 - 0.000522121579171.T.Ni \\ &+ 0.016910379084767.Ni^3 - 1.157795038899461.Ni^4 \\ &- 0.000001076437280.Ni^6\end{aligned} \tag{3.22}$$

After inserting this equation in the model, the other four model parameters have been calibrated with all nine experimental data together, based on the same approach mentioned earlier. The results for calibrated parameters and the corresponding model outputs for all the experimental cases have been reported in Tables 3.5 and 3.6, respectively. As can be observed, there are some discrepancies between the model responses and experimental data due to very high uncertainties in this problem. These large uncertainties probably result from very high influences of the applied parameters on model responses, missing physics in the precipitation model, and high uncertainties in the experimental measurements. For example, model shows very high sensitivity to interfacial energy according to Equation 3.3 since the nucleation rate exponentially changes with the cube of this parameter's value. Therefore, a slight change in interfacial energy can considerably alter

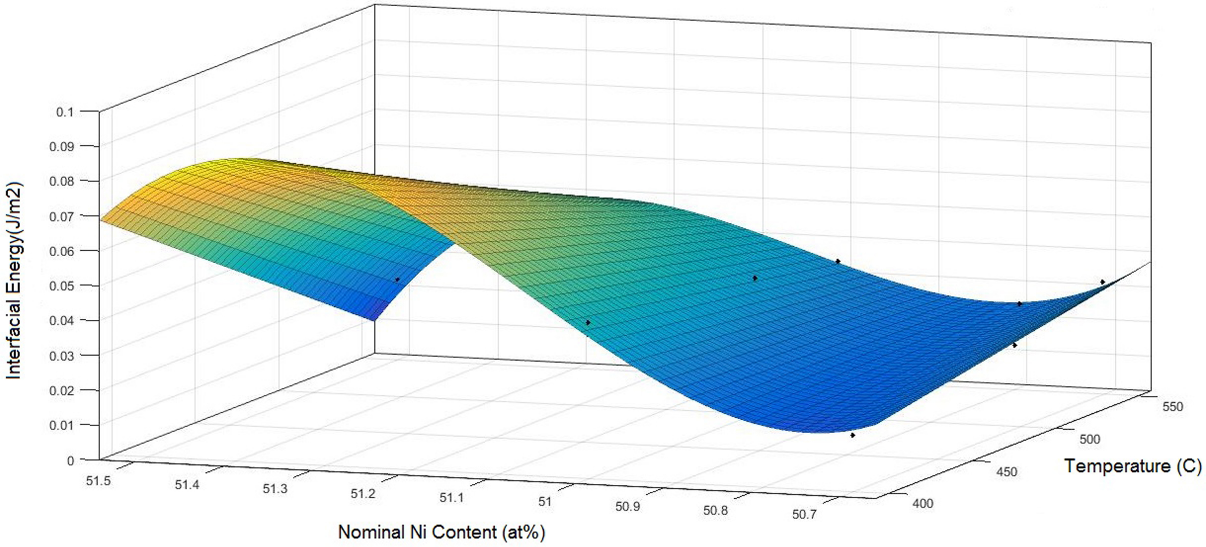


Figure 3.7: The relationship of the interfacial energy with nominal Ni Composition and aging temperature

the model responses. Moreover, as mentioned in Section 3.3.1, the precipitation model is based on classical theory of nucleation and growth that includes many simplification assumptions and incomplete physics. Significant amount of experimental uncertainties also result from the TEM analysis since it is restricted to a very small 2-D image as a representative of the whole bulk material. Moreover, very small particles may not be completely visible in the microscopic images that can considerably affect the reported mean precipitate sizes unlike the precipitate volume fractions and the matrix chemical compositions. The reproducibility of the applied MCMC approach has also been checked against the synthetic data obtained from the calibrated model (results in Table 3.6) in order to ensure that no noticeable uncertainties result from the calibration approach in this case study.

3.4.3 Co-kriging surrogate modelling for precipitate volume fraction

Remarkable differences between the obtained model results and the experimental data has been an important reason to use co-kriging technique for more precise prediction of the experimental precipitation responses (as the high-fidelity model) throughout the entire design space of the pre-

Table 3.5: Optimal values of the model parameters and their uncertainties after MCMC calibration against all the experimental data together

χ	$N_0(1/m^3)(\times 10^{18})$	ξ	H
2.179±1.293	39238.286±29309.092	0.491±0.298	4.058±3.018

Table 3.6: Comparison of model results and experimental data obtained through MCMC calibration of four model parameters against all the experimental data together after using the interfacial energy equation in the model

	Model Input			Model Output			Experiment		
#	$\bar{c}_{Ni}(\%)$	$T(^oC)$	$t(s)$	$\bar{c}'_{Ni}(\%)$	$V_f(\%)$	$\bar{d}(nm)$	$\bar{c}'_{Ni}(\%)$	$V_f(\%)$	$\bar{d}(nm)$
1	50.7	400	5400	50.27±0.177	6.2±2.32	18.4±12.8	50.50±0.029	3.6±0.50	35.0±5.0
2	50.7	500	3600	50.10±0.005	8.7±0.09	32.0±18.1	50.49±0.029	3.8±0.50	100.0±10.0
3	50.7	550	5400	50.17±0.047	7.9±0.07	93.6±44.7	50.48±0.029	4.0±0.50	400.0±20.0
4	50.8	550	5400	50.17±0.046	9.3±0.03	59.9±25.6	50.53±0.029	5±0.50	360.0±20.0
5	51	400	3600	50.38±0.300	9.0±3.96	33.2±15.8	50.48±0.030	9.4±0.50	25.0±5.0
6	51	400	5400	50.15±0.188	12.1±2.35	38.8±19.0	50.49±0.030	9.3±0.50	35.0±5.0
7	51	500	3600	50.10±0.002	13.0±0.06	51.0±26.4	50.50±0.030	9.1±0.50	100.0±10.0
8	51	550	5400	50.17±0.044	12.1±0.05	82.7±36.1	50.51±0.030	9±0.50	400.0±20.0
9	51.5	550	5400	50.17±0.046	19.3±0.03	59.7±27.8	50.64±0.032	16±0.50	400.0±20.0

cipitation process in Ni-Ti SMAs. As mentioned in Section 3.3.3, this is performed by building a co-kriging surrogate model on the experimental data points based on the Kennedy and O'Hagan approach [12]. In this approach, a linear correlation between low- and high-fidelity models is considered to take advantage of the physics included in the precipitation model for better predictions of experimental behavior.

This method can be applied for all the three precipitation model results in this work; however, precipitate volume fraction has been selected as an example in this section to show the corresponding co-kriging prediction. According to the information observed in Table 3.6, nine data points for each low- and high-fidelity models have been utilized to create the 4-D co-kriging surrogate model in terms of Nominal Ni composition and heat treatment temperature and time. For better visualization, the mean response of the co-kriging surrogate model (Blue surfaces) and this response plus

95% confidence interval borders (green surfaces) have been projected in 3-D space for each two model inputs in Figures 3.8(a-c) and 3.8(d-f), respectively. Consequently, it can be stated that the experimental precipitate volume fractions are situated between the borders shown in 3.8(d-f) with 95% probability, which is a very important finding to reduce the cost in robust design of SMAs precipitation hardening.

3.5 Summary and Conclusion

In this work, the NKW precipitation model in MatCalc© for Ni-Ti SMAs has been probabilistically calibrated against experimental data [64] for three model objectives, including mean precipitate size, precipitate volume fraction, and average Ni content in matrix. A MCMC approach in the context of Bayesian statistics has been applied for this multi-objective problem to calibrate the influential parameters in this model and quantify their uncertainties.

First, the model parameters have been calibrated against each experimental data individually in order to find a relationship for interfacial energy in terms of nominal Ni composition and aging temperature, which is very difficult to do experimentally. After inserting the obtained polynomial equation for interfacial energy in the model, the other influential parameters have been calibrated with all available experimental data together. After the uncertainty propagation from the parameters to the model results, some discrepancies have been observed between the model results and the corresponding experimental data due to high uncertainties in both the model structure and the data. For more precise prediction of precipitation behavior, the co-kriging surrogate modelling in the framework of Kennedy and O'Hagan approach [12] has been applied in this work. As a manner of illustration, the experimental precipitate volume fractions and their uncertainties have been predicted through the construction of a co-kriging surrogate model in terms of nominal Ni composition and heat treatment temperature and time in a 4-D design space. This result is very important in the case of robust design since the response of the high-fidelity experimental behavior which is very close to the real response of the system can be identified with 95% confidence.

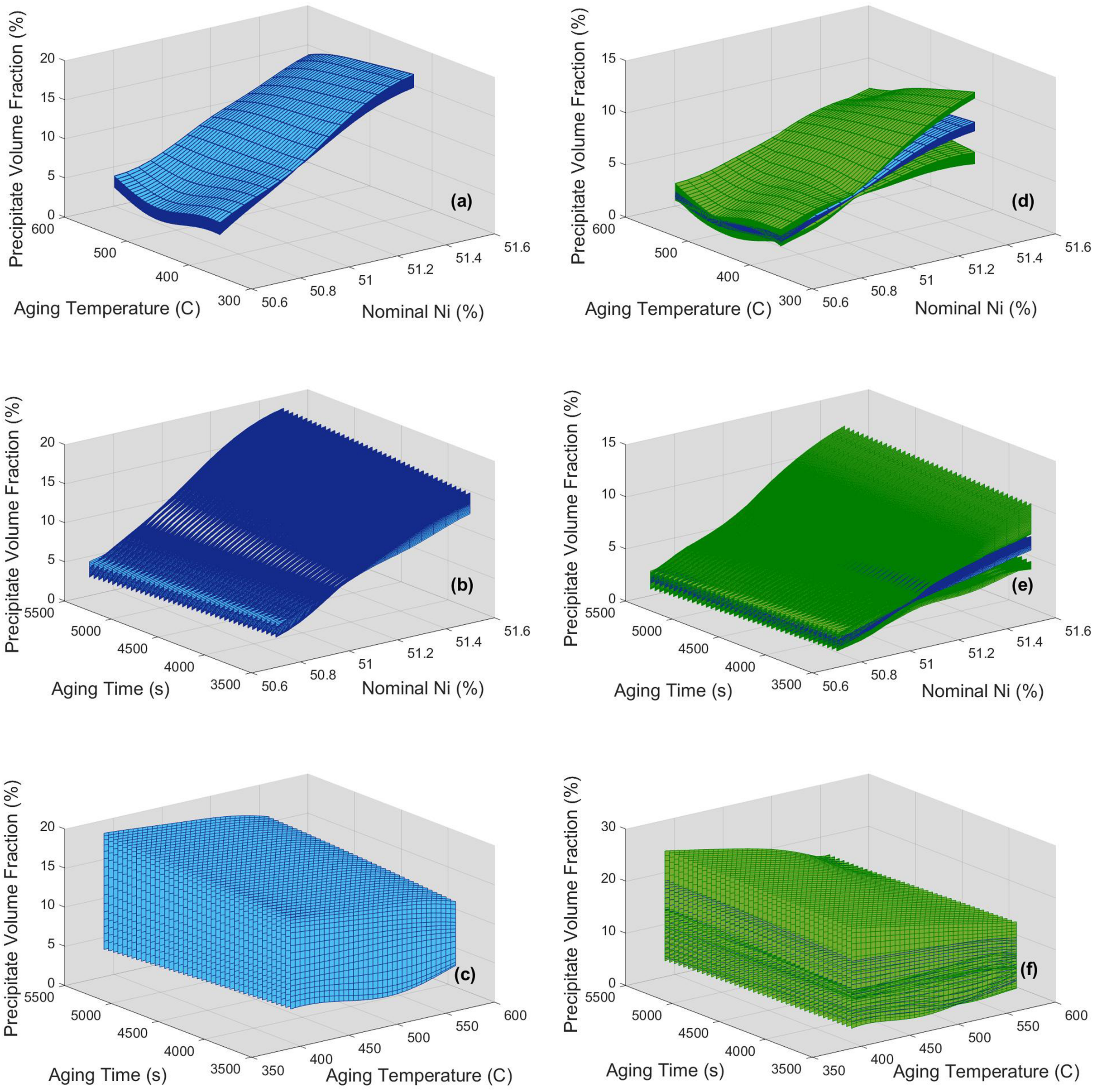


Figure 3.8: a-c) the mean response of the co-kriging surrogate model (Blue surfaces) and d-f) mean response besides 95% confidence interval borders (green surfaces)

4. UNCERTAINTY QUANTIFICATION OF THE PARAMETERS AND PREDICTIONS OF A PHENOMENOLOGICAL CONSTITUTIVE MODEL FOR THERMALLY INDUCED PHASE TRANSFORMATION IN NI-TI SHAPE MEMORY ALLOYS [*]

4.1 Overview

In this work, a probabilistic calibration approach has been adopted to quantify the uncertainties of the parameters and predictions of a phenomenological Shape Memory Alloys (SMAs) constitutive model which has been used to predict the thermally induced phase transformation of Ni-Ti SMAs. Furthermore, the impact of the adopted calibration method, which enables the determination of the uncertainty bounds of the model predictions, on the design of robust engineering applications has been discussed. To this end prior to the probabilistic model calibration, a design of experiments (DOE) has been performed in order to identify the most influential parameters on the response of the system and thus reduce the dimensionality of the problem. Subsequently, uncertainty quantification (UQ) of the influential parameters has been carried out through Bayesian Markov Chain Monte Carlo (MCMC). The assessed uncertainties in the model parameters has been then propagated to the model predictions using an approximate approach based on the variance-covariance matrix of the MCMC-calibrated model parameters and then an explicit propagation of uncertainty through MCMC-based sampling. The determined 95% Bayesian confidence intervals of the model predictions, by the latter methods, has been demonstrated and compared. Additionally, good agreement between the experimentally measured and model predicted SMA hysteresis loops has been observed where the experimental data are situated within the predicted 95% Bayesian confidence intervals. Finally, the application of the MCMC-based UQ/UP approach in decision making for experimental design has also been shown by comparing the information that can be gained by performing multiple repetitive experiments under identical thermo-mechanical

---

[*] Republished with permission of IOP Publishing, from "Uncertainty quantification of the parameters and predictions of a phenomenological constitutive model for thermally induced phase transformation in Ni-Ti shape memory alloys", Pejman Honarmandi, Alex Solomou, Raymundo Arroyave, Dimitris Lagoudas, 27, 3, 2019; permission conveyed through Copyright Clearance Center, Inc.

conditions versus experiments under different conditions.

4.2 Introduction

Nickel-Titanium (Ni-Ti) alloys are one of the most common shape memory alloys (SMAs), which are widely used in different engineering applications, such as biomedical devices and implants, microelectromechanical systems, sensors and actuators, seismic protection tools, and aerospace products and structures [86, 87]. Such applications are enabled by either the shape memory effect or super-elasticity, which in turn are the result of a reversible thermoelastic martensitic transformation that is triggered by applying thermal and/or mechanical loads [88, 89].

The macroscopic thermal actuation response of SMAs enables actuators with high specific weight ratio compared to the conventional alternatives [90,91]. Thermal actuation can occur during thermal cycles under iso-baric condition where inelastic strain is recovered through the forward and reverse martensitic phase transformation induced during cooling and heating [92]. However, the forward and reverse transformations take place at different temperatures due to energy dissipation during this cyclic thermal process [93], which results in a hysteresis loop in strain-temperature space.

Under the framework of Integrated Computational Materials Engineering, completing the loop along the process-structure-property-performance chain through multiscale modeling is a necessary (albeit not sufficient) condition for the efficient and robust design of materials for specific applications. In the case of Ni-Ti SMAs, different thermodynamics and kinetics modeling are usually applied to connect process and structure, such as CALculation of PHAse Diagram (CALPHAD) [59, 94], precipitation of secondary phases [60, 95], and phase field models [66, 67, 96]. In regard to the structure-property linkages, a number of constitutive models have been proposed over the recent decades to predict the thermo-mechanical response of Ni-Ti SMAs. These models can be categorized into phenomenological [87–89, 97–103] or micromechanical-based models [92, 104–107] which have been reviewed thoroughly by Patoor et al. [93] and Lagoudas et al. [108]. Generally, micro-mechanical models are high-fidelity and expensive models that use microstructural information to predict the macroscopic behavior of SMAs while phenomenological models are cheap and fast since the free energy associated with a non-microscopic homogenized material volume is taken into account for the prediction of the macroscopic responses [99].

Despite a substantial number of constitutive models for the response of SMAs, there are a very few studies associated with analysis, calibration, and uncertainty quantification (UQ) of the model parameters and subsequent uncertainty propagation (UP) from these parameters to the model responses. In this regard, the sensitivity of the design outputs to the uncertainty of the design input variables has been evaluated for SMA morphing structures by Oehler et al [109] through an iterative procedure based on a quasi-Monte Carlo simulation. Martowicz et al [110] also performed some preliminary work on the sensitivity analysis and uncertainty quantification of a finite element based numerical model aided by meta-modeling which predicts the phase transformation response of a SMA bumper due to external thermal and mechanical stimuli. In this work, the influential input variables and the uncertainty of the output variables have been determined using central finite difference approach and Monte Carlo simulation, respectively. In another work, Enemerk et al [111] have applied an adaptive Markov Chain Monte Carlo (MCMC) approach to quantify the parameter uncertainties and correlations of a thermo-mechanical model which predicts the cyclic response of helical springs made of pseudo-elastic SMAs. Generally, it should be noted that lack of thorough sensitivity and uncertainty analysis in computational modeling is a general issue in different areas of materials science and engineering. For this reason, our developed UQ and UP frameworks have been applied for plastic flow behavior in TRIP steels [69] as well as calculation of phase diagrams (CALPHAD) modeling [70] to spread the use of these frameworks and show their importance in computational materials design.

Generally, uncertainties can result from different sources, including natural uncertainty (NU) from the random nature of the system, model parameter uncertainty (MPU) due to lack of knowledge and data about the model parameters, and Model structure uncertainty (MSU) as a result of incomplete or simplified physics in the model [3]. For example, the cyclic thermomechanical responses of an SMA material with the same chemistry and processing history can be slightly different during exact replications of an iso-baric experiment due to the inherent randomness (NU) in the system. In addition, there are different physical parameters in the phenomenological- or micromechanical-based models, such as transformation temperatures, which have some uncertainties (MPU) in their experimental measurements. It should be noted that physical models are usually incomplete and unable to capture all the physics in the system, including the existing thermo-mechanical models. For example, MSU in most of these models can be attributed to the absence of smooth phase transformations at austenite/martensite start and finish temperatures, independency of the applied stress to the maximum transformation strain at low stress levels and/or independency of the transformation hysteresis area to the applied stress in the prediction of the cyclic thermo-mechanical behavior of SMAs. Although the thermo-mechanical model in the current work (described in section 2) addresses the above-mentioned modeling deficiencies, the model assumptions can still result in some MSU. The consideration of the quadratic dependency of Gibbs free energy of phases to the applied stress and linear relationship between the applied stress and transformation temperatures and between the rates of martensitic volume fraction and the transformation strain are some of the main assumptions in this model. The importance of UQ and UP, which tend to be ignored in deterministic approaches towards model parameterization, is noticeably highlighted in the case of robust design where it is sought that the system response be relatively independent of the uncertainties of its variables/parameters under the operating conditions. These inverse and forward uncertainty analyses are discussed in more detail as follows.

The estimation of model parameters in order the model outputs to predict the available (experimental) data, is an inverse problem, which can be solved through deterministic or probabilistic approaches [112]. Deterministic calibration such as the least squares approach yields a single best estimate based on the minimization of the squared difference between the model results and data, while probabilistic calibration provides probability distributions and uncertainty bounds for the model parameters. It is worth noting that in some cases different combinations of the model parameters usually lead to similar model outputs, but, by definition, deterministic calibration only arrives at one best estimate. Probabilistic calibration, on the other hand, provides an ensemble of possible combinations of model parameter values in the form of probability distributions [113].

In different engineering fields, deterministic approaches have commonly been used to calibrate the model parameters leading to deterministic model predictions. However, it is widely recognized that probabilistic calibration approaches can result in more informative predictions because they enable the quantification of the uncertainties of the model parameters which can be subsequently propagated to the model results, thus offering model predictions with defined uncertainty/confidence bounds [114].To this end, provided the design requirements of a particular application, the latter models can be used to perform robust design by designing the components associated with the application such that they meet the design requirements despite the variability on the model predictions due to the uncertainties of the model parameters. On the other hand, the deterministic models cannot lead to robust designs due to the lack of information regarding the uncertainties of their parameters and predictions.

Among probabilistic approaches, Bayesian-based MCMC algorithms have become the most commonly used tools due to the fast development in computing capabilities [4]. These sampling techniques are more robust and simpler compared to the traditional analytic and numerical approaches to solve high-dimensional intractable integrals in the probabilistic calibration process [5, 6]. In addition, MCMC methods can be used to sample from complex multivariate distributions that are usually very hard to sample using other sampling techniques such as inversion or rejection sampling [6].

MCMC sampling, however, is an expensive method, particularly in high-dimensional cases where the parameter convergence in the calibration process can be very time-consuming. A reduction in the dimensionality of the parameter space through sensitivity analysis (SA), however, can significantly improve the MCMC calibration efficiency. Sensitivity analysis (SA) is often applied to determine the impact of the parameters' variability/uncertainty on the total variability/uncertainty of the model results [7] in order to find the insensitive model parameters that can be eliminated from the calibration process. In this regard, design of experiments (DOE) is the most common approach, which requires a description of experiments (such as complete factorial design (CFD) offered by Fisher [8] or fractional factorial design (FFD) to consider possible combinations of pre-defined parameters' levels) besides a subsequent analysis of the experiments (such as ANalysis Of VAriance, ANOVA).

In robust design, the final results of the model and their corresponding uncertainties are the main interest. Therefore, a UP technique should be applied by considering the trade-off between cost and information loss to propagate uncertainties from the model parameters to the model outputs [115]. This forward problem can approximately be solved through different methods with different cost and precision, e.g., forward model analysis of the parameters' optimal combinations, statistical second moment methods, polynomial chaos expansion, etc.

In the present work, the parameter SA and MCMC calibration have been performed for the phenomenological constitutive model proposed by Lagoudas et al. [99]. This model has been selected due to the consideration of three important features in SMAs which are usually ignored in this type of modeling: 1) the smooth transitions in thermo-mechanical responses of SMAs during forward and reverse phase transformations that take place in a range of temperatures and/or mechanical loadings, 2) the dependency of maximum transformation strain after full transformation on the magnitude of the applied stress, and 3) the definition of a critical driving force in terms of the magnitude of the applied stress and the transformation direction [99]. In section 2, this thermo-mechanical model is described thoroughly. In section 3, parameter SA using the combination of CFD and ANOVA are explained in details and subsequently utilized to perform the sensitivity assessment of the model parameters. In section 4, the applied MCMC sampling algorithm is described and used to probabilistically calibrate the influential model parameters, and then the parameter uncertainties are propagated to the model outcomes using different UP approaches. In section 5, the information gain of two experimental designs are compared together using the relative entropies obtained after the sequential MCMC training of the model parameters with two different synthetic data-sets. This is followed by the summary and conclusion in section 6.

4.3 Description of the Constitutive Thermo-Mechanical Model

In this section, the phenomenological SMA constitutive model proposed by Lagoudas et al. [99] is described in detail. The capability to predict the thermo-mechanical responses under general temperature-loading paths for wide ranges of SMAs with different processing histories makes this model powerful and usable for different applications. In the previous models [100, 116, 117], this capability could be achieved by high modeling cost due to increased complexity or parameter numbers.

In this model, a total Gibbs free energy is defined for the SMA material during the phase transformation in terms of the thermos-elastic contributions of austenite (G^A), martensite (G^M) and their interaction (G^{mix}) as follows,

$$G(\sigma, T, \varepsilon^t, \xi, g^t) = (1-\xi)G^A(\sigma, T) + \xi G^M(\sigma, T) + G^{mix}(\sigma, \varepsilon^t, g^t) \tag{4.1}$$

where, the external state variables σ and T are the applied stress tensor and the absolute temperature, and the internal state variables ϵ^t, ξ, and g^t account for the inelastic strain produced during forward/reverse martensitic transformation, the total volume fraction of martensite which accounts for all martensitic variants, and the transformation hardening energy, respectively. Moreover, G^A, G^M, and G^{mix} according to Lagoudas et al. [99] can be written as follows,

$$\begin{aligned} G^{A/M} &= -\frac{1}{2\rho}\sigma : S^{A/M}\sigma - \frac{1}{\rho}\sigma : \alpha(T - T_0) \\ &+ c^{A/M}\left[(T - T_0) - T\ln\left(\frac{T}{T_0}\right)\right] - s_0^{A/M}T + u_0^{A/M} \end{aligned} \tag{4.2}$$

$$G^{mix}(\sigma, \varepsilon^t, g^t) = -\frac{1}{\rho}\sigma : \varepsilon^t + \frac{1}{\rho}g^t \tag{4.3}$$

The constant parameters ρ and α are the alloy density and the thermal extension tensor, and the phase-dependent parameters S, c, s_0, and u_0 corresponds to the compliance tensor, specific heat, specific entropy, and specific internal energy, respectively. It should be noted that the superscripts A and M denote quantities corresponding to the austenitic and martensitic phases while the notation $(. : .)$ shows the inner product of two second-order tensors.

The present phenomenological SMA model captures the effective response of the SMA material without considering the detailed microscale behavior of the martensitic variants. Hence any change in the microstructural state of the material is depicted macroscopically by the change of the scalar quantity ξ. To this end the evolution equations of internal state variables ϵ^t and g^t are defined in terms of rate of ξ.The evolution equation of ϵ^t is given as [99],

$$\dot{\varepsilon}^t = \Lambda^t \dot{\xi} \ , \ \ \Lambda^t = \begin{cases} \Lambda^t_{fwd} \ , \ \dot{\xi} > 0 \\ \Lambda^t_{rev} \ , \ \dot{\xi} < 0 \end{cases} \tag{4.4}$$

where, Λ^t represents the transformation direction tensor, which is expressed for both directions as,

$$\Lambda^t_{fwd} = \frac{3}{2} H^{cur} \Big(\frac{\sigma^{'}}{\bar{\sigma}} \Big) \ , \ \ \Lambda^t_{rev} = \frac{\epsilon^t}{\xi} \tag{4.5}$$

$\bar{\sigma} = \sqrt{(3/2)\sigma' : \sigma'}$ is the effective stress where $\sigma^{'}$ is the deviatoric stress, and H^{cur} is the maximum transformation strain after full martensitic transformation as a function of the effective stress,

$$H^{cur}(\bar{\sigma}) = H_{sat}(1 - e^{-k\bar{\sigma}}) \tag{4.6}$$

where, H_{sat} is a saturated value for the maximum transformation strain at sufficiently high value of the effective stress, and the parameter k determines the exponential rate of H^{cur} variation from 0 to H_{sat} [92, 99].

In the same manner, the evolution equation of g^t is given as,

$$\dot{g}^t = f^t \dot{\xi} \ , \ \ f^t = \begin{cases} f^t_{fwd} \ , \ \dot{\xi} > 0 \\ f^t_{rev} \ , \ \dot{\xi} < 0 \end{cases} \tag{4.7}$$

where, f^t is a direction-dependent hardening function, which has been introduced in this model in order to capture the smooth transitions from the elastic to transformation regimes and vice versa. The expression of f^t for the forward and reverse transformation is given as follows,

$$\begin{cases} f^t_{fwd}(\xi) = \frac{1}{2}a_1(1 + \xi^{n_1} - (1-\xi)^{n_2}) + a_3 \\ f^t_{rev}(\xi) = \frac{1}{2}a_2(1 + \xi^{n_3} - (1-\xi)^{n_4}) - a_3 \end{cases} \tag{4.8}$$

where, n_i can take values in the range of (0, 1]. The value of n_i closer to zero corresponds to more smooth transition at the beginning or the end of the forward phase transformation (in the case of $i = 1$ or 2, i.e., n_1 or n_2) or their counterparts for the reverse phase transformation (in the case of $i = 3$ or 4, i.e., n_3 or n_4).

Provided Equations 4.1 to 4.8, the first and second law of thermodynamics can be utilized to derive the remaining model equations. Hence, at a local point in the material, the conservation of energy or first law of thermodynamics can be expressed as,

$$\rho\dot{u} = \sigma : \dot{\varepsilon} - \text{div}(q) + \rho r \tag{4.9}$$

where, $\dot{u}$, q, and r denote the internal energy rate, the heat flux vector, and the internal heat generation rate, respectively. In the same manner, the second law of thermodynamics at this local point can also be written as the Clausius-Planck inequality [118],

$$\rho\dot{s} + \frac{1}{T}\text{div}(q) - \frac{\rho r}{T} \geq 0 \tag{4.10}$$

By multiplying the two sides of the inequality 4.10 by T and substituting div(q) by its equivalent obtained from Equation 4.9, the following inequality can be obtained,

$$\rho\dot{s}T + \sigma : \dot{\varepsilon} - \rho\dot{u} \geq 0 \tag{4.11}$$

It should be noted that Gibbs free energy can be defined in terms of internal energy using the Legendre transformation, as follows,

$$G = u - \frac{1}{\rho}\sigma : \varepsilon - sT \tag{4.12}$$

Hence by differentiating Equation 4.12 in respect to time and substitute the obtained expression of $\rho \dot{s} T$ in Equation 4.11 the following thermodynamic constraint is derived,

$$-\rho \dot{G} - \dot{\sigma} : \varepsilon - \rho s \dot{T} \geq 0 \tag{4.13}$$

Using the chain rule, the second law of thermodynamics can turn into,

$$-\rho\Big(\partial_{\sigma} G : \dot{\sigma} + \partial_{T} G : \dot{T} + \partial_{\varepsilon^t} G : \dot{\varepsilon}^t + \partial_{\xi} G : \dot{\xi} + \partial_{g^t} G : \dot{g}^t\Big) - \dot{\sigma} : \varepsilon - \rho s \dot{T} \geq 0 \tag{4.14}$$

In inequality 4.14, $-\rho\partial_{\xi}G = p$, $-\rho\partial_{\varepsilon^t}G = \sigma$, and $-\rho\partial_{g^t}G = -1$ are three generalized thermodynamic forces. Following the Coleman and Noll approach [119] which states that the inequality 4.14 must hold true for every thermodynamic path, including the paths where all the variables are fixed in time except one, which is evolving, the total infinitesimal strain and entropy can be calculated and expressed,respectively as follows,

$$\varepsilon = -\rho\partial_{\sigma}G = S\sigma + \alpha(T - T_0) + \varepsilon^t \tag{4.15}$$

$$s = -\partial_{T}G = \frac{1}{\rho}\alpha : \sigma + c\ln\left(\frac{T}{T_0}\right) + s_0 \tag{4.16}$$

From the same process, the following inequality can be obtained,

$$p\dot{\xi} + \sigma : \dot{\varepsilon}^t - \dot{g}^t \geq 0 \tag{4.17}$$

where, the generalized thermodynamic force p can be obtained by taking derivative of total G in Equation 4.1 with respect to ξ,

$$\begin{aligned} p &= -\rho\partial_{\xi}G = G^M(\sigma, T) - G^A(\sigma, T) \\ &= \frac{1}{2}\sigma : \Delta S\sigma + \sigma : \Delta\alpha(T - T_0) - \rho\Delta c\left[(T - T_0) - T\ln\left(\frac{T}{T_0}\right)\right] + \rho\Delta s_0 T - \rho\Delta u_0 \end{aligned} \tag{4.18}$$

As observed in this equation, p is proportional to $G = G^M - G^A$, that is the difference between the Gibbs free energy of pure martensite and pure austenite phase.Substituting Equations 4.4 and 4.7 into Equation 4.17 results in,

$$(\sigma : \Lambda^t + p - f^t)\dot{\xi} = \pi^t \dot{\xi} \geq 0 \tag{4.19}$$

where, π^t is the total thermodynamic force which must be positive during the forward transformation (when $\dot{\xi} > 0$) and negative during the reverse transformation to hold the second law of thermodynamics (when $\dot{\xi} < 0$). In this model, this thermodynamic driving force must reach a stress-dependent critical thermodynamic driving force (Y^t) in order to start and continue the phase transformation. Different values of Y^t for forward and reverse phase transformation lead to a hyper-surface for each transformation direction, as follows,

$$\Phi^t(\sigma, T, \xi) = 0 \ \ , \ \ \Phi^t(\sigma, T, \xi) = \begin{cases} \Phi^t_{fwd} = \pi^t_{fwd} - Y^t_{fwd} \ \ , \ \ \dot{\xi} > 0 \\ \Phi^t_{rev} = -\pi^t_{rev} - Y^t_{rev} \ \ , \ \ \dot{\xi} < 0 \end{cases} \tag{4.20}$$

where,

$$Y^t(\sigma) = \begin{cases} Y^t_{fwd} = Y^t_0 + D\sigma : \Lambda^t_{fwd} \ \ , \ \ \dot{\xi} > 0 \\ Y^t_{rev} = Y^t_0 + D\sigma : \Lambda^t_{rev} \ \ , \ \ \dot{\xi} < 0 \end{cases} \tag{4.21}$$

In Equation 4.21, Y^t_0 is a constant, and the model parameter D indicates the dependency of the critical value for driving force on the applied stress.

In the case where the model is used to capture the response of an SMA subjected to one thermal cycle under iso-baric conditions (σ=cons), Equation 4.15 is utilized. Hence, for the considered iso-baric conditions for any specific T, the value of the developed total strain is obtained by calculating the value of ϵ^t and ξ during the phase transformations. That is achieved by solving the ordinary differential Equation 4.4 while considering that during the phase transformations, the thermo-mechanical state of the material should stay on the surface defined by Equation 4.20. To this end for a full thermal cycle where the temperature ranges from A_f to M_f and vice versa, the hysteresis loop of the material can be obtained.

It is worth noting that the model parameters ΔS, $\Delta\alpha$, Δc, H_{sat}, k, n_1, n_2, n_3, n_4, $\rho\Delta s_0$, $\rho\Delta u_0$, a_1, a_2, a_3, Y_0^t, and D can be considered for calibration. Among these parameters, ΔS are obtained by $E^{A/M}$ and $\nu^{A/M}$; and the materials properties $\rho\Delta s_0$, $\rho\Delta u_0$, a_1, a_2, a_3, Y_0^t, and D are associated with $E^{A/M}$, M_s, M_f, A_s, A_f, $C^{A/M}$, and H_{sat} through relations reported in [99]. Note that $E^{A/M}$ and $\nu^{A/M}$ denote the Young's modulus and Poisson's ratios of austenite/martensite phases respectively while $C^{A/M}$ denote the slopes of the lines that define the austenitic/martensitic transformation temperatures ($A_{s/f}/M_{s/f}$) in the stress-temperature phase diagram [92, 99]. Therefore, assuming $\Delta\alpha$ and Δc as fixed parameters, $E^{A/M}$, M_s, M_f, A_s, A_f, $C^{A/M}$, H_{sat}, k, n_1, n_2, n_3, and n_4 are the candidates for the model calibration against the existing experimental data.

4.4 Sensitivity Analysis of the Model Parameters

Before model calibration, SA is usually required to reduce the dimensionality of the parameter space with minimum possible loss of information about the model outcome in order to save calibration time and cost as a result of faster convergence of the model parameters. In this work, we carried a DOE based on a complete factorial design (CFD) coupled to ANOVA in order to identify the parameters most highly correlated to the output of the model.

4.4.1 Complete factorial design

CFD is an experimental design method that considers all possible combinations of R predefined levels for N factors (parameters in the case of parameter sensitivity analysis) in a given system/model, which is equivalent to R^N level-factor combinations. In the case of parameter experimental design, model response should be obtained for each level-parameter combination for the sake of sensitivity assessment.

Based on the expert's intuition, 14 parameters have been considered as candidates for model calibration. Two plausible values have been selected for each parameter as lower and upper levels in the context of two-level CFD. These levels are considered as the applied initial values $\pm 10\%$ of the range of the parameters in order to hold the following temperature constraints for their design combinations,

$$M_f < M_s < A_s < A_f \tag{4.22}$$

Therefore, $2^{14} = 16384$ level-parameter combinations can be constructed during CFD whose responses are determined through the difference between the hysteresis loop obtained from running the model with each parameter combination and a reference hysteresis loop obtained using the average value of lower and upper levels of the parameters. It is worth noting that any reference curve can be applied in so far as it is consistent for all the responses. Tschopp et al. [120] has evaluated various approaches to determine the similarity/difference of two images by comparing two vectors which represent the image features. In our case, these vectors were constructed based on the coordinate of spatial points on the resulting hysteresis loops,

$$d_{SE} = \sum_i (X_i - Y_i)^2 \tag{4.23}$$

Now, the CFD results can be applied for SA of the given model parameters through ANOVA.

4.4.2 Analysis of variance

A 14-way ANOVA has been performed in this work using "anovan" function in Matlab Machine Learning toolbox, where N is the number of independent Model Parameters–see Appendix A.

ANOVA is a statistical hypothesis testing technique based on the deviation associated with each individual level of each factor or each level combination of each interaction between factors in the system from the overall mean of responses. The resulting response for each one of 2^{14} factorial design combinations has been used in the 14-way ANOVA to evaluate the sensitivity of 14 model parameters introduced in Section 4.3. The ANOVA results associated with each parameter are shown in Table 4.1. In this table, the model parameters have been ranked based on their corresponding p-values in descending order, which indicate the parameters from the highest to the lowest sensitivity. According to the significance level 0.05, the first eight parameters have been selected for further uncertainty analysis in order to reduce the computational cost for the model calibration that will be discussed in Section 4.5.1.It is worth noting that the values considered as upper and lower levels in the ANOVA process, should be chosen carefully before the initiation of the process, since they may affect the estimated sensitivity of the model to some of these parameters. For example, it seems that the selected lower values of the levels for the parameter C^M result in a much lower influence of this parameter on the model results compared to the parameter C^A. Furthermore, another aspect which may affect the results of the sensitivity analysis are the loading conditions under which the considered model is evaluated during this process. In the current example, the model sensitivity has been determined using iso-baric loading paths. In these conditions, the sensitivity of the model on the value of the E^M and E^A is minimized since the response of the model under the selected conditions is dictated mainly due the thermally induced phase transformation. In the current work, we selected the iso-baric loading path due to the focus on the uncertainty quantification of the thermal phase transformation related model parameters.

4.5 Bayesian Uncertainty Quantification and Propagation of the Model Parameters

4.5.1 Parameter estimation and uncertainty quantification using Markov Chain Monte Carlo-Metropolis Hastings algorithm

In order to sample the parameter space, one can use MCMC sampling approaches [6]. Among these approaches, Metropolis-Hastings (M-H) sampling is one of the most common applied tools to probabilistically solve the inverse problem associated to the probabilistic parameter calibration given experimental data. In the case of model calibration, this technique updates the existing prior knowledge about the parameters to their posterior information based on the available data from the model responses.

In the present work, MCMC toolbox in Matlab has been used to probabilistically calibrate the eight sensitive model parameters identified in Section 4.4.2. First, the prior parameter information, including the initial values (θ^0), valid ranges, and probability density functions (PDFs) for the parameters are considered. Then, a new parameter vector/candidate θ^{cand} is randomly sampled

Table 4.1: ANOVA table results showing the model parameters ranked based on their p-values in descending order

Source	Sum sq.	d.f.	Mean sq.	F	Prob>F
H_{sat}	0.65941	1	0.65941	800.62	5.3041e-172
$A_f(K)$	0.33278	1	0.33278	404.05	8.5195e-89
$M_s(K)$	0.08907	1	0.08907	108.14	3.0049e-25
$M_f(K)$	0.04402	1	0.04402	53.45	2.7735e-13
$C^A(MPa/K)$	0.03488	1	0.03488	42.35	7.8628e-11
$k(MPa^{-1})$	0.02462	1	0.02462	29.90	4.6223e-08
$E^M(GPa)$	0.01029	1	0.01029	12.50	4.0821e-04
$A_s(K)$	0.00323	1	0.00323	3.93	0.0476
$E^A(GPa)$	0.00018	1	0.00018	0.21	0.6443
$C^M(MPa/K)$	0.00008	1	0.00008	0.09	0.7612
n_1	0	1	0	0	1
n_2	0	1	0	0	1
n_3	0	1	0	0	1
n_4	0	1	0	0	1
Error	13.4819	16369	0.00082		
Total	14.6805	16383			

from a proposal posterior PDF (q) which is defined as a multivariate Gaussian distribution in the toolbox with a mean value at θ^0 and an arbitrary variance-covariance matrix. The candidate is accepted/rejected using a probabilistic criterion defined based on the M-H ratio that is expressed as follows,

$$MH = \frac{P(\theta^{cand})P(D|\theta^{cand})}{P(\theta^0)P(D|\theta^0)} \frac{q(\theta^0|\theta^{cand})}{q(\theta^{cand}|\theta^0)} \tag{4.24}$$

In this equation, the first ratio is called Metropolis ratio which is the ratio of the posterior probability of θ^{cand} to θ^0 given data (D). In the context of the Bayes' theorem, the posterior probabilities are proportional to the prior probability times the likelihood—probability of the data given model results evaluated with the candidate model parameters—for each case. We considered the likelihood to be a Gaussian distribution centered at the data whose error is considered as the distribution variance. In the case of multiple data sets, it should be noted that the product of the likelihoods obtained from each data set can yield the total likelihood if data sets are assumed to be conditionally independent.

The Metropolis ratio can be applied for the acceptance/rejection of the candidate when the proposal distribution is symmetric; however, this is not usually the case since the probability of moving from θ^0 to θ^{cand} is not necessarily equal to the probability of the reverse move. For this reason, Hastings introduced the second ratio in Equation 4.24 to involve the above-mentioned non-symmetric property of the proposal distribution [6,121]. Now, $min(MH \times 100, 100)$ is considered as the acceptance probability of the candidate [121]. According to this criterion, If the candidate is accepted, then $\theta^1 = \theta^{cand}$; otherwise, $\theta^1 = \theta^0$. The sampling of the new candidate and its acceptance/rejection process are sequentially repeated n times by finding the M-H ratio using the new sampled candidate and the previous accepted parameter during each iteration. At the end of the procedure, n samples of the parameter vector ($\left\{\theta^0, ..., \theta^n\right\}$) are generated to represent a multivariate posterior distribution for the model parameters.

Before the convergence of the parameter samples, samples generated during the "burn-in period" — in which the different samples are causally correlated — must be removed before further analysis of the parameters. After the removal of the burn-in period, the mean values and the square root of diagonal elements in variance-covariance matrix of the remaining samples in the convergence region can be introduced as the plausible optimal values and uncertainties of the model parameters.

In our calibration work, initial values and their ranges have been determined based on the expert's beliefs, and an uniform (non-informative) probability distribution has been selected for each model parameter due to the absence of knowledge about the parameters' PDF. In addition, three experimental datasets for transformation strain-temperature hysteresis loop under different iso-baric conditions have simultaneously been utilized together to calibrate the model. It is worth noting that the experimental iso-baric response of SMAs has been obtained using SMA dog-bone specimens. The specimens were uniaxially loaded at the stress levels under iso-thermal conditions above the A_f temperature. Subsequently, they have been thermally cycled between a temperature below the M_f and a temperature above the A_f at a constant stress [122]. The calibration has been performed through the calculation of the difference between various iso-baric hysteresis curves obtained from each parameter sample and the corresponding experimental data using equation 4.23. For the sake of the parameter calibration, these differences (discrepancies) resulting from three given iso-baric conditions ($[M(\theta)]_{1\times 3}$) are simultaneously compared with a zero vector ($D = [0\ \ 0\ \ 0]$) in the likelihood function.

In this work, the likelihood variance σ^2 (the squared error of the distances between model and experimental hysteresis curves from zero) is set as an unknown hyper-parameter in the calibration process and updated in parallel with the other model parameters during MCMC sampling. The only difference is that an inverse-gamma distribution is defined as the non-informative prior PDF for this hyper-parameter, which yields the same form of distribution for the posterior PDF after applying the Bayes' theorem since the inverse-gamma distribution is the corresponding conjugate prior density for exponential likelihood functions. Therefore, σ^2 is sampled from an inverse gamma posterior distribution during MCMC process [52]. In addition, it should be noted that an adaptive proposal distribution has been employed in this calibration work. In this scheme, an arbitrary positive covariance (V_0) is selected in the beginning of MCMC sampling, and then it is adapted in each next iteration using the covariance obtained from the ensemble of previous samples in the MCMC chain [51, 123].

In this work, a constrained MCMC approach has also been developed in order to consider the constraints for the transformation temperatures mentioned in Equation 4.22. For this purpose, the likelihood is penalized by assigning an infinitely high value for the distance between model and experimental hysteresis curves in the cases that the constraints for the parameters are not satisfied. In these cases, the induced penalty results in a likelihood close to zero, which in turn results in an infinitesimally small acceptance probability for new model parameter candidates that violate constraints.

In the calibration of the sensitive model parameters, high number of samples (i.e., 200,000) are generated to ensure the convergence of the parameters and the uniqueness of the optimal response in the parameter space. The multi-variate posterior distribution after MCMC sampling can be assessed through joint and marginal frequency distributions. In Figure 4.1, some of the parameter joint distributions have been shown in 2D color graphs as examples, where the colors represent different density of samples in the parameter space, increasing from blue to red spectra. These types of graphs can also demonstrate the correlation between each pair of parameters qualitatively. For instance, negative/positive linear correlations can be inferred from the negative/positive slopes of the elliptical shapes in the graphs. The linearity of the sample distribution in the parameter space can be determined quantitatively through Pearson correlation coefficient ($-1 < \rho < 1$) that is expressed as,

$$\rho_{X,Y} = Corr(X,Y) = \frac{cov(X,Y)}{\sigma_X \sigma_Y} \tag{4.25}$$

where, σ_X and σ_Y are the standard deviations for parameters X and Y, respectively, which equal the square root of the corresponding diagonal elements in 8×8 variance-covariance matrix. $cov(X,Y)$ is also the covariance of X and Y obtained from the corresponding non-diagonal element in the variance-covariance matrix. Generally, a value of ρ closer to either -1 or 1 implies a higher linear correlation between X and Y, whereas a value closer to zero suggests their lower linear correlation. The negative or positive sign refers to the negative or positive linear correlation. The linear coefficients for all pairs of parameters in the thermo-mechanical model have been listed in Table 4.2. As can be observed in this table, most of the pair parameters are linearly uncorrelated although 6 out of 28 parameter pairs show some weak or moderate correlations with $|\rho|$ greater than 0.2. In addition, the parameters C^A and A_f (Figure 4.1(d)) and the parameters E^M and C^A (Figure 4.1(f)) show the highest and the lowest linear correlation, respectively, regardless of the correlation signs.

As mentioned above, the burn-in period at the beginning of MCMC sampling must be removed before the calibration and UQ of the model parameters, which can be found using the cumulative mean distributions for the parameters, as shown in Figure 4.2. In these figures, there are some noisy trends at the beginning of MCMC sampling for the cumulative mean values of all the parameters,

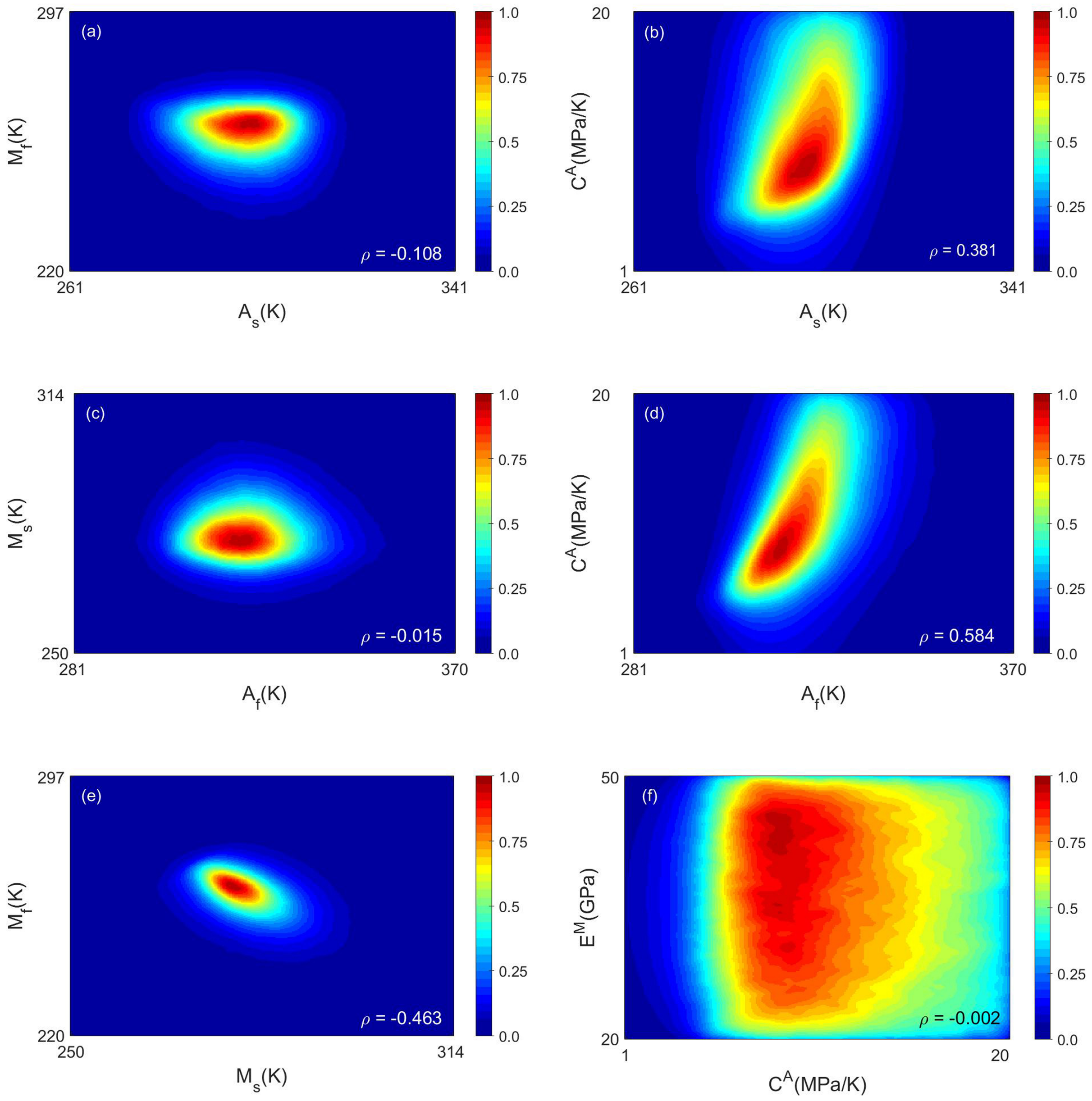


Figure 4.1: Joint frequency distributions for some of the pair model parameters in the form of 2D color graphs with normalized color bars

Table 4.2: Linear correlation coefficient for each pair of model parameters

	$A_s(K)$	$A_f(K)$	$M_s(K)$	$M_f(K)$	$C^A(MPa/K)$	$E^M(GPa)$	H_{sat}	$k(MPa^{-1})$
$A_s(K)$	1	-0.0721	0.2948	-0.1080	0.3813	-0.0167	-0.1330	-0.0378
$A_f(K)$	-0.0721	1	-0.0150	-0.0822	0.5836	-0.0553	0.0249	-0.0672
$M_s(K)$	0.2948	-0.0150	1	-0.4628	0.1156	-0.0547	-0.0465	-0.0415
$M_f(K)$	-0.1080	-0.0822	-0.4628	1	-0.1280	-0.0289	-0.1355	0.0160
$C^A(MPa/K)$	0.3813	0.5836	0.1156	-0.1280	1	-0.0015	0.0476	-0.0405
$E^M(GPa)$	-0.0167	-0.0553	-0.0547	-0.0289	-0.0015	1	0.2771	0.0091
H_{sat}	-0.1330	0.0249	-0.0465	-0.1355	0.0476	0.2771	1	-0.2103
$k(MPa^{-1})$	-0.0378	-0.0672	-0.0415	0.0160	-0.0405	0.0091	-0.2103	1

but they reach almost plateaus after about 63,000 generations that show the convergence values (mean values) for the parameters. Therefore, the first 63,000 samples are extracted from the total MCMC sample collection before further analysis. The mean and the square root of the diagonal elements in the variance-covariance matrix of the remaining samples in the convergence region can be introduced as the plausible optimal values and the uncertainty of the model parameters, which are listed in Table 4.3.

Table 4.3: Plausible optimal values and uncertainties (standard deviations) of the model parameters after MCMC calibration besides their initial values

	$A_s(K)$	$A_f(K)$	$M_s(K)$	$M_f(K)$	$C^A(MPa/K)$	$E^M(GPa)$	H_{sat}	$k(MPa^{-1})$
Initial Values	307.0	318.0	300.0	270.0	9.0	40.0	0.034	0.02
After MCMC Calibration	296.6±8.5	322.6±11.3	280.4±6.6	259.9±8.5	11.8±4.1	35.6±8.6	0.0517±0.0044	0.0595±0.0237

After removal of the burn-in period, the marginal posterior frequency distribution of the parameters can be plotted using the remaining MCMC samples as observed in Figure 4.3. These plots suggest truncated skewed Gaussian distributions for the marginal posterior frequency distributions; although they are flatter in the case of the parameters E^M, k and C^A, and more close to a Gaussian distribution for the parameter H_{sat}.

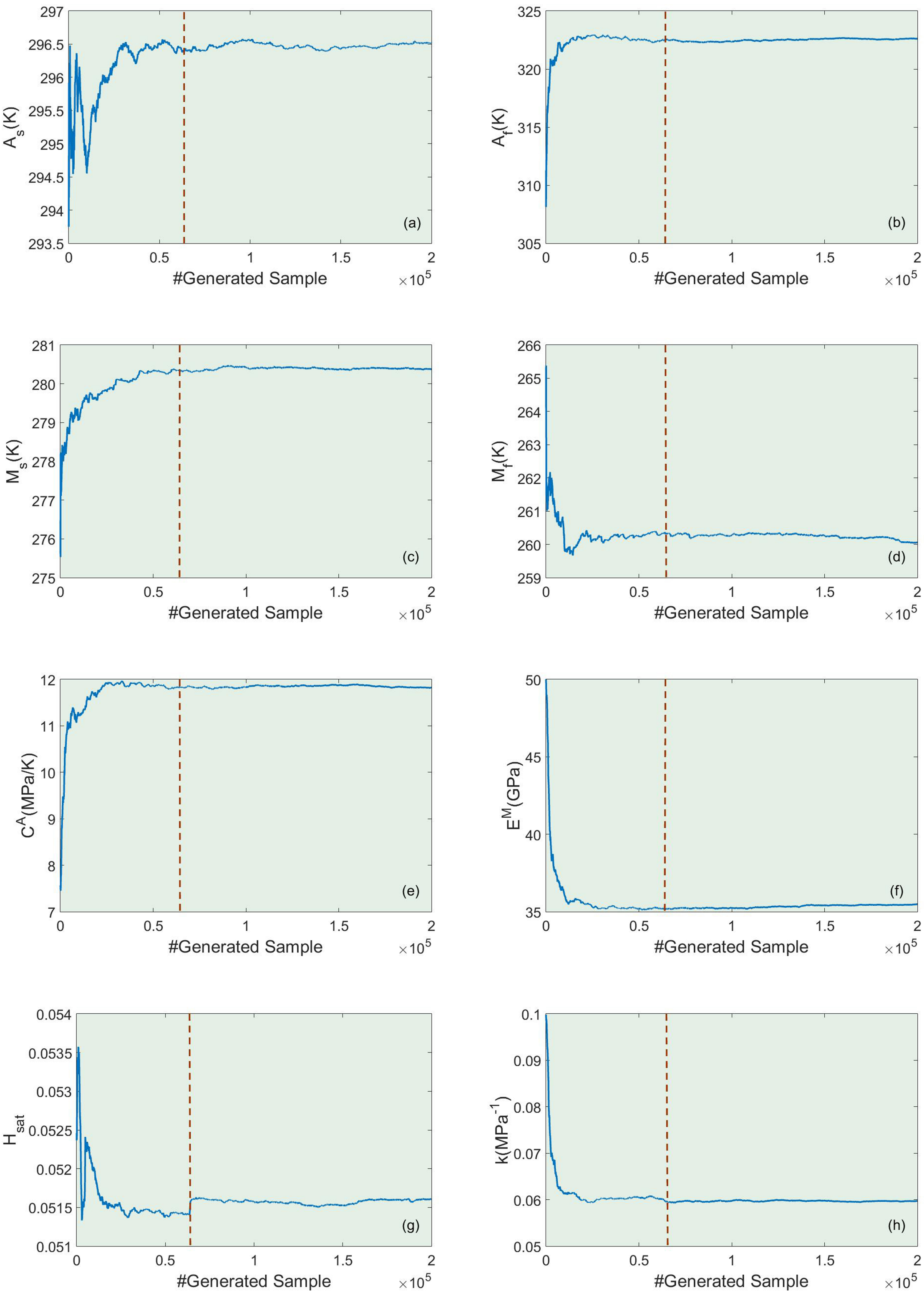


Figure 4.2: Cumulative mean distribution plot for each model parameter after MCMC sampling

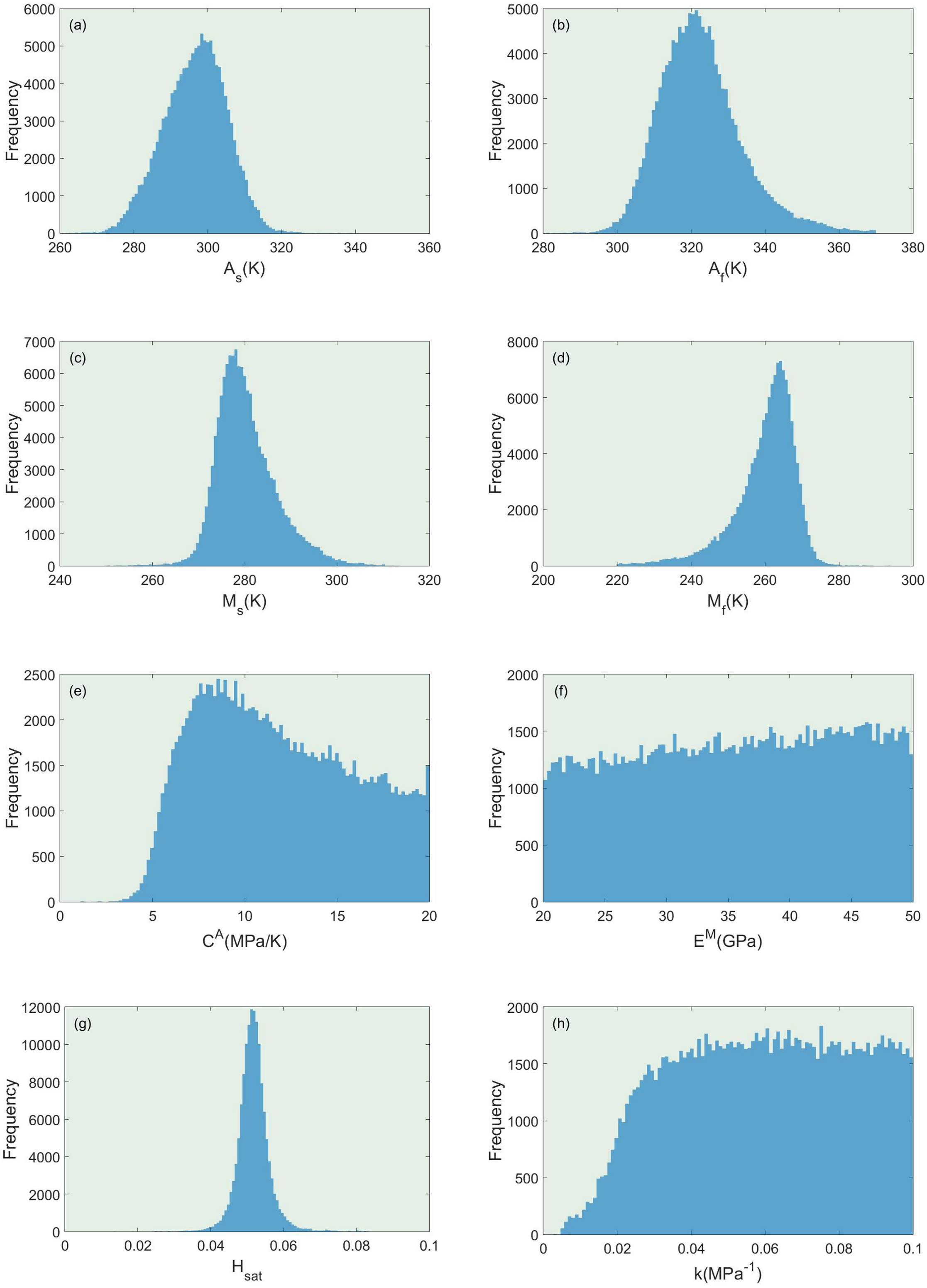


Figure 4.3: marginal posterior frequency distribution plot for each model parameter after the removal of the burn-in period from MCMC samples

4.5.2 Uncertainty propagation from the model parameters to the model outcomes

As mentioned earlier, only the final results of the model and the corresponding uncertainties are considered important in the context of the robust design. The reason is that such information can help the designers to reduce the influence of the system performance by the existing uncertainties in the outcomes, regardless of how many variables and models have been involved. Therefore, an appropriate UP technique is essential to propagate uncertainties from the model parameters/input variables to the outputs with a relatively high efficiency and a minimum or at least quantifiable loss of information [115]. Among different UP methods, first/second order second moment (FOSM/SOSM) approaches are probably the most common approximation procedures in engineering [124]. In this work, FOSM approach has been used to perform non-linear propagation of the variance-covariance matrix of the model parameters.

Let θ and $\bar{\theta}$ vectors denote the optimal MCMC samples for the parameters (after removal of the burn-in period) and their mean values, respectively, and $M(\theta)$ denote the model objective (output) function. If the objective function is estimated by a first order Taylor expansion at $\bar{\theta}$,

$$M(\theta) \approx M(\bar{\theta}) + \sum_{i=1}^{N} \frac{\partial M(\theta)}{\partial \theta_i}(\theta_i - \bar{\theta}_i) \tag{4.26}$$

The expected values associated with first and second moment of the above first order Taylor expansion approximates the mean value ($\bar{M}$) and variance (σ_M^2) of the objective function M [124] as follows,

$$\bar{M} = E[M(\theta)] \approx M(\bar{\theta}) \tag{4.27}$$

$$\sigma_M^2 = E[(M(\theta) - \bar{M})^2] \approx \sum_{i=1}^{N} |\frac{\partial M(\theta)}{\partial \theta_i}|^2 \sigma_i^2 + \sum_{i=1}^{N}\sum_{j\neq i}^{N} \frac{\partial M(\theta)}{\partial \theta_i}\frac{\partial M(\theta)}{\partial \theta_j}\sigma_{ij} \tag{4.28}$$

where, σ_i^2 and σ_{ij} are the variance of the parameter θ_i and the covariance between the parameters θ_i and θ_j, respectively. The approximations in Equations 4.27 and 4.28 have been derived thoroughly by Kriegesmann [125]. It is worth noting that Equation 4.28 can also be expressed in terms of the variance-covariance matrix of the model parameters (V) [53] as follows,

$$\sigma_M^2 \approx g^T V g \tag{4.29}$$

where, g is a column vector which consists of all partial derivative elements, i.e.,

$$g = (\frac{\partial M(\theta)}{\partial \theta_1}, ..., \frac{\partial M(\theta)}{\partial \theta_N})^T \tag{4.30}$$

The uncertainties of parameters reported in Table 4.3 have been propagated to the uncertainty of the transformation strain along the hysteresis curves by applying the variance-covariance matrix obtained from the optimal MCMC samples (without burn-in period) in the above-mentioned FOSM approach. The results have been shown in Figure 4.4 for different iso-baric conditions. In these figures, the blue/red solid and dashed lines correspond to the model results during cooling/heating process (forward/reverse martensitic transformation) and their experimental counterparts, respectively. In addition, the blue/red shaded regions indicate 95% Bayesian confidence interval (95% BCI) for model results during cooling/heating. It should be noted that the interval is equivalent to $\varepsilon^t(T) \pm 2\sigma_{\varepsilon^t}$, where $\varepsilon^t(T)$ and σ_{ε^t} can be calculated through Equations 4.27 and 4.28, respectively.

The average Euclidean distances between the mean strain values of the hysteresis curves obtained from the model (solid lines in Figure 4.4) and their corresponding experimental values (dashed lines in Figure 4.4) have been calculated as 0.0083, 0.0059 and 0.0032 for the iso-baric conditions 100, 150 and 200 MPa, respectively. As can be observed, these results show small discrepancies (good agreements) between the model and experimental hysteresis loops, or at least the experimental data are located inside the 95% BCIs based on Figure 4.4. Although it seems that the given thermo-mechanical model underestimates the slope of the curves during forward and reverse transformation and their convergence temperatures (where cooling and heating curves meet each other), there are no discrepancies between model results and experimental data for the maximum transformation strain (H^{cur}) in each iso-baric condition. The difference in the curves' slopes can

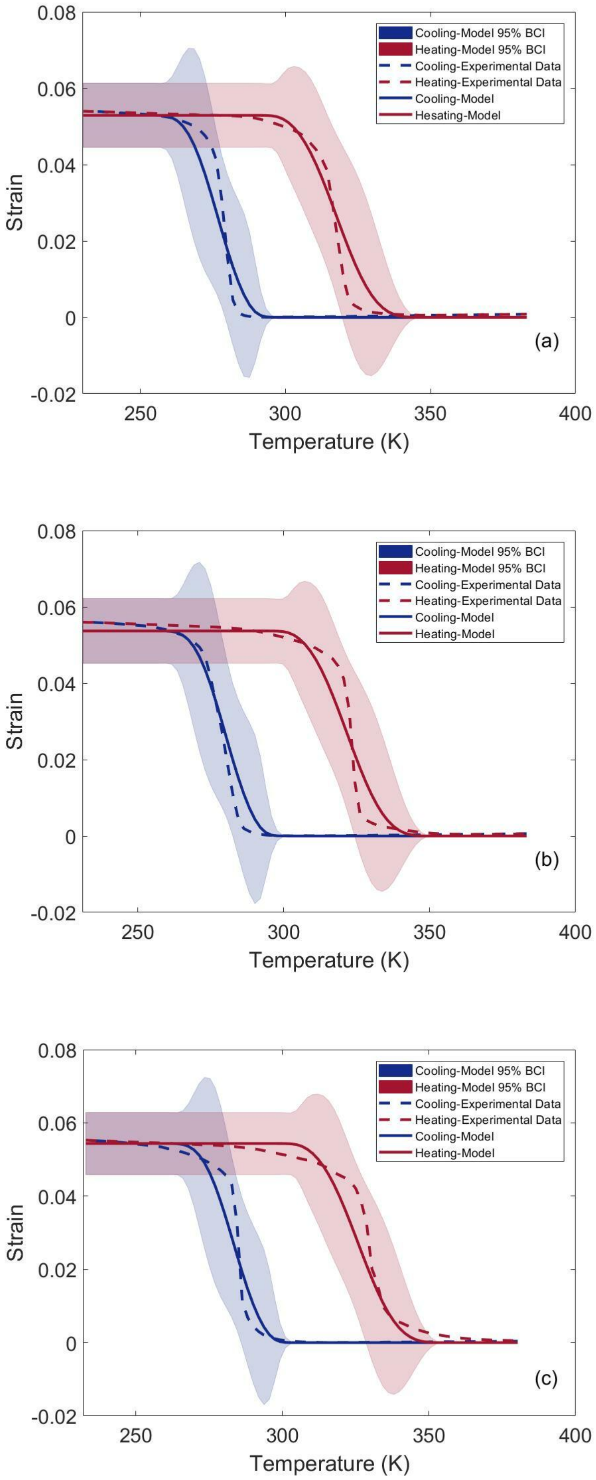


Figure 4.4: 95% BCIs for the model hysteresis curves obtained from the FOSM approach for different iso-baric conditions, a) 100, b) 150, and c) 200 MPa, besides their corresponding experimental counterparts

be attributed to the slight effects of other parameters that are not considered in the calibration, the missing physics in the model, and the uncertainties in the experimental data all together.

Another important feature in theses graphs are the unrealistic humps around the transformation temperatures, which can be related to the deficiency of the partial derivatives in the applied UP approach (Equation 4.28) at where the changes are sudden. For this reason, a direct UP has been performed using the model forward analysis of optimal MCMC samples to find more rational regions for 95% BCIs. After running the model for all the optimal samples, 2.5% of the resulting hysteresis curves have been eliminated from each one of the top and down margins to obtain the mentioned intervals. As shown in Figure 4.5, this approach yields smoother and more precise boundaries for 95% BCIs with no humps. Assuming the 95% Bayesian confidence intervals obtained from the direct UP as the ground truth, the average relative error of the FOSM approach has been calculated around 21%; which is based on an average of the relative differences between the corresponding upper and lower bounds of Bayesian confidence intervals obtained in Figures 4.4 and 4.5 at different temperatures and iso-baric conditions. However, these UP approaches are different in cost. The cost of the direct UP approach depends on the applied number of parameter samples which can slightly influence the accuracy of uncertainty bands obtained. Despite the wide range of cost for the direct UP approach, it can be stated this approach costs at least 30 times more than the FOSM approach in most applications. Therefore, there is a trade-off between cost and accuracy for the application of these UP approaches. The expensive direct UP approach is required in the case that the accuracy of the quantified uncertainty is highly important; otherwise, the FOSM method can be applied to reduce the UP cost.

4.6 MCMC Application in Decision Making for Experimental Design

In this section, the goal is to respond the experimentalists' question about which experimental design can gain more knowledge about the system: performing experiments at the same conditions or at different conditions in design input space. For this purpose, it is assumed that the probabilistically calibrated model in Section 4.5.1 can yield the ground truth and its natural uncertainty bounds at any design condition (i.e., any iso-baric condition). According to this assumption, two

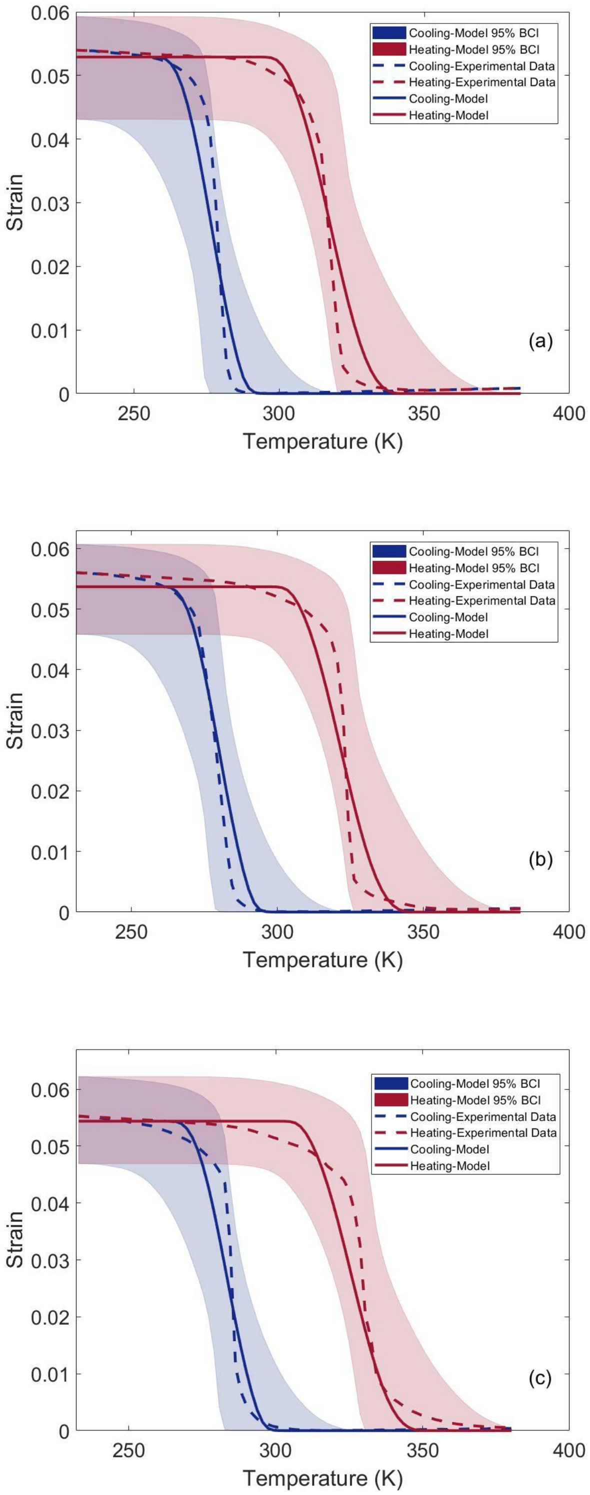


Figure 4.5: 95% BCIs for the model hysteresis curves obtained from the direct UP approach for different iso-baric conditions, a) 100, b) 150, and c) 200 MPa, besides their corresponding experimental counterparts

different synthetic experimental sets are designed through sampling the hysteresis curves from the model uncertainty bounds obtained at any iso-baric condition of interest. The first set contains three random samples (replicas) from the same experimental condition (i.e., σ = 150 MPa), while the second set considers three random samples from three different conditions (i.e., σ = 175, 250, and 300 MPa). After the generation of synthetic data sets, the calibrated results obtained in Table 4.3 for the influential parameters are considered as the parameter priors which are separately updated against each set of synthetic data using the MCMC technique in a sequential way of data training. In this way of training, the parameter posterior distribution obtained after each training is considered as the prior distribution for the next training. The relative entropy (Kullback-Leibler (K-L) divergence) is calculated to find how diverged the probability distributions are relatively, i.e., what is the distance between the prior and final posterior probability distributions of the parameters after the three sequential MCMC calibrations in each case. The calculated K-L divergence can be introduced as a comparison measure for the amount of information gained from each synthetic experimental sets. Generally, K-L divergence for continuous probability distributions P and Q is defined as,

$$D_{KL}(P||Q) = \int_{-\infty}^{+\infty} p(x) log(\frac{p(x)}{q(x)}) dx \tag{4.31}$$

where, $p(x)$ and $q(x)$ are the densities of P and Q at any given random point x, respectively. In the case of multivariate Gaussian distribution, the integration in Equation 4.31 can be obtained as follows,

$$D_{KL}(\mathcal{N}_1||\mathcal{N}_2) = \frac{1}{2}\Big[ln(\frac{|\Sigma_2|}{|\Sigma_1|}) - d + tr(\Sigma_2^{-1}\Sigma_1) + (\mu_2 - \mu_1)^T\Sigma_2^{-1}(\mu_2 - \mu_1)\Big] \tag{4.32}$$

The equivalent normal distributions obtained for parameter prior and posterior distributions are compared for each set of experimental design using Equation 4.32. The K-L divergence values for synthetic experimental sets 1 and 2 are calculated around 3.8 and 4.4, respectively. The higher value of the K-L divergence for synthetic experimental set 2 indicates a greater difference between prior and final posterior distributions in this case, which implies that the consideration of different conditions in experimental design can gain more information about the system rather than experimental replicas in the same condition.

It is also important for experimentalists to know how much information they gain by adding each experiment. In our work, the K-L divergences obtained after each training sequence are 1.65, 0.64 and 0.60 for synthetic experimental set 1 and 3.65, 0.80 and 0.67 for synthetic experimental set 2. It is worth noting that the addition of these K-L divergences should not necessarily equal the overall K-L divergence already obtained for each set since the random generated data can not only diverges, but also converges the parameter mean values to the mean values of the parameter prior probability used in the beginning of the trainings. As can be observed in both synthetic experimental designs, the K-L divergence reduces as the training continues with more synthetic data. This trend can be attributed to a dramatic decrease in the rate of the parameter uncertainty reduction from the prior to the posterior distributions during the sequential MCMC training. This decreasing rate has the main contribution in the reduction of the prior distribution divergence from the posterior distribution as the trainings proceed. Moreover, the difference in the mean values of the prior and posterior distribution also influences the K-L divergence after each training, as shown in Equation 4.32. Although this difference is directly affected by how far the randomly generated data are from each other in the sequence, it generally tends to reduce by adding more synthetic data. The reason being the prior distributions become more peaked and thus more dominant than likelihood during the MCMC sampling which result in less changes in the parameter mean values after each training.

4.7 Summary and Conclusion

In this work, the calibration and UQ of the sensitive parameters in a thermo-mechanical model have been performed against three different iso-baric experimental data simultaneously using a constrained MCMC-MH algorithm in the context of Bayesian statistics. Compared to the few prior preliminary works for sensitivity and uncertainty assessment of SMAs thermo-mechanical models through central finite difference, Monte Carlo and MCMC approaches, this work proposes a global variance-based sensitivity analysis known as ANOVA as well as a thorough MCMC analysis of the uncertainty for model parameters and its propagation to model responses. It should be noted that the applied model is able to predict the thermo-mechanical responses of SMAs with various chemistries and processing histories under general temperature-loading paths. Besides the generality of this model in the prediction of SMA thermo-mechanical responses, it is less complex with the least possible number of parameters compared to the previous similar models, which makes the model calibration process less expensive. However, it should be noted that just the iso-baric conditions with cyclic varying temperatures are considered in this work to calibrate the model that makes its accurate predictions limited to the mentioned conditions. In order to generalize the application and predictability of this model, the parameter calibrations should be performed with much more experimental data obtained from different thermo-mechanical conditions, such as the iso-thermal conditions with cyclic varying stresses or the conditions where applied stresses and temperatures are both changing. In this work, the iso-baric conditions are used as examples to propose a UQ framework that can be applied to provide probability distributions in the form of uncertainty bands for the thermo-mechanical responses of the applied Ni-Ti SMA using the propagation of uncertainties from the probabilistically calibrated model parameters. This is very important since predicting probablistically rather than deterministically can provide confidences in the robust design of these SMAs for their specified applications.

Eight sensitive parameters have been found before the calibration using a DOE approach which includes CFD and N-way ANOVA. After MCMC sampling of these eight parameters, their convergence have been checked by the joint frequency distributions and the cumulative mean plots, which have also been utilized to identify the qualitative correlation of each pair parameters and the burn-in period, respectively. In addition, the linear correlation between each two parameters has been calculated through Pearson coefficient. These coefficients suggest that most pair parameters are linearly uncorrelated, although a few of them show weak or moderate linear correlations.

After removal of the samples belonging to the burn-in period, skewed Gaussian distributions have been obtained for the marginal posterior frequency distributions of the model parameters, which are truncated in the range defined for each parameter. Moreover, it should be noted that the mean and square root of the diagonal elements in the variance-covariance matrix of the remaining optimal samples after the elimination of the burn-in period provide the plausible optimal values and uncertainties of the selected parameters, respectively.

At the end, the parameters' uncertainties have been propagated to the uncertainties of the transformation strain along the hysteresis curves using two various UP techniques in order to find 95% BCIs during cooling/heating process. In this regard, FOSM approach results in some unrealistic humps around the transformation temperatures. For this reason, a direct UP technique has been performed through the forward analysis of optimal MCMC samples in order to achieve more rational and precise uncertainty bands that are very important in the context of robust design. The hysteresis curves obtained from the mean values of the parameters are in good agreements with their experimental counterparts for different given iso-baric conditions, or at least it can be stated that the experimental data are situated in 95% BCIs. Although the slopes and the convergence temperatures of the forward and reverse transformation curves are not predicted precisely which can be the deficiencies of the applied model, the maximum transformation strain associated with each iso-baric condition is very close to its corresponding data.

In this work, It has been also shown that the Markov Chain Monte Carlo approach can be applied for decision making in experimental design. Assuming the calibrated model can predict the ground truth and the natural uncertainties in different experimental conditions, two synthetic experimental sets have been sampled from the model response uncertainty bounds obtained at the same and different iso-baric conditions, respectively. After sequential MCMC updates of parameters' posterior distributions with each set, their information gains are compared together through the calculation of K-L divergence for each case which indicate the difference between prior and final posterior distributions of the model parameters. A higher information gain has been obtained from the synthetic experimental set 2 that are sampled from different experimental conditions (different iso-baric conditions). It has also been shown that the K-L divergence and its rate reduce as the sequential training proceeds in both synthetic experimental designs.

Generally, this work introduces a complete framework for UQ of model parameters and subsequent UP from model parameters to outputs as a guideline for material design.

5. BAYESIAN UNCERTAINTY QUANTIFICATION AND INFORMATION FUSION IN CALPHAD-BASED THERMODYNAMIC MODELING*

5.1 Overview

Calculation of phase diagrams is one of the fundamental tools in alloy design—more specifically under the framework of Integrated Computational Materials Engineering. Uncertainty quantification of phase diagrams is the first step required to provide confidence for decision making in property- or performance-based design. As a manner of illustration, a thorough probabilistic assessment of the CALPHAD model parameters is performed against the available data for a Hf-Si binary case study using a Markov Chain Monte Carlo sampling approach. The plausible optimum values and uncertainties of the parameters are thus obtained, which can be propagated to the resulting phase diagram. Using the parameter values obtained from deterministic optimization in a computational thermodynamic assessment tool (in this case Thermo-Calc) as the prior information for the parameter values and ranges in the sampling process is often necessary to achieve a reasonable cost for uncertainty quantification. This brings up the problem of finding an appropriate CALPHAD model with high-level of confidence which is a very hard and costly task that requires considerable expert skill. A Bayesian hypothesis testing based on Bayes' factors is proposed to fulfill the need of model selection in this case, which is applied to compare four recommended models for the Hf-Si system. However, it is demonstrated that information fusion approaches, i.e., Bayesian model averaging and an error correlation-based model fusion, can be used to combine the useful information existing in all the given models rather than just using the best selected model, which may lack some information about the system being modelled.

---

*Reprinted with permission from "Bayesian uncertainty quantification and information fusion in CALPHAD-based thermodynamic modeling" by Pejman Honarmandi, Thien Duong, Seyede Fatemeh Ghoreishi, Douglas Allaire, Raymundo Arroyave, 2019. *Acta Materialia*, 164:636-647, Copyright 2019 by Elsevier B.V. or its licensors or contributors

5.2 Introduction

Uncertainty quantification (UQ) and its propagation (UP) across multi-scale model/experiment chains are key elements of decision-based materials design in the framework of Integrated Computational Materials Engineering (ICME), where databases, multi-scale modeling and simulation tools, and experiments are integrated with the aim of time reduction in design and manufacturing of materials/products [126]. In this context, the understanding and quantification of uncertainties can provide a confidence measure for the applicability of models for decision making in materials design.

Generally, UQ incorporates the detection of uncertainty sources and the development of corresponding appropriate mathematical approaches to calculate the error bounds for any quantity of interest in models [1, 2, 127]. The uncertainty can arise from different sources, which are categorized as natural uncertainty (NU) due to the random nature of a physical system, model parameter uncertainty (MPU) resulting from the lack of sufficient and/or accurate data for model parameters, propagated uncertainty (PU) in the case of multi-scale modeling, and model structure uncertainty (MSU) owing to any simplifications, assumptions, and/or incomplete physics in the model [3]. UP also refers to the determination of model output uncertainties based on the uncertainties of its input variables for either individual or multi-scale models [1, 128, 129].

Despite the importance of UQ/UP, there are comparably few works applied to problems in (computational) materials science. Chernatynskiy et al. [2] have reviewed these efforts, focusing on UQ problems related to atomic-scale as well as multi-scale simulations. For over a decade, at the electronic/atomic scale, probabilistic parameterization in first principles calculations based on density function theory (DFT) [130, 131] remains as some of the most significant works on UQ applied to materials problems. Aldegunde et al. [132] have recently used a machine learning-informed Bayesian approach to quantify uncertainties associated to the prediction, via cluster expansions, of the thermodynamic properties of alloys. In that work, the resulting uncertainties are associated to the model parameters as well as the structure of the models themselves, which results from poorly converged cluster expansions due to lack of training data. In another work, the diffusivity error in molecular dynamic simulations of a Ni/Al nano-laminate bilayer has been assessed using a Markov Chain Monte Carlo (MCMC) method in the context of Bayesian statistics [133]. In regard to UQ in multi-scale simulations, Volker et al. [134] have used DFT results for a phase field model to provide a connection between atomic- and meso-scale descriptions of ferro-electric materials, and also applied a sensitivity analysis to identify the most significant parameters of the free energy functional and to quantify the uncertainties in the simulations. Moreover, Liu et al. [135] have utilized a Bayesian stochastic approach for probabilistic parameter calibration and a stochastic projection with polynomial chaos expansions for the propagation of their uncertainties across multi-scale constitutive models that link the microstructure, property, and performance of random heterogeneous composite materials. It is worth mentioning that there is also some UQ work for multi-scale modeling of plasticity (deformation) in poly-crystalline materials [136–138].

While there are emerging several efforts centered on UQ/UP applied to materials simulations, UQ applied to CALculation of PHAse Diagrams (CALPHAD) methodologies remains poorly explored. This is unfortunate as CALPHAD-based thermodynamic descriptions are the foundation of most proposed ICME/alloy design frameworks. UQ in the calculation of equilibrium of phase diagrams is crucial since any small variations in chemical composition and/or temperature due to their uncertainties can alter the (predicted) stability, volume fraction, and chemical composition of microstructural phases and may affect the materials properties considerably. For example, if the stability of a phase should be suppressed in the design to prohibit its detrimental effects on the final product, the design space for this particular materials system should lie outside the boundaries representing the onset of stability of such detrimental phases but also out of their uncertainty bounds in order to have more confidence in the design.

Focused work on UQ applied to CALPHAD-based thermodynamic assessments remains scarce [5, 139–143]. However, Stan and Reardon [144] and Otis and Liu [145] have recently proposed approaches that constitute increasingly sophisticated approaches to UQ applied to CALPHAD. In the former study [144], a fuzzy logic weighted genetic algorithm (GA) was applied as a sampling tool in a Bayesian-based framework to find the free energy parameters' posterior probability distribution given some uncertain thermodynamic data, and then the parameters' uncertainties were propagated to the resulting phase diagram by sampling from the posterior probability distribution. In the later work [145], an ICME-directed multi-scale (or rather multi-level) modeling has been introduced that links high-throughput first-principles calculations to CALPHAD modeling. The most relevant parameters in the sublattice-based CALPHAD modeling of that work were determined using both Akaike Information Criterion (AIC) and F-test, and a Markov chain Monte Carlo (MCMC) sampler was used to quantify the parameters' posterior probability distribution.

Furthermore, Duong et al. [146] have analyzed the uncertainties in the case of a multi-scale modeling for pseudo-binary Ti_2AlC-Cr_2AlC MAX phases, which links first principles calculations and the CALPHAD method. In that study, the overall uncertainties have been reflected on the calibrated thermodynamic parameters through an MCMC approach, and then propagated to the quasi-binary phase diagram. Earlier, Duong et al. [70] employed UQ to investigate the effect of uncertainties in the phase stability of the U-Nb system. However, a detailed and thorough analysis of uncertainty and selection of models with appropriate parameters is still lacking in CALPHAD modeling.

In this work, a thorough evaluation of the uncertainty is performed for not only CALPHAD model parameters, but also for their resulting phase stability diagrams using the Hf-Si binary system as a case study. It should be noted that calculation of Hf-Si binary phase diagram and its uncertainties is of great importance since adding Hafnium to Niobium silicide based alloys (as promising turbine airfoil materials with high operating temperature) increases their strength, fracture toughness, and oxidation resistance significantly [147–149]. In this regard, an MCMC-Metropolis Hastings algorithm and a forward analysis of parameters' posterior samples are applied for the quantification of the parameters' uncertainties and their propagation to resulting phase diagrams, respectively. In this paper, it is also shown how to select the most relevant CALPHAD model given experimental data for a system based on Bayesian hypothesis testing. For this analysis, four independently generated CALPHAD models for the Hf-Si system are considered.

Since finding the appropriate number of parameters in CALPHAD models—or rather, the appropriate model parameter set—is a hard and skillful task, it is necessary to propose systematic approaches that can provide sufficiently reliable results with no need for the manual identification of the best predicting model.While we do not claim that the search for improved CALPHAD models as new experiments/simulations or model-forms become available is render no longer necessary, the proposed information fusion framework provides a robust strategy to continuously improve CALPHAD descriptions without necessarily completely discarding previous assessments.

In our work, Bayesian model averaging (BMA) and an error correlation-based model fusion (CMF) approach are used to combine all the given model results together in different ways with their own specific purposes. In BMA, each model has some probability of being true and the fused estimate is a weighted average of the models. This method is extremely useful in the case of model-building process based on a weighted average over the models' responses, and/or less risk (more confidence) in design based on broader uncertainty bands provided by a weighted average over the uncertainties of the models' responses. On the other hand, the information fusion technique based on the correlations between the model deviations is applied for the purpose of acquiring more precise estimations and lower uncertainties compared to results obtained from each applied individual model.

5.3 CALPHAD Model Description

Four sets of models describing Gibbs free energies of the binary system are considered in the current work. Here we consider a family of models that have been carefully assessed/fitted to available data to avoid artifacts that are not to be expected in this system, such as inverted miscibility gaps. In each of these sets, the intermetallic compounds are described using the line-compound formalism [150], which reads:

$$G^{Hf_k Si_l} = \frac{k}{k+l} {}^0G_{Hf}^{HCP} + \frac{l}{k+l} {}^0G_{Si}^{Diamond} + a + bT \tag{5.1}$$

where k and l are the compound numbers, ${}^0G_{Hf}^{HCP}$ and ${}^0G_{Si}^{Diamond}$ are the chosen energy references corresponding to the energies of pure HCP-Hf and diamond-Si as given in the SGTE database

[151], a and b are model parameters, and T is temperature (in Kelvin).

The liquid phase is described using the sub-regular solid solution model as follows:

$$G^{Liq} = \sum_{i=1}^{N} x_i {}^0G_i^{Liq} + RT \sum_{i=1}^{N} x_i \ln x_i + \sum_{i=1}^{N} x_i \sum_{j \neq i}^{N} x_j \sum_{n=0}^{M} {}^nL_{ij}(x_i - x_j)^n \tag{5.2}$$

where x_i is the mole fraction of the constituent i (either Si or Hf), ${}^0G_i^{liq}$ is the constituent energy taken, again, from the SGTE database [151], T is temperature (in Kelvin), R is the gas constant, and ${}^nL_{ij}$ is given as:

$${}^{\nu}L_{ij} = {}^{\nu}a_{ij} + {}^{\nu}b_{ij}T \tag{5.3}$$

where ${}^{\nu}a_{ij}$ and ${}^{\nu}b_{ij}$ are model parameters, which describe the interactions between the constituents beyond those of ideal mixing.

The terminal phases, namely HCP-Hf and diamond-Si, are described by either the line-compound formalism [150] or the sub-regular solid solution model.

In the current work, four models with 17, 20, 28, and 30 parameters identified as models 1 through 4 are chosen for Hf-Si binary system by expert opinion to show how the model selection and model fusion approaches work in this case study. The parameters associated with the liquid, intermetallic compounds and the end members are listed in Table 5.1 for all the given models.

5.4 Uncertainty Quantification Methodology

5.4.1 Markov Chain Monte Carlo-Metropolis Hastings algorithm

In our work, the MCMC Metropolis Hastings toolbox in Matlab has been utilized for probabilistic calibration of the parameters in the applied CALPHAD models. After the introduction of prior knowledge for parameters (initial values (vector θ^0), lower and upper bounds, and probability density functions (PDFs) in this algorithm, parameter vectors are randomly sampled from a non-stationary proposal posterior probability distribution, which is an arbitrary multivariate Gaussian distribution with a mean value at θ^0 for the first sampling or the last accepted parameter vector dur-

Table 5.1: The parameters of each given model that describe the Gibbs free energy of the phases in the system (G and L denote the parameters for the line-compound and the sub-regular solid solution models, respectively)

Model	Phase	Parameter Specification	Model	Phase	Parameter Specification
1	$HCP-Hf$	${}^0L^{HCP}_{Hf,Si} : {}^0a^{HCP}_{Hf,Si}$	2	$HCP-Hf$	${}^0G^{HCP}_{Hf,Si} : {}^0a^{HCP}_{Hf,Si} - {}^0b^{HCP}_{Hf,Si}T$
	$BCC-Hf$	${}^0L^{BCC}_{Hf,Si} : {}^0a^{BCC}_{Hf,Si}$		$BCC-Hf$	${}^0G^{BCC}_{Hf,Si} : {}^0a^{BCC}_{Hf,Si} - {}^0b^{BCC}_{Hf,Si}T$
	Hf_2Si	${}^0G^{Hf_2Si}_{Hf,Si} : {}^0a^{Hf_2Si}_{Hf,Si} - {}^0b^{Hf_2Si}_{Hf,Si}T$		Hf_2Si	${}^0G^{Hf_2Si}_{Hf,Si} : {}^0a^{Hf_2Si}_{Hf,Si} - {}^0b^{Hf_2Si}_{Hf,Si}T$
	Hf_3Si_2	${}^0G^{Hf_3Si_2}_{Hf,Si} : {}^0a^{Hf_3Si_2}_{Hf,Si} - {}^0b^{Hf_3Si_2}_{Hf,Si}T$		Hf_3Si_2	${}^0G^{Hf_3Si_2}_{Hf,Si} : {}^0a^{Hf_3Si_2}_{Hf,Si} - {}^0b^{Hf_3Si_2}_{Hf,Si}T$
	Hf_5Si_3	${}^0G^{Hf_5Si_3}_{Hf,Si} : {}^0a^{Hf_5Si_3}_{Hf,Si} - {}^0b^{Hf_5Si_3}_{Hf,Si}T$		Hf_5Si_3	${}^0G^{Hf_5Si_3}_{Hf,Si} : {}^0a^{Hf_5Si_3}_{Hf,Si} - {}^0b^{Hf_5Si_3}_{Hf,Si}T$
	Hf_5Si_4	${}^0G^{Hf_5Si_4}_{Hf,Si} : {}^0a^{Hf_5Si_4}_{Hf,Si} - {}^0b^{Hf_5Si_4}_{Hf,Si}T$		Hf_5Si_4	${}^0G^{Hf_5Si_4}_{Hf,Si} : {}^0a^{Hf_5Si_4}_{Hf,Si} - {}^0b^{Hf_5Si_4}_{Hf,Si}T$
	$HfSi$	${}^0G^{HfSi}_{Hf,Si} : {}^0a^{HfSi}_{Hf,Si} - {}^0b^{HfSi}_{Hf,Si}T$		$HfSi$	${}^0G^{HfSi}_{Hf,Si} : {}^0a^{HfSi}_{Hf,Si} - {}^0b^{HfSi}_{Hf,Si}T$
	$HfSi_2$	${}^0G^{HfSi_2}_{Hf,Si} : {}^0a^{HfSi_2}_{Hf,Si} - {}^0b^{HfSi_2}_{Hf,Si}T$		$HfSi_2$	${}^0G^{HfSi_2}_{Hf,Si} : {}^0a^{HfSi_2}_{Hf,Si} - {}^0b^{HfSi_2}_{Hf,Si}T$
	$liquid$	${}^0L^{liq}_{Hf,Si} : {}^0a^{liq}_{Hf,Si} - {}^0b^{liq}_{Hf,Si}T$		$liquid$	${}^0L^{liq}_{Hf,Si} : {}^0a^{liq}_{Hf,Si} - {}^0b^{liq}_{Hf,Si}T$
	Si	${}^0L^{Si}_{Hf,Si} : {}^0a^{Si}_{Hf,Si}$		Si	${}^0G^{Si}_{Hf,Si} : {}^0a^{Si}_{Hf,Si} - {}^0b^{Si}_{Hf,Si}T$
3	$HCP-Hf$	${}^0L^{HCP}_{Hf,Si} : {}^0a^{HCP}_{Hf,Si} - {}^0b^{HCP}_{Hf,Si}T$	4	$HCP-Hf$	${}^0L^{HCP}_{Hf,Si} : {}^0a^{HCP}_{Hf,Si} - {}^0b^{HCP}_{Hf,Si}T$
		${}^1L^{HCP}_{Hf,Si} : {}^1a^{HCP}_{Hf,Si} - {}^1b^{HCP}_{Hf,Si}T$			${}^1L^{HCP}_{Hf,Si} : {}^1a^{HCP}_{Hf,Si} - {}^1b^{HCP}_{Hf,Si}T$
	$BCC-Hf$	${}^0L^{BCC}_{Hf,Si} : {}^0a^{BCC}_{Hf,Si} - {}^0b^{BCC}_{Hf,Si}T$		$BCC-Hf$	${}^0L^{BCC}_{Hf,Si} : {}^0a^{BCC}_{Hf,Si} - {}^0b^{BCC}_{Hf,Si}T$
		${}^1L^{BCC}_{Hf,Si} : {}^1a^{BCC}_{Hf,Si} - {}^1b^{BCC}_{Hf,Si}T$			${}^1L^{BCC}_{Hf,Si} : {}^1a^{BCC}_{Hf,Si} - {}^1b^{BCC}_{Hf,Si}T$
	Hf_2Si	${}^0G^{Hf_2Si}_{Hf,Si} : {}^0a^{Hf_2Si}_{Hf,Si} - {}^0b^{Hf_2Si}_{Hf,Si}T$		Hf_2Si	${}^0G^{Hf_2Si}_{Hf,Si} : {}^0a^{Hf_2Si}_{Hf,Si} - {}^0b^{Hf_2Si}_{Hf,Si}T$
	Hf_3Si_2	${}^0G^{Hf_3Si_2}_{Hf,Si} : {}^0a^{Hf_3Si_2}_{Hf,Si} - {}^0b^{Hf_3Si_2}_{Hf,Si}T$		Hf_3Si_2	${}^0G^{Hf_3Si_2}_{Hf,Si} : {}^0a^{Hf_3Si_2}_{Hf,Si} - {}^0b^{Hf_3Si_2}_{Hf,Si}T$
	Hf_5Si_3	${}^0G^{Hf_5Si_3}_{Hf,Si} : {}^0a^{Hf_5Si_3}_{Hf,Si} - {}^0b^{Hf_5Si_3}_{Hf,Si}T$		Hf_5Si_3	${}^0G^{Hf_5Si_3}_{Hf,Si} : {}^0a^{Hf_5Si_3}_{Hf,Si} - {}^0b^{Hf_5Si_3}_{Hf,Si}T$
	Hf_5Si_4	${}^0G^{Hf_5Si_4}_{Hf,Si} : {}^0a^{Hf_5Si_4}_{Hf,Si} - {}^0b^{Hf_5Si_4}_{Hf,Si}T$		Hf_5Si_4	${}^0G^{Hf_5Si_4}_{Hf,Si} : {}^0a^{Hf_5Si_4}_{Hf,Si} - {}^0b^{Hf_5Si_4}_{Hf,Si}T$
	$HfSi$	${}^0G^{HfSi}_{Hf,Si} : {}^0a^{HfSi}_{Hf,Si} - {}^0b^{HfSi}_{Hf,Si}T$		$HfSi$	${}^0G^{HfSi}_{Hf,Si} : {}^0a^{HfSi}_{Hf,Si} - {}^0b^{HfSi}_{Hf,Si}T$
	$HfSi_2$	${}^0G^{HfSi_2}_{Hf,Si} : {}^0a^{HfSi_2}_{Hf,Si} - {}^0b^{HfSi_2}_{Hf,Si}T$		$HfSi_2$	${}^0G^{HfSi_2}_{Hf,Si} : {}^0a^{HfSi_2}_{Hf,Si} - {}^0b^{HfSi_2}_{Hf,Si}T$
	$liquid$	${}^0L^{liq}_{Hf,Si} : {}^0a^{liq}_{Hf,Si} - {}^0b^{liq}_{Hf,Si}T$		$liquid$	${}^0L^{liq}_{Hf,Si} : {}^0a^{liq}_{Hf,Si} - {}^0b^{liq}_{Hf,Si}T$
		${}^1L^{liq}_{Hf,Si} : {}^1a^{liq}_{Hf,Si} - {}^1b^{liq}_{Hf,Si}T$			${}^1L^{liq}_{Hf,Si} : {}^1a^{liq}_{Hf,Si} - {}^1b^{liq}_{Hf,Si}T$
	Si	${}^0L^{Si}_{Hf,Si} : {}^0a^{Si}_{Hf,Si} - {}^0b^{Si}_{Hf,Si}T$			${}^2L^{liq}_{Hf,Si} : {}^2a^{liq}_{Hf,Si} - {}^2b^{liq}_{Hf,Si}T$
		${}^1L^{Si}_{Hf,Si} : {}^1a^{Si}_{Hf,Si} - {}^1b^{Si}_{Hf,Si}T$		Si	${}^0L^{Si}_{Hf,Si} : {}^0a^{Si}_{Hf,Si} - {}^0b^{Si}_{Hf,Si}T$
					${}^1L^{Si}_{Hf,Si} : {}^1a^{Si}_{Hf,Si} - {}^1b^{Si}_{Hf,Si}T$

ing the next parameter sampling. The sampled parameter vector in each iteration can be accepted or rejected through a criterion in the context of the Bayesian statistic, known as the Metropolis Hastings (M-H) ratio:

$$\text{M-H} = \frac{P(\theta^{cand})P(D|\theta^{cand})}{P(\theta^{i})P(D|\theta^{i})} \frac{q(\theta^{i}|\theta^{cand})}{q(\theta^{cand}|\theta^{i})} \tag{5.4}$$

where θ^{i}, θ^{cand}, and D are the last accepted sample of parameter vector, the new sample of parameter vector as a candidate, and the given data, respectively. It should be noted that some synthetic data sampled from the Hf-Si calculated phase diagram in [147] has been used as D in this work.

Based on Bayes' theorem, the joint (posterior) probability in each case is proportional to the prior probability times the likelihood, $P(\theta)P(D|\theta)$. The likelihood is the conditional probability of the data, D, given the parameter vector, θ, and in this work is considered as a Gaussian distribution centered at the data, D, with variances, σ^2, determined by the data errors. Here, the variances are assumed to be the same for all the data and updated as a hyper-parameter during MCMC sampling since the data error is unknown. In this case, the variance samples are generated from an inverse gamma posterior PDF resulting from the introduction of a non-informative inverse gamma prior PDF for the variance σ^2 [52, 70]. The second ratio in Equation 5.4 is the Hastings ratio, which considers the asymmetric effect of the proposal probability distribution in the acceptance/rejection criterion of the parameter vector.

After the calculation of the M-H ratio for each sampling iteration, if $\min\{\text{M-H}, 1\} \times 100$ is greater than 1, the candidate for the parameter vector is accepted as the new sample; otherwise, the candidate may still be accepted with a probability of $\min\{\text{M-H}, 1\} \times 100$ [121]. In the case of the candidate rejection, the last accepted parameter vector is repeated in the MCMC chain as the new sample. At the end of the MCMC sampling process, a chain of the parameter vectors is obtained as $\left\{\theta^{0}, ..., \theta^{N}\right\}$ whose mean values and variance-covariance matrix after removal of "burn-in period" demonstrate the plausible optimum values of the parameters and their overall reflected uncertainties, respectively. It should be noted that the parameters' convergence during

MCMC sampling algorithm is defined as the convergence of the parameters' cumulative mean values towards almost constant values, which is equivalent to the fluctuation of the parameters' samples around their mean values across MCMC sampling [6].

5.4.2 Bayesian model selection/averaging

Jeffreys [14] proposed a hypothesis testing method in the context of Bayesian statistics that evaluates the evidence in favor of a scientific theory or hypothesis. Models are considered as hypotheses in the case of Bayesian model selection (BMS). In BMS, non-nested models (models with at least one uncommon parameter) can also be compared together, which is very difficult or sometimes impossible through frequentist approaches [152]. Generally, models/hypotheses with greater posteriors are more favored by the evidence and considered as better models/hypotheses. Accordingly, the Bayesian comparison criterion for models/hypotheses is defined as the posterior odds of each two applied models/hypotheses given data. For instance, posterior odds of model one (M_1) to model two (M_2) given the data (D) is expressed as follows:

$$\frac{P(M_1|D)}{P(M_2|D)} = \frac{P(D|M_1)}{P(D|M_2)}\frac{P(M_1)}{P(M_2)} = B_{12}\lambda_{12} \tag{5.5}$$

Bayes' factor is the ratio of the marginal likelihoods and usually suggests which model/hypothesis is more favored by the evidence (data):

$$P(D|M_k) = \int P(D|\theta_k, M_k)P(\theta_k|M_k)d\theta_k \tag{5.6}$$

where θ_k is the parameter vector under M_k. Therefore, the key element in BMS is the calculation of the marginal likelihoods offered by Equation 5.6, which is challenging in the case of high dimensional θ_k. There are different methods to approximate these integrals, which include Laplace's method, the Schwarz criterion, simple Monte Carlo, importance sampling, adaptive Gaussian quadrature, and simulating approaches from the parameters' posterior [152]. In our work, the MCMC sampling method is used to simulate from the parameters' posterior for the integral estimation.

In the Monte Carlo method, the integral in Equation 5.6 is approximated as:

$$P(D|M_k) \approx \frac{1}{N}\sum_{i=1}^{N} P(D|\theta_k^{(i)}, M_k) \tag{5.7}$$

where $\theta_k^{(i)}$s are samples from the parameters' prior $P(\theta_k|M_k)$. This approximation is equivalent to the average of the likelihood values obtained from the sampled parameters. However, this approach can be fairly inefficient in the case of a more concentrated posterior distribution compared to the prior distribution since the likelihood values for most of the sampled $\theta_k^{(i)}$ will be very small and the approximation result will be considerably weighted by a few samples with large likelihood values. For this reason, importance sampling techniques are usually applied [152]:

$$P(D|M_k) \approx \frac{1}{N}\sum_{i=1}^{N} \frac{P(D|\theta_k^{(i)}, M_k)P(\theta_k^{(i)}|M_k)}{P^*(\theta_k^{(i)})} \tag{5.8}$$

$$P(D|M_k) \approx \frac{1}{N}\sum_{i=1}^{N} \frac{P(D|\theta_k^{(i)}, M_k)w_k^{(i)}}{w_k^{(i)}} \tag{5.9}$$

where $P^*(\theta_k)$ is a suitable distribution that enables the sampling of the important parameter values for the sake of more efficient sampling. In Equation 5.9, $w_k^{(i)} = \frac{P(\theta_k^{(i)}|M_k)}{P^*(\theta_k^{(i)})}$, which can provide a weighted average of likelihood values obtained using the sampled parameters from $P^*(\theta_k)$.

In the case that the importance sampling function is proportional to the posterior probability density function of the parameters in model k, i.e., $P^*(\theta_k) = P(D|\theta_k, M_k)P(\theta_k|M_k)$, Equation 5.9 is turned into:

$$P(D|M_k) \approx \left\{\frac{1}{N}\sum_{i=1}^{N} P(D|\theta_k^{(i)}, M_k)^{-1}\right\}^{-1} \tag{5.10}$$

MCMC is usually used to sample $\theta_k^{(i)}$s from $(P^*(\theta_k))$. Equation 5.10 shows a harmonic mean of likelihood values obtained from the MCMC sampled parameters after the removal of the burn-in period.

In addition to model selection, Bayes' factors can be used to determine the posterior probability distribution of the competing models given data, which can be defined as their associated weight in the context of Bayesian model averaging (BMA). Generally, BMA can provide a combined inference from all the competing models that can be very useful for model-building process or less risky predictions in design. In this approach, the posterior density (weight) associated with each model can be obtained as follows:

$$P(M_k|D) = \frac{P(D|M_k)P(M_k)}{\sum_{i=1}^{K} P(D|M_i)P(M_i)} = \frac{B_{k1}\lambda_{k1}}{\sum_{i=1}^{K} B_{i1}\lambda_{i1}} = \frac{B_{k1}}{\sum_{i=1}^{K} B_{i1}} \tag{5.11}$$

where K is the total number of models. In this equation, M_1 is considered as reference for the calculation of all the Bayes' factors, and all the prior odds are also 1. The posterior densities of any quantity of interest (Δ) existing in all the competing models can be combined together through BMA as [152–154]:

$$P(\Delta|D) = \sum_{i=1}^{K} P(\Delta|D, M_i)P(M_i|D) \tag{5.12}$$

5.4.3 Applied error correlation-based model fusion

In the case of multiple uncertain sources of information (e.g., different models for the same problem), there is a need to integrate all the sources to produce more reliable results [9]. In practice, there are several approaches for fusing information from multiple models. BMA is a model fusion technique that has some benefits in robust design. Other available techniques are fusion under known correlation [155–158], and the covariance intersection method [159]. The key distinction of BMA over other model fusion approaches is the assumption of statistical independence among models, which may be incorrect in some cases and can lead to potentially serious misconceptions regarding confidence in quantity of interest estimates.

A fundamental claim in this work is that *any model can provide potentially useful information to a given task.* We thus seek to take into account all potential information any given model may provide and fuse unique information from the available models. Our fusion goal then is to identify dependencies, via estimated correlations, among the model discrepancies. With these estimated correlations, the models are fused following standard practice for the fusion of normally distributed data.

Under the case of known correlations between the discrepancies of models, the fused mean and variance at a design point, $\mathbf{x}$, are shown to be [157]

$$E[\hat{f}(\mathbf{x})] = \frac{\mathbf{e}^\top \tilde{\Sigma}(\mathbf{x})^{-1} \mu(\mathbf{x})}{\mathbf{e}^\top \tilde{\Sigma}(\mathbf{x})^{-1} \mathbf{e}} \tag{5.13}$$

$$\text{Var}\left(\hat{f}(\mathbf{x})\right) = \frac{1}{\mathbf{e}^\top \tilde{\Sigma}(\mathbf{x})^{-1} \mathbf{e}} \tag{5.14}$$

where $\mathbf{e} = [1, \ldots, 1]^\top$, $\mu(\mathbf{x}) = [\mu_1(\mathbf{x}), \ldots, \mu_S(\mathbf{x})]^\top$ is the vector of mean values given S models, and $\tilde{\Sigma}(\mathbf{x})^{-1}$ is the inverse of the covariance matrix between the models given as

$$\tilde{\Sigma} = \begin{bmatrix} \sigma_1^2 & \rho_{12}\,\sigma_1\sigma_2 & \cdots & \rho_{1S}\,\sigma_1\sigma_S \\ \rho_{12}\,\sigma_1\sigma_2 & \sigma_2^2 & \cdots & \rho_{2S}\,\sigma_2\sigma_S \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1S}\,\sigma_1\sigma_S & \rho_{2S}\,\sigma_2\sigma_S & \cdots & \sigma_S^2 \end{bmatrix} \tag{5.15}$$

where ρ_{ij} is the correlation coefficient between the deviations of models i and j at point $\mathbf{x}$, and σ_i^2 is the variance of model i at point $\mathbf{x}$.

To estimate the correlations between the model deviations in the case that they are unknown, we use the reification process defined in [10, 11], which refers to the process of treating each model, in turn, as ground truth. This means that we assume the data generated by the reified model represents the true quantity of interest. These data are used to estimate the correlation between the errors of the different models. The process is repeated for each model. The detailed process of estimating the correlation between the errors of two models can be found in [10, 11].

Following Thomison et al. [11], assuming model i is chosen to reify, the correlation coefficients between the models i and j, for $j = 1, \ldots, i-1, i+1, \ldots, S$, are given as:

$$\rho_{ij}(\mathbf{x}) = \frac{\sigma_i^2(\mathbf{x})}{\sigma_i(\mathbf{x})\sigma_j(\mathbf{x})} = \frac{\sigma_i(\mathbf{x})}{\sqrt{(\mu_i(\mathbf{x}) - \mu_j(\mathbf{x}))^2 + \sigma_i^2(\mathbf{x})}} \tag{5.16}$$

where $\mu_i(\mathbf{x})$ and $\mu_j(\mathbf{x})$ are the mean values of models i and j respectively at design point $\mathbf{x}$, and $\sigma_i^2(\mathbf{x})$ and $\sigma_j^2(\mathbf{x})$ are the variances at point $\mathbf{x}$. The first subscript under the correlation coefficient denotes which model has been reified. Since the only information we have regarding which model we believe to be more realistic is the variance of each, in addition to reifying model i to estimate the correlation, we also reify information source j and estimate $\rho_{ji}(\mathbf{x})$. We then estimate the correlation between the errors as the variance-weighted average of the two correlation coefficients as follows:

$$\bar{\rho}_{ij}(\mathbf{x}) = \frac{\sigma_j^2(\mathbf{x})}{\sigma_i^2(\mathbf{x}) + \sigma_j^2(\mathbf{x})}\rho_{ij}(\mathbf{x}) + \frac{\sigma_i^2(\mathbf{x})}{\sigma_i^2(\mathbf{x}) + \sigma_j^2(\mathbf{x})}\rho_{ji}(\mathbf{x}) \tag{5.17}$$

These correlations can then be used to estimate the mean and variance of the fused estimate from Equations 5.13 and 5.14.

5.5 Results and Discussion

In CALPHAD modeling, it is a hard task, even for CALPHAD experts, to know what formalism should be used to describe the Gibbs free energy of each phase in the system and to what order the polynomial term in Equation 5.2 should be expanded in the case of the sub-regular solid solution model consideration in order to obtain the closest phase diagram to the available data after parameter optimization. For this reason, different models may be suggested for a system with different formalisms for the Gibbs free energy description of the phases and different orders of polynomial expansions for the sub-regular solid solution models. In the context of Bayesian statistics, the best model among the selected candidates according to the available data can be identified through the calculation of the Bayes' factor based on Equation 5.5.

Each of the models are probabilisticly calibrated against the available calculated data for the phase diagram using one million MCMC samples of the parameter vectors. For each model, the deterministically optimized values of the parameters obtained from the PARROT module in Themo-

Calc software (through least squares) and ± 3 times these values are considered as the initial values (θ_0) and the ranges of the parameters, respectively. This prior information enables much faster convergence of the model parameters during MCMC sampling. However, it is assumed that the deterministic optimization in Thermo-Calc is global, which is not necessarily true as we discuss later. It should be noted that the compositions of stable phases estimated from the convex hull of the phase Gibbs energies at any given temperature are compared with the corresponding available data through the likelihood during the sampling process. Here, the results obtained for model 2 are discussed in detail to show how MCMC is used in this work for probabilistic calibration of the model parameters. The same approach is applied for the calibration of the other three model parameter sets.

During MCMC sampling, there are some homogeneous fluctuations around plausible optimum values of parameters after convergence. However, this means the cumulative means of parameter samples should converge to constant values. Therefore, plotting cumulative mean distributions of parameters can help identify the convergence regions. As examples, two of these plots are shown in Figure 5.1. According to this figure, smooth changes towards constant values are observed for both parameters after 30,000 parameter generations, which correspond to the parameter convergence regions. Therefore, the first 30,000 generated samples are considered as the burn-in period (red shaded regions) and discarded from the total number of parameter samples. Since all the parameters generally converge at the same time, the burn-in period is assumed to be the same for the other model parameters (the other 18 parameters in model 2). To show that the parameters in other models also converge to their optimum values, the cumulative means of parameter a_{liq} are plotted for the other models as dotted lines in Figure 5.1(a).

After the removal of the burn-in period, the mean values of the remaining parameter samples (970,000 samples) and the diagonal terms of their variance-covariance matrix are considered as the optimal plausible mean values and the overall reflected uncertainties (variances) for the model parameters, respectively. The initial values of the model parameters obtained through a deterministic optimization and their probabilistic values after MCMC calibration are listed in Table 5.2.

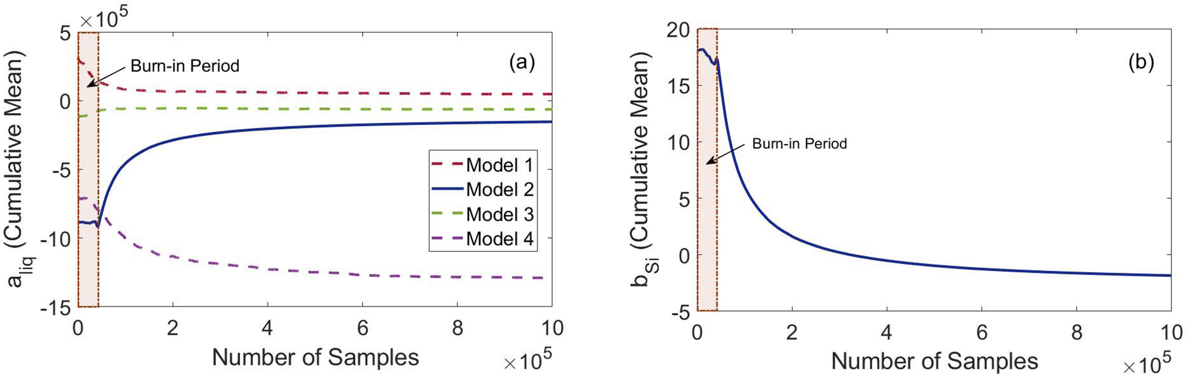


Figure 5.1: Cumulative mean plots of the parameters a) a_{liq} and b) b_{Si} in model 2. The red shaded regions show the number of parameters belonging to the burn-in period

Table 5.2: Initial values for the parameters in model 2 and their plausible optimum values and overall uncertainties after MCMC calibration

Phase	**Parameter Specification**	**Deterministically Calibrated Initial Values**	**Probabilistically Calibrated Values after MCMC**
$HCP - Hf$	${}^{0}G^{HCP}_{Hf,Si} : {}^{0}a^{HCP}_{Hf,Si} - {}^{0}b^{HCP}_{Hf,Si}T$	${}^{0}G^{HCP}_{Hf,Si} : -455420 + 221.79T$	${}^{0}G^{HCP}_{Hf,Si} : -29070 \pm 16377 + 15.08 \pm 8.10T$
$BCC - Hf$	${}^{0}G^{BCC}_{Hf,Si} : {}^{0}a^{BCC}_{Hf,Si} - {}^{0}b^{BCC}_{Hf,Si}T$	${}^{0}G^{BCC}_{Hf,Si} : -45254 + 18.56T$	${}^{0}G^{BCC}_{Hf,Si} : 10437 \pm 1845 - 6.17 \pm 1.10T$
Hf_2Si	${}^{0}G^{Hf_2Si}_{Hf,Si} : {}^{0}a^{Hf_2Si}_{Hf,Si} - {}^{0}b^{Hf_2Si}_{Hf,Si}T$	${}^{0}G^{Hf_2Si}_{Hf,Si} : -454090 + 90.46T$	${}^{0}G^{Hf_2Si}_{Hf,Si} : -31675 \pm 6672 - 3.13 \pm 1.60T$
Hf_3Si_2	${}^{0}G^{Hf_3Si_2}_{Hf,Si} : {}^{0}a^{Hf_3Si_2}_{Hf,Si} - {}^{0}b^{Hf_3Si_2}_{Hf,Si}T$	${}^{0}G^{Hf_3Si_2}_{Hf,Si} : -481270 + 84.19T$	${}^{0}G^{Hf_3Si_2}_{Hf,Si} : -32102 \pm 7168 - 5.78 \pm 1.61T$
Hf_5Si_3	${}^{0}G^{Hf_5Si_3}_{Hf,Si} : {}^{0}a^{Hf_5Si_3}_{Hf,Si} - {}^{0}b^{Hf_5Si_3}_{Hf,Si}T$	${}^{0}G^{Hf_5Si_3}_{Hf,Si} : -454950 + 77.55T$	${}^{0}G^{Hf_5Si_3}_{Hf,Si} : - - 29246 \pm 6805 - 6.27 \pm 1.52T$
Hf_5Si_4	${}^{0}G^{Hf_5Si_4}_{Hf,Si} : {}^{0}a^{Hf_5Si_4}_{Hf,Si} - {}^{0}b^{Hf_5Si_4}_{Hf,Si}T$	${}^{0}G^{Hf_5Si_4}_{Hf,Si} : -522680 + 100.88T$	${}^{0}G^{Hf_5Si_4}_{Hf,Si} : -33133 \pm 7614 - 5.60 \pm 1.80T$
$HfSi$	${}^{0}G^{HfSi}_{Hf,Si} : {}^{0}a^{HfSi}_{Hf,Si} - {}^{0}b^{HfSi}_{Hf,Si}T$	${}^{0}G^{HfSi}_{Hf,Si} : -503250 + 98.26T$	${}^{0}G^{HfSi}_{Hf,Si} : -32910 \pm 7595 - 5.47 \pm 1.90T$
$HfSi_2$	${}^{0}G^{HfSi_2}_{Hf,Si} : {}^{0}a^{HfSi_2}_{Hf,Si} - {}^{0}b^{HfSi_2}_{Hf,Si}T$	${}^{0}G^{HfSi_2}_{Hf,Si} : -468950 + 132.66T$	${}^{0}G^{HfSi_2}_{Hf,Si} : -32518 \pm 7259 + 0.36 \pm 2.43T$
$liquid$	${}^{0}L^{liq}_{Hf,Si} : {}^{0}a^{liq}_{Hf,Si} - {}^{0}b^{liq}_{Hf,Si}T$	${}^{0}L^{liq}_{Hf,Si} : -886370 - 9.47T$	${}^{0}L^{liq}_{Hf,Si} : -119005 \pm 13962 + 16.72 \pm 3.06T$
Si	${}^{0}G^{Si}_{Hf,Si} : {}^{0}a^{Si}_{Hf,Si} - {}^{0}b^{Si}_{Hf,Si}T$	${}^{0}G^{Si}_{Hf,Si} : -114810 + 68.22T$	${}^{0}G^{Si}_{Hf,Si} : 6842 \pm 794 - 2.72 \pm 0.32T$

The marginal and joint posterior frequency distributions of the parameters can also be plotted to evaluate the convergence and dispersion of samples in the parameter space. Two examples of marginal posterior frequency histograms for parameters are plotted in Figure 5.2. As can be observed in this figure, the parameter posteriors are almost in the form of Gaussian distributions. Since all the parameters show a similar form for their marginal posterior distributions, a distribution close to a multivariate Gaussian distribution is expected as the joint distribution for all the model parameters (20 parameters in the case of model 2). Figure 5.3 shows the joint frequency distributions for two examples of a pair of parameters in model 2. Figure 5.3 can also offer a qualitative representation of the correlation between the applied two parameters in each case. For example, the linearity and direction of the red regions show the strength and negativity/positivity of the correlation between parameters, quantified through the Pearson coefficient [56]:

$$\rho_{x,y} = \frac{cov(x,y)}{\sigma_x \sigma_y} \tag{5.18}$$

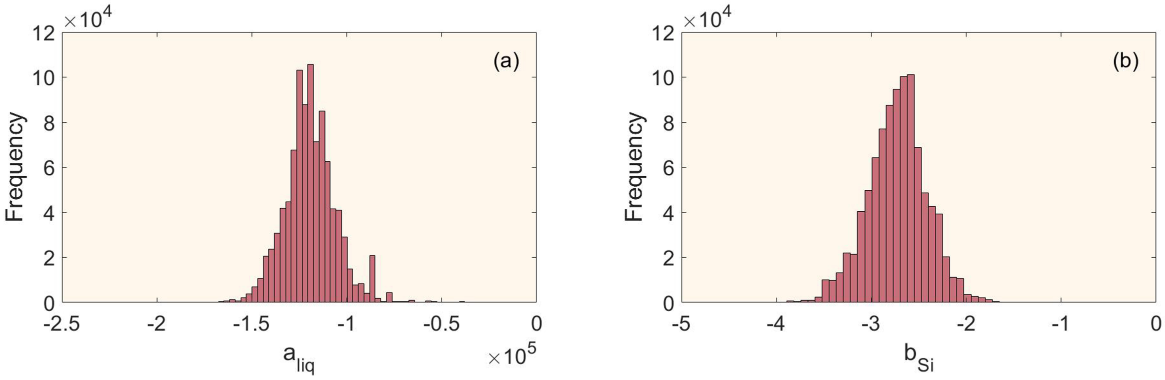


Figure 5.2: Marginal posterior frequency distributions for the parameters a) a_{liq} and b) b_{Si} in model 2

The same MCMC approach is applied to probabilistically calibrate the other three model parameter sets. Then, the overall uncertainties are propagated from parameters to phase diagram for

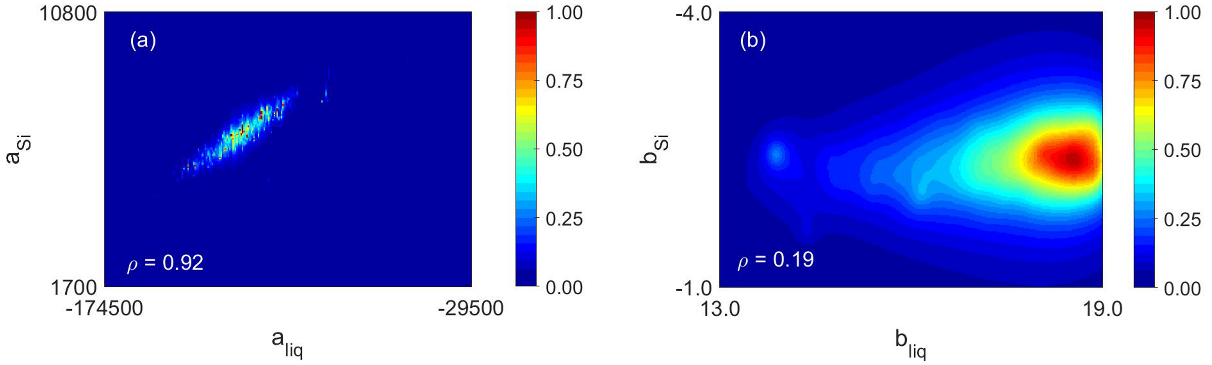


Figure 5.3: Joint frequency distributions between the pair parameters a) (a_{Si},a_{liq}) and b) (b_{Si},b_{liq}) in model 2 besides their linear correlation coefficients and normalized colour bars

each model through the model forward analysis. This process is performed using the forward calculation of the phase diagram by the last 5,000 MCMC parameter samples as an ensemble of the whole convergence region. To find 95% Bayesian credible intervals (BCI), 2.5% of the resulting samples associated with liquidus and transformation lines are discarded from above and also below the total obtained uncertainty band at any specified composition. These results are shown for each model in Figure 5.4. In this figure, red lines and shaded areas are the results obtained from the optimum (mean) values of the parameters and 95% BCI for each model, respectively. It should be noted that these uncertainty bands are not sharp around eutectic points since the eutectic compositions are predicted slightly different for each parameter sample. A cross-sectional probability distribution can be achieved at any specific Si content along the green uncertainty interval, while there is a fixed (composition independent) cross-sectional probability distribution throughout any one of the blue or red uncertainty intervals. The red shaded regions in all models show unstable results with large uncertainties that result from the difficulty to predict the intermediate high temperature Hf_5Si_3 phase through CALPHAD modeling. Moreover, it is clear in Figure 5.4 that there is a very good agreement between the results obtained from model 2 and the data with a very small uncertainty band.

For more clear demonstration of UP in this work, the propagation of uncertainty from the model

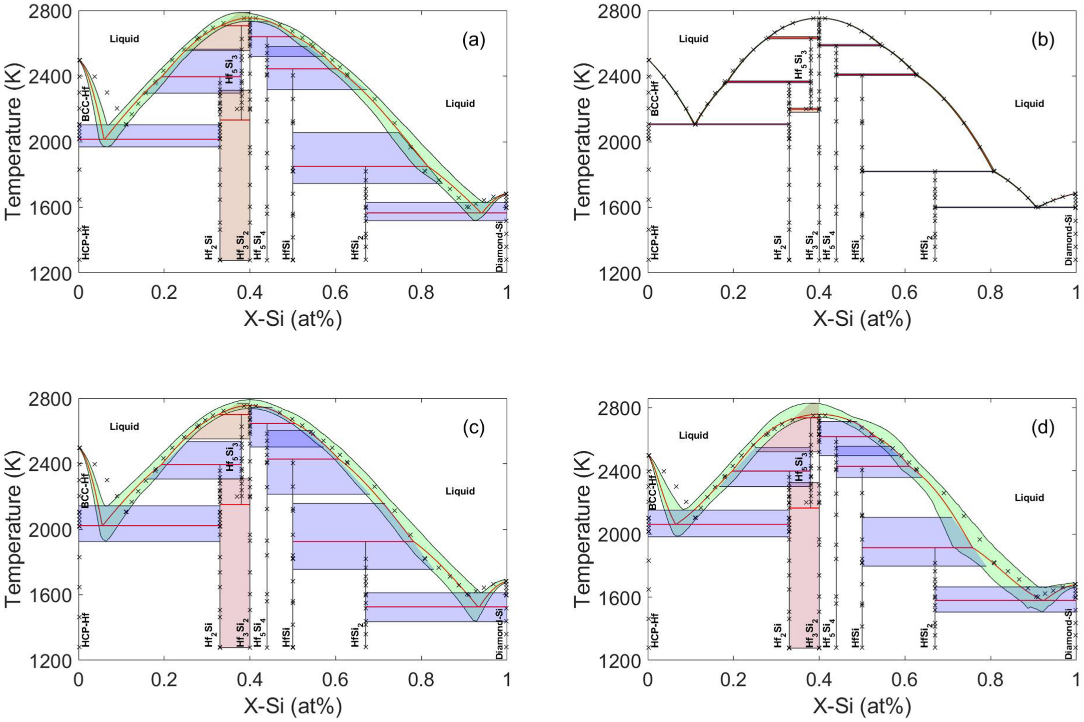


Figure 5.4: Optimum Hf-Si phase diagrams and their 95% BCIs obtained from models 1-4 (a-d) after uncertainty propagation of the MCMC calibrated parameters in each case

parameters to the single-phase Gibbs free energies to phase diagram are illustrated in Figure 5.5. In this figure, two isothermal examples of Gibbs free energy-Si composition curves are plotted for liquid and two different intermetallic phases in two models (i.e., models 1 and 2) to show how the uncertainties of single-phase Gibbs free energies and phase diagram can be related together. This relationship can clearly be observed in Figure 5.5 through the consistency between the 95% BCIs of the phase equilibria obtained from parameter uncertainty propagations to Gibbs free energy of the single phases and resulting convex hulls at any specific temperature and their counterparts obtained from parameter uncertainty propagations to the corresponding phase diagram boundaries.

It is worth noting that the mean responses and uncertainty intervals in the Gibbs free energy graphs and phase diagrams have probably been affected by MSU. Generally, there are some statistical approaches such as Kennedy and O'Hagan approach [12] to quantify MSU in the presence of true data [3]. However, model 2 shows very small uncertainties and close mean value responses to the data after MCMC calibration that implies a very small MSU in this case. Therefore, larger models, i.e., models 3 and 4, are expected to have even less MSU due to fitting better to the data in the case that the global optimum for parameters are obtained during calibration (although the mean values obtained from models 3 and 4 are not close to the data in some parts of the phase diagrams due to the inability of the parameter global optimization in Thermo-calc software). For this reason, the large uncertainties for the phase diagrams obtained from models 3 and 4 are mostly attributed to MPU.

In the context of BMS/BMA, the weight of each one of the applied models can be obtained by calculating their corresponding posteriors given data using Equations 5.10 and 5.11. Based on these calculations, the weights associated with models 1 to 4 are 0.1352, 0.5938, 0.1331, and 0.1379, respectively. Model 2 thus has three times the weight of the other models, which otherwise have similar Bayesian importance, which is consistent with the phase diagram results in Figure 5.4.

Traditionally, it has been generally expected that a better fit to experimental data can be achieved by simply increasing the number of parameters of the free energies. However, our results show that this is not always the case. The results hold, for example, when comparing model 1 to model 2, as

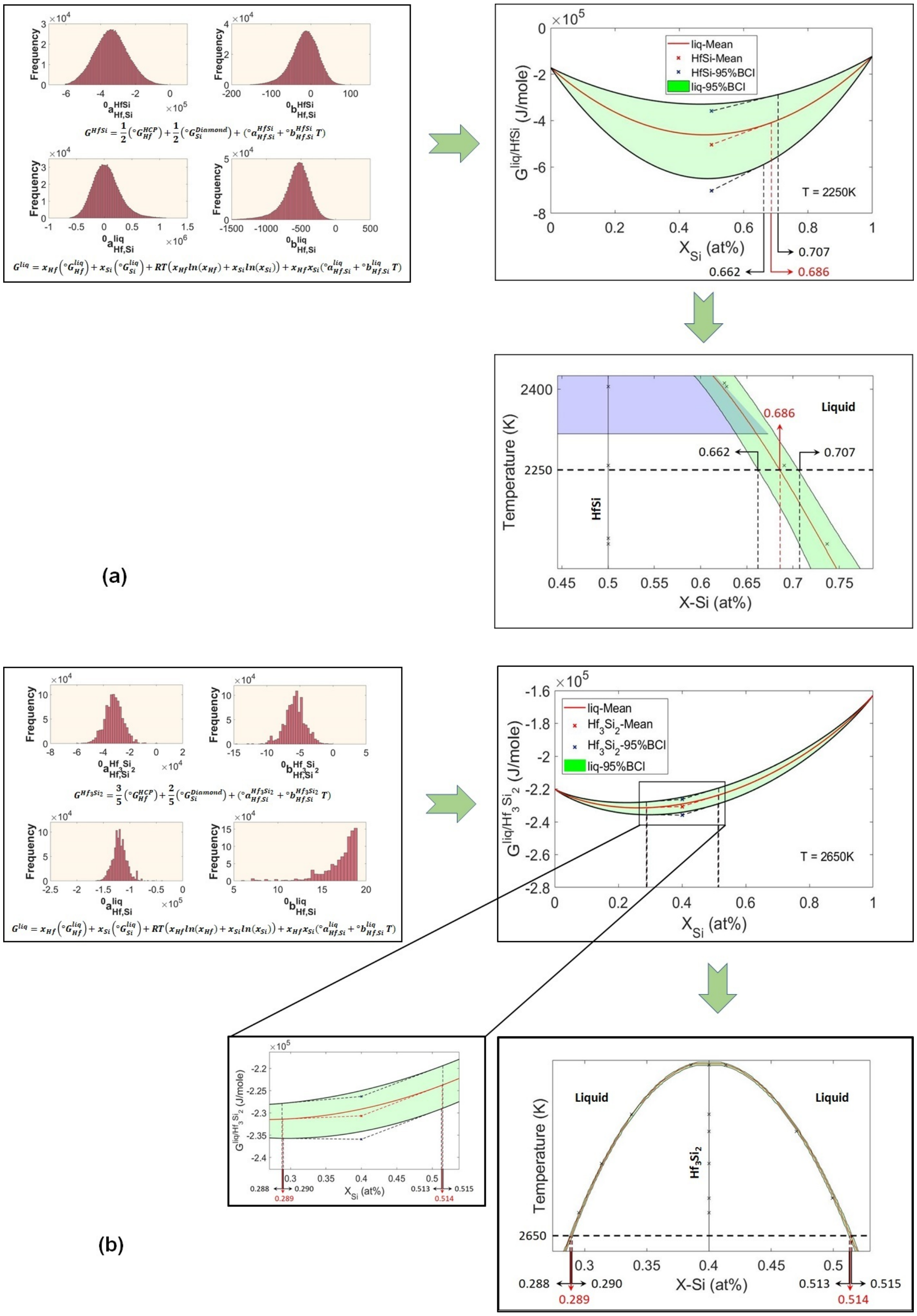


Figure 5.5: Gibbs free energy-Si composition graphs with their 95% BCIs obtained from the parameter uncertainty propagation for a) liquid and HfSi intermetallic phases in model 1 at 2250K and b) liquid and Hf_3Si_2 intermetallic phases in model 2 at 2650K besides their corresponding phase diagram regions to show the uncertainty propagation from the parameters to single-phase Gibbs free energies to phase diagram

increasing the complexity of the free energy functions clearly results in a much better fit to the data and a much narrower uncertainty bound. As the complexity of the models are increased, however, it is apparent that the uncertainty bounds become much worse. In this case, we have found that the most parsimonious model (i.e., simplest) that still has sufficient freedom to explain all the available data is clearly superior.

These results, although counter-intuitive to a degree, may be explained by imagining the fitness landscape of a particular model as a multi-dimensional space, with many local minima corresponding to combinations of parameters that result in a fit with lower error relative to the neighborhood in model parameter space. Schematically, one can visualize this as shown in Figure 5.6. The simplest model (blue) is depicted as having a relatively simple landscape with a sub-optimal fitness: the model is too simple to explain all the available data and the large uncertainty bounds are associated with the fitness error as a result of (small) variations in the values of individual parameters. Very complex models (red) have a high number of local minima as there are many combinations of parameters that result in somewhat-optimal fits. The fitting process, however, can get stuck in these local minima and since the MCMC results are inevitably biased by the parameter ranges defined based on the deterministically-attained values—to sample the model space at a reasonable cost—there is a significant probability that the global minimum is outside the parameter ranges, which results in an MCMC chain that converges to a sub-optimal region during MCMC sampling. The large uncertainty bounds in this case could be ascribed to the large sensitivity of such models to small deviations in the values of the parameters. The best model (green) has the right amount of complexity to fit all the available data without having too much model parameter uncertainty. There is thus a valid argument towards *model parsimony*: to make the model as simple as possible, but not more.

It must be stated, however, that while there is a clear superiority in model 2 in this case, the value of the inferior models should not be discounted outright as they may provide valuable information that can contribute to a more robust quantification of uncertainty. Based on Equation 5.12, weighted averages are taken over the cross-sectional probability distributions of all four phase dia-

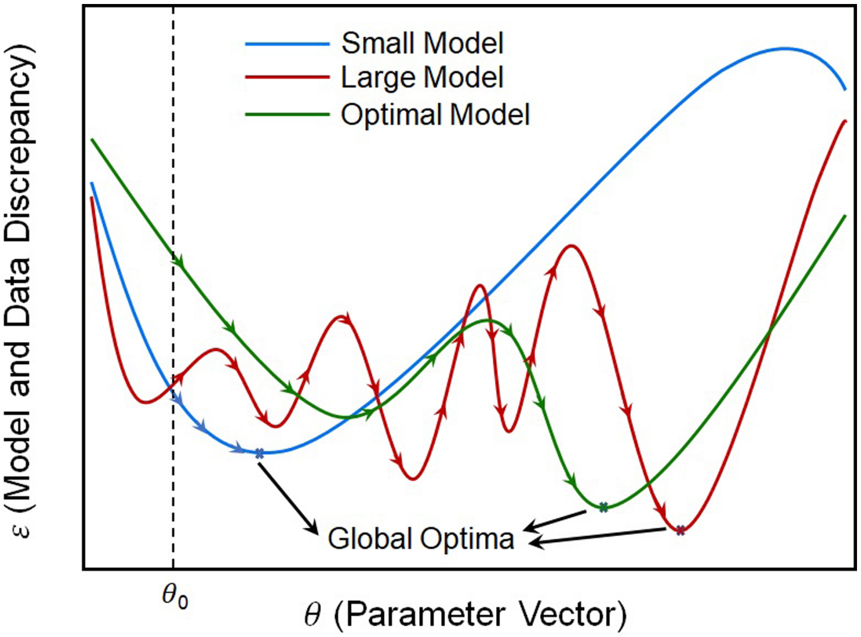


Figure 5.6: 2D schematic plots of the discrepancy between model and data in the parameter space for different model sizes (different number of parameters)

grams in Figure 5.4 at around 1000 Si compositions between 0 and 1 with a step size 0.001 for the green region and at just one arbitrary composition for each of the blue/red regions (since the cross-sectional probability distributions are fixed and composition independent along these regions) to calculate the average model and its uncertainty bands throughout the whole phase diagram, as shown in Figure 5.7. For each individual model and weighted average model, the cross-sectional probability distributions obtained through kernel density estimation (KDE) are also demonstrated for all the transformation line uncertainty bands (the blue shaded regions with composition independent distributions) labeled from 1 to 8 and at some random compositions for the liquidus uncertainty band (across the green shaded region) labeled from 9 to 13 in Appendix B for Figure 5.7. As can be observed, 95% BCIs of the average model (black lines in probability distribution plots) is broader than each one of the applied models and covers all the uncertainty ranges offered by the models 1 to 4. This is why BMA can provide more confidence in the results in the context of robust design.

As observed in Figure 5.7, the phase diagram resulting from the mean value of probability

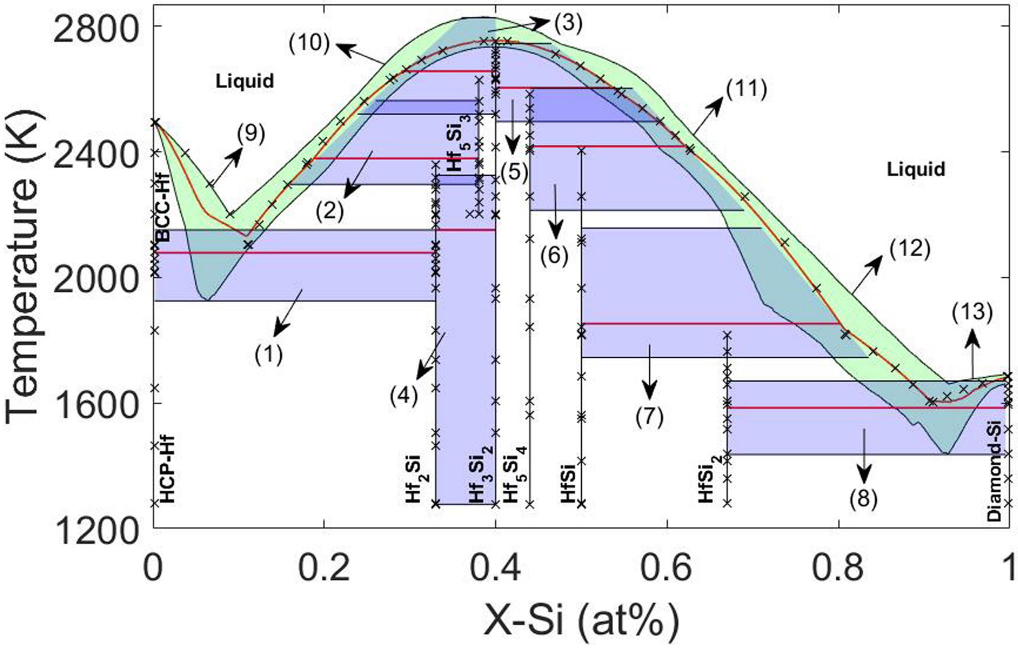


Figure 5.7: Mean values and 95% BCIs at different compositions/regions in Hf-Si phase diagram obtained after BMA

distributions at different compositions and regions (red lines) leads to two unrealistic results (indentations) around the eutectic points at approximately $X_{Si} = 0.1$ and 0.9. According to Figure 5.7, this issue can be attributed to the difference of models in the prediction of eutectic point compositions, which results in averages of some hypo-eutectic points from model 2 with some hyper-eutectic points from models 1, 3 and 4 at small composition ranges. In order to solve this issue, the posterior modes of the probability distributions can be introduced as the optimum phase diagram rather than their mean values. The result is shown by red lines in Figure 5.8. It is worth noting that the posterior mode of the probability distributions in the average model exactly corresponds to the posterior modes of the probability distributions in model 2, as indicated by black and purple lines in distribution plots of Appendix B for Figure 5.7, respectively. Therefore, the best model results can be considered as the optimum results for the average model, but with broader uncertainties, contributed by the inferior models.

Generally, the dependency of the model precision to the number of parameters defined for the

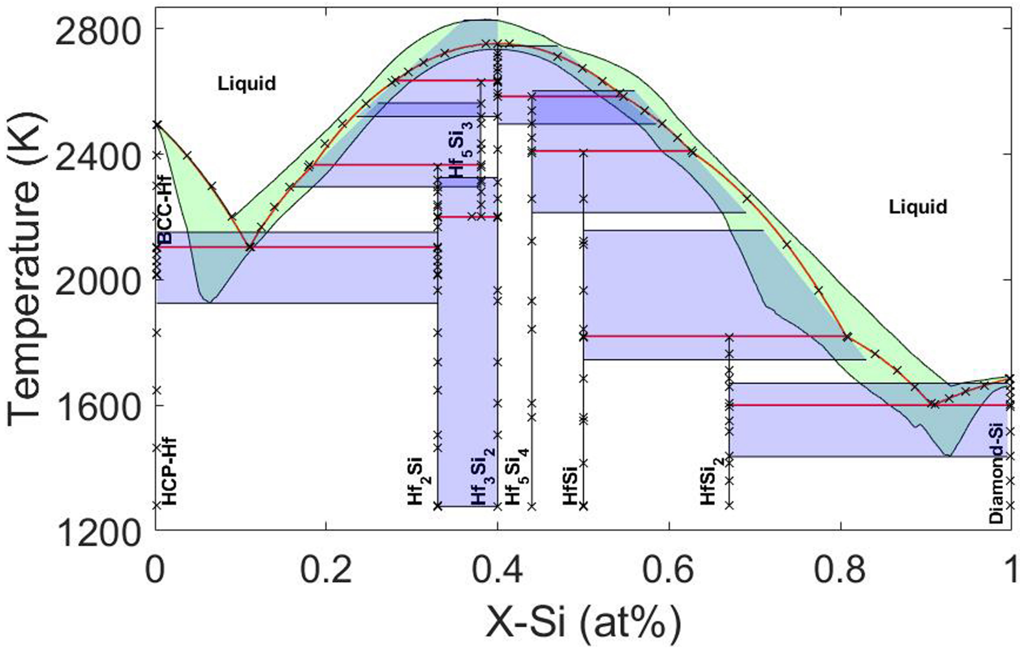


Figure 5.8: Posterior modes and 95% BCIs at different compositions/regions in Hf-Si phase diagram obtained after BMA

existing phases in the system has made CALPHAD modeling a hard task that requires a lot of experience. Therefore, the lack of a systematic approach towards thermodynamic assessments means that model search is essentially carried out in a trial-and-error mode or by relying on (considerable) expert opinion. While the BMA approach already put forward in this work provides a robust estimate of the (weighed) output of a set of models, one drawback is that BMA assumes that models are statistically independent from each other. In this specific case, since the three suboptimal models have almost as much weight at the optimal model, the uncertainties carried out by the former over-estimate (in a very conservative manner) the fused uncertainty. All models (at least in this case) have some degree of correlation as (i) they are describing the same underlying ground truth; (ii) have common model structures. We thus propose to exploit the statistical correlations between models as a strategy to arrive at an improved fused prediction of phase stability.

To perform this model fusion, the equivalent normal distributions at different compositions-/regions are required to be calculated for each of the four models. Then, three and four models are fused based on the CMF approach, whose results are shown in Figure 5.9. It should be noted that models 1, 3, and 4 with low precision and high uncertainties are first chosen for model fusion to examine whether the resulting fused model can be closer to the data and reduce the uncertainties. Figure 5.9(a) shows that the approach can provide a phase diagram in much better agreement with data and less uncertainties compared to phase diagrams obtained from each one of the applied models individually. This result implies that random CALPHAD models can be fused together to find a reasonable estimation for phase diagram instead of trial-and-error to find the best predicting model. In addition, it is obvious that better predictions can be achieved as shown in Figure 5.9(b) if model 2 (the best model) is also involved in the model fusion. However, it is very hard to compare the uncertainty of the fused model and model 2 based on the resulting phase diagrams. For this reason, the information (Shannon) entropies are calculated to quantify the model uncertainties.

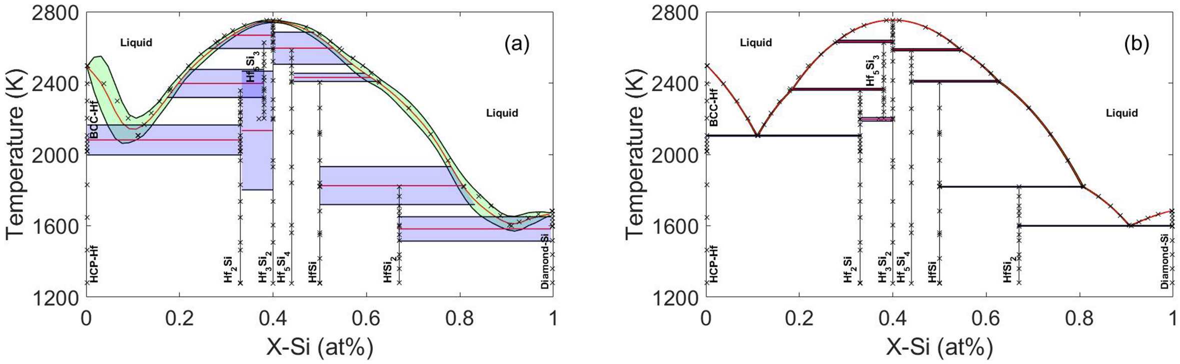


Figure 5.9: Error correlation-based model fusions of a) three models (1, 3, and 4) and b) all four models

Shannon entropy can be utilized as a measure of uncertainty or missing information, which is determined as a weighted average (expected value) of information content gained from all the possible outcomes of an event. The information content obtained from an event outcome i is defined as the negative logarithm of its probability, $-log(p_i)$. In this regard, it should be noted that event outcomes with lower probabilities convey more information since their occurrence are more surprising. In the case of r discrete possible outcomes for an event, entropy can be defined as the following summation [160]:

$$H = E[-log(p_i)] = -\sum_{i=1}^{r} p_i log(p_i) \tag{5.19}$$

Here, the transformation temperature resulting from each individual or fused model at a specific composition/region is considered as an event. For each individual or fused model, the average of all the entropies associated with probability distributions over 95% BCIs in different specified compositions/regions are introduced as the average entropy of that model:

$$H_M = \frac{1}{m}\sum_{j=1}^{m} H_j \tag{5.20}$$

where m is the total number of probability distributions in different specified compositions-/regions, which is the same for all the models. The average entropies of the models are listed in Table 5.3. It can be observed that model 2 has lower average entropy/uncertainty than model 1, 3 or 4, as expected.

The BMA fused model shows less average entropy/uncertainty compared to these three models. The lower uncertainty of the BMA fused model can be attributed to its concentrated distribution around the results of model 2, which gives more certainty to the total distributions over 95% BCIs. However, the average model still has broader uncertainty bounds compared to each of the individual models, which can provide more confidence in robust design. In other words, broader uncertainties can give more assurance about the presence of specific microstructural phases corresponding to a phase diagram region of interest by more shrinking the safe design space of composition-temperature in that region. In addition, it seems that the BMA model can incorporate some additional information about the uncertainties in the phase diagram compared to model 2 (the best model).

The incorporation of uncertainty from the individual models through CMF can be more optimal than BMA due to the consideration of model error correlations, as the given thermodynamic models clearly are statistically correlated. According to Table 5.3, it is clear that the fused models can yield better predictions with less average entropies/uncertainties than the individual models used in each fusion case.

Table 5.3: Average entropy as a measure of uncertainty for each individual and fused model

	Model 1	Model 2	Model 3	Model 4	Bayesian Average Model	Fused Model from 3 sources	Fused Model from 4 sources
H_M	1.8450	0.6553	1.9044	1.9876	1.4916	1.7923	0.6510

5.6 Summary and Conclusion

Due to the importance of uncertainty quantification in CALPHAD, an MCMC sampling approach is utilized in this work for the probabilistic calibration of CALPHAD model parameters against the available data in the case of the Hf-Si binary system. Considering the vast high-dimensional parameter space in CALPHAD modeling, applying some prior information about the parameter values and ranges from Thermo-Calc optimization module is often required to achieve parameter convergence with a reasonable cost. However, choosing an appropriate CALPHAD model with a sufficient number of parameters is a challenging task.

Therefore, a systematic approach is required to find an optimal model rather than a trial-and-error approach. For this purpose, Bayesian hypothesis testing (or Bayesian model selection) based on Bayes' factors is proposed in this work and applied in a case study for the Hf-Si system to show how the best model can be chosen from a list of independently generated models. However, our work suggests to use information fusion approaches to smartly combine the given individual models into a fused model rather than just the application of the best model that may lose some useful information. BMA and an error correlation-based model fusion are used for our Hf-Si case study to show different beneficial purposes of these information fusion approaches.

The average model obtained from BMA shows larger 95% confidence intervals compared to any one of the individual models, which can provide more confidence for robust design but is likely too conservative. On the other hand, the error correlation-based technique can provide closer results to data with less uncertainties than the individual models used for the fusion. The uncertainty reductions through this fusion approach are also verified through the comparison of the average entropies (as a measure of uncertainty) obtained for the individual and fused models. Therefore, random CALPHAD models can smartly be fused together to find reasonable predictions for phase diagrams with no need to go through the cumbersome task of identifying the best CALPHAD models.

6. UNCERTAINTY QUANTIFICATION AND PROPAGATION IN CALPHAD MODELING*

6.1 Overview

Design is about making decisions bounded by a quantifiable degree of certainty. In the context of alloy design, Integrated Computational Materials Engineering (ICME) provides the framework whereby performance requirements are ultimately transformed into alloy/processing specifications through the combination of (complex) computational models connecting process-structure-property-performance relationships and experiments. Most ICME approaches consider the models used as deterministic and thus do not provide the means to make alloy design decisions with proper confidence measures. At the root of ICME lie CALPHAD models that describe the thermodynamics and phase stability of phases under specific thermodynamic boundary conditions. To date, the vast majority of efforts within the CALPHAD community have been deterministic in that thermodynamic models and the resulting thermodynamic properties and phase diagram features do not explicitly account for the uncertainties inherent in the model formulation or in the experimental/computational data used. In this contribution, we provide an overview of the state of the field. We review major efforts thus far and we then provide a (brief) tutorial on basic concepts of uncertainty quantification and propagation (UQ/UP) in CALPHAD. We discuss the major features of frequentist and Bayesian interpretations of uncertainty and proceed with a discussion of recent case studies in which UQ has been used to parameterize models for the thermodynamic properties of phases. We follow our discussion by presenting frameworks and demonstrating the propagation of uncertainty in thermodynamic properties and phase diagram predictions and briefly discuss how we can use Bayesian frameworks for rigorous model selection as well as for model fusion. We close our contribution by providing context for what has been done and what remains to be accomplished in order to fully embrace the management of uncertainty in CALPHAD modelling, a foundational element of ICME.

---

*Republished with permission of IOP Publishing, from "Uncertainty quantification and propagation in CALPHAD modelling", Pejman Honarmandi, Noah Paulson, Raymundo Arroyave, Marius Stan, 27, 3, 2019; permission conveyed through Copyright Clearance Center, Inc.

6.2 Introduction

Uncertainty quantification (UQ) and its propagation (UP) across multi-scale model/experiment chains are key elements of decision-based [1,161] materials design in the framework of Integrated Computational Materials Engineering (ICME) [162] where databases, multi-scale computations, and experiments are integrated with the aim of reducing the time and cost of design and manufacturing of materials and products [163]. In any materials design effort within the ICME framework, understanding and quantifying uncertainties can provide the materials/product designers with a confidence measure for the models used in the decision-making process in materials innovation.

Uncertainty quantification (UQ) is the process of identifying different sources of uncertainty and developing the corresponding mathematical representations for the error bounds of quantities of interest (QoI) [1,2,164]. Uncertainty in QoIs arise from different sources, including: the random/stochastic nature of the physical system in question; uncertainty in the values of the variables used to parameterize models; lack of sufficient or accurate data to parameterize the models; and incomplete knowledge of the model structure representative of the physical model due to excessive simplifications, wrong assumptions and/or incomplete representation of the underlying physics/chemistry in the model [3]. Different sources of uncertainty can be further classified as aleatoric or epistemic. Uncertainty Propagation (UP), on the other hand, refers to the determination of the distribution of values in model outputs–i.e. their uncertainty–arising from the uncertainty in the inputs and/or parameters of the models [1,127]. This propagated uncertainty thus represents a measure of the impact of unavoidable aleatoric/epistemic model uncertainties in the distribution of values of QoIs that ultimately should inform decisions in the materials development cycle.

For over a decade, uncertainty quantification (UQ), its propagation (UP), and the overall uncertainty management (UM) have been recognized as essential components to any effort in which simulations assist the discovery and design of materials. Relative to their importance in the overall materials innovation ecosystem [161], surprisingly few examples [2,130,131,165,166] exist in which state-of-the-art UQ/UP/UM approaches have been used in computational materials modeling and simulation.

Perhaps the general area in the field of computational materials science that has experienced a comparatively slow rate of development–in the context of UQ/UP/UM–that is at odds with its importance in the entire ICME edifice is the CALculation of PHAse Diagrams (CALPHAD) method [167]. Developed in the 1970s, the CALPHAD method [168] is based on the systematic encoding of alloy phase stability in terms of Gibbs energy functions assessed, initially, with the use of experimental evidence for phase stability as well as thermo-chemical measurements. Experimental thermochemical data has, for the past two decades, been enriched through the incorporation of quantum mechanical "first-principles" calculations of thermodynamic properties of phases [169, 170]. Gibbs free energy functions so-assessed are then encoded in self-consistent databases [171].

The quality of CALPHAD models is ultimately limited not only by the quality of the thermochemical and phase stability data available but also by the intrinsic uncertainty associated with the optimal model structure to represent the thermodynamic behavior of phases. On the experimental front, a major challenge is the fact that in many systems it is exceedingly difficult to guarantee that the phase state observed and the corresponding thermochemical measurements truly correspond to thermodynamic equilibrium states [70]. Even conceding that the states under investigation correspond (or are close) to equilibrium conditions, major sources of uncertainty arise due to limitations in the resolution of the experimental techniques. CALPHAD assessments rely more and more on first-principles calculations to estimate the thermodynamic properties of phases. However, it is clear that in some regions of the chemical space there remain considerable discrepancies between calculations and experiments [172], putting into question the reliability of first-principles calculations to supplement unavailable/inaccessible thermodynamic information. In addition, being based in quantum mechanics, first-principle calculations are considered highly accurate and precise and are rarely accompanied by uncertainty quantification.

Uncertainty with regards to the correct form of the model used to describe the thermodynamics of phases constitutes a major challenge [173] Discrepancy in the context can arise from different assumptions about the nature of the underlying structure of the phase in question and the proper mathematical representation of the Gibbs free energy of the phases. Typical examples of model form inconsistencies leading to uncertainty are: the use of competing models to describe the liquid phase in intermetallic-forming systems–regular solution vs associated liquid models [174]–; the use of different numbers of sublattices to describe the site occupancy in solid phases—the use of three or five sublattices to describe the sigma phase [175]–; or the use of different orders of expansion in Redlich-Kister polynomials that describe the excess Gibbs energy in solid solutions [176, 177].

Clearly, both aleatoric and epistemic sources contribute significantly to the uncertainty in CALPHAD-based thermodynamic descriptions of phases. Uncertainty in CALPHAD models has important consequences in the predicted state of a material system under (close to) equilibrium conditions. Uncertainty becomes increasingly important when extrapolating knowledge about the thermodynamic behavior of low-order systems to higher order systems for which little experimental information is available. Recently, the exploration of the high entropy alloy space using CALPHAD-based calculations and databases has highlighted the importance of uncertainty quantification in thermodynamic descriptions of phases [178].

Thankfully, while scarce, several studies have addressed issues related to UQ/UP in thermodynamic models of phases. Early on, Stan and Reardon [144] presented a Bayesian framework to identify the effect of uncertainty in the values of parameters in thermodynamic models on the predicted phase diagram of systems. They used a fuzzy logic-weighted algorithm as a sampling tool to propagate the uncertainty from model parameters to phase boundaries in binary systems. In later work, Otis and Liu [145] introduced an ICME-directed multi-level framework that linked first-principles calculations to CALPHAD modeling. They proposed to determine the most relevant model parameters through the use of the Akaike Information Criterion (AIC) [179] and more traditional F-tests, while Markov Chain Monte Carlo (MCMC) sampling was used to sample the (stochastic) distribution of the model parameters in order to determine the posterior probability distribution of the (predicted) thermodynamic quantities. Duong et al. [70] used MCMC sampling to carry out Bayesian uncertainty quantification of the model parameters that describe the phase stability in the U-Nb system.

Understanding how the uncertainty in thermodynamic model parameters propagates to predictions of phase stability is necessary, but not sufficient, for creating a truly stochastic/probabilistic measure of uncertainty in the phase stability of multi-component systems. Duong et al. [146] studied the effect of uncertainty in CALPHAD model parameters, first-principles calculations, experimental information and even synthesis conditions on the prediction of the equilibrium states along the Ti2AlC-Cr2AlC pseudo-binary system. As a result of a full account of all uncertainty sources, the prediction of phase stability in this system was deemed stochastic as opposed to deterministic and the presence or absence of a given phase at specific (synthesis) conditions was represented as a probability measure.

More recently, Honarmandi et al. [180] demonstrated the Bayesian probabilistic assessment of CALPHAD model parameters through the use of a MCMC approach. They proceeded to use Monte Carlo sampling to propagate (non-uniform) uncertainties in predicted phase stabilities. Moreover, they introduced Bayesian hypothesis testing–using the models' Bayes factor–in order to evaluate the quality of different models. Rather than considering only the propagated uncertainty of individual models, they used Bayesian Model Averaging (BMA) for ensemble probabilistic phase stability predictions. Finally, they carried out correlation-based model fusion to take into account all potential information any given model may provide and fused unique information from the available models. In this way, one can identify dependencies among model discrepancies via measures of statistical correlations between models, allowing a more accurate measure of propagated uncertainty in phase stability computations.

It is also critical to specifically address the details of uncertainty quantification with regards to ensuring the thermodynamic consistency between models, robustness to outliers and the weighting of data sets. Recently Paulson et al. [181] investigated these questions through a Bayesian framework for model selection, calibration and uncertainty quantification and straightforward modifications to the Bayesian Likelihood and Prior distributions. Most importantly, this work demonstrated the possibility of automatically weighting the measurements of various authors in model parameter calibration and providing uncertainty distributions for these weights.

In the present contribution, we review major concepts associated to the quantification and propagation of uncertainty in CALPHAD thermodynamic models. We discuss frequentist and Bayesian perspectives of uncertainty measures and how uncertainty is propagated across thermodynamic models. We finalize the paper by discussing the outlook and outstanding issues related to uncertainty management in CALPHAD models and its importance in ICME-based frameworks for alloy design.

6.3 Uncertainty Quantification

Uncertainty quantification (UQ) of thermodynamic models and phase diagrams requires an understanding of the connections between the uncertainty in data and in the resultant models (and model parameters). Data uncertainty comes in two forms: aleatoric and epistemic [182]. Aleatoric uncertainty describes uncontrollable and random variations in a process which cause statistical scatter in measured output metrics. In this regard, it can be stated that not any two identical measurements of a physical phenomenon result in the same response due to inherent randomness of the phenomenon; however, these measurements may have different occurrence frequencies that can be described in the form of a probability distribution. In general, aleatoric uncertainty may be measured, but not corrected given current technologies. In contrast, epistemic uncertainty refers to sources of error in measurements or modeling due to inaccurate and/or insufficient knowledge about the system that are potentially knowable and controllable and are typically expressed in the form of a systematic bias. In experimental measurements, epistemic uncertainty may come from a variety of sources, including environmental conditions, equipment calibration and human error; while, in modelling, it may result from the lack of information or precession about the parameters known as model parameter uncertainty (MPU) and missing physics, assumptions and/or simplifications in the model structures known as model structure uncertainty (MSU). In some cases, it may be possible to correct for these factors and reduce the epistemic uncertainty [180, 181]. This is desirable whenever possible as it may reduce uncertainty in the final model. In the remainder of this section we will discuss forward and inverse propagation of uncertainty. Furthermore, we will introduce Frequentist and Bayesian methods for uncertainty quantification.

6.3.1 Forward vs. inverse uncertainty propagation

In most UQ tasks, we desire to understand the uncertainty in a model and its predictions. Model prediction uncertainty is obtained in a process called forward uncertainty propagation (UP). In contrast, uncertainty in the model itself may be described by the uncertainty of the model's parameters. Obtaining parameter uncertainties given the data is called inverse uncertainty propagation. For both tasks the first step is to select a physical or machine-learning model which appropriately represents the phenomena of interest.

In forward UP, uncertainty in the model parameters is passed through to the model predictions. Model parameter uncertainties may be estimated from the literature, previous experience or the best intuition of the researcher. In the simplest case, parameter uncertainties are expressed as hard boundaries. For example, the uncertainty in the Debye temperature for Hafnium metal may be represented as a uniform distribution between 145 and 165K [183]. If more information is available, sophisticated parameter distributions may be assumed, such as Normal or Exponential distributions. Furthermore, multivariate distributions may be employed to express known covariances between parameters.

Traditionally, parameter uncertainties yield model prediction uncertainties through analytical approaches. These approaches usually take advantage of surrogate modelling to propagate the uncertainties. In this regard, first order second moment (FOSM) and second order second moment (SOSM) approaches have widely been used to perform nonlinear uncertainty propagations. In these UP approaches, the uncertainties of the model responses resulting from the parameter uncertainties are estimated through the first and second order Taylor expansion of the model at the mean values of the parameters [124, 184].

Polynomial Chaos Expansion (PCE) and Kriging (Gaussian process (GP) regression) are two other common used methods which have recently received plentiful attention in engineering applications. Building the surrogate model in PCE includes the selection of a set of orthonormal polynomials using, for example, least-angle regression, which is followed by the determination of the coefficients through an optimization approach, such as least squares [185]. Kriging is a stochastic fitting approach that builds the surrogate model based on a weighted distance-based correlation between the errors of the given data points obtained from the primary model function at random parameter values [81, 185].

Sometimes parameter uncertainties are only available as discrete samples from the true parameter uncertainties and must be propagated through numerical methods. This is common when parameter uncertainties are obtained from Monte Carlo sampling methods [52]. In this case, the model is sampled by evaluating the model at the available discrete parameter values. Confidence intervals may then be constructed using the model prediction samples to quantify the uncertainty. Analytical UP techniques are usually so cheap that is very useful for the propagation of uncertainties within the expensive models, but they are not as precise as the numerical methods.

Inverse UP, also traditionally known as (inverse) UQ, describes the process of computing model parameter uncertainties given the observation of data [186]. Typically, inverse uncertainty propagation methods result in fully-dimensional parameter uncertainties including the covariances between parameters. The uncertainties may be either analytical distributions or merely samples depending on the method employed. Furthermore, inverse uncertainty propagation methods may be used to estimate the uncertainties in the data itself. Inverse uncertainty propagation is performed via either Frequentist or Bayesian methods as discussed in detail in the next section.

6.3.2 Frequentist vs. Bayesian methods for uncertainty quantification

Uncertainty quantification in thermodynamic modelling can be performed in the context of either frequentist or Bayesian statistical paradigms. Although both views have their own strengths and weaknesses, their inferences should be consistent and correlated. The way that the probability is interpreted in these frameworks is the basis of their differences, which is discussed in more detail in the remainder of this section.

6.3.2.1 Frequentist inference

In frequentist approaches, the probability is defined as the frequency at which a particular outcome occurs during a long-term repetition of a specific measurement or observation. Within the context of this paradigm, it is generally assumed that existing data are a random sample from a larger population (a random variable), while parameters are fixed and usually unknown in the larger population [187]. Therefore, frequentists believe that there is only one true set of the parameter values whose noises result in the available data or observation samples. It is clear that more available data can provide better estimation of these true values. In this paradigm, the average ($\bar{\theta}$) and ensemble covariance matrix ($\hat{E}$) of the "best parameter values" or "parameter estimates" obtained from the available sample of data/observations are introduced as an estimate of the true values of the parameters and their accuracy, i.e.

$$\bar{\theta} = \frac{1}{N_D} \sum_{i=1}^{N_D} \hat{\theta}_i \tag{6.1}$$

$$\hat{E} = \frac{1}{N_D} \sum_{i=1}^{N_D} (\hat{\theta}_i - \bar{\theta})(\hat{\theta}_i - \bar{\theta})^T \tag{6.2}$$

where N_D is the number of the available data and $\hat{\theta}_i$ is the parameter estimate obtained from the $i^t h$ data in the sample [188]. Maximum likelihood approaches (MLE) are the most commonly used estimators for frequentist inference, where the parameter values that maximize the likelihood function, i.e., are most likely to yield the data, are considered as the parameter estimates [187]. Let $L(\theta|D) \equiv P(D|\theta)$ denote the likelihood function where θ and D are the parameters and data, respectively. Then, the parameter estimates are mathematically obtained as follows,

$$\hat{\theta}_{MLE} = arg_{\theta} max \big[L(\theta|D) \big] \tag{6.3}$$

In the frequentist framework, hypothesis testing is usually performed based on the calculation of the p-value for two hypotheses known as the null and alternative hypotheses. P-values indicate how likely the sampled data would be observed on condition that the null hypothesis is true. In frequentist hypothesis testing, the null hypothesis is rejected in the case that the calculated p-value is less than a significant level (α) that is usually selected as 0.01, 0.05 or 0.10. One of the main advantages of this hypothesis testing approach is the objectivity conferred through global agreement on the statistical inference obtained from the p-value.

6.3.2.2 Bayesian inference

In the Bayesian statistical view, the probability is defined as a degree of belief for the occurrence of a particular outcome determined based on the state of knowledge (prior beliefs) in combination with newly acquired data (evidence) [187]. Since beliefs are subjective, Bayesian probabilities are always introduced in conditional forms. In this paradigm, observed data are considered fixed, while parameters are treated as random variables with assumed known prior distributions. The prior (current) knowledge about the parameters should be updated upon acquiring new data under the Bayesian inference framework. In this regard, the information update is performed using Bayes' rule from a prior to a posterior distribution given the observed data. This theorem is expressed as,

$$P(\theta|M, D) = \frac{P(D|\theta, M)P(\theta|M)}{P(D|M)} \tag{6.4}$$

where $P(\theta|M)$ is the probability having the parameter θ given the model M, which is known as the prior probability that represents the prior knowledge about the parameter θ; $P(D|M)$ is the probability having the observed data D given the model M, which is a normalizing constant called the evidence; $P(D|\theta, M)$ is the likelihood function that is the conditional probability of D given the model M with the parameter θ. In other words, the likelihood shows how likely it is to obtain the data D in the case that the model M with the parameter values θ is applied for the system. $P(\theta|M, D)$ is the posterior probability in the form of a conditional probability of θ given the model M and the data D.

Figure 6.1 shows a clearer illustration of the Bayesian update of the model parameters with the given data. Bayesian inferences are generally based on the posterior probability rather than just the likelihood which is the primary inference element in the frequentist paradigm. Technically, the posterior is the solution to the inverse uncertainty propagation problem. As observed in Equation 6.4, the posterior probability is proportional to the prior probability times the likelihood–in other words, the likelihood is modified by the prior beliefs.

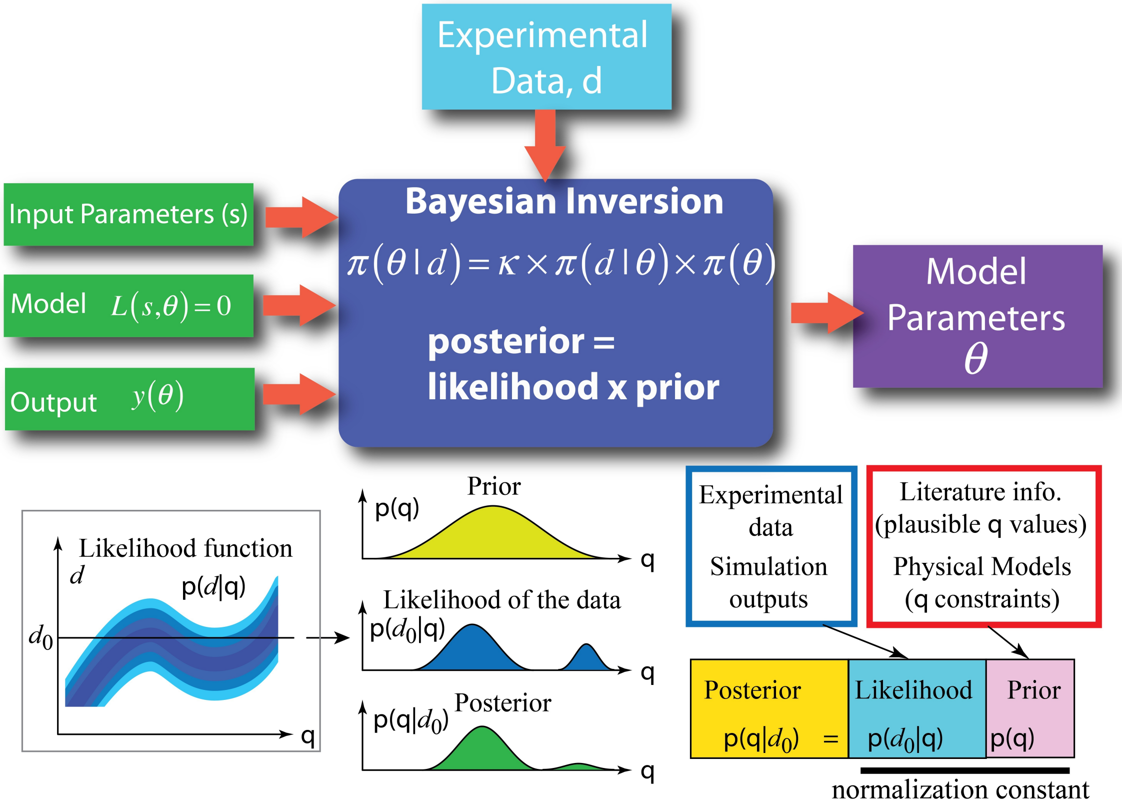


Figure 6.1: Bayesian framework to update the prior information of the model parameters given the data through the likelihood function

The prior probability distribution can be either informative with a specific defined distribution (e.g. normal distribution) or non-informative (e.g. uniform distribution) in the case that no or very little knowledge is available about the parameters. Generally, Bayesian inference is criticized for being subjective due to expert opinions in most cases for the selection of the prior distributions; even in the presence of data for the parameters, a decision should be made about the form of the prior distribution. However, the ability to incorporate the prior knowledge for better statistical inference has its worth when a reasonable prior distribution is selected. Therefore, the prior determination is very important in Bayesian inference since incorrect prior information sometimes leads to a wrong inference or a longer time to infer correctly.

The likelihood form is also important in the Bayesian framework. It is usually defined as a function of the difference between the data and the corresponding model results that are known as the residuals. The likelihood can be determined through either formal or informal approaches which have their own advantages and disadvantages. In formal methods, a statistical model is assumed for the residuals, which can be used to statistically drive the corresponding likelihood function [189]. The form of this function is usually determined using some assumptions about the residual distribution, correlation, and/or heteroscedasticity. These explicit assumptions in the formal approaches are advantageous since their validity can be confirmed by the posterior probability of the parameters given the data. However, it should be noted that the mentioned assumptions for the residuals are not correct in many cases. For this reason, informal approaches have been proposed to define the likelihood functions. Generalized likelihood uncertainty estimation (GLUE) [190] is one of the most common informal approaches, where the likelihood has a very general functional form with high flexibility and shows a fuzzy measure of belief about the model ability to capture the data. Unlike the formal approaches, there are no underlying models for residuals distribution in the informal determination of the likelihood function. Therefore, the assumptions in these approaches are mainly implicit and cannot be verified through the posterior probability distributions. However, it is worth noting that a generalized formal likelihood function has been introduced by Schoups et al. [191], which includes no assumptions about the residuals. The hyper-parameters of this generalized function can be calibrated against the data obtained for the residual distribution through numerous runs of the model with the parameter values that are sampled from the prior probability distribution.

As mentioned earlier, the state of knowledge about the parameters can be updated as soon as the new data is observed. In this sequential process, the posterior distribution obtained from the previous data training is considered as the prior probability distribution for the new training. In Bayesian approaches, more data can result in more precise inference with reduced uncertainties. The probability of the data given the model ($P(D|M)$) is a normalization constant in the denominator of the Bayes' rule, which is expressed as

$$P(D|M) = \int P(D|\theta, M)P(\theta|M)d\theta \tag{6.5}$$

$P(D|M)$ can be considered as a marginal likelihood that indicates how likely the data D is given the model M. Solving the above integral is important for the calculation of Bayes' factor that is used in Bayesian model selection (BMS) and Bayesian model averaging (BMA) (further discussed in section 4). However, this integration suffers from the curse of dimensionality in the case of high-dimensional parameter spaces. However, different numerical approaches are applied to tackle the multi-dimensionality problems, e.g. Laplace's methods, the Schwarz criterion, importance sampling methods, Quadrature Methods, and approaches that samples from the posterior probability distribution.

Bayesian hypothesis testing, proposed by Jeffreys [14] in 1935, help us determine to what extent the evidence is in favor of a hypothesis. Assigning probabilities to the hypotheses in this testing is criticized by frequentists since the hypotheses are philosophically either true or wrong. On the other hand, knowing the hypothesis probabilities is required for decision making. From the Bayesian point of view, the null hypotheses are not considered true and tested for the evidence in their favor in contrast to the frequentist p-value based hypothesis testing where the null hypothesis is assumed to be true and tested for possible rejection using a significance level. The main problem in the frequentist point of view is that the strength of the evidence in favor of the null hypothesis cannot be obtained when no rejection takes place, whereas the Bayesian counterpart can provide this strength through the calculation of Bayes' factor with no need to reject the null hypothesis. Moreover, Bayesian hypothesis testing provides the ability to incorporate the prior information and uncertainty which are missing in the frequentist approach. Another advantage of Bayesian hypothesis testing is the capability of having multiple alternative hypotheses instead of just one. In model selection where competing models act as hypotheses, models are not required to be nested, which means they can include different sets of parameters [152].

Generally, Bayesian inference for the model parameters given the data is characterized by most plausible values ($\langle\theta\rangle$) and the covariance matrix (C^{θ}) of the parameters using the posterior probability distribution, as follows,

$$\langle\theta\rangle = \int \theta P(\theta|D) d\theta \tag{6.6}$$

$$C^{\theta} = \int (\theta - \langle\theta\rangle)(\theta - \langle\theta\rangle)^T P(\theta|D) d\theta \tag{6.7}$$

The above-mentioned integrations can be performed analytically, when there are closed-form solutions. For example, using conjugate priors can result in closed form expressions for the posterior probability distributions. However, the absence of the closed form solutions in most real cases is the reason to use the numerical approaches to solve the integrals. Monte Carlo integration is one of the strong numerical approaches to estimate these integrals [52, 186, 188]. In this method, the estimations are expressed in the form of the following summations,

$$\langle\theta\rangle \approx \frac{1}{N}\sum_{i=1}^{N} \theta_i \tag{6.8}$$

$$C^{\theta} \approx \frac{1}{N}\sum_{i=1}^{N} (\theta_i - \langle\theta\rangle)(\theta_i - \langle\theta\rangle)^T \tag{6.9}$$

where θ_i are the parameter samples from the posterior probability distribution, $P(\theta|D)$. In this regard, direct or rejection sampling methods can be applied for simple cases. However, this is not usually the case due to the complexity resulting from the multi-dimensionality and unknown form of the posterior distributions in the practical situations. In these cases, Markov Chain Monte Carlo (MCMC) techniques are commonly used to sample from the parameter posterior distributions. Among MCMC techniques, Gibbs' sampler and Metropolis-Hastings algorithm are the most common used sampling methods.

Gibbs sampling begins with an initial guess about the parameter values, $\theta^0 = \{\theta_1^0, ..., \theta_n^0\}$. Then, the parameter values are sequentially sampled one by one from their corresponding conditional probability distributions, $P(\theta_i|\theta_1^z, ..., \theta_{(i-1)}^z, \theta_{(i+1)}^{(z-1)}, ..., \theta_n^{(z-1)}, D)$, to generate the parameter vector θ^z after n samplings. This is a way to go around the problem of sampling directly from the posterior probability distribution, since these conditional distributions can analytically be determined in most cases. The Gibbs sampling stops when the multivariate distribution for the sampled parameter vectors becomes stationary.

In the Metropolis-Hastings algorithm, a new candidate for the parameter vector is randomly sampled from a proposal distribution (q) in each iteration. Unlike Gibbs sampling where all the samples are accepted, the candidates can be accepted or rejected through a probabilistic criterion based on the Metropolis-Hastings ratio that is expressed as

$$MH = \frac{P(\theta^{cand})P(D|\theta^{cand})}{P(\theta^{z-1})P(D|\theta^{z-1})} \frac{q(\theta^{z-1}|\theta^{cand})}{q(\theta^{cand}|\theta^{z-1})} \tag{6.10}$$

where the first ratio compares the posterior probability of a sampled candidate (θ^{Cand}) for the parameter vector with its counterpart for the initial (θ^0) or the last accepted ($\theta^{(z-1)}$) parameter vector, which is known as Metropolis ratio. It should be noted that $P(D)$ is a constant that can be canceled out in this ratio. The second ratio is known as Hastings ratio that shows the effect of asymmetric proposal distribution by comparing the probability of a forward move from $\theta^{(z-1)}$ to θ^{Cand} with its reverse move in the MCMC chain. It is worth noting that the hastings ratio equals 1 in the case of symmetric proposal distributions. In each iteration, min(MH,1) is considered as the acceptance probability of the new candidate. When a candidate is accepted, $\theta^z = \theta^{Cand}$; otherwise, $\theta^z = \theta^{(z-1)}$. Again, the above sampling process continues until the MCMC chain converges to a stationary distribution.

Generally, Gibbs sampling method is a specific case of Metropolis-Hastings algorithm on the condition that the proposal distribution is considered as the conditional posterior probability distribution. In this case, the probability of acceptance is always 1.

Other advanced MC and MCMC based sampling methods have been developed to address the weaknesses of the previously discussed approaches in sampling multi-modal distributions or distributions with pronounced degeneracy between parameters. Two of the most notable such approaches are the MultiNest [192] and Affine Invariant Ensemble Samplers [193].

6.4 Uncertainty of Thermodynamic Data

Fundamental thermodynamic properties such as the specific heat, enthalpy, entropy and Gibbs free energy of materials form the basis of numerous important technological applications, including the prediction of phase stability and microstructure evolution during processing. To therefore obtain uncertainty in the phase diagram using the calculation of phase diagrams (CALPHAD) approach, uncertainty in the parameters of the models of the thermodynamic properties of the possible phases must be obtained. Parameter uncertainties in the form of probability distribution are obtained via inverse uncertainty propagation protocols through frequentist or Bayesian approaches as discussed in Section 6.3. The first consideration for the development of thermodynamic property models is the reliability and uncertainties of the experimental or simulated thermodynamic data itself, as this may have an outsized effect on the final model. Once data is obtained and processed, a model is selected, fit to the data, and probability distributions for the model parameters are computed. In the remainder of this section, we will address questions regarding various data sets and the models and how they relate to the quantification of uncertainty.

It is crucial to understand, and where possible, reduce, the uncertainty in thermodynamic data as it may lead to thermodynamic property models that drastically depart from the experimentally observed material's behaviour [194]. While experimental data exhibits aleatoric uncertainty, this is rarely the cause of poor thermodynamic property models. Instead, epistemic uncertainty plays a critical role, biasing datasets in ways that may be misleading or unpredictable. Two common forms of epistemic uncertainty in thermodynamic data are the compositional impurity and the use of historical temperature scales (e.g. if a temperature was performed prior to the introduction of the ITS-90 standard) [195]. Fortunately, with care and attention these errors can be corrected and the epistemic uncertainty reduced. It is also possible to correct for known biases of certain measurement techniques (e.g. a DFT functional may have a known energy prediction bias for a class of materials [196]). In most cases, however, the source of epistemic uncertainty is unknown and therefore impossible to reduce. Examples of such uncertainties include unknown biases inherent to a measurement technique, poor instrument calibration or human error.

It is important to obtain a reasonable estimate of the combined aleatoric and epistemic uncertainty for data points or data sets. These estimated uncertainties, typically given by a standard deviation from the mean, may serve as effective "weighting factors" in the determination of model parameter distributions in both Bayesian and Frequentist approaches. It is possible to estimate these uncertainties through a variety of means. Typically, the CALPHAD expert assigns weights based on knowledge of measurement approaches, research group reputation, or trial and error [197]. As this approach is subjective and not necessarily reliable, the authors have recently proposed automated approaches to weight thermodynamic datasets [181]. Beyond the weighting of datasets, ensuring robustness to outliers and thermodynamic consistency between models are important in the CALPHAD approach and have significant impacts on the uncertainty of the resultant models. The authors have recently developed a Bayesian methodology to address each of these problems [181]. This framework is briefly introduced in the next section.

6.4.1 Bayesian framework

As discussed in Section 6.3, Bayes' Theorem and modern sampling algorithms enable robust approaches to model selection, calibration and uncertainty quantification. We perform model selection by comparing competing models via their marginal Likelihoods. Monte Carlo algorithms perform inverse uncertainty propagation to obtain the Posterior distribution of the model parameters given the observation of data. We then propagate sampled parameters from the Posterior forward to the model prediction to obtain uncertainty intervals.

In CALPHAD, it is necessary to analyse thermodynamic data from different sources and to ensure consistency between property models. For example, the specific heat is simply the derivative of the enthalpy. As such, the independent calibration of specific heat and enthalpy models using their respective measurements may not incorporate all relevant data and would not necessarily result in thermodynamically consistent models. Furthermore, the resultant models may have unnecessarily large uncertainty intervals due to the inefficient use of the data. Therefore, in the CALPHAD approach we employ common parameters (e.g. mathematical function coefficients) for related models whenever possible. In a Bayesian framework we achieve this by defining a common set of Prior distributions for parameters shared between multiple models. Furthermore, we simultaneously calibrate the Posterior distribution for parameters across all models via analytical or numerical Bayesian methods [181].

Beyond ensuring thermodynamic consistency between models, it is critical to evaluate the quality of the data itself. This step is necessary as data sets sometimes conflict with each other to a degree that makes it impossible to ensure thermodynamic self-consistency. Currently, it is the responsibility of the researcher to identify outliers and to assign weights to datasets. From an uncertainty quantification perspective, it is desirable to handle these tasks in a more automated and transparent way.

Instead of manually removing outliers from consideration, the issue can be addressed in the Bayesian framework by modifying the likelihood function to better accommodate scatter in the data [52]. The use of the Normal distribution for the likelihood assumes that the data is also normally distributed around the model prediction. This assumption is reasonable if each datum represents the mean of multiple measurements and if all data comes from the same source. In some cases, however, individual data points may significantly deviate from the mean behaviour. One reasonable way to account for non-Normally distributed data is to use a distribution for the Likelihood with more weight in the tails, such as Student's t-distribution with few degrees of freedom. Alternately, the Likelihood may be expressed as the weighted sum of two independent Normal distributions, one to represent the bulk of the data and the other to represent the outliers.

We may also weight datasets in an automated manner by making modifications to the Likelihood function [198]. We start by assuming that the distribution of each data point around the mean in the Likelihood is represented by the error reported in the original publication. As such, the most natural way to correct these errors is by applying a multiplicative weighting factor to each data point. In the Bayesian framework, we treat these weighting factors, herein referred to as hyperparameters, the same as all other model parameters and allow the sampler to explore their Posterior distribution. While it is possible to assign each data point an independent hyperparameter, this would dramatically increase the number of parameters in the Bayesian analysis at great computational expense. Instead, we assume that all points in a dataset should be rescaled by the same factor and assign a unique hyperparameter to each dataset. This approach allows the framework to automatically find consistency between varied datasets (potentially measuring different quantities) and to give reasonable and coherent estimates of each data point's combined aleatoric and epistemic uncertainty.

6.4.2 Case study: Hafnium

In this section, we present the application of the proposed Bayesian framework to the development of thermodynamic property models for elemental Hafnium (Hf) with quantified uncertainty. Hf is a tetravalent transition metal with applications in nuclear energy, plasma welding, super-alloys and electronics [199–202]. It is suitable as a case study because its high melting and vaporization temperatures ($\sim$2500K and $\sim$4900K respectively), difficulty of separation from the chemically similar Zirconium (Zr) and susceptibility to oxidation at high temperatures lead to the data problems previously mentioned such as the presence of outliers and datasets with systematic errors. Further details of this case study are available in our prior work [181].

We develop thermodynamic property models for the specific heat, enthalpy, entropy and Gibbs energy in the alpha, beta and liquid phases of Hf. Twenty sets of specific heat and enthalpy measurements were extracted from the literature and corrected for temperature scale [195] and Zr content [183, 203]. Furthermore, errors were transformed to standard errors (1 standard deviation assuming the data is normally distributed) according to the GUM standard [204] when reported and estimated to be +/- %5 when not available.

We identify a diverse pool of models for each temperature range (associated with the alpha, beta and liquid phases). For the alpha phase, the fundamental component of the specific heat and enthalpy models is either an Einstein or a Debye [205, 206] term, depending on the physics of the solid phase, to account for quantum effects. Polynomial terms or a bent cable model [207] are then added to account for higher temperature behaviour. For the beta and liquid phases, polynomial model forms are employed as we choose not to extrapolate to low temperatures. The distribution of the data is modelled using a Student's-t Likelihood with $2+10^{-6}$ degrees of freedom to impart robustness to outliers. The standard deviation parameter in the Likelihood for each data point is given by the reported (or estimated) standard error rescaled by the Bayesian hyperparameter (as described in Section 6.4.1). We employ Uniform and Exponential Priors for model parameters and hyperparameters, respectively. For each phase and candidate model family, the MultiNest [192] algorithm is used to sample the Posterior and compute the marginal Likelihood for all relevant specific heat and enthalpy measurements. For each phase, we select the optimal model according to the Bayes' Factors.

Through the application of this Bayesian framework, we identified that, for the case of the specific heat formulation, the Debye plus quartic polynomial model is optimal for the alpha phase, the quadratic polynomial model is optimal for the beta phase and the linear model is optimal for the liquid phase. As an example, Figure 6.2 displays Posterior samples for pairs of liquid phase model parameters and hyperparameters in addition to univariate distributions for all parameters.

Figure 6.2 reveals that model parameters C0, C1 and C2 are all correlated with each other while the hyperparameters are relatively uncorrelated with each other and with the model parameters. This indicates that weights determined by the automated weighting procedure described in Section 6.4.1 are relatively insensitive to each other and, perhaps more surprisingly, to fluctuations in the model's parameterization. Looking at the values of the hyperparameters themselves, we can see that the Bayesian framework highly weights the measurements of Korobenko [208], Cagran [209] and Rosner-Kuhn [210] while assigning the measurements of Paradis [211] a low weight. Figures 6.3 and 6.4, which display the model predictions and experimental datasets confirm that the model prediction aligns closely with the prioritized datasets while effectively ignoring the Paradis measurements. Furthermore, this confirms the observation that the inclusion of hyperparameters for dataset weighting makes the Bayesian framework find consistency between datasets, even between

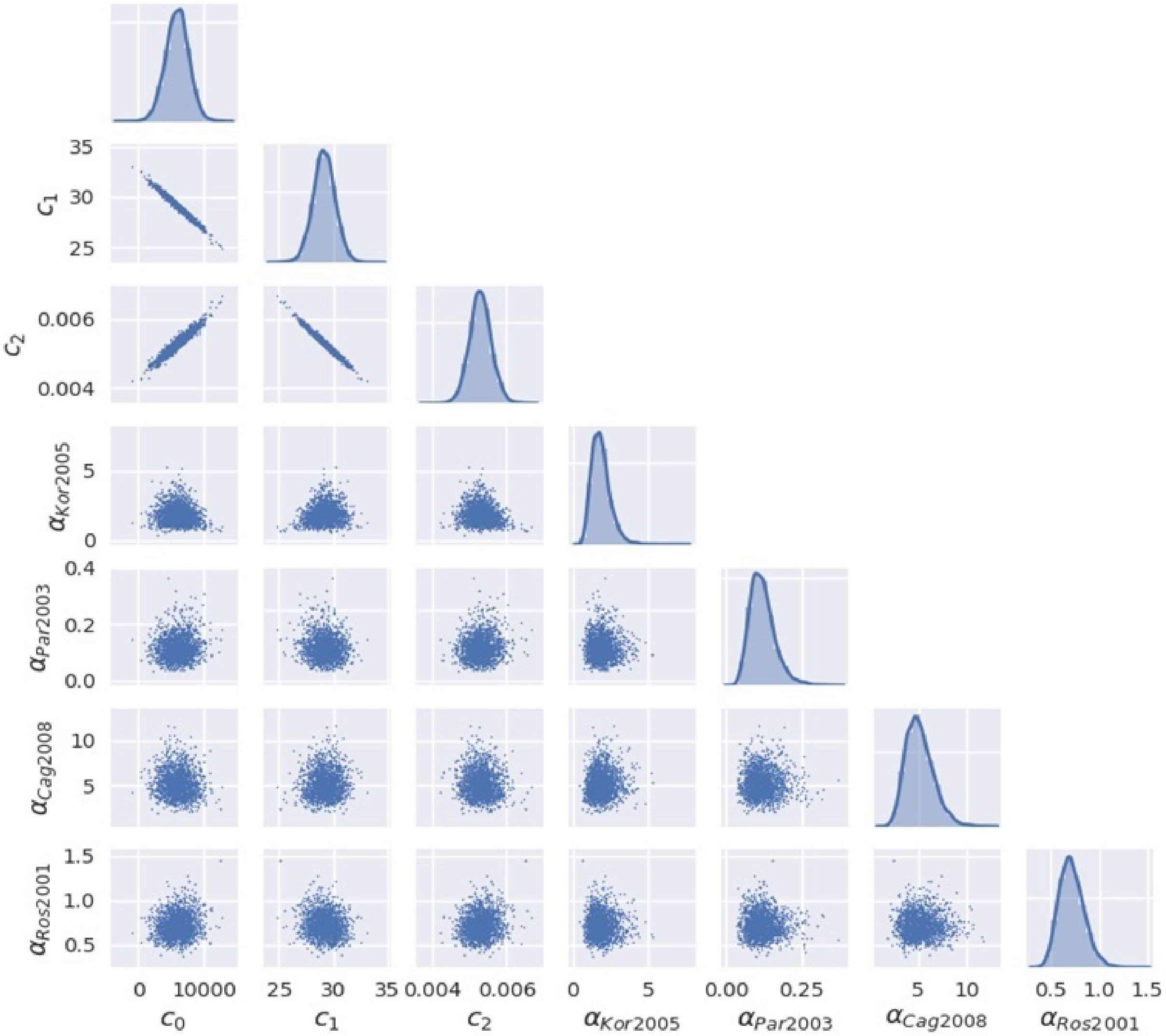


Figure 6.2: A corner plot is presented for the Hafnium liquid phase model parameters and hyperparameters. Posterior samples are plotted for pairs of parameters and the univariate Posterior distribution is plotted for each parameter

disparate thermodynamic quantities (in this case specific heat and enthalpy).

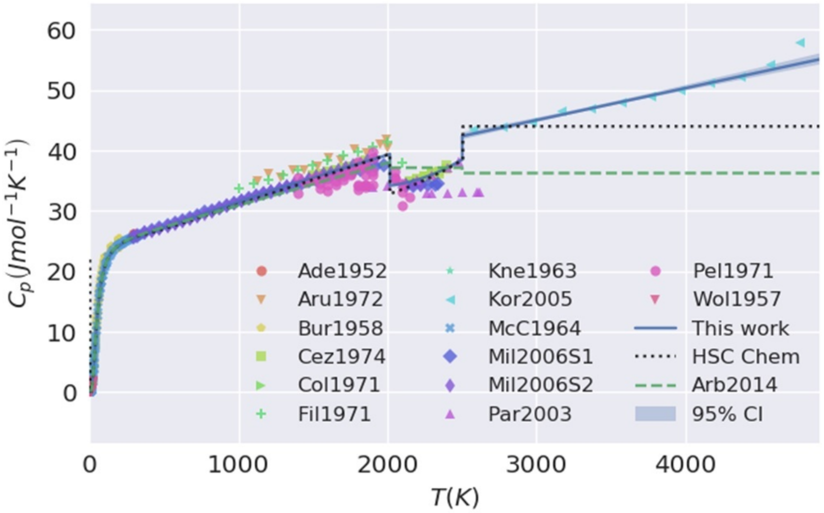


Figure 6.3: The alpha, beta and liquid phase specific heat models for Hafnium are plotted versus temperature alongside the experimental data and the results of two previous assessments

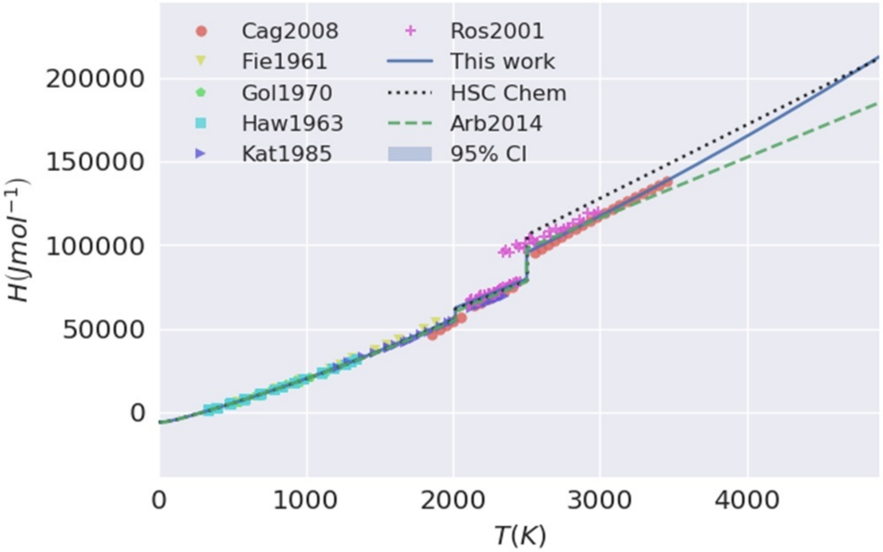


Figure 6.4: The alpha, beta and liquid phase enthalpy models for Hafnium are plotted versus temperature alongside the experimental data and the results of two previous assessments

Figures 6.3 and 6.4 show the experimental measurements and model predictions for the alpha, beta and liquid phases for specific heat and enthalpy, respectively. Ninety-five percentile uncertainty intervals are plotted for all model predictions. In general, the availability of large quantities of data constrain the model predictions through tight Posterior parameter distributions. This effect is evident in both Figures 6.3 and 6.4. The uncertainty intervals are visible in the liquid-phase specific heat model prediction as only four datasets were available to the Bayesian analysis–resulting in a less constrained Posterior distribution. Figure 6.2 corroborates this observation; for example, the C1 parameter (the intercept of the linear model for specific heat model) has a standard deviation of $\sim$1 J mol^{-1} K^{-1} which is easily propagated to the model prediction uncertainty interval.

To compare the results of our Bayesian framework with other model building strategies, we add to the figures two previous sets of models; the first from an expert (Arblaster [183]), and the second from commercial software (HSC-Chemistry [212]). We find it informative that both HSC-Chemistry and the work of Arblaster employ linear models for the enthalpy of the liquid phase. This requires the associated specific heat models to be constant as the specific heat is the derivative of enthalpy with respect to temperature. Unfortunately, this is not consistent with the specific heat measurements. As the Bayesian framework presented in this work simultaneously incorporates the specific heat and enthalpy measurements in the model selection and calibration of both models, our model is consistent with both sets of measurements.

6.5 Challenges in CALPHAD Modeling

Generally, large sparsity of the experimental data due to high measurement uncertainties, uncertainty quantification of the high-dimensional model parameters and the resulting phase diagram, and the choice of thermodynamic model with appropriate number of parameters are some of the major challenges in CALPHAD, which will be addressed in this section.

The presence of uncertain sparse experimental data for the phase diagram in most studied multicomponent systems leads to large uncertainties in the CALPHAD model parameters and the resulting phase diagrams. The high uncertainties are not usually desirable in materials design due to lack of confidence in equilibrium compositions and fractions of thermodynamic phases in the system that may considerably influence the materials properties. For this reason, thermochemical data obtained from first-principles density functional theory (DFT) calculations are used to provide better alignments in the phase diagram with less uncertainties. Duong et al. [70] work is one of the examples that has considered the thermo-chemical and experimental data together to probabilistically estimate the phase diagram of U-Nb binary system. However, any availability of new information about low component systems, specially DFT predictions, can cause difficulties in modification of the modeling of their higher component systems; since the CALPHAD model of the high component systems are strongly connected to the lower component subsystem counterparts. To solve this problem, Shang et al. [213] proposed an automated extensible self-optimization approach (ESPEI) for the phase equilibrium in CALPHAD modeling. In this technique, the model parameters are determined for each phase in the system through a two-step optimization–first with the DFT thermo-chemical data and then the experimental data. Otis et al. [145] has incorporated the uncertainty quantification of the appropriate CALPHAD model parameters into ESPEI framework in the direction of the ICME design. This work attempts to address the challenges associated with the model parameter selection and uncertainty quantification in CALPAD. Stan and Reardon [144], Duong et al. [70] and Honarmandi et al. [180] works are also recognized as other major studies in the literature with regard to UQ and UP in CALPHAD, with the later work also incorporating model selection and fusion to target the challenge related to the identification of the appropriate CALPHAD model. All the above-mentioned works are discussed further in Section 6.5.2.

6.5.1 Uncertainty quantification in CALPHAD: Fundamentals

Bayesian UQ in CALPHAD has schematically been demonstrated in Figure 6.5. As observed in this figure, the CALPHAD model parameters (θ) describing the molar Gibbs free energy of each phase (ϕ) in the system at any specific temperature (T), pressure (P) and composition (x_i^ϕ) are probabilistically calibrated against the available calculated and/or experimental data by solving the inverse problem introduced through the Bayes' rule to find the parameter posterior probability distribution ($\pi(\theta|d, M)$ which is the same as $P(\theta|D, M)$ in Equation 6.4). As mentioned in section

6.2, this inverse problem is usually solved using sampling techniques, such as MCMC approaches. The identification of the parameter posterior distribution results in the quantification of the parameter uncertainties and the Molar Gibbs energy of the phases. In the case of model selection, the evidence ($\pi(d|M)$ which is the same as $P(D|M)$ in Equation 6.4 and 6.5) obtained for different competing models can be compared through Bayes' factor.

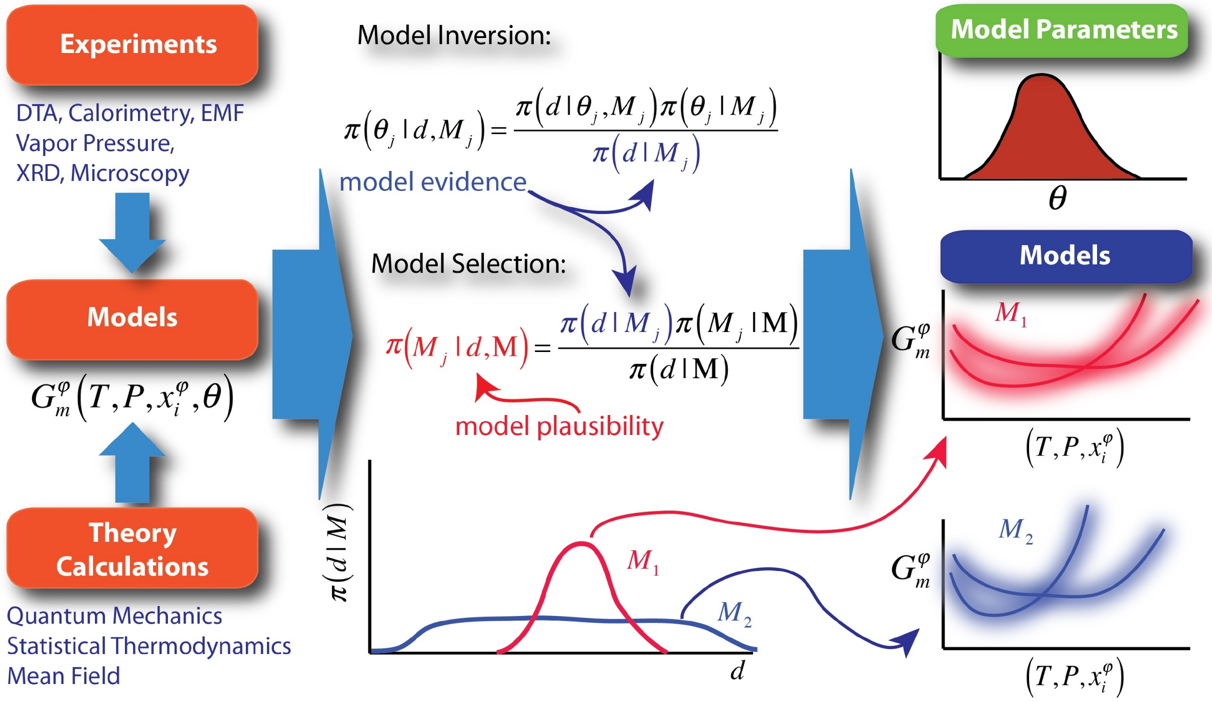


Figure 6.5: Schematic illustration of Bayesian UQ in CALPHAD

As shown in Figure 6.6, the uncertainties of the parameters and the molar Gibbs free energy of the phases can be propagated across the CALPHAD model which minimizes the total Gibbs free energy of the system subject to some compositional constraints in order to determine the uncertainty of the equilibrium phase diagram. UP can also be performed to quantify the uncertainty of any other quantity of interest, such as the lattice parameters.

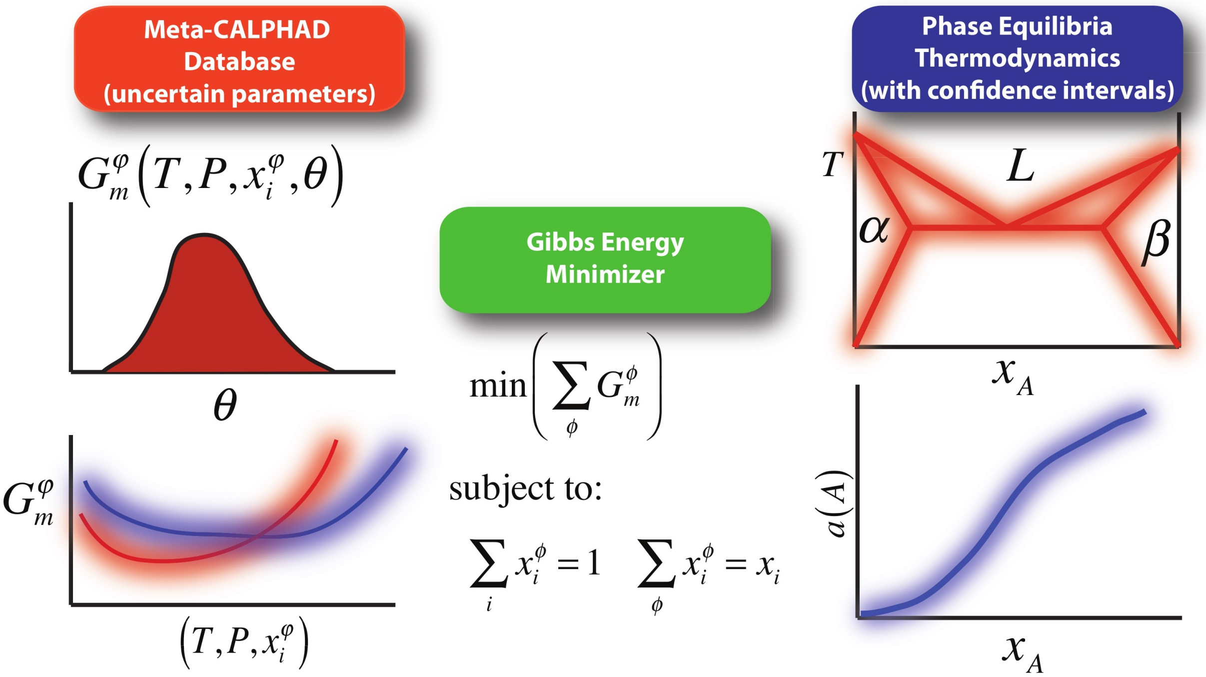


Figure 6.6: Schematic illustration of UP in CALPHAD

6.5.2 Uncertainty quantification in CALPHAD: An overview of the recent literature

As pioneers in CALPHAD Bayesian uncertainty quantification, Stan and Reardon [144] have evaluated the uncertainties of UO_2-PuO_2 and UO_2-BeO binary phase diagrams based on the available data for these systems. In this work, enthalpy and melting temperature of the individual phases are the model parameters that are probabilistically calibrated against the data for each system. In this regard, a multi-objective genetic algorithm (GA) is applied to sample from the multivariate posterior probability distribution of the model parameters in the context of the Bayesian inference, as discussed in Section 6.3. In the GA context, the fitness of each particular member (parameter-vector sample) of the population is determined through taking an average over all the corresponding objective values rescaled by a fuzzy logic weighting from 0 to 1. In this scheme, 0 belongs to the worst objective outcome of the population and 1 corresponds to the objectives whose values fall into the uncertainty bands of the experimental data. The GA continues until the fitness of all the members of the population become 1 or a constant value with no improvement. The last population at this state is considered to find the uncertainty bands of the model parameters. The parameter uncertainties have also been propagated to the phase diagrams (as shown in Figure 6.7), using the forward calculation of the phase diagrams with a large number of parameter vectors sampled from the estimated parameter posterior distributions [144].

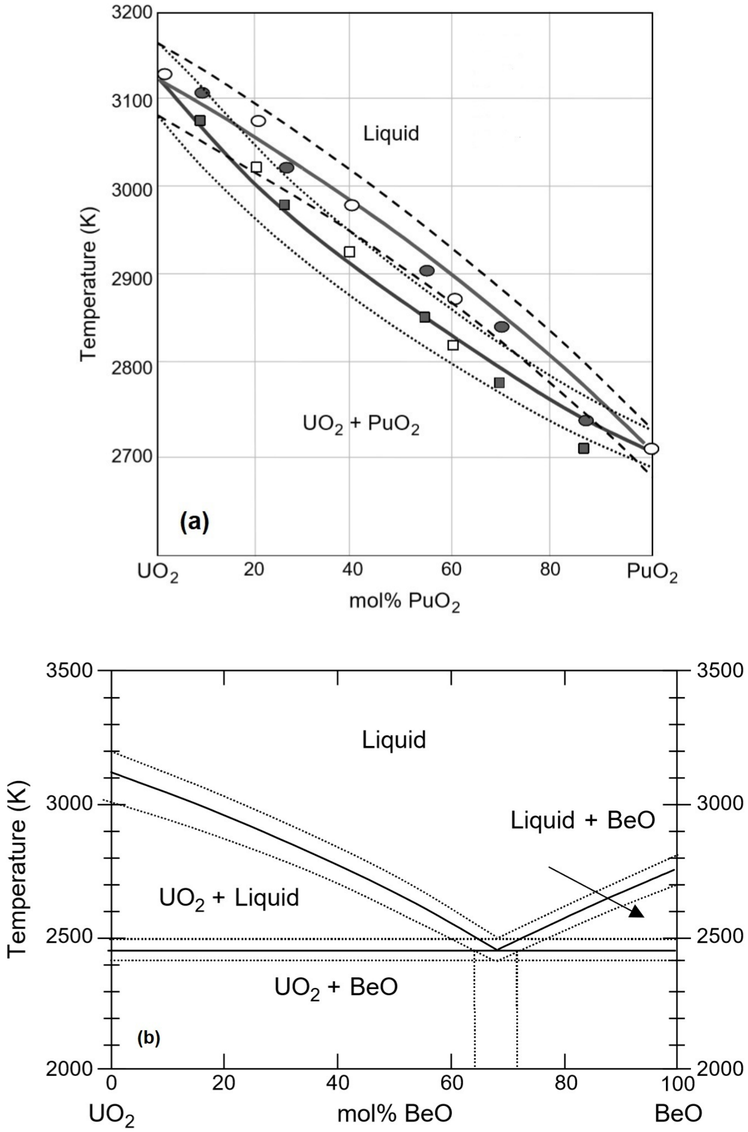


Figure 6.7: a) UO_2-PuO_2 and b) UO_2-BeO binary phase diagrams with their uncertainty bands. Reprinted with permission from [144].

In another important work, Otis et al. [145] have developed an automated high-throughput modeling in CALPHAD with the capability of quantifying the parameter uncertainties at the same time. However, the selection of the model parameters is not absolutely objective and expert opinion is still required to be incorporated into this sublattice-based modeling technique due to the challenge associated with a high degree of freedom in the parameter selection. In this work, the number of parameters in the modelling of pure elements, end-members or stoichiometric phases is identified through the Akaike Information Criterion (AIC) that is a multi-variate parameter scoring approach with a penalizing factor for overfitting. However, a univariate scoring approach (such as an F-test) has been proposed to select from the rescaled interaction parameters of mixing in sublattices in the general case that multi-phases are present, since AIC is not applicable in the presence of numerous interaction parameters. After the identification of the model with appropriate number of parameters, an MCMC sampling approach is applied to find the parameter posterior distribution which represents the uncertainty of the model parameters. The developed UQ approach has been tested with a model for the excess Gibbs free energy in a binary case that is expressed as

$$G_{mf}^{E} = x_A x_B (H_{ex} - T S_{ex} + L_{ex} T^2) \tag{6.11}$$

The prior probability distributions for the parameters H_{ex}, S_{ex} and L_{ex} have been updated to their posterior distribution using 10 synthetic data points. The results of the MCMC sampling are been shown in Figure 6.8 in the forms of joint and marginal distributions for the parameters. The initial values of the parameters and their updated values with 95% credible intervals have also been reported in this figure.

This approach has also been applied for the UQ of the Ni-Al binary system with magnetic properties to show the capability of this automated workflow for the rapid construction of CALPHAD databases [145].

As mentioned earlier in this section, Duong et al. [70] have integrated thermo-chemical and experimental data to probabilistically calibrate a sub-regular solid solution CALPHAD model for the U-Nb binary system. An adaptive MCMC sampling technique based on the Metropolis-Hastings

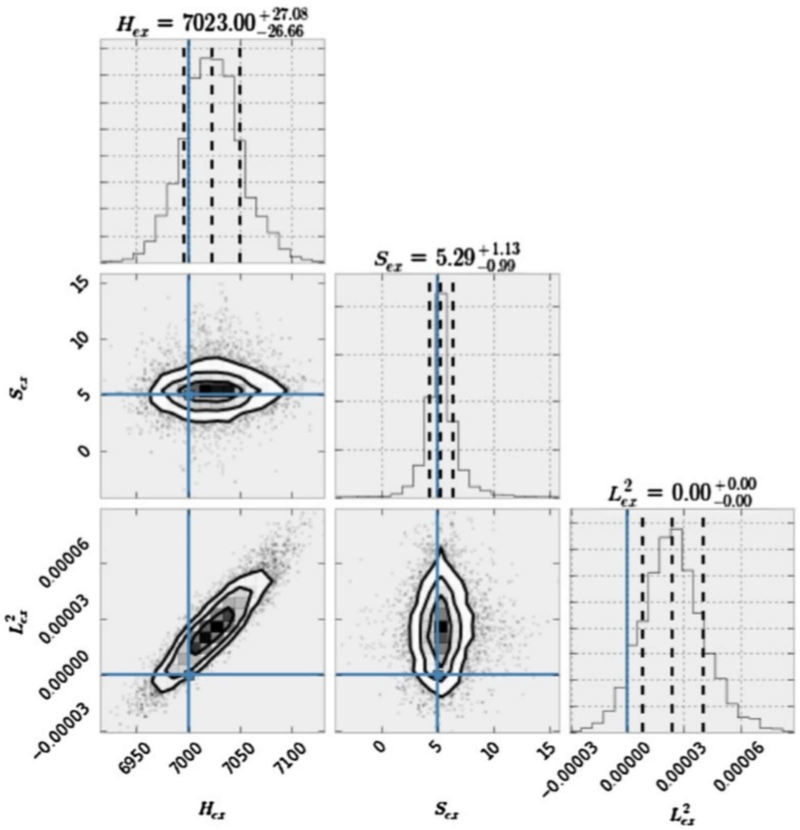


Figure 6.8: Joint and marginal parameter distributions after MCMC sampling, initial values of the parameters (blue solid lines) and their updated mean values with 95% credible intervals (black dashed lines). Reprinted with permission from [145].

algorithm has been employed for this purpose. In this MCMC approach, the deterministically optimized parameter values obtained from the PARROT module in Thermo-Calc software are used as initial values for the model parameters, $\theta^0 = \theta^{opt}$. A uniform distribution has also been considered for the parameter prior distribution which is bounded between ($\theta^{opt} - 3\theta^{opt}$) and ($\theta^{opt} + 3\theta^{opt}$). Moreover, the likelihood function has been considered Gaussian with unknown error due to lack of knowledge about the data uncertainties. Here, the errors of the data are assumed to be the same as a constant hyper-parameter which is updated like the model parameters during MCMC sampling. It should also be noted that the proposal distribution is an adaptive multivariate Gaussian distribution, where its initial arbitrary covariance matrix is adapted in each iteration of MCMC sampling based on the covariance matrix obtained from the previous generated parameter samples. After quantifying the parameter uncertainties using this MCMC sampling, the analytical FOSM approach has been applied in order to propagate the uncertainties to the Gibbs free energy formation of phase γ and the phase diagram for this system, as shown in Figure 6.9 [70].

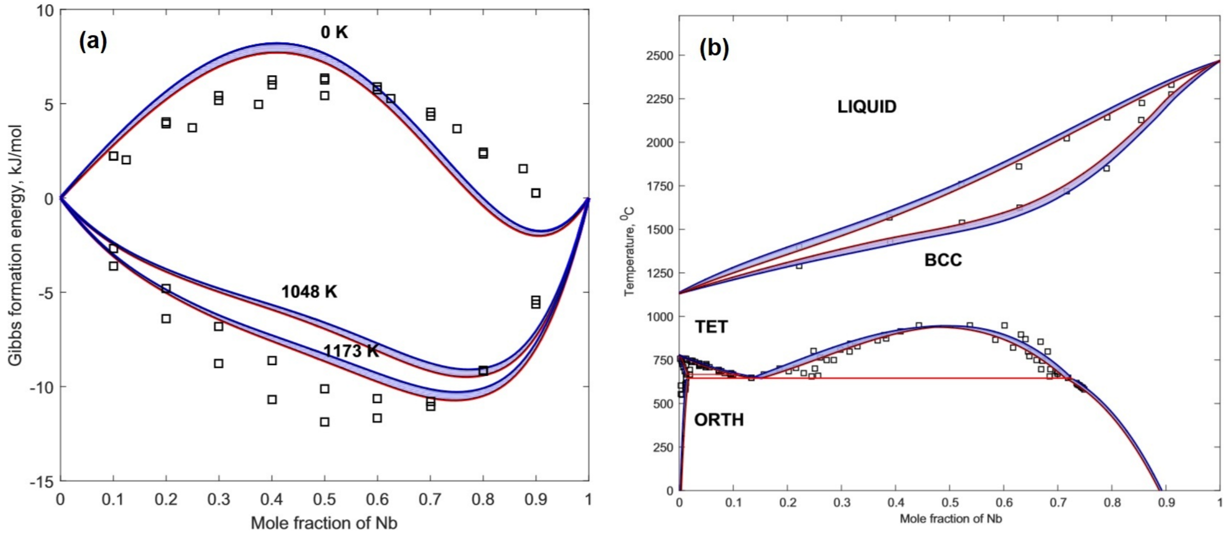


Figure 6.9: 95% credible intervals of a) the Gibbs free energy formation of phase γ and b) the phase diagram of U-Nb system with their corresponding data points. Reprinted with permission from [70].

Following the Duong et al. [70] work, Honarmandi et al. [180] have recently used the same MCMC approach for the uncertainty quantification of the Hf-Si binary phase diagram and developed this approach further for the calculation of the Bayes' factor which is an important measure for Bayesian model selection (BMS) and Bayesian model averaging (BMA). In other words, not only the Challenges for the uncertainty quantification and propagation in CALPHAD have been addressed in this work, but the challenges for model selection and highly objective estimation of phase diagrams have also been studied using Bayesian hypothesis testing and information fusion approaches with an underlying goal towards automated design under the ICME framework [180]. This work is discussed in more detail in the following case study.

6.5.2.1 Case study: UQ in Hf-Si phase diagram

In this case study based on the Honarmandi et al. [180] work, the MCMC-Metropolis Hastings algorithm explained in [70] is applied for an expert recommended CALPHAD model in the case of Hf-Si binary system to show how the uncertainties of the model parameters are quantified based on the given data in the context of Bayesian statistics. This case study also shows how the parameter uncertainties can precisely propagate to phase diagrams.

The CALPHAD model in this example has 20 parameters based on line-compound formalism which includes 2 parameters for each phase in the system (i.e. $^0a_\phi$ and $^0b_\phi$ which are referred to as a_ϕ and b_ϕ in this section). Synthetic data extracted from the calculated phase diagram in [147] has also been considered for this probabilistic calibration.

1000000 parameter-vector samples have been generated during MCMC sampling with exactly the same prior, likelihood and proposal distributions as [70] to ensure about true parameter convergences. After removal of the burn-in period (unnecessary samples in the beginning of MCMC sampling before parameter convergence) using the parameter cumulative mean plots, the average values and the square root of the diagonal elements in the variance-covariance matrix of the remaining samples have been introduced as the most plausible values and uncertainties of the model parameters, respectively. For better visualization of the MCMC simulation, the marginal and joint posterior frequency distributions can be plotted for the model parameters. Marginal posterior distributions are useful to show the distribution form of the model parameters. The marginal distributions of the parameters are almost Gaussian in this case study, which can be observed for one of the liquid phase parameters (a_{liq}) in Figure 6.10(a) as an example. Joint frequency distributions show the density of MCMC samples in the parameter space, which can graphically represent the parameter convergences as well as their qualitative correlations. For example, one of these plots has been shown in Figure 6.10(b). The convergence of the given pair parameters towards the red region can be observed in this figure. In addition, this plot suggests a positive correlation between the two given parameters. A positive /negative correlation indicates a positive/negative change in the value of one parameter as the other increases. Linear correlation between pair parameters can also be quantified by the Pearson linear coefficient expressed as,

$$\rho_{x,y} = \frac{cov(x,y)}{\sigma_x \sigma_y} \tag{6.12}$$

where σ_x, σ_y and $cov(x,y)$ are the standard deviation of x, the standard deviation of y and the covariance of x and y, respectively. The correlation coefficient $\rho_{x,y}$ can change from -1 to 1. Here, -1 or 1 corresponds to a perfect negative or positive linear correlation, while 0 means no correlation between the two parameters. As observed in Figure 6.10(b), the linear coefficient for the two given parameters is 0.95 implying a high positive linear correlation between these two parameters.

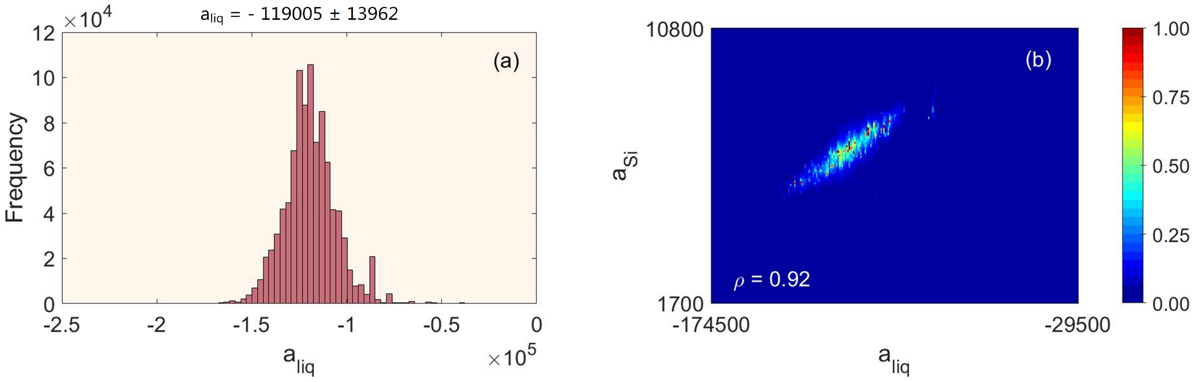


Figure 6.10: Examples of a) marginal and b) joint posterior frequency distributions of the parameters after MCMC sampling. Reprinted with permission from [180].

The obtained uncertainties of the model parameters have been propagated to the phase diagram using the forward model analysis of the last 5000 MCMC generated samples as an ensemble of all converged samples. In this scheme, 95% Bayesian credible intervals (BCI) can be determined by discarding 2.5% of the resulting samples from the above and below the total variation band at any specified point in the phase diagram. The resulting probabilistic phase diagram for the Hf-Si system has been shown in Figure 6.11. It is worth noting that the cross-sectional probability distributions are the same along the horizontal lines in the phase diagram, while they are composition dependent throughout the liquidus line.

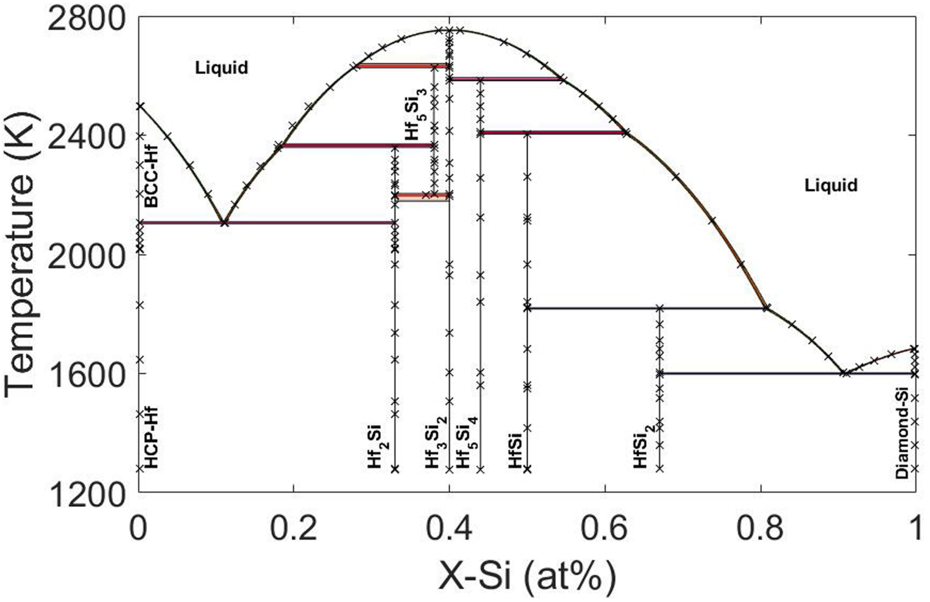


Figure 6.11: Hf-Si phase diagram with 95% credible intervals obtained after MCMC calibration of the CALPHAD model with 20 parameters against synthetic data shown by black crosses. Reprinted with permission from [180].

6.5.2.2 Case study: Bayesian model selection and information fusion for the Hf-Si system

In this section, the model in the previous section and three other models that have randomly been generated in [180] for the Hf-Si system are considered to show the applications of Bayesian model selection and information fusion approaches in CALPHAD. The number of parameters in these models are 17, 20, 28 and 30 which are referred as models 1, 2, 3 and 4, respectively. In this regard, it should be noted that the intermetallic phases in these models have been described by the line-compound formalism, while the end members are defined by either the line-compound or sub-regular solid solution model. The additional three models have also been calibrated against the same synthetic data through the MCMC approach explained in the previous section. The probabilistic phase diagrams after the propagation of the parameter uncertainties are shown in Figure 6.12. It is clear that these models have poor predictions with very high uncertainties for the phase diagram.

Before proceeding our discussion about the application of information fusion approaches in

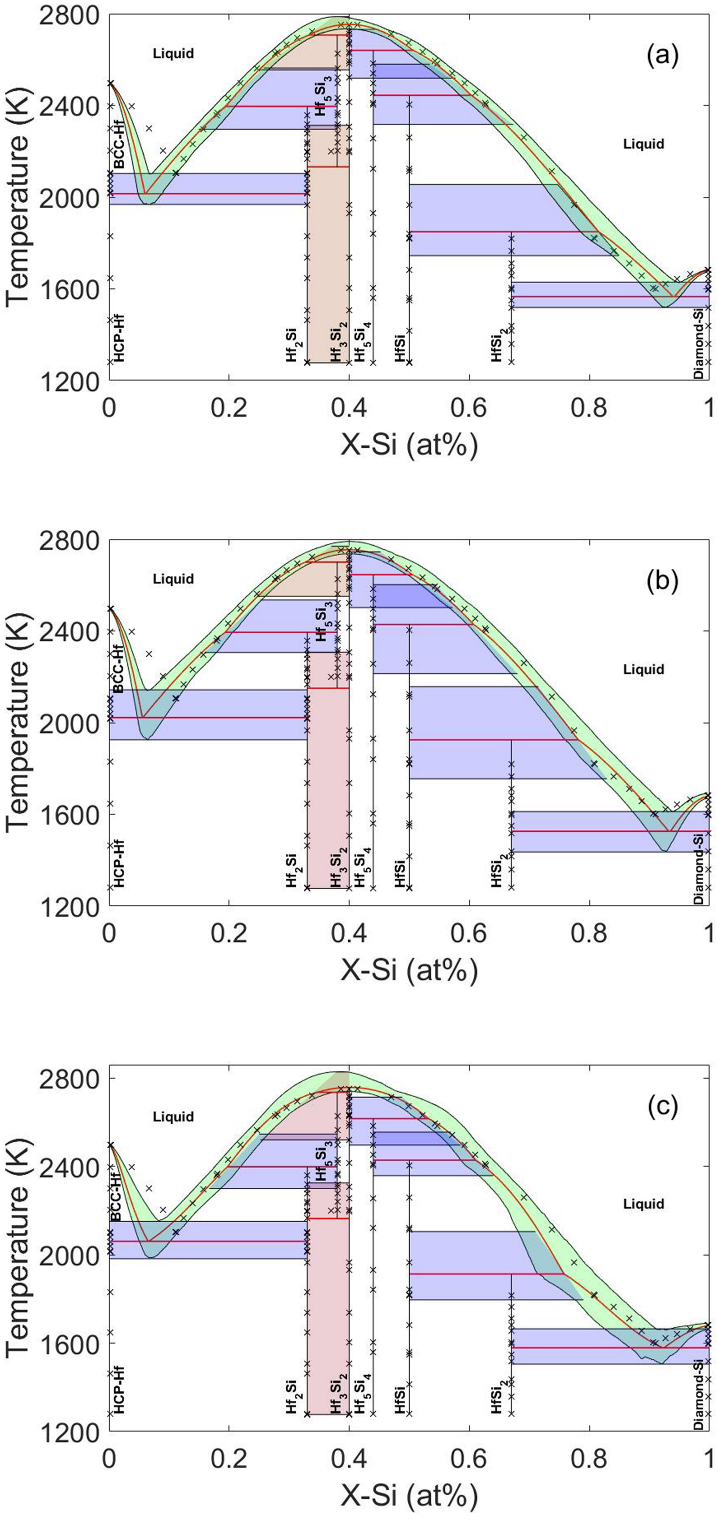


Figure 6.12: Hf-Si phase diagrams with 95% credible intervals obtained after MCMC calibration of the CALPHAD models with a) 17, b) 28 and c) 30 parameters against synthetic data shown by black crosses. Reprinted with permission from [180].

CALPHAD modeling, it is worth noting that Gibbs free energies of phases are the intermediate variables which connect the thermodynamic parameters to phase diagram. Therefore, uncertainties can be propagated from the parameters to Gibbs free energies and then to phase diagram, as shown in Figures 6.13(a) and (b) for model 1 and 2 at two different iso-thermal conditions. As expected, perfect agreements are generally observed between the mean values and uncertainties of phase equilibrium compositions obtained from the convex hull of Gibbs free energy plots and the corresponding phase diagrams at any arbitrary temperature. The parameter uncertainties can also be propagated in the same way to the first and second derivatives of Gibbs free energy with respect to composition. The uncertainty quantification of these derivatives is also important since the first derivatives indicate the elemental chemical potentials in phases which are used for the identification of the equilibrium conditions in the system; and the second derivatives determine the stability of the single phases in the miscibility gap (if any) to any small compositional fluctuations which can imply two different transformation mechanisms in these phase diagram regions, i.e. either nucleation and growth mechanism or spinodal decomposition.

Although it is intuitively expected to predict more precise phase diagrams as the number of the parameters in the model increases, the comparison of the phase diagram obtained from model 2 (20 parameters) in Figure 6.11 with the ones obtained from the models 3 and 4 (28 and 30 parameters) in Figures 6.12(b) and (C) implies that higher number of parameters is not necessarily useful in practice. In order to prove that unsuccessful parameter convergences are not the reasons for these contradictory results, the parameter convergence during MCMC sampling have been checked for all the models through the cumulative mean plots of their parameter samples. The plot for one of the parameters in liquid has been shown in Figure 6.14 as an example. It is worth noting that MCMC convergence appears as constant fluctuations around a constant value for each model parameter, which is equivalent to the convergence of the parameter cumulative mean values to constant values, as observed in Figure 6.14. Therefore, the contradictions do not result from the unsuccessful parameter convergences, but where they converge to. This is clarified further below.

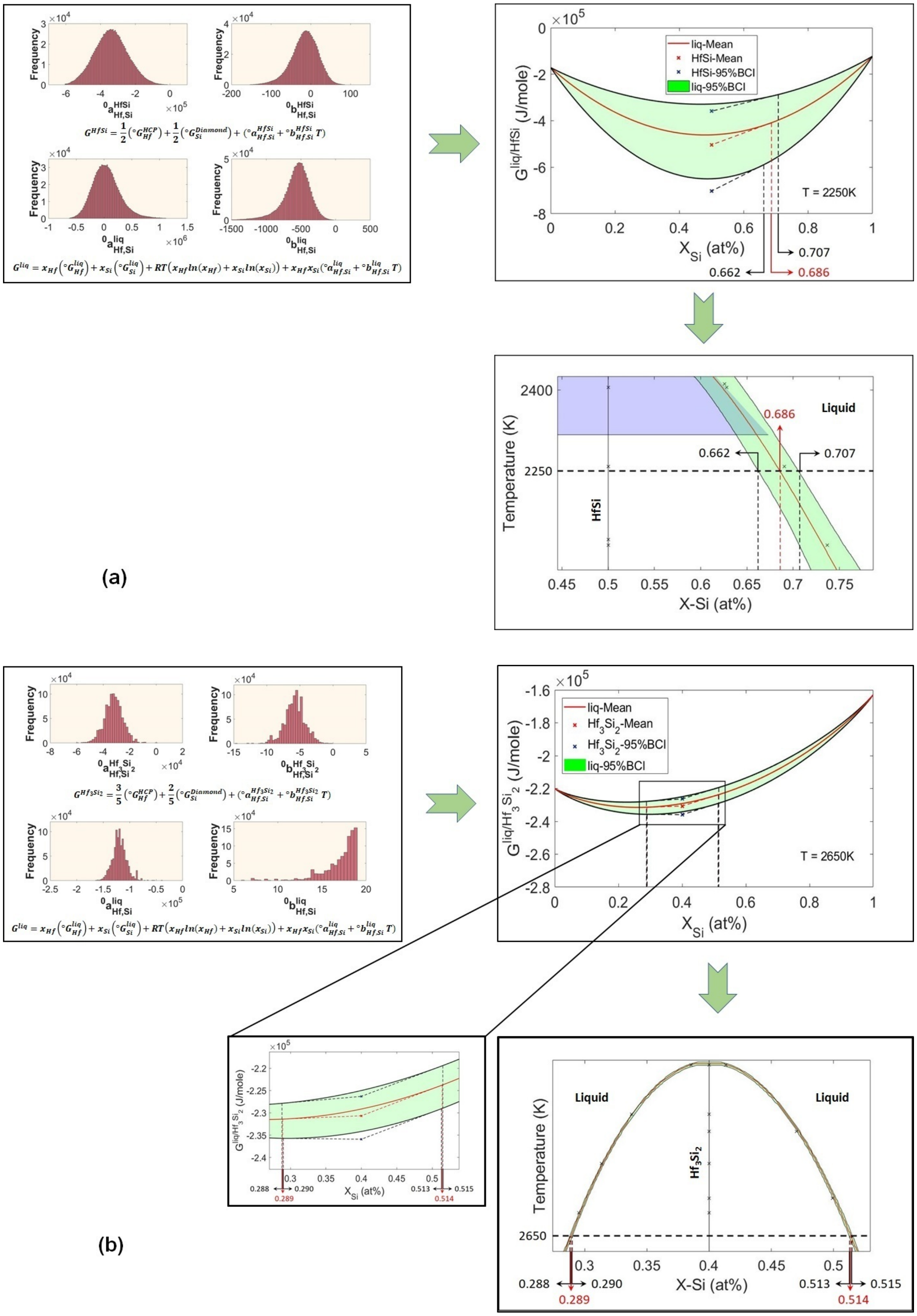


Figure 6.13: Two examples of the uncertainty propagation from the model parameters to Gibbs free energy curves to phase diagram for model 1 and 2 at two different iso-thermal conditions. Reprinted with permission from [180].

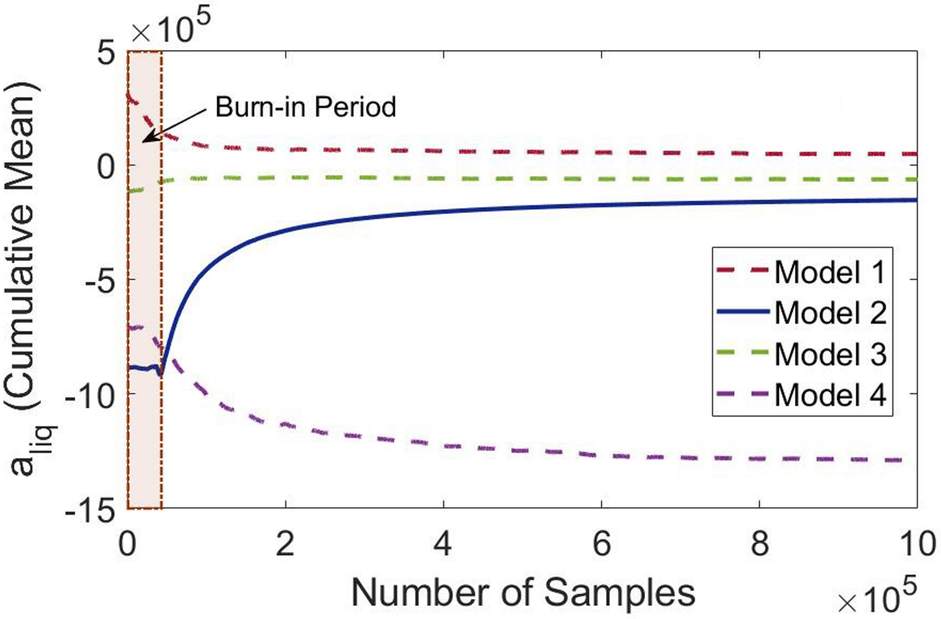


Figure 6.14: Cumulative mean plots for the parameter samples of one of the liquid parameters that exists in all the applied models. Reprinted with permission from [180].

Model 1 with 17 parameters is too simple to capture the correct phase diagram based on the given data. In this case, the large uncertainties have been attributed to the error of fitness. As expected, increasing the number of the model parameters from 17 to 20 in model 2 significantly improves the prediction of the phase diagram. However, higher number of parameters in model 3 or 4 results in so many local minima for the residual (the differences between results and the data) in the high dimensional parameter space, making this global optimization (calibration) problem very difficult or sometimes impossible to solve by the available deterministic or probabilistic optimization approaches. Therefore, the possibility of falling into a local minimum is extremely high during the deterministic calibration of large models in Thermo-Calc software. Since the parameter ranges are obtained based on these deterministically optimized parameter values in the applied MCMC approach, there is a very high probability for the global minimum to be outside the parameter ranges. Here, the high sensitivity of the large models to small variations of the parameters has been introduced as a reason for the high uncertainties in their resulting phase diagrams. Consequently, this study verifies the idea of model parsimony which emphasizes making the model as simple as possible, but not too simple.

In this case study, the best model (i.e. model 2) can visually be detected through the resulting phase diagrams in Figures 6.11 and 6.12. However, there might be cases that all or some of the models show comparably good predictions making the qualitative identification of the best model very difficult. Therefore, a systematic quantitative approach is necessary in these cases for model selection. Honarmandi et al. [180] have proposed Bayesian model selection based on the Bayes' factor to quantify to what extent the given models are favored by the data. Technically, Bayesian model selection is performed by comparing the posterior probabilities of the models given the data in the form of the posterior odds for each pair of models, as follows

$$\frac{P(M_i|D)}{P(M_j|D)} = \frac{P(D|M_i)}{P(D|M_j)}\frac{P(M_i)}{P(M_j)} = B_{ij}\lambda_{ij} \tag{6.13}$$

Since there are usually no prior preferences for any of the given models, prior odds (λ_{ij}) are assumed to be 1. Therefore, the posterior odds equal the ratio of the marginal likelihoods known as Bayes' factor. For the calculation of the marginal likelihoods, the integral in Equation 6.5 should be solved for each given model. In this regard, a harmonic average of likelihood values obtained during MCMC parameter sampling of each model has been suggested as a numerical solution [152].

In regard to the identification of the best CALPHAD model with an appropriate number of parameters, an MCMC Metropolis Hastings technique can be applied to sample from all the feasible models based on Equation 6.13. However, faster approaches are required for the prediction of phase diagrams in the direction of the automated design under ICME framework. Information fusion techniques can be candidate solutions for this challenge based on the fact that any feasible CALPHAD model can potentially provide some useful information about the system.

Honarmandi et al. [180] have applied BMA and an error correlation-based model fusion (CMF) to show how to smartly combine the existing information in the models to obtain a fused model with more informative results compared to each applied individual model (including the best model). In BMA, Bayes' factors are converted into the weights for the applied models, and then the fused model is constructed based on a weighted average of the results obtained from the given models. In this approach, the weight of models are defined using their posterior probabilities given the data which can be calculated through Bayes' factors as follows [152]

$$P(M_k|D) = \frac{P(D|M_k)\pi(M_k)}{\sum_{i=1}^{K} P(D|M_i)\pi(M_i)} = \frac{B_{k1}\lambda_{k1}}{\sum_{i=1}^{K} B_{i1}\lambda_{i1}} = \frac{B_{k1}}{\sum_{i=1}^{K} B_{i1}} \tag{6.14}$$

where K is the number of competing models and M_1 is a model reference that is used to calculate the Bayes' factors in this ratio. Based on Equation 6.14, the weight of models 1 to 4 are 0.1474, 0.5572, 0.1451, and 0.1504, respectively. It is clear that model 2 as the best model has a weight much more than the other models. The BMA fused model obtained from a weighted average of the resulting phase diagrams has been shown in Figure 6.15. In this figure, the posterior mode (peak) of the probability distributions at each composition has been considered rather than the mean values of the BMA model to prevent the unrealistic behaviour of the fused phase diagram around the eutectic points due to the difference in eutectic compositions predicted by the given models. In Figure 6.15, the examples of the probability distributions of the individual and fused models at different compositions/regions indicate that the posterior mode of the distributions in the fused model corresponds to the posterior mode of the distributions in model 2 (shown by black and purple distributions, respectively).

Therefore, the BMA fused model provides broader 95% credible intervals compared to the best model (model 2) which can offer more confidence for robust design applications. However, the BMA approach considers all the given CALPHAD models as statistically uncorrelated, which might not be true in this case due to the consideration of the same underlying structure for the CALPHAD models that attempts to predict the same ground truth. For this reason, the CMF approach has been used for fusion of the models. In this approach, the correlations between model errors are taken into account as a covariance matrix ($\tilde{\Sigma}$)) in this case in order to reduce the uncertainties for more precise predictions. The mean value and variance of the fused model at any arbitrary point ($\mathbf{x}$) is calculated as follows [10].

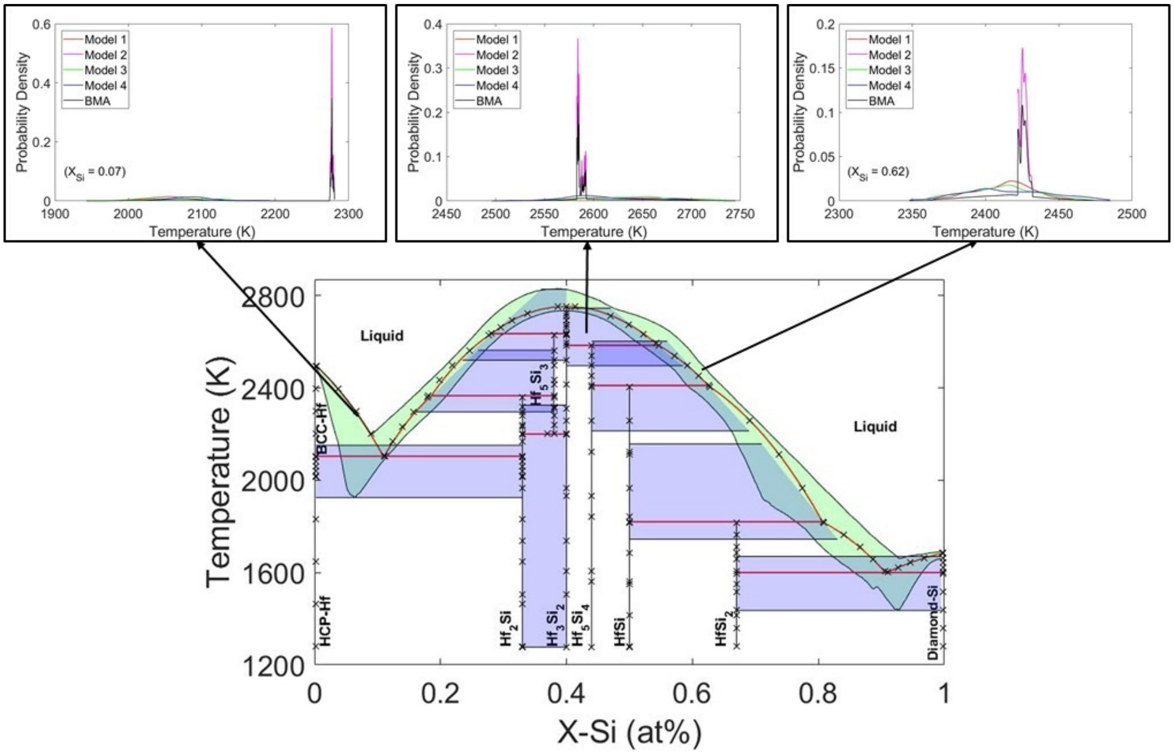


Figure 6.15: The posterior mode and 95% credible intervals of the BMA fused model for Hf-Si phase diagram. Reprinted with permission from [180].

$$E[\hat{f}(\mathbf{x})] = \frac{\mathbf{e}^\top \tilde{\Sigma}(\mathbf{x})^{-1} \mu(\mathbf{x})}{\mathbf{e}^\top \tilde{\Sigma}(\mathbf{x})^{-1} \mathbf{e}} \tag{6.15}$$

$$\text{Var}\left(\hat{f}(\mathbf{x})\right) = \frac{1}{\mathbf{e}^\top \tilde{\Sigma}(\mathbf{x})^{-1} \mathbf{e}} \tag{6.16}$$

where $e = [1, ..., 1]^T$ is an $r \times 1$ vector of ones, $\mu(\mathbf{x}) = [\mu_1(\mathbf{x}), ..., \mu_r(\mathbf{x})]^T$ denotes the vector of mean values for all the given models at point $\mathbf{x}$, and $\tilde{\Sigma}(\mathbf{x})^{-1}$ denotes the inverse of the $r \times r$ covariance matrix which can be determined using the mean values ($\mu_i(\mathbf{x})$) and standard deviations ($\sigma_i(\mathbf{x})$) of the applied models at point $\mathbf{x}$.

In this case study, the uncertainty at any composition/region of each individual phase diagram has been defined as an equivalent Gaussian distribution. In this context, the fused models obtained from the combination of the three models in Figure 6.12 and all the four models have been shown in Figures 6.16(a) and (b), respectively. While the phase diagram calculated by the fusion of the three models does not capture all the features of the phase diagram (especially at eutectic points), it can provide a phase diagram closer to the data with smaller uncertainty bands that can cover all available data for the given system. This result is very important for incorporating CALPHAD into automated ICME-based design, since it shows that the CMF approach can be applied for random CALPHAD models to obtain an appropriate phase diagram in the context of design, rather than a lengthy traditional trial-and-error approach to identify the best CALPHAD model. Even in the presence of model 2, the fused model (shown in Figure 6.16(b)) suggests more precise prediction compared to any individual model, which has been verified by the calculation of an average information (Shannon) entropy over 95% credible intervals of each phase diagram. Here, the average entropies have been considered to compare the uncertainties resulting from the individual models with their corresponding fused model. This result also implies that the CMF fused model that includes the best model can be more precise than the best model itself that may lose some information about the system [180].

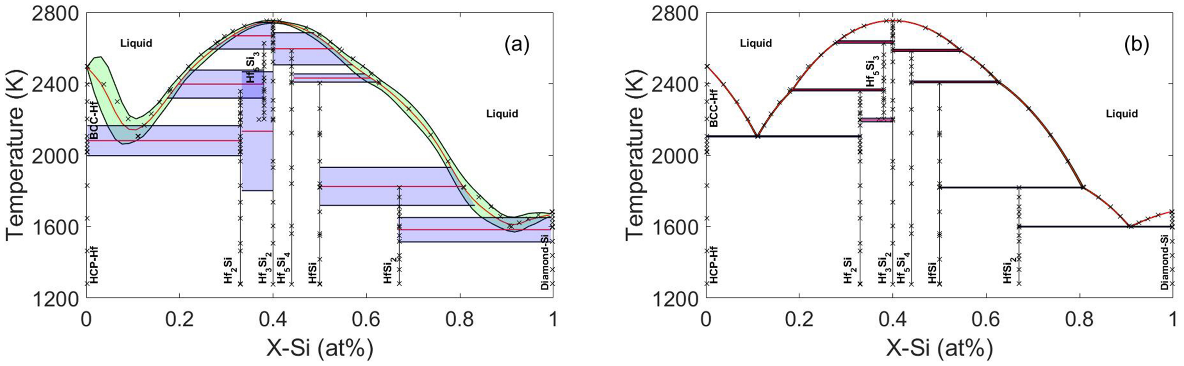


Figure 6.16: CMF fused models with their 95% credible intervals obtained from the fusion of a) models 1, 3 and 4, and b) all four models. Reprinted with permission from [180].

6.6 What to Do about (Uncertain) DFT Data?

For well over a decade, DFT calculations of the thermochemical properties of phases have become an integral element of CALPHAD assessments [169]. As such, they have provided invaluable information [170, 214, 215] that is particularly useful in the absence of sufficient thermochemical data to properly ground the scale of the Gibbs energies of the phases being considered for possible equilibria. There are in fact close to 4,000 manuscripts published in the past two decades in which DFT calculations have been used to assist in the assessment of the phase stability of materials systems–see for example [216–220].

While powerful, DFT calculations have been shown to have important limitations. When compared with CALPHAD-based assessments, for example, computed lattice stabilities of metastable phases tend to have significant disagreements with respect to their CALPHAD-derived counterparts [221]. Traditional DFT methods are also limited in their ability to compute the energetics of systems with strong electron correlation effects, making their incorporation into CALPHAD-based thermodynamic assessments problematic [222]. Even when computing the thermochemical properties of stable systems, DFT-based predictions are often seen to be in conflict with carefully acquired calorimetric data [172].

Given their general usefulness but acknowledging their limitations, the question then is what is the real value of DFT-based calculations in the context of CALPHAD assessments? In the context of UQ, the answer to this question remains easy to answer: DFT calculations are but one additional source of information about the (elusive) ground truth of the phase stability. This means that the methods presented above to infer the thermodynamic properties of phases from available, incomplete, sparse and most certainly uncertain (pun intended) phase stability and thermodynamic data are amenable to treating DFT data without further modification. In a sense, we can consider DFT data as additional information that can be used to build thermodynamically consistent CALPHAD models through the use of rigorous Bayesian inference methods. Recent examples [70, 146] by the some of the present authors attest to the viability of this approach.

6.7 Summary and Conclusion

Uncertainty quantification (UQ) of thermodynamic models and phase diagrams requires an understanding of the connections between the uncertainty in data and in the resultant models (and model parameters). After examining both aleatoric and epistemic uncertainty sources [182], we conclude that both have an important impact on phase diagram models. Although aleatoric uncertainty may cause statistical scatter in measured output metrics and the epistemic uncertainty may introduce systematic biases, both sources provide key information for evaluating the overall uncertainty in phase stability diagrams.

Going beyond traditional, analytical approaches, we examined frequentist and Bayesian statistical paradigms. Although both methodologies have specific strengths and weaknesses, their inferences can be consistent and correlated. In addition, both paradigms allow for forward and inverse uncertainty propagation. In the context of CALPHAD assessments, the inverse problem is critical and inverse uncertainty propagation methods result in fully dimensional parameter uncertainties including the covariances between parameters. Furthermore, inverse uncertainty propagation methods may be used to estimate the uncertainties in the data itself.

In this work, we discuss estimates of the combined aleatoric and epistemic uncertainty for data points or data sets. These estimated uncertainties, typically given by as standard deviation from the mean, may serve as effective "weighting factors" in the determination of model parameter distributions in both Bayesian and Frequentist approaches. While acknowledging that CALPHAD experts often assign weights based on expert knowledge of measurement approaches, research group reputation, or trial and error, the methodology we propose expands the type and quality of thermodynamic information used as input in model development and optimization.

We conclude that it is necessary to analyse thermodynamic data from different sources and to ensure consistency between property models. For example, the model of specific heat–a derivative of enthalpy with respect to temperature–must be consistent with the enthalpy and free energy model. To optimally represent the phase stability, all thermodynamic property models must satisfy the well-established thermodynamic relationships.

Through the application of this Bayesian framework to Hafnium, we identified that, for the case of the specific heat formulation, the Debye plus quartic polynomial model is optimal for the alpha phase, the quadratic polynomial model is optimal for the beta phase and the linear model is optimal for the liquid phase.

The analysis of binary phase diagrams revealed a large sparsity of the experimental data due to high measurement uncertainties and limited uncertainty quantification of the high-dimensional model parameters. The choice of thermodynamic model with appropriate number of parameters is one of the major challenges in CALPHAD assessments. Using Hf-Si as a case study and the MCMC-Metropolis Hastings algorithm as optimization method, the Bayesian framework can propagate uncertainties to the phase diagram. Moreover, we have shown how a Bayesian framework enables model selection through the evaluation of Bayesian evidence for competing models, given the available data. Finally, it has been shown how it is possible to fuse CALPHAD predictions in a robust manner to obtain reasonable phase diagram uncertainties.

The discussions presented in this work provide a snapshot of the field as it stands today. We are certain that with an ever-increasing emphasis on the development of simulation-assisted alloy development approaches, those that account for aleatoric/epistemic uncertainty in predictions will necessarily become ubiquitous. In the case of alloy design, quantification, propagation and management of uncertainty of CALPHAD models will certainly become a necessary (albeit not sufficient) condition for robust decision-making and it is our hope that this important aspect of materials simulation will become one of many tools available to those who seek to make the premise of ICME/MGI a reality.

7. SUMMARY AND FUTURE WORK

UQ is one of the most important tasks towards the fulfillment of ICME/MGI promises for materials design. For this reason, this work intends to address how to quantify the overall uncertainties of the parameters and to propagate them to the system outcomes in different cases of computational materials modeling; and how to perform model selection and probabilistic model fusion in the case that multiple competing models exist.

7.1 Summary

This dissertation begins by the discussions about the necessity of uncertainty quantification in robust and reliability-based design which is followed by the explanations of different UQ fundamental concepts, including the sources of uncertainty in computational modeling, model parameter selection, Frequentist and Bayesian approaches for inverse UQ of model parameters with their advantages and disadvantages, analytical and numerical methods for forward UP from model parameters to responses with their advantages and disadvantages. In addition, the fundamentals of Bayesian model selection and different model fusion approaches when several models try to describe the same physical phenomenon.

A stress-strain curve model have been implemented for TRIP steels, and then its influential parameters have probabilistically been calibrated using two different procedures of MCMC data training, i.e. sequential and simultaneous data training. The corresponding calibration results have been compared through the marginal and joint parameter probability distributions in addition to optimal stress-strain curves with propagated uncertainties obtained at both training and test experimental input conditions. This comparisons have implied no significant differences between these two types of data training. However, the fairly high discrepancies between the calibrated model results and the corresponding experimental data have brought the idea of the probabilistic estimation of uncertain input data. Therefore, the micro-structural phase volume fractions have been estimated using the same MCMC framework (while the model with fixed sequentially calibrated parameters has been considered) in order to reduce the epistemic uncertainties associated with the volume fraction measurements. The reproducibility of the MCMC approach has also been verified by the parameter probabilistic recalibration with the synthetic data sampled from the simultaneously calibrated model results. This means no considerable uncertainties result from the applied MCMC approach.

A multi-objective MCMC approach has also been developed for the probabilistic calibration of an Ni-Ti precipitation model in MatCalc© software against the available experimental data. Here, a empirical function has been found for matrix/precipitate interfacial energy in terms of nominal composition and aging temperature using the optimal values obtained for the interfacial energy after the MCMC calibration of the most influential model parameters with each given experimental data individually. After replacing the interfacial energy parameter by this relationship, the other parameters have probabilistically been determined through the multi-objective MCMC calibration with all the given experimental data simultaneously. Again, large discrepancies and uncertainties between the calibrated model results and the corresponding experimental data have reduced by applying the co-kriging model fusion approach that probabilistically combines the model results and experimental data for more precise prediction of precipitation behavior in Ni-Ti binary system.

In another work, a constrained MCMC approach has been proposed to probabilistically calibrate the most influential parameters a thermo-mechanical model under iso-baric conditions. For this purpose, a distance based metric has been utilized to compare the transformation strain-temperature curves obtained from the model with their experimental counterparts in the MCMC likelihood function. In this work, a systematic design of experiment that combines CFD and ANOVA methods has been applied to identify the most influential parameters. Moreover, UP from the model parameters to responses have been performed through the FOSM and forward analysis approaches separately. While FOSM approach showed some unrealistic uncertainty bounds around the transformation temperatures, the forward analysis approach provided very smooth uncertainty bounds with more precision. However, there is a trade-off between cost and precision which can be considered in order to select any of these UP approaches. Generally, there are good agreements between the hysteresis curves obtained from the model and their corresponding experimental data (especially at the maximum transformation strains), or at least the model 95% BCIs contain the experimental data; although the slopes and the temperatures–where the forward and reverse transformation curves meet–have not been predicted precisely. The application of the MCMC technique for decision making in experimental design has also been shown in this work. This purpose has been fulfilled through the comparison of the information gained by the sequential MCMC calibrations of the model parameters with two synthetic data sets which have been sampled from the uncertainty bounds of the model responses at the same and different iso-baric conditions. The mentioned comparison has been carried out by the calculation of K-L convergence for each experimental design which determines the contrast between prior and final posterior probability distribution of the parameters in each case. A higher value of K-L divergence obtained for the experimental design at different iso-baric conditions implies more information gain in this case. It should be noted that the K-L divergence values and their rates decrease as one-by-one sequential MCMC training continues in each experimental design case.

As the most important highlight of this dissertation, the challenge corresponding to the model selection in CALPHAD has been addressed. In this regard, the plausible optimal values of the parameters and their uncertainties in four independently generated CALPHAD models for Hf-Si binary system have been quantified through the same MCMC calibration against synthetic data which have been sampled from the phase diagram calculated by Zhang et al. [218]. It should be noted that the deterministically optimized parameters obtained from Thermo-Calc software are considered as the initial parameter values in the MCMC calibration for faster convergence of the parameter values. After the probabilistic calibrations, the parameter uncertainties have been propagated using the forward analysis to find the uncertainties in phase diagram. Bayes' factor in the context of Bayesian hypothesis testing has been applied to find the best model out of the given CALPHAD models. The results have shown that models with higher number of parameters do not necessarily have better predictions. This has been attributed to the incapibility of PARROT module in Thermo-Calc to find the global optimum in very high dimensional cases. In other words, increasing the number of parameters results in better prediction as far as Thermo-Calc has a good estimation of the global optimum in the parameter space. However, information fusion approaches–such as BMA and CMF–have been suggested in this work to solve the issue of finding the most appropriate CALPHAD model which may also lose some useful information about the system. BMA generates an average with larger uncertainties compared to any individual given model which can provide more confidence in the case of robust design, but it is probably too conservative. In contrast, CMF provides more precise predictions with lower uncertainties by considering correlations between the model errors. This reduction in uncertainties has been confirmed by the comparison of the average entropies associated with the phase diagrams obtained from each individual and fused models. It has also been shown that CMF can be applied on random CALPHAD models to produce a reasonable fused phase diagram with the capability of being used in robust design. As a result, this work was a big step toward making the CALPHAD modeling more objective rather than being subjective to the expert opinions.

In the end, this dissertation has been closed by a literature review on UQ/UP in CALPHAD modeling that is one of the fundamental tools in materials design under the ICME framework.

7.2 Future Work

The calibration of expensive high-fidelity models, such as phase field models, has still remained a big challenge in the field of materials science and engineering. The calibration of these types of models are almost impossible with the conventional deterministic and probabilistic approaches due to high computational cost. Therefore, an optimal search approach is required to efficiently combine the exploration and exploitation in the parameter space. In this regard, an efficient global optimization (EGO) method [81] can be one of the most promising solutions employing surrogate models (such as GP) over the discrepancy between the model results and the corresponding experimental observations to perform an efficient exploration and exploitation using the expected improvement metric. This calibration can also be performed in a probabilistic way based on the Kennedy and O'Hagan approach [13].

This dissertation has mostly been focused on the illustration of the Bayesian UQ/UP framework and different model fusion approaches for the sake of materials design under the ICME loop of process-structure-property-performance. However, these approaches have only been applied on individual low-fidelity (cheap) models and not on a chain of models. For future work, it is important to propagate the uncertainties across a chain of models, particularly chains that include expensive models. For example, the propagation of uncertainties from thermodynamic parameters to Gibbs free energies and phase diagrams to materials microstructures through CALPHAD and phase field models can be one of the important task in materials design. Since phase field modeling is very expensive, the development of an efficient UP approach seems to be required. In this regard, a UP approach based on probability measure optimized importance weights are being developed by Sanghvi et al. [223] in order to efficiently propagate the uncertainties of Gibbs free energy of phases to microstructural features across a phase field model [224] developed for Mg_2Sn-Mg_2Si quasi-binary system.

Another ongoing work is the application of information fusion approaches (such as co-kriging or CMF) for more precise and more efficient multi-objective efficient optimization which can be very useful for materials discovery in the case of Ni-Ti based SMAs. In this work, process-structure-property in this class of materials are connected together through a chain of models that includes a precipitation model and a thermo-mechanical/micro-mechanical model. The fusion of the results obtained from the mentioned chain of models and the experiments can provide faster optimization with less number of experiments to run compared to common multi-objective EGO with GP surrogate modeling, which can reduce the cost of optimization/materials discovery. Another important application of information fusion approaches is efficient queries from different sources of information with different fidelities and costs in the case of knowledge gradient (KG) or entropy-based optimization, which can be considered for future work.

REFERENCES

[1] J. H. Panchal, S. R. Kalidindi, and D. L. McDowell, “Key computational modeling issues in integrated computational materials engineering,” *Computer-Aided Design*, vol. 45, no. 1, pp. 4–25, 2013.

[2] A. Chernatynskiy, S. R. Phillpot, and R. LeSar, “Uncertainty quantification in multi-scale simulation of materials: A prospective,” *Annual Review of Materials Research*, vol. 43, pp. 157–182, 2013.

[3] H. Choi, D. L. McDowell, J. K. Allen, D. Rosen, and F. Mistree, “An inductive design exploration method for robust multi-scale materials design,” *Journal of Mechanical Design*, vol. 130, no. 3, p. 031402, 2008.

[4] S. Jackman, “Estimation and inference via bayesian simulation: An introduction to markov chain monte carlo,” *American Journal of Political Science*, vol. 44, no. 2, pp. 375–404, 2000.

[5] W. Olbricht, N. D. Chatterjee, and K. Miller, “Bayes estimation: A novel approach to derivation of internally consistent thermodynamic data for minerals, their uncertainties, and correlations. part I: Theory,” *Physics and Chemistry of Minerals*, vol. 21, no. 1-2, pp. 36–49, 1994.

[6] S. M. Lynch, *Introduction to applied Bayesian statistics and estimation for social scientists*. Springer Science & Business Media, 2007.

[7] A. Saltelli, P. Annoni, I. Azzini, F. Campolongo, M. Ratto, and S. Tarantola, “Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index,” *Computer Physics Communications*, vol. 181, no. 2, pp. 259–270, 2010.

[8] R. A. Fisher, *The design of experiments*. Oliver And Boyd; Edinburgh; London, 1937.

[9] T. J. Dasey and J. J. Braun, "Information fusion and response guidance," *Lincoln Laboratory Journal*, vol. 17, no. 1, 2007.

[10] D. Allaire and K. Willcox, "Fusing information from multifidelity computer models of physical systems," in *2012 15th International Conference on Information Fusion*, pp. 2458–2465, July 2012.

[11] W. D. Thomison and D. L. Allaire, "A model reification approach to fusing information from multi-fidelity information sources," in *19th AIAA Non-Deterministic Approaches Conference*, p. 1949, 2017.

[12] M. C. Kennedy and A. O'Hagan, "Predicting the output from a complex computer code when fast approximations are available," *Biometrika*, vol. 87, no. 1, pp. 1–13, 2000.

[13] M. C. Kennedy and A. O'Hagan, "Bayesian calibration of computer models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 63, no. 3, pp. 425–464, 2001.

[14] H. Jeffreys, "Some tests of significance, treated by the theory of probability," in *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 31, pp. 203–222, Cambridge University Press, 1935.

[15] W. Dan, W. Zhang, S. Li, and Z. Lin, "A model for strain-induced martensitic transformation of TRIP steel with strain rate," *Computational Materials Science*, vol. 40, no. 1, pp. 101–107, 2007.

[16] A. Grajcar, "Hot-working in the γ+ α region of TRIP-aided microalloyed steel," *Archives of Materials Science and Engineering*, vol. 28, no. 12, pp. 743–750, 2007.

[17] B. De Cooman, "Structure–properties relationship in TRIP steels containing carbide-free bainite," *Current Opinion in Solid State and Materials Science*, vol. 8, no. 3-4, pp. 285–303, 2004.

[18] H. N. Han, C.-S. Oh, G. Kim, and O. Kwon, "Design method for trip-aided multi-phase steel based on a microstructure-based modelling for transformation-induced plasticity and

mechanically induced martensitic transformation," *Materials Science and Engineering: A*, vol. 499, no. 1-2, pp. 462–468, 2009.

[19] E. Girault, P. Jacques, P. Harlet, K. Mols, J. Van Humbeeck, E. Aernoudt, and F. Delannay, "Metallographic methods for revealing the multiphase microstructure of TRIP-assisted steels," *Materials Characterization*, vol. 40, no. 2, pp. 111–118, 1998.

[20] N. Van Dijk, A. Butt, L. Zhao, J. Sietsma, S. Offerman, J. Wright, and S. Van der Zwaag, "Thermal stability of retained austenite in TRIP steels studied by synchrotron X-ray diffraction during cooling," *Acta Materialia*, vol. 53, no. 20, pp. 5439–5447, 2005.

[21] O. Grässel, L. Krüger, G. Frommeyer, and L. Meyer, "High strength Fe–Mn–(Al, Si) TRIP/TWIP steels developmentâĂŤpropertiesâĂŤapplication," *International Journal of plasticity*, vol. 16, no. 10-11, pp. 1391–1409, 2000.

[22] C. Herrera, D. Ponge, and D. Raabe, "Design of a novel mn-based 1 GPa duplex stainless TRIP steel with 60% ductility by a reduction of austenite stability," *Acta Materialia*, vol. 59, no. 11, pp. 4653–4664, 2011.

[23] O. Matsumura, Y. Sakuma, and H. TakeChi, "Enhancement of elongation by retained austenite in intercritical annealed 0.4 C-1.5 Si-o.8 Mn steel," *Transactions of the Iron and Steel Institute of Japan*, vol. 27, no. 7, pp. 570–579, 1987.

[24] S. Oliver, T. Jones, and G. Fourlaris, "Dual phase versus TRIP strip steels: Comparison of dynamic properties for automotive crash performance," *Materials Science and Technology*, vol. 23, no. 4, pp. 423–431, 2007.

[25] F. Lani, Q. Furnémont, T. Van Rompaey, F. Delannay, P. Jacques, and T. Pardoen, "Multiscale mechanics of TRIP-assisted multiphase steels: Ii. micromechanical modelling," *Acta Materialia*, vol. 55, no. 11, pp. 3695–3705, 2007.

[26] J.-Y. Liu, H. Lu, J.-M. Chen, J.-F. Jullien, and T. Wu, "Simulation of mechanical behavior of multiphase TRIP steel taking account of transformation-induced plasticity," *Computational Materials Science*, vol. 43, no. 4, pp. 646–654, 2008.

[27] J. Bouquerel, K. Verbeken, and B. De Cooman, "Microstructure-based model for the static mechanical behaviour of multi-phase steels," *Acta Materialia*, vol. 54, no. 6, pp. 1443–1456, 2006.

[28] A. Perlade, O. Bouaziz, and Q. Furnemont, "A physically based model for TRIP-aided carbon steels behaviour," *Materials Science and Engineering: A*, vol. 356, no. 1-2, pp. 145–152, 2003.

[29] L. Delannay, P. Jacques, and T. Pardoen, "Modelling of the plastic flow of trip-aided multi-phase steel based on an incremental mean-field approach," *International Journal of Solids and Structures*, vol. 45, no. 6, pp. 1825–1843, 2008.

[30] V. F. Zackay, E. R. Parker, D. Fahr, and R. Busch, "The enhancement of ductility in high-strength steels," *ASM transactions quarterly*, vol. 60, no. 2, pp. 252–259, 1967.

[31] G. Olson and M. Cohen, "Kinetics of strain-induced martensitic nucleation," *Metallurgical Transactions A*, vol. 6, no. 4, p. 791, 1975.

[32] G. B. Olson and M. Azrin, "Transformation behavior of TRIP steels," *Metallurgical Transactions A*, vol. 9, no. 5, pp. 713–721, 1978.

[33] Y. Tomita and T. Iwamoto, "Computational prediction of deformation behavior of TRIP steels under cyclic loading," *International Journal of Mechanical Sciences*, vol. 43, no. 9, pp. 2017–2034, 2001.

[34] S. Thibaud, N. Boudeau, and J.-C. Gelin, "TRIP steel: Plastic behaviour modelling and influence on functional behaviour," *Journal of Materials Processing Technology*, vol. 177, no. 1-3, pp. 433–438, 2006.

[35] M. Gallagher and J. Doherty, "Parameter estimation and uncertainty analysis for a watershed model," *Environmental Modelling & Software*, vol. 22, no. 7, pp. 1000–1020, 2007.

[36] A. Kanso, G. Chebbo, and B. Tassin, "Application of MCMC–GSA model calibration method to urban runoff quality modeling," *Reliability Engineering & System Safety*, vol. 91, no. 10-11, pp. 1398–1405, 2006.

[37] W. J. Browne, D. Draper, *et al.*, "A comparison of bayesian and likelihood-based methods for fitting multilevel models," *Bayesian Analysis*, vol. 1, no. 3, pp. 473–514, 2006.

[38] E. Galindo-Nava, A. Perlade, and P. Rivera-Díaz-del Castillo, "A thermostatistical theory for solid solution effects in the hot deformation of alloys: an application to low-alloy steels," *Modelling and Simulation in Materials Science and Engineering*, vol. 22, no. 1, p. 015009, 2013.

[39] G. Haidemenopoulos, N. Aravas, and I. Bellas, "Kinetics of strain-induced transformation of dispersed austenite in low-alloy TRIP steels," *Materials Science and Engineering: A*, vol. 615, pp. 416–423, 2014.

[40] P. Rivera-Díaz-del Castillo, K. Hayashi, and E. Galindo-Nava, "Computational design of nanostructured steels employing irreversible thermodynamics," *Materials Science and Technology*, vol. 29, no. 10, pp. 1206–1211, 2013.

[41] K. Irvine, T. Gladman, and F. Pickering, "The strength of austenitic stainless steels," *The Journal of the Iron and Steel Institute*, vol. 207, no. 7, pp. 1017–1028, 1969.

[42] R. Rodriguez, I. Gutierrez, *et al.*, "Unified formulation to predict the tensile curves of steels with different microstructures," in *Materials Science Forum*, vol. 426, pp. 4525–4530, Trans Tech Publications Ltd., Zurich-Uetikon, Switzerland, 2003.

[43] M. Azuma, N. Fujita, M. Takahashi, T. Senuma, D. Quidort, and T. Lung, "Modelling upper and lower bainite trasformation in steels," *ISIJ International*, vol. 45, no. 2, pp. 221–228, 2005.

[44] U. Kocks and H. Mecking, "Physics and phenomenology of strain hardening: The FCC case," *Progress in Materials Science*, vol. 48, no. 3, pp. 171–273, 2003.

[45] E. Galindo-Nava and P. Rivera-Díaz-del Castillo, "Modelling plastic deformation in BCC metals: Dynamic recovery and cell formation effects," *Materials Science and Engineering: A*, vol. 558, pp. 641–648, 2012.

[46] E. Galindo-Nava and P. Rivera-Díaz-del Castillo, "Thermostatistical modelling of hot deformation in FCC metals," *International Journal of Plasticity*, vol. 47, pp. 202–221, 2013.

[47] S. Li, P. Honarmandi, R. Arróyave, and P. Rivera-Diaz-del Castillo, "Describing the deformation behaviour of TRIP and dual phase steels employing an irreversible thermodynamics formulation," *Materials Science and Technology*, vol. 31, no. 13, pp. 1658–1663, 2015.

[48] S. Singh and H. Bhadeshia, "Estimation of bainite plate-thickness in low-alloy steels," *Materials Science and Engineering: A*, vol. 245, no. 1, pp. 72–79, 1998.

[49] c. H. Young and H. Bhadeshia, "Strength of mixtures of bainite and martensite," *Materials Science and Technology*, vol. 10, no. 3, pp. 209–214, 1994.

[50] M. Laine, "MCMC toolbox for Matlab," *Last Visited August*, vol. 22, p. 2013, 2013.

[51] H. Haario, E. Saksman, J. Tamminen, *et al.*, "An adaptive Metropolis algorithm," *Bernoulli*, vol. 7, no. 2, pp. 223–242, 2001.

[52] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian data analysis*, vol. 2. CRC press Boca Raton, FL, 2014.

[53] J. Tellinghuisen, "Statistical error propagation," *The Journal of Physical Chemistry A*, vol. 105, no. 15, pp. 3917–3921, 2001.

[54] P. Jacques, Q. Furnémont, A. Mertens, and F. Delannay, "On the sources of work hardening in multi-phase steels assisted by transformation-induced plasticity," *Philosophical Magazine A*, vol. 81, no. 7, pp. 1789–1812, 2001.

[55] X. Mo, F. Pappenberger, K. Beven, S. Liu, A. De Roo, and Z. Lin, "Parameter conditioning and prediction uncertainties of the LISFLOOD-WB distributed hydrological model," *Hydrological Sciences Journal*, vol. 51, no. 1, pp. 45–65, 2006.

[56] D. J. Battle, P. Gerstoft, W. S. Hodgkiss, W. Kuperman, and P. L. Nielsen, "Bayesian model selection applied to self-noise geoacoustic inversion," *The Journal of the Acoustical Society of America*, vol. 116, no. 4, pp. 2043–2056, 2004.

[57] X. Ren, N. Miura, J. Zhang, K. Otsuka, K. Tanaka, M. Koiwa, T. Suzuki, Y. I. Chumlyakov, and M. Asai, “A comparative study of elastic constants of Ti–Ni-based alloys prior to martensitic transformation,” *Materials Science and Engineering: A*, vol. 312, no. 1-2, pp. 196–206, 2001.

[58] J. Michutta, C. Somsen, A. Yawny, A. Dlouhy, and G. Eggeler, “Elementary martensitic transformation processes in Ni-rich NiTi single crystals with $\mathrm{Ni_4Ti_3}$ precipitates,” *Acta Materialia*, vol. 54, no. 13, pp. 3525–3542, 2006.

[59] E. Povoden-Karadeniz, D. Cirstea, P. Lang, T. Wojcik, and E. Kozeschnik, “Thermodynamics of Ti–Ni shape memory alloys,” *Calphad*, vol. 41, pp. 128–139, 2013.

[60] L. Johnson and R. Arroyave, “An inverse design framework for prescribing precipitation heat treatments from a target microstructure,” *Materials & Design*, vol. 107, pp. 7–17, 2016.

[61] G. Fan, W. Chen, S. Yang, J. Zhu, X. Ren, and K. Otsuka, “Origin of abnormal multi-stage martensitic transformation behavior in aged Ni-rich Ti–Ni shape memory alloys,” *Acta Materialia*, vol. 52, no. 14, pp. 4351–4362, 2004.

[62] J. Khalil-Allafi, A. Dlouhy, and G. Eggeler, “$\mathrm{Ni_4Ti_3}$-precipitation during aging of NiTi shape memory alloys and its influence on martensitic phase transformations,” *Acta Materialia*, vol. 50, no. 17, pp. 4255–4274, 2002.

[63] J. Kim and S. Miyazaki, “Effect of nano-scaled precipitates on shape memory behavior of Ti-50.9 at.% Ni alloy,” *Acta Materialia*, vol. 53, no. 17, pp. 4545–4554, 2005.

[64] E. Y. Panchenko, Y. I. Chumlyakov, I. V. Kireeva, A. V. Ovsyannikov, H. Sehitoglu, I. Karaman, and Y. H. J. Maier, “Effect of disperse $\mathrm{Ni_4Ti_3}$ particles on the martensitic transformations in titanium nickelide single crystals,” *The Physics of Metals and Metallography*, vol. 106, no. 6, pp. 577–589, 2008.

[65] N. Zhou, C. Shen, M.-X. Wagner, G. Eggeler, M. Mills, and Y. Wang, “Effect of $\mathrm{Ni_4Ti_3}$ precipitation on martensitic transformation in Ti–Ni,” *Acta Materialia*, vol. 58, no. 20, pp. 6685–6694, 2010.

[66] C.-b. KE, S.-s. CAO, M. Xiao, and X.-p. ZHANG, “Modeling of $\mathrm{Ni_4Ti_3}$ precipitation during stress-free and stress-assisted aging of bi-crystalline NiTi shape memory alloys,” *Transactions of Nonferrous Metals Society of China*, vol. 22, no. 10, pp. 2578–2585, 2012.

[67] C. Ke, S. Cao, and X. Zhang, “Phase field simulation of coherent precipitation of $\mathrm{Ni_4Ti_3}$ particles during stress-assisted aging of a porous NiTi alloy,” *Modelling and Simulation in Materials Science and Engineering*, vol. 23, no. 5, p. 055008, 2015.

[68] R. Kampmann and R. Wagner, “Kinetics of precipitation in metastable binary alloys-theory and application to Cu-1.9 at% Ti and Ni-14 at% Al,” in *Decomposition of Alloys: the Early Stages*, pp. 91–103, Pergamon, 1984.

[69] P. Honarmandi and R. Arroyave, “Using bayesian framework to calibrate a physically based model describing strain-stress behavior of trip steels,” *Computational Materials Science*, vol. 129, pp. 66–81, 2017.

[70] T. C. Duong, R. E. Hackenberg, A. Landa, P. Honarmandi, A. Talapatra, H. M. Volz, A. Llobet, A. I. Smith, G. King, S. Bajaj, *et al.*, “Revisiting thermodynamics and kinetic diffusivities of uranium–niobium with bayesian uncertainty analysis,” *Calphad*, vol. 55, pp. 219–230, 2016.

[71] R. C. Sommitsch, E. Kozeschnik, G. Wasle, and B. Buchmayr, “A precipitation model for multi-phase systems in Nickel-base superalloy,” in *T. Chandra, et al. (Eds.), Proceedings of THERMEC 2000-International Conference on Processing and Manufacturing of Advanced Materials*, pp. 1–6, Elsevier Science, 2000.

[72] M. Hillert and L.-I. Staffansson, “The regular solution model for stoichiometric phases and ionic melts.,” *Acta Chemica Scandinavica*, vol. 24, pp. 3618–3626, 1970.

[73] E. Kozeschnik, J. Svoboda, and F. D. Fischer, “Modified evolution equations for the precipitation kinetics of complex phases in multi-component systems,” *Calphad*, vol. 28, no. 4, pp. 379–382, 2004.

[74] R. W. Balluffi, S. Allen, and W. C. Carter, *Kinetics of materials*. John Wiley & Sons, 2005.

[75] Q. Du, W. J. Poole, and M. A. Wells, "A mathematical model coupled to CALPHAD to predict precipitation kinetics for multicomponent Aluminum alloys," *Acta Materialia*, vol. 60, no. 9, pp. 3830–3839, 2012.

[76] O. R. Myhr and Ã. Grong, "Modelling of non-isothermal transformations in alloys containing a particle distribution," *Acta Materialia*, vol. 48, no. 7, pp. 1605–1615, 2000.

[77] Q. Du, B. Holmedal, J. Friis, and C. D. Marioara, "Precipitation of non-spherical particles in Aluminum alloys part II: Numerical simulation and experimental characterization during aging treatment of an Al-Mg-Si alloy," *Metallurgical and Materials Transactions A*, vol. 47, no. 1, pp. 589–599, 2016.

[78] B. Calderhead, "A general construction for parallelizing Metropolis-Hastings algorithms," *Proceedings of the National Academy of Sciences*, vol. 111, no. 49, pp. 17408–17413, 2014.

[79] J. E. Valenzuela del Río and D. Mavris, "Gaussian process surrogate model for levering similar trends across concepts," *AIAA Journal*, vol. 53, no. 4, pp. 1002–1015, 2015.

[80] S. Remes, M. Heinonen, and S. Kaski, "Latent correlation gaussian processes," *ArXiv Preprint ArXiv:1702.08402*, 2017.

[81] D. R. Jones, M. Schonlau, and W. J. Welch, "Efficient global optimization of expensive black-box functions," *Journal of Global Optimization*, vol. 13, no. 4, pp. 455–492, 1998.

[82] J. Sacks, W. J. Welch, T. J. Mitchell, and H. P. Wynn, "Design and analysis of computer experiments," *Statistical Science*, vol. 4, no. 4, pp. 409–423, 1989.

[83] A. I. Forrester, A. Sóbester, and A. J. Keane, "Multi-fidelity optimization via surrogate modelling," in *Proceedings of the Royal Society of London a: Mathematical, Physical and Engineering Sciences*, vol. 463, pp. 3251–3269, The Royal Society, 2007.

[84] L. Leifur and K. Slawomir, *Simulation-driven aerodynamic design using variable-fidelity models*. World Scientific, 2015.

[85] I. Couckuyt, T. Dhaene, and P. Demeester, “ooDACE toolbox,” *Advances in Engineering Software*, vol. 49, no. 3, pp. 1–13, 2012.

[86] G. Kang and Q. Kan, “Thermomechanical cyclic deformation of shape-memory alloys,” in *Cyclic Plasticity of Engineering Materials*, pp. 405–530, John Wiley & Sons, Ltd, 2017.

[87] Q. Kan and G. Kang, “Constitutive model for uniaxial transformation ratchetting of superelastic NiTi shape memory alloy at room temperature,” *International Journal of Plasticity*, vol. 26, no. 3, pp. 441–465, 2010.

[88] C. Yu, G. Kang, Q. Kan, and X. Xu, “Physical mechanism based crystal plasticity model of NiTi shape memory alloys addressing the thermo-mechanical cyclic degeneration of shape memory effect,” *Mechanics of Materials*, vol. 112, pp. 1–17, 2017.

[89] C. Yu, G. Kang, and Q. Kan, “A macroscopic multi-mechanism based constitutive model for the thermo-mechanical cyclic degeneration of shape memory effect of NiTi shape memory alloy,” *Acta Mechanica Sinica*, vol. 33, no. 3, pp. 619–634, 2017.

[90] S. K. Bhaumik, K. V. Ramaiah, and C. N. Saikrishna, “Nickel-titanium shape memory alloy wires for thermal actuators,” in *Micro and Smart Devices and Systems*, pp. 181–198, Springer, New Delhi, 2014.

[91] S. Dilibal, H. Sahin, E. Dursun, and E. D. Engeberg, “NickelâĂŞtitanium shape memory alloy-actuated thermal overload relay system design,” *Electrical Engineering*, vol. 99, no. 3, pp. 923–930, 2017.

[92] T. Baxevanis, A. Cox, and D. C. Lagoudas, “Micromechanics of precipitated near-equiatomic Ni-rich NiTi shape memory alloys,” *Acta Mechanica*, vol. 225, no. 4, pp. 1167–1185, 2014.

[93] E. Patoor, D. C. Lagoudas, P. B. Entchev, L. C. Brinson, and X. Gao, “Shape memory alloys, part I: General properties and modeling of single crystals,” *Mechanics of Materials*, vol. 38, no. 5, pp. 391–429, 2006.

[94] M. Frost, P. Sedlák, M. Sippola, and P. Šittner, “Thermomechanical model for NiTi shape memory wires,” *Smart Materials and Structures*, vol. 19, no. 9, p. 094010, 2010.

[95] G. Tapia, L. Johnson, B. Franco, K. Karayagiz, J. Ma, R. Arroyave, I. Karaman, and A. Elwany, “Bayesian calibration and uncertainty quantification for a physics-based precipitation model of nickel–titanium shape-memory alloys,” *Journal of Manufacturing Science and Engineering*, vol. 139, no. 7, p. 071002, 2017.

[96] I. Steinbach and O. Shchyglo, “Phase-field modelling of microstructure evolution in solids: Perspectives and challenges,” *Current Opinion in Solid State and Materials Science*, vol. 15, no. 3, pp. 87–92, 2011.

[97] L. Haller, B. Nedjar, Z. Moumni, I. Vedinaş, and E. Trană, “A thermomechanical model accounting for the behavior of shape memory alloys in finite deformations,” *Continuum Mechanics and Thermodynamics*, vol. 28, no. 4, pp. 957–975, 2016.

[98] D. Jiang, S. Kyriakides, and C. M. Landis, “Propagation of phase transformation fronts in pseudoelastic NiTi tubes under uniaxial tension,” *Extreme Mechanics Letters*, vol. 15, pp. 113–121, 2017.

[99] D. Lagoudas, D. Hartl, Y. Chemisky, L. Machado, and P. Popov, “Constitutive model for the numerical analysis of phase transformation in polycrystalline shape memory alloys,” *International Journal of Plasticity*, vol. 32-33, pp. 155–183, 2012.

[100] A. Saleeb, S. Padula Ii, and A. Kumar, “A multi-axial, multimechanism based constitutive model for the comprehensive representation of the evolutionary response of SMAs under general thermomechanical loading conditions,” *International Journal of Plasticity*, vol. 27, no. 5, pp. 655–687, 2011.

[101] M. Bodaghi, A. Damanpack, M. Aghdam, and M. Shakeri, “A robust three-dimensional phenomenological model for polycrystalline SMAs: Analytical closed-form solutions,” *International Journal of Engineering Science*, vol. 82, pp. 1–21, 2014.

[102] Z. Song and H.-H. Dai, “Closed-form solutions for inhomogeneous states of a slender 3-D SMA cylinder undergoing stress-induced phase transitions,” *International Journal of Engineering Science*, vol. 88, pp. 40–63, 2015.

[103] P. Zhu, P. Feng, Q.-P. Sun, J. Wang, and H.-H. Dai, “Determining the up-down-up response through tension tests of a pre-twisted shape memory alloy tube,” *International Journal of Plasticity*, vol. 85, pp. 52–76, 2016.

[104] R. Mirzaeifar, R. DesRoches, A. Yavari, and K. Gall, “A micromechanical analysis of the coupled thermomechanical superelastic response of textured and untextured polycrystalline NiTi shape memory alloys,” *Acta Materialia*, vol. 61, no. 12, pp. 4542–4558, 2013.

[105] C. Yu, G. Kang, Q. Kan, and D. Song, “A micromechanical constitutive model based on crystal plasticity for thermo-mechanical cyclic deformation of NiTi shape memory alloys,” *International Journal of Plasticity*, vol. 44, pp. 161–191, 2013.

[106] C. Yu, G. Kang, and Q. Kan, “A micromechanical constitutive model for anisotropic cyclic deformation of super-elastic NiTi shape memory alloy single crystals,” *Journal of the Mechanics and Physics of Solids*, vol. 82, pp. 97–136, 2015.

[107] R. Heinen and S. Miro, “Micromechanical modeling of NiTi shape memory alloys including austenite, r-phase, and martensite,” *Computer Methods in Applied Mechanics and Engineering*, vol. 229, pp. 44–55, 2012.

[108] D. C. Lagoudas, P. B. Entchev, P. Popov, E. Patoor, L. C. Brinson, and X. Gao, “Shape memory alloys, part II: Modeling of polycrystals,” *Mechanics of Materials*, vol. 38, no. 5-6, pp. 430–462, 2006.

[109] S. Oehler, D. Hartl, R. Lopez, R. Malak, and D. Lagoudas, “Design optimization and uncertainty analysis of SMA morphing structures,” *Smart Materials and Structures*, vol. 21, no. 9, p. 094016, 2012.

[110] A. Martowicz, J. Bryła, and T. Uhl, “Uncertainty quantification for the properties of a structure made of SMA utilising numerical model,” in *Proceedings of the Conference on Noise*

and Vibration Engineering ISMA 2016 & 5th Edition of the International Conference on Uncertainly in Structural Dynamics USD, 2016.

[111] S. Enemark, I. F. Santos, and M. A. Savi, “Modelling, characterisation and uncertainties of stabilised pseudoelastic shape memory alloy helical springs,” *Journal of Intelligent Material Systems and Structures*, vol. 27, no. 20, pp. 2721–2743, 2016.

[112] M. Hill, “Methods and guidelines for effective model calibration. USGS water-resources investigations report 98-4005,” *US Geological Survey: Reston, VA, USA*, 1998.

[113] T. Larssen, R. B. Huseby, B. J. Cosby, G. Høst, T. Høgåsen, and M. Aldrin, “Forecasting acidification effects using a bayesian calibration and uncertainty propagation approach,” *Environmental Science & Technology*, vol. 40, no. 24, pp. 7841–7847, 2006.

[114] Y. Kim, K. Kim, C. Park, and I. Kim, “Deterministic vs. stochastic calibration of energy simulation model for an existing building,” in *The 2nd Asia conference of International Building Performance Simulation Association*, pp. 28–29, 2014.

[115] R. Ghanem and J. Red-Horse, “Propagation of probabilistic uncertainty in complex physical systems using a stochastic finite element approach,” *Physica D: Nonlinear Phenomena*, vol. 133, no. 1-4, pp. 137–144, 1999.

[116] P. Popov and D. C. Lagoudas, “A 3-D constitutive model for shape memory alloys incorporating pseudoelasticity and detwinning of self-accommodated martensite,” *International Journal of Plasticity*, vol. 23, no. 10-11, pp. 1679–1720, 2007.

[117] Y. Chemisky, A. Duval, E. Patoor, and T. B. Zineb, “Constitutive model for shape memory alloys including phase transformation, martensitic reorientation and twins accommodation,” *Mechanics of Materials*, vol. 43, no. 7, pp. 361–376, 2011.

[118] A. Paglietti, “The mathematical formulation of the local form of the second principle of thermodynamics,” in *Annales de l’IHP Physique théorique*, vol. 27, pp. 207–219, 1977.

[119] B. D. Coleman and W. Noll, "The thermodynamics of elastic materials with heat conduction and viscosity," *Archive for Rational Mechanics and Analysis*, vol. 13, no. 1, pp. 167–178, 1963.

[120] M. A. Tschopp and E. Hernandez-Rivera, "Quantifying similarity and distance measures for vector-based datasets: histograms, signals, and probability distribution functions," Tech. Rep. ARL-TN-0810, US Army Research Laboratory Aberdeen Proving Ground United States, US Army Research Laboratory Aberdeen Proving Ground United States, 2017.

[121] L. J. Billera and P. Diaconis, "A geometric interpretation of the Metropolis-Hastings algorithm," *Statistical Science*, pp. 335–339, 2001.

[122] A. Cox, B. Franco, S. Wang, T. Baxevanis, I. Karaman, and D. Lagoudas, "Predictive modeling of the constitutive response of precipitation hardened Ni-rich NiTi," *Shape Memory and Superelasticity*, vol. 3, no. 1, pp. 9–23, 2017.

[123] H. Haario, M. Laine, A. Mira, and E. Saksman, "Dram: efficient adaptive MCMC," *Statistics and Computing*, vol. 16, no. 4, pp. 339–354, 2006.

[124] M. Putko, A. Taylor III, P. Newman, and L. Green, "Approach for uncertainty propagation and robust design in CFD using sensitivity derivatives," in *15th AIAA Computational Fluid Dynamics Conference*, p. 2528, 2001.

[125] B. Kriegesmann, *Probabilistic design of thin-walled fiber composite structures*. PhD thesis, Technische Informationsbibliothek und Universitätsbibliothek Hannover (TIB), 2012.

[126] N. R. Council and Others, *Integrated computational materials engineering: A transformational discipline for improved competitiveness and national security*. National Academies Press, 2008.

[127] S. Ghoreishi and D. Allaire, "Adaptive uncertainty propagation for coupled multidisciplinary systems," *AIAA Journal*, pp. 3940–3950, 2017.

[128] S. F. Ghoreishi, *Uncertainty analysis for coupled multidisciplinary systems using sequential importance resampling*. PhD thesis, Texas A and M University, 2016.

[129] S. F. Ghoreishi and D. L. Allaire, "Compositional uncertainty analysis via importance weighted gibbs sampling for coupled multidisciplinary systems," in *18th AIAA Non-Deterministic Approaches Conference*, p. 1443, 2016.

[130] J. J. Mortensen, K. Kaasbjerg, S. L. Frederiksen, J. K. Nørskov, J. P. Sethna, and K. W. Jacobsen, "Bayesian error estimation in density-functional theory," *Physical Review Letters*, vol. 95, no. 21, p. 216401, 2005.

[131] F. Hanke, "Sensitivity analysis and uncertainty calculation for dispersion corrected density functional theory," *Journal of Computational Chemistry*, vol. 32, no. 7, pp. 1424–1430, 2011.

[132] M. Aldegunde, N. Zabaras, and J. Kristensen, "Quantifying uncertainties in first-principles alloy thermodynamics using cluster expansions," *Journal of Computational Physics*, vol. 323, pp. 17–44, 2016.

[133] F. Rizzi, M. Salloum, Y. Marzouk, R.-G. Xu, M. Falk, T. Weihs, G. Fritz, and O. M. Knio, "Bayesian inference of atomic diffusivity in a binary Ni/Al system based on molecular dynamics," *Multiscale Modeling & Simulation*, vol. 9, no. 1, pp. 486–512, 2011.

[134] B. Völker, P. Marton, C. Elsässer, and M. Kamlah, "Multi-scale modeling for ferroelectric materials: A transition from the atomic level to phase-field modeling," *Continuum Mechanics and Thermodynamics*, vol. 23, no. 5, pp. 435–451, 2011.

[135] W. K. Liu, L. Siad, R. Tian, S. Lee, D. Lee, X. Yin, W. Chen, S. Chan, G. B. Olson, L.-E. Lindgen, *et al.*, "Complexity science of multiscale materials via stochastic computations," *International Journal for Numerical Methods in Engineering*, vol. 80, no. 6-7, pp. 932–978, 2009.

[136] B. Kouchmeshky and N. Zabaras, "Microstructure model reduction and uncertainty quantification in multiscale deformation processes," *Computational Materials Science*, vol. 48, no. 2, pp. 213–227, 2010.

[137] M. Koslowski and A. Strachan, “Uncertainty propagation in a multi-scale model of nanocrystalline plasticity,” *Reliability Engineering & System Safety*, vol. 96, no. 9, pp. 1161–1170, 2011.

[138] S. Salehghaffari, M. Rais-Rohani, E. Marin, and D. Bammann, “A new approach for determination of material constants of internal state variable based plasticity models and their uncertainty quantification,” *Computational Materials Science*, vol. 55, pp. 237–244, 2012.

[139] E. Königsberger, “Improvement of excess parameters from thermodynamic and phase diagram data by a sequential bayes algorithm,” *Calphad*, vol. 15, no. 1, pp. 69–78, 1991.

[140] N. D. Chatterjee, K. Miller, and W. Olbricht, “Bayes estimation: A novel approach to derivation of internally consistent thermodynamic data for minerals, their uncertainties, and correlations. part II: Application,” *Physics and Chemistry of Minerals*, vol. 21, no. 1-2, pp. 50–62, 1994.

[141] D. V. Malakhov, “Confidence intervals of calculated phase boundaries,” *Calphad*, vol. 21, no. 3, pp. 391–400, 1997.

[142] N. D. Chatterjee, R. Krüger, G. Haller, and W. Olbricht, “The bayesian approach to an internally consistent thermodynamic database: Theory, database, and generation of phase diagrams,” *Contributions to Mineralogy and Petrology*, vol. 133, no. 1-2, pp. 149–168, 1998.

[143] J. Acker and K. Bohmhammel, “Optimization of thermodynamic data of the Ni–Si system,” *Thermochimica Acta*, vol. 337, no. 1-2, pp. 187–193, 1999.

[144] M. Stan and B. Reardon, “A bayesian approach to evaluating the uncertainty of thermodynamic data and phase diagrams,” *Calphad*, vol. 27, no. 3, pp. 319–323, 2003.

[145] R. A. Otis and Z.-K. Liu, “High-throughput thermodynamic modeling and uncertainty quantification for icme,” *JOM*, vol. 69, no. 5, pp. 886–892, 2017.

[146] T. C. Duong, A. Talapatra, W. Son, M. Radovic, and R. Arróyave, “On the stochastic phase stability of Ti 2 AlC-Cr 2 AlC,” *Scientific Reports*, vol. 7, no. 1, p. 5138, 2017.

[147] J.-C. Zhao, B. Bewlay, M. Jackson, and Q. Chen, “Hf-Si binary phase diagram determination and thermodynamic modeling,” *Journal of Phase Equilibria*, vol. 21, no. 1, pp. 40–45, 2000.

[148] Y. Yang, Y. Chang, J.-C. Zhao, and B. Bewlay, “Thermodynamic modeling of the Nb–Hf–Si ternary system,” *Intermetallics*, vol. 11, no. 5, pp. 407–415, 2003.

[149] J.-C. Zhao, B. Bewlay, and M. Jackson, “Determination of Nb–Hf–Si phase equilibria,” *Intermetallics*, vol. 9, no. 8, pp. 681–689, 2001.

[150] M. Hillert, “The compound energy formalism,” *Journal of Alloys and Compounds*, vol. 320, no. 2, pp. 161–176, 2001.

[151] A. Dinsdale, “SGTE data for pure elements,” *Calphad*, vol. 15, no. 4, pp. 317–425, 1991.

[152] R. E. Kass and A. E. Raftery, “Bayes factors,” *Journal of the American Statistical Association*, vol. 90, no. 430, pp. 773–795, 1995.

[153] J. A. Hoeting, D. Madigan, A. E. Raftery, and C. T. Volinsky, “Bayesian model averaging: A tutorial,” *Statistical Science*, vol. 14, no. 4, pp. 382–401, 1999.

[154] T. M. Fragoso, W. Bertoli, and F. Louzada, “Bayesian model averaging: A systematic review and conceptual classification,” *International Statistical Review*, vol. 86, no. 1, pp. 1–28, 2018.

[155] S. Geisser, “A bayes approach for combining correlated estimates,” *Journal of the American Statistical Association*, vol. 60, no. 310, pp. 602–607, 1965.

[156] P. A. Morris, “Combining expert judgments: A bayesian approach,” *Management Science*, vol. 23, no. 7, pp. 679–693, 1977.

[157] R. L. Winkler, “Combining probability distributions from dependent information sources,” *Management Science*, vol. 27, no. 4, pp. 479–488, 1981.

[158] S. F. Ghoreishi and D. L. Allaire, "A fusion-based multi-information source optimization approach using knowledge gradient policies," in *2018 AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, p. 1159, 2018.

[159] S. J. Julier and J. K. Uhlmann, "A non-divergent estimation algorithm in the presence of unknown correlations," in *Proceedings of the 1997 American Control Conference (Cat. No. 97CH36041)*, vol. 4, pp. 2369–2373, IEEE, 1997.

[160] J. F. Wellmann and K. Regenauer-Lieb, "Uncertainties have a meaning: Information entropy as a quality measure for 3-d geological models," *Tectonophysics*, vol. 526, pp. 207–216, 2012.

[161] D. L. McDowell and S. R. Kalidindi, "The materials innovation ecosystem: A key enabler for the materials genome initiative," *Mrs Bulletin*, vol. 41, no. 4, pp. 326–337, 2016.

[162] J. Allison, "Integrated computational materials engineering: A perspective on progress and future steps," *Jom*, vol. 63, no. 4, p. 15, 2011.

[163] X. Liu, D. Furrer, J. Kosters, and J. Holmes, "Vision 2040: A roadmap for integrated, multiscale modeling and simulation of materials and systems," 2018.

[164] W. Chen, L. Baghdasaryan, T. Buranathiti, and J. Cao, "Model validation via uncertainty propagation and data transformations," *AIAA Journal*, vol. 42, no. 7, pp. 1406–1415, 2004.

[165] F. Rizzi, R. Jones, B. Debusschere, and O. Knio, "Uncertainty quantification in md simulations of concentration driven ionic flow through a silica nanopore. i. sensitivity to physical parameters of the pore," *The Journal of Chemical Physics*, vol. 138, no. 19, p. 194104, 2013.

[166] R. E. Jones, F. Rizzi, B. Boyce, J. A. Templeton, and J. Ostien, "Plasticity models of material variability based on uncertainty quantification techniques," tech. rep., Sandia National Lab.(SNL-NM), Albuquerque, NM (United States), 2017.

[167] G. B. Olson and C. Kuehmann, "Materials genomics: From CALPHAD to flight," *Scripta Materialia*, vol. 70, pp. 25–30, 2014.

[168] N. Saunders and A. P. Miodownik, *CALPHAD (calculation of phase diagrams): A comprehensive guide*, vol. 1. Elsevier, 1998.

[169] P. E. Turchi, I. A. Abrikosov, B. Burton, S. G. Fries, G. Grimvall, L. Kaufman, P. Korzhavyi, V. R. Manga, M. Ohno, A. Pisch, *et al.*, "Interface between quantum-mechanical-based approaches, experiments, and calphad methodology," *Calphad*, vol. 31, no. 1, pp. 4–27, 2007.

[170] S. Bigdeli, *Developing the third generation of Calphad databases: What can ab-initio contribute?* PhD thesis, KTH, Computational Thermodynamics, 2017. QC 20171006.

[171] W. Cao, S.-L. Chen, F. Zhang, K. Wu, Y. Yang, Y. Chang, R. Schmid-Fetzer, and W. Oates, "PANDAT software with PanEngine, PanOptimizer and PanPrecipitation for multi-component phase diagram calculation and materials property simulation," *Calphad*, vol. 33, no. 2, pp. 328–342, 2009.

[172] G. Kim, S. V. Meschel, P. Nash, and W. Chen, "Experimental formation enthalpies for intermetallic phases and other inorganic compounds," *Scientific Data*, vol. 4, p. 170162, Oct. 2017.

[173] M. Palumbo, T. Abe, S. G. Fries, and A. Pasturel, "First-principles approach to phase stability for a ternary σ phase: Application to Cr-Ni-Re," *Physical Review B*, vol. 83, p. 144109, Apr 2011.

[174] A. Halimi and M. S. Ferah, "Thermodynamic description of systems Cd-Te, Hg-Te and Cd-Hg-Te using the model of associated liquid solution," *International Journal of Microstructure and Materials Properties*, vol. 3, pp. 77–85, 2008.

[175] O. Grånäs, P. A. Korzhavyi, A. Kissavos, and I. A. Abrikosov, "Theoretical study of the Mo–Ru sigma phase," *Calphad*, vol. 32, no. 1, pp. 171–176, 2008.

[176] T. Uesugi, S. Miyamae, and K. Higashi, "Enthalpies of solution in Ti–X (X = Mo, Nb, V and W) alloys from first-principles calculations," *Materials Transactions*, vol. 54, no. 4, pp. 484–492, 2013.

[177] A. D. Pelton, "The polynomial representation of thermodynamic properties in dilute solutions," *Metall and Materi Trans B*, vol. 28, pp. 869–876, Oct. 1997.

[178] A. Abu-Odeh, E. Galvan, T. Kirk, H. Mao, Q. Chen, P. Mason, R. Malak, and R. ArrÃşyave, "Efficient exploration of the high entropy alloy composition-phase space," *Acta Materialia*, vol. 152, pp. 41–57, June 2018.

[179] S. Hu, "Akaike information criterion," *Center for Research in Scientific Computation*, vol. 93, 2007.

[180] P. Honarmandi, T. C. Duong, S. F. Ghoreishi, D. Allaire, and R. Arroyave, "Bayesian uncertainty quantification and information fusion in CALPHAD-based thermodynamic modeling," *Acta Materialia*, vol. 164, pp. 636–647, Feb. 2019.

[181] N. H. Paulson, E. Jennings, and M. Stan, "Bayesian strategies for uncertainty quantification of the thermodynamic properties of materials," *ArXiv Preprint ArXiv: 1809.07365*, 2018.

[182] A. D. Kiureghian and O. Ditlevsen, "Aleatory or epistemic? Does it matter?," *Structural Safety*, vol. 31, pp. 105–112, Mar. 2009.

[183] J. Arblaster, "Thermodynamic properties of hafnium," *Journal of Phase Equilibria and Diffusion*, vol. 35, no. 4, pp. 490–501, 2014.

[184] W. Chen, R. Jin, and A. Sudjianto, "Analytical uncertainty propagation via meta-models in simulation-based design under uncertainty," in *10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, p. 4356, 2004.

[185] R. Schöbi, P. Kersaudy, B. Sudret, and J. Wiart, *Combining polynomial chaos expansions and kriging*. PhD thesis, ETH Zurich, Switzerland; Orange Labs Research, 2014.

[186] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, "emcee: The MCMC hammer," *Publications of the Astronomical Society of the Pacific*, vol. 125, no. 925, pp. 306–312, 2013.

[187] P. C. Austin, C. D. Naylor, and J. V. Tu, “A comparison of a Bayesian vs. a frequentist method for profiling hospital performance,” *Journal of Evaluation in Clinical Practice*, vol. 7, no. 1, pp. 35–45, 2001.

[188] S.-K. Au, “Connecting bayesian and frequentist quantification of parameter uncertainty in system identification,” *Mechanical Systems and Signal Processing*, vol. 29, pp. 328–342, 2012.

[189] G. E. Box and G. C. Tiao, *Bayesian inference in statistical analysis*, vol. 40. John Wiley & Sons, 2011.

[190] K. Beven and A. Binley, “The future of distributed models: Model calibration and uncertainty prediction,” *Hydrological Processes*, vol. 6, no. 3, pp. 279–298, 1992.

[191] G. Schoups and J. A. Vrugt, “A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-gaussian errors,” *Water Resources Research*, vol. 46, no. 10, 2010.

[192] F. Feroz, M. Hobson, E. Cameron, and A. Pettitt, “Importance nested sampling and the MultiNest algorithm,” *ArXiv Preprint ArXiv:1306.2144*, 2013.

[193] B. Farr and W. M. Farr, “Kombine: A kernel-density-based, embarrassingly parallel ensemble sampler,” 2015. In Prep.

[194] A. Davydov and U. R. Kattner, “Thermodynamic assessment of the Co-Mo system,” *Journal of Phase Equilibria*, vol. 20, no. 1, pp. 5–16, 1999.

[195] R. N. Goldberg and R. Weir, “Conversion of temperatures and thermodynamic properties to the basis of the international temperature scale of 1990 (technical report),” *Pure and Applied Chemistry*, vol. 64, no. 10, pp. 1545–1562, 1992.

[196] K. Lejaeghere, G. Bihlmayer, T. Björkman, P. Blaha, S. Blügel, V. Blum, D. Caliste, I. E. Castelli, S. J. Clark, A. Dal Corso, *et al.*, “Reproducibility in density functional theory calculations of solids,” *Science*, vol. 351, no. 6280, p. aad3000, 2016.

[197] H. Lukas, E. T. Henig, and B. Zimmermann, “Optimization of phase diagrams by a least squares method using simultaneously different types of data,” *Calphad*, vol. 1, no. 3, pp. 225–236, 1977.

[198] Y.-Z. Ma and A. Berndsen, “How to combine correlated data setsâĂŤa bayesian hyperparameter matrix method,” *Astronomy and Computing*, vol. 5, pp. 45–56, 2014.

[199] H. W. Keller, J. M. Shallenberger, D. A. Hollein, and A. C. Hott, “Development of hafnium and comparison with other pressurized water reactor control rod materials,” *Nuclear Technology*, vol. 59, no. 3, pp. 476–482, 1982.

[200] Y.-l. Wang, X. Xiong, X.-j. Zhao, G.-d. Li, Z.-k. Chen, and W. Sun, “Structural evolution and ablation mechanism of a hafnium carbide coating on a C/C composite in an oxyacetylene torch environment,” *Corrosion Science*, vol. 61, pp. 156–161, 2012.

[201] J. Hou, J. Guo, Y. Wu, L. Zhou, and H. Ye, “Effect of hafnium on creep behavior of a corrosion resistant nickel base superalloy,” *Materials Science and Engineering: A*, vol. 527, no. 6, pp. 1548–1554, 2010.

[202] R. Hegde, D. Triyoso, S. Samavedam, and B. White Jr, “Hafnium zirconate gate dielectric for advanced gate stack applications,” *Journal of Applied Physics*, vol. 101, no. 7, p. 074113, 2007.

[203] J. Arblaster, “Thermodynamic properties of zirconium,” *Calphad*, vol. 43, pp. 32–39, 2013.

[204] I. BIPM, I. IFCC, and I. ISO, “IUPAP, and OIML, 2008,âĂIJevaluation of measurement dataâĂŤguide to the expression of uncertainty in measurement,âĂİ joint committee for guides in metrology,” tech. rep., Technical Report No. JCGM 100, 2008.

[205] S. Ringberg, “Thermodynamic models and data for pure elements and other end-members of solutions,” *Calphad*, vol. 19, no. 4, pp. 437–447, 1995.

[206] G. Grimvall, *Thermophysical properties of materials*. Elsevier, 1999.

[207] I. Roslyakova, B. Sundman, H. Dette, L. Zhang, and I. Steinbach, “Modeling of gibbs energies of pure elements down to 0 K using segmented regression,” *Calphad*, vol. 55, pp. 165–180, 2016.

[208] V. N. Korobenko, O. Polyakova, and A. I. Savvatimskii, “Heat capacity of liquid hafnium from the melting point to the boiling point at atmospheric pressure,” *High Temperature*, vol. 43, no. 1, pp. 38–44, 2005.

[209] C. Cagran, T. Hüpf, B. Wilthan, and G. Pottlacher, “Selected thermophysical properties of Hf-3% Zr from 2200 K to 3500 K obtained by a fast pulse-heating technique.,” *High Temperatures–High Pressures*, vol. 37, no. 3, 2008.

[210] M. Rösner-Kuhn, K. Drewes, H. Franz, and M. G. Frohberg, “Enthalpy measurements of the solid high-temperature β-phase of titanium and zirconium by levitation drop calorimetry,” *Journal of Alloys and Compounds*, vol. 316, no. 1-2, pp. 175–178, 2001.

[211] P.-F. Paradis, T. Ishikawa, and S. Yoda, “Non-contact measurements of the thermophysical properties of hafnium-3 mass% zirconium at high temperature,” *International Journal of Thermophysics*, vol. 24, no. 1, pp. 239–258, 2003.

[212] A. Roine, “Outokumpu HSC chemistry for windows: Chemical reaction and equilibrium software with extensive thermochemical database,” *Pori: Outokumpu Research OY*, 2002.

[213] S.-L. Shang, Y. Wang, and Z.-K. Liu, “ESPEI: Extensible, self-optimizing phase equilibrium infrastructure for magnesium alloys,” 2010.

[214] Z.-K. Liu, “First-principles calculations and CALPHAD modeling of thermodynamics,” *Journal of Phase Equilibria and Diffusion*, vol. 30, no. 5, pp. 517–534, 2009.

[215] G. Cacciamani, “An introduction to the CALPHAD method and the compound energy formalism (CEF),” *Tecnologia em Metalurgia, Materiais e MineracÌğaÌČo*, vol. 13, no. 1, pp. 16–24, 2016.

[216] R. Mathieu, N. Dupin, J.-C. Crivello, K. Yaqoob, A. Breidi, J.-M. Fiorani, N. David, and J.-M. Joubert, “CALPHAD description of the Mo–Re system focused on the sigma phase modeling,” *Calphad*, vol. 43, pp. 18–31, 2013.

[217] A. Berche, J.-C. Tédenac, and P. Jund, “Ab-initio calculations and CALPHAD description of Cr–Ge–Mn and Cr–Ge–Si,” *Calphad*, vol. 49, pp. 50–57, 2015.

[218] C. Zhang and M. C. Gao, “CALPHAD modeling of high-entropy alloys,” in *High-Entropy Alloys*, pp. 399–444, Springer, 2016.

[219] M. Schick, B. Hallstedt, A. Glensk, B. Grabowski, T. Hickel, M. Hampl, J. Gröbner, J. Neugebauer, and R. Schmid-Fetzer, “Combined ab initio, experimental, and CALPHAD approach for an improved thermodynamic evaluation of the Mg–Si system,” *Calphad*, vol. 37, pp. 77–86, 2012.

[220] R. Arroyave, D. Shin, and Z. Liu, “Modification of the thermodynamic model for the Mg–Zr system,” *Calphad*, vol. 29, no. 3, pp. 230–238, 2005.

[221] A. E. Kissavos, S. Shallcross, V. Meded, L. Kaufman, and I. A. Abrikosov, “A critical test of ab initio and CALPHAD methods: The structural energy difference between bcc and hcp molybdenum,” *Calphad*, vol. 29, no. 1, pp. 17–23, 2005.

[222] S. Bajaj, A. Landa, P. Söderlind, P. E. Turchi, and R. Arróyave, “The U–Ti system: Strengths and weaknesses of the CALPHAD method,” *Journal of Nuclear Materials*, vol. 419, no. 1-3, pp. 177–185, 2011.

[223] M. Sanghvi, P. Honarmandi, V. Attari, T. Duong, R. Arroyave, and D. L. Allaire, “Uncertainty propagation via probability measure optimized importance weights with application to parametric materials models,” in *AIAA Scitech 2019 Forum*, p. 0967, 2019.

[224] S.-i. Yi, V. Attari, M. Jeong, J. Jian, S. Xue, H. Wang, R. Arroyave, and C. Yu, “Strain-induced suppression of the miscibility gap in nano-structured Mg 2 Si–Mg 2 Sn solid solutions,” *Journal of Materials Chemistry A*, vol. 6, no. 36, pp. 17559–17570, 2018.

[225] D. C. Montgomery, *Design and analysis of experiments*. John Wiley & Sons, 2017.

APPENDIX A

TWO-WAY ANOVA

In the case of two-way ANOVA where there are L observations and two multi-level factors A and B, system response can be defined as a Means model as follows [225],

$$y_{ijl} = \mu_{ij} + \varepsilon_{ijl} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijl} \ , \ \begin{cases} i = 1, ..., I \\ j = 1, ..., J \\ l = 1, ..., L \end{cases} \tag{A.1}$$

where, I and J are the number of levels associated with factors A and B, and τ_i, β_j, and $(\tau\beta)_{ij}$ denote the effects of the level i of factor A, level j of factor B, and their interaction, respectively. ε is the residual or the random error which is considered as an independent normally distributed function with a fixed variance, $\mathcal{N}(0, \sigma^2)$. In this context, the following hypotheses are statistically evaluated to find the most sensitive factors and interactions in the system [225],

$$\begin{aligned} H_0 &: \ \tau_1 = \tau_2 = ... = \tau_I = 0 \\ H_1 &: \ \text{at least one } \tau_i \neq 0 \end{aligned} \tag{A.2}$$

$$\begin{aligned} H_0 &: \ \beta_1 = \beta_2 = ... = \beta_J = 0 \\ H_1 &: \ \text{at least one } \beta_j \neq 0 \end{aligned} \tag{A.3}$$

$$\begin{aligned} H_0 &: \ (\tau\beta)_{ij} = 0 \ \text{ for all } i \text{ and } j \\ H_1 &: \ \text{at least one } (\tau\beta)_{ij} \neq 0 \end{aligned} \tag{A.4}$$

In each case, H_0 and H_1 are null and alternative hypothesis, respectively. It should be noted that a component (either a factor or an interaction) is identified sensitive when the corresponding null hypothesis is rejected after hypothesis testing. In this regard, a variance partitioning into the components of the system is required; however, this partitioning can be performed on the total sum of squares instead due to the proportionality between total variance and total sum of squares.

Let $y_{i..}$, $y_{.j.}$, $y_{ij.}$, and $y_{...}$ denote the sum of all the factorial design responses which includes the effects of the level i of the factor A, the level j of the factor B, the combination of the level i for A and the level j for B, and all the level-factor combinations, respectively. Moreover, let $\bar{y}_{i..}$, $\bar{y}_{.j.}$, $\bar{y}_{ij.}$, and $\bar{y}_{...}$ denote the corresponding averages as follows [225],

$$y_{i..} = \sum_{j=1}^{J}\sum_{l=1}^{L} y_{ijl} \ \ , \ \ \bar{y}_{i..} = \frac{y_{i..}}{JL} \tag{A.5}$$

$$y_{.j.} = \sum_{i=1}^{I}\sum_{l=1}^{L} y_{ijl} \ \ , \ \ \bar{y}_{.j.} = \frac{y_{.j.}}{IL} \tag{A.6}$$

$$y_{ij.} = \sum_{l=1}^{L} y_{ijl} \ \ , \ \ \bar{y}_{ij.} = \frac{y_{ij.}}{L} \tag{A.7}$$

$$y_{...} = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{l=1}^{L} y_{ijl} \ \ , \ \ \bar{y}_{...} = \frac{y_{...}}{IJL} \tag{A.8}$$

Total sum of squares (SS_T) can be decomposed to the sum of squares of the system components, as follows,

$$\begin{aligned}
SS_T = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{l=1}^{L}(y_{ijl} - \bar{y}_{...})^2 &= \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{l=1}^{L} \Big[(\bar{y}_{i..} - \bar{y}_{...}) + (\bar{y}_{.j.} - \bar{y}_{...}) \\
&\quad + (\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...}) + (y_{ijk} - \bar{y}_{ij.})\Big]^2 \\
&= JL\sum_{i=1}^{I}(\bar{y}_{i..} - \bar{y}_{...})^2 + IL\sum_{j=1}^{J}(\bar{y}_{.j.} - \bar{y}_{...})^2 \\
&\quad + L\sum_{i=1}^{I}\sum_{j=1}^{J}(\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2 \\
&\quad + \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{l=1}^{L}(y_{ijl} - \bar{y}_{ij.})^2 \\
&= SS_A + SS_B + SS_{AB} + SS_E
\end{aligned} \tag{A.9}$$

where, SS_A, SS_B, SS_{AB}, and SS_E are the sum of squares associated with A, B, A-B interaction, and the random error, respectively. It is worth noting that SS_E can be obtained by the subtraction of SS_A, SS_B, and SS_{AB} from SS_T. In Equation A.9, the first equality results from the addition and subtraction of the same terms on the right side, and the second equality is a consequence of the polynomial expansion of the function where the six cross products becomes zero. If these sum of squares are divided by their corresponding degrees of freedom, the mean squares (MS) are obtained. The expected values of these mean squares when the levels are random variables can provide relationships to estimate the total deviations of all the levels associated with A, B, and A-B interaction, as follows,

$$E(MS_A) = E\Big(\frac{SS_A}{I-1}\Big) = \sigma^2 + \frac{JL\sum_{i=1}^{I}\tau_i^2}{I-1} \tag{A.10}$$

$$E(MS_B) = E\Big(\frac{SS_B}{J-1}\Big) = \sigma^2 + \frac{IL\sum_{j=1}^{J}\beta_j^2}{J-1} \tag{A.11}$$

$$E(MS_{AB}) = E\Big(\frac{SS_{AB}}{(I-1)(J-1)}\Big) = \sigma^2 + \frac{L\sum_{i=1}^{I}\sum_{j=1}^{J}(\tau\beta)_{ij}^2}{(I-1)(J-1)} \tag{A.12}$$

$$E(MS_E) = E\Big(\frac{SS_E}{IJ(L-1)}\Big) = \sigma^2 \tag{A.13}$$

It should be noted that the mean square of each factor or interaction can be used in the above equations in the problems that the levels are fixed. In the case that the null hypothesis about the insensitivity of a factor or an interaction is true, the total corresponding effects of the component becomes zero; therefore, the corresponding expected value in Equations A.10, A.11 or/and A.12 equals the expected value for the mean squares of the random error, i.e., its variance (σ^2). On the other hand, if there are any effects from the system components, the corresponding expected values are greater than σ^2. It is clear that bigger ratio of the expected value of a component to σ^2 is equivalent to more sensitivity of the component. This is also true for the ratio of the corresponding mean squares, which is called F-ratio. However, a certain criterion is required to fully reject the null hypotheses for the identification of the sensitive components in the system. For this reason, a F-distribution is created from anyone of the null hypotheses based on the numerator and denominator degrees of freedom of F-ratio to calculate the corresponding p-value. Generally, the area underneath the F-distribution confined between the F-ratio value and infinity is introduced as the p-value. Different significance levels can be set as the criteria to reject the null hypotheses, i.e., 0.01, 0.05, or 0.1. For any system component, a p-value less than the significance level rejects the corresponding null hypothesis, and demonstrates the sensitivity of the given component.

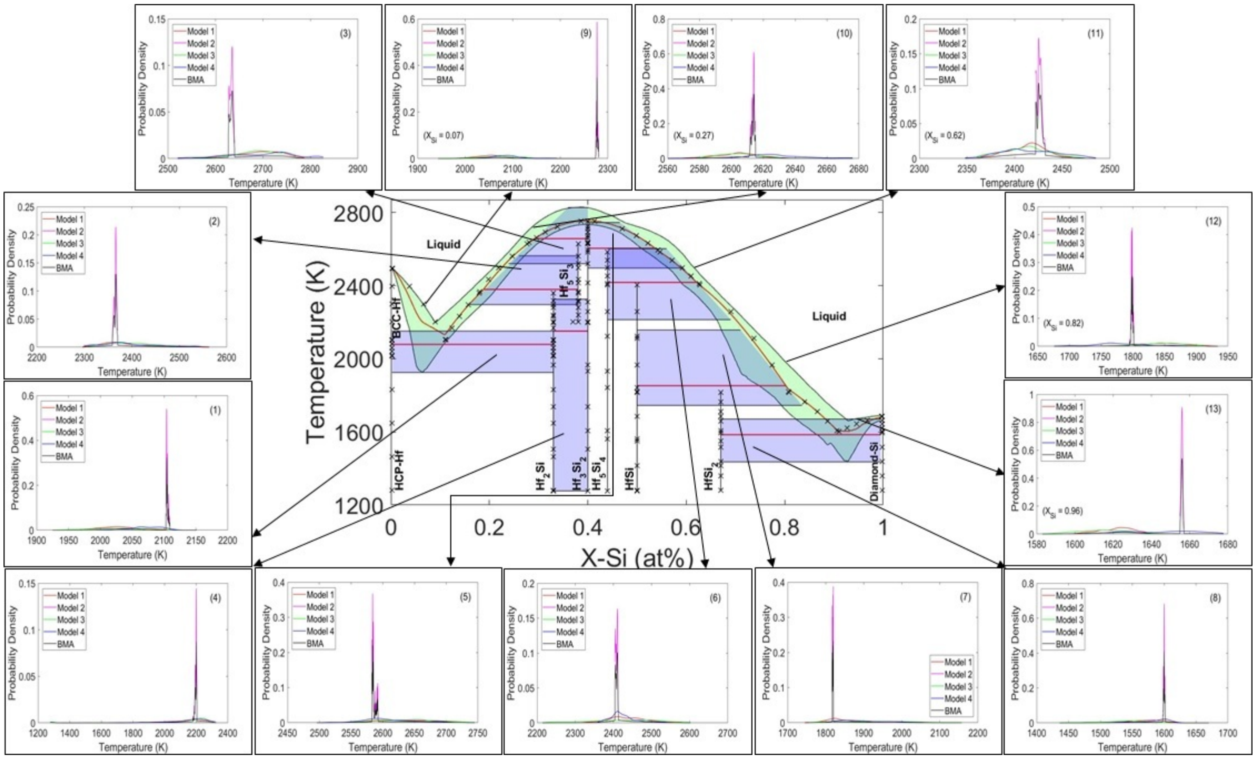
In the case that there is just one observation/response for each level-factor combinations (generally for the models and simulations where their response is deterministic such as our thermo-mechanical model), the variance of the random error (σ^2) can not be determined by the expected mean squares of the random error in Equation A.13 since L equals 1. Under this condition, no clear differentiation can be established between the effect of the interaction and the random error in Equation A.12. Therefore, the interaction effect in Equation A.1 must be considered zero in order to be able to estimate σ^2 using Equation A.12; otherwise, the effect of each individual factor can not be estimated through Equations A.10 and A.11. ANOVA table for this case can be observed in Table A.1 [225].

Table A.1: Two-way ANOVA table for the case of one observation/response per level-factor combination [225].

Source of Variation	Sum of Squares	Degrees of Freedoms	Mean Square	Expected Mean Square	F-Ratio
A	$J\sum_{i=1}^{I}(\bar{y}_{i.} - \bar{y}_{..})^2$	$I-1$	MS_A	$\sigma^2 + \frac{J\sum\tau_i^2}{I-1}$	$\frac{MS_A}{MS_E}$
B	$I\sum_{j=1}^{J}(\bar{y}_{.j} - \bar{y}_{..})^2$	$J-1$	MS_B	$\sigma^2 + \frac{I\sum\beta_j^2}{J-1}$	$\frac{MS_B}{MS_E}$
Residual or AB	Subtraction	$(I-1)(J-1)$	MS_E	$\sigma^2 + \frac{\sum\sum(\tau\beta)_{ij}^2}{(I-1)(J-1)}$	
Total	$\sum_{i=1}^{I}\sum_{j=1}^{J}(y_{ij} - \bar{y}_{..})^2$	$IJ-1$			

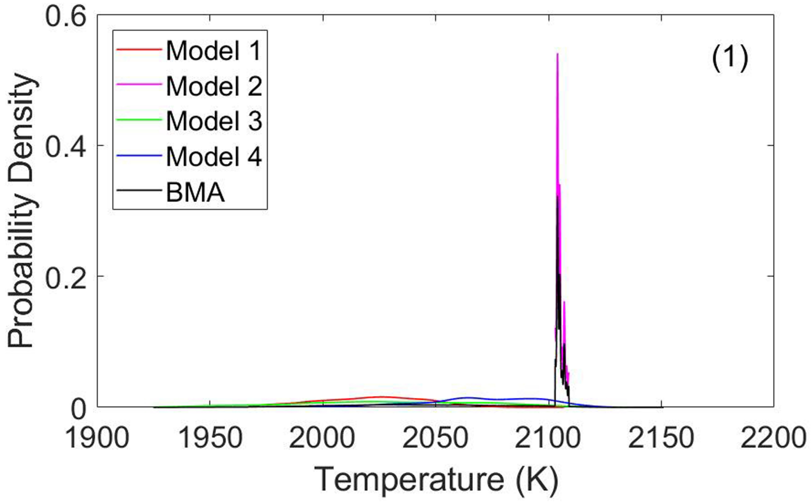
APPENDIX B

SUPPLEMENTARY MATERIAL (CHAPTER 5)

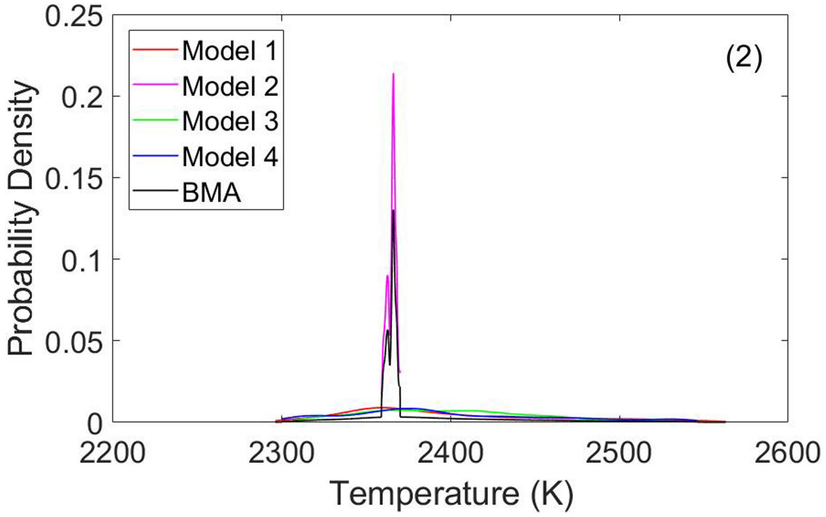


For each individual model and weighted average model, cross-sectional probability distributions obtained through kernel density estimation are demonstrated in the above figure for all the transformation line uncertainty bands (the blue shaded regions with composition independent distributions) and at some random compositions for the liquidus uncertainty band (across the green shaded region) that have been labeled from 1 to 13 in Figure 5.7.

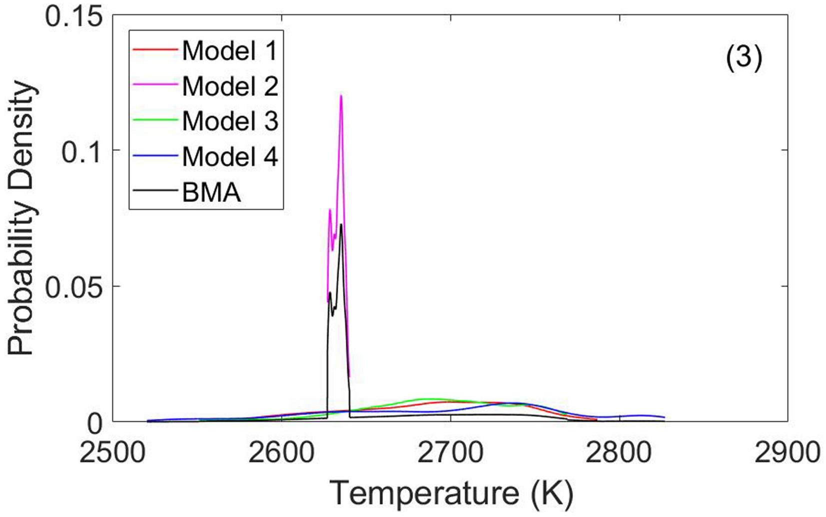
For better clarity, theses probability distributions are also shown individually as follows.



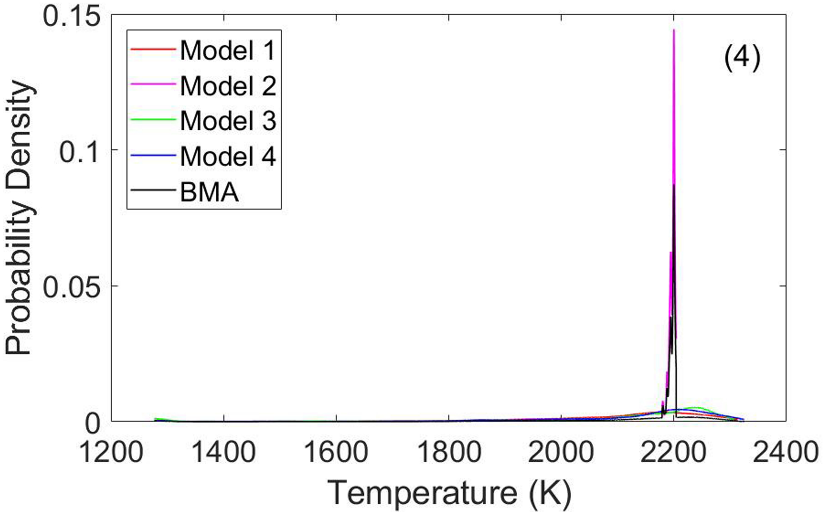
(1)
Model 1
Model 2
Model 3
Model 4
BMA
Probability Density
0.6
0.4
0.2
0
1900
1950
2000
2050
2100
2150
2200
Temperature (K)

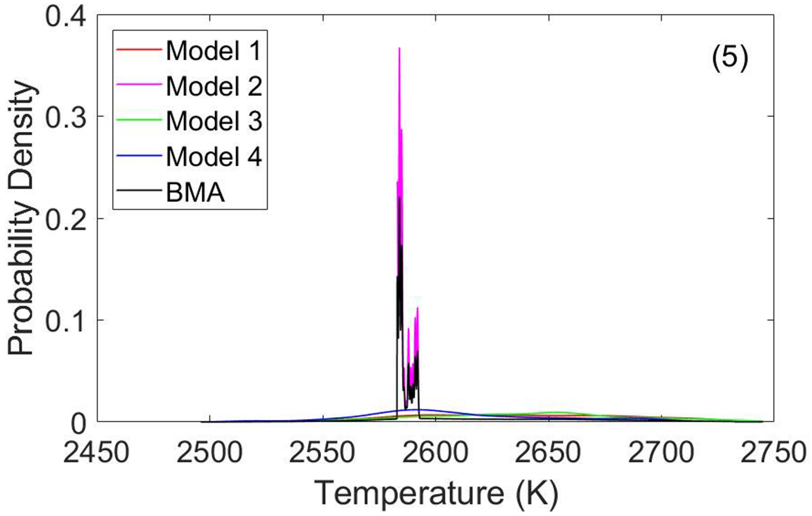


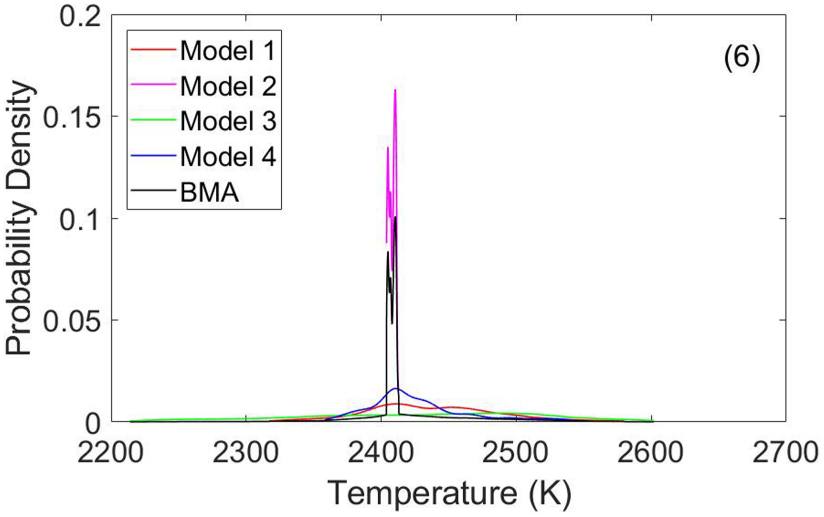
(2)
Model 1
Model 2
Model 3
Model 4
BMA
Probability Density
0.25
0.2
0.15
0.1
0.05
0
2200
2300
2400
2500
2600
Temperature (K)

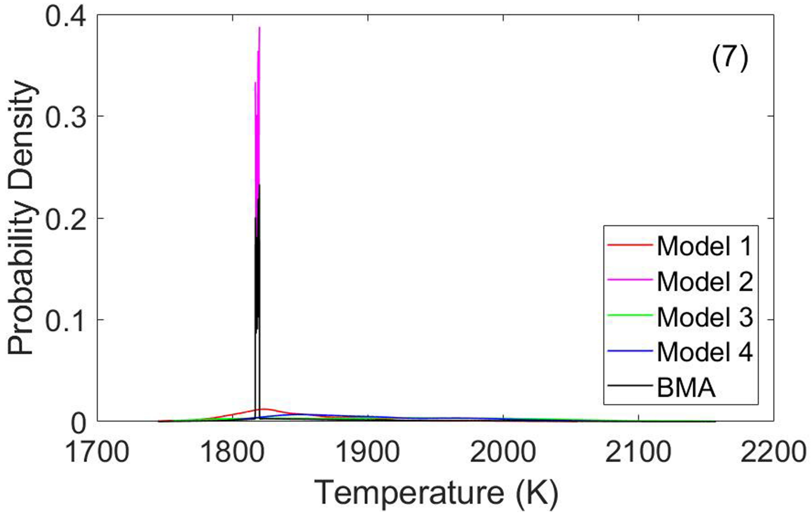


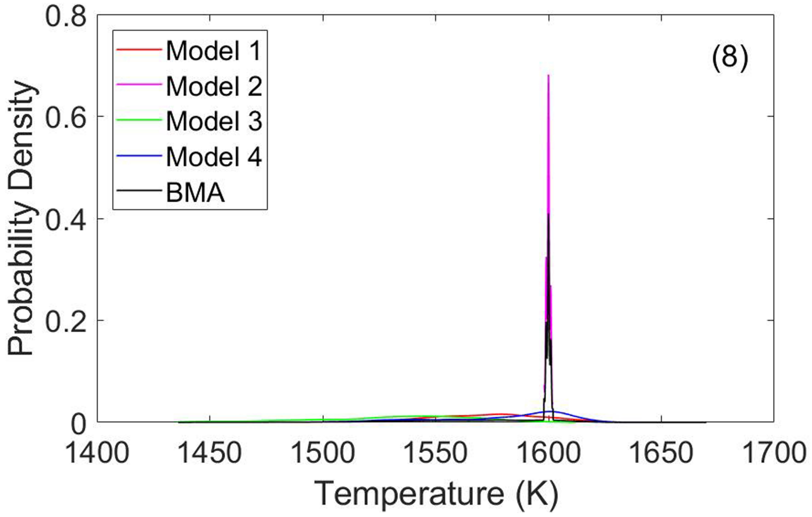
(3)
Model 1
Model 2
Model 3
Model 4
BMA
Probability Density
0.15
0.1
0.05
0
2500
2600
2700
2800
2900
Temperature (K)

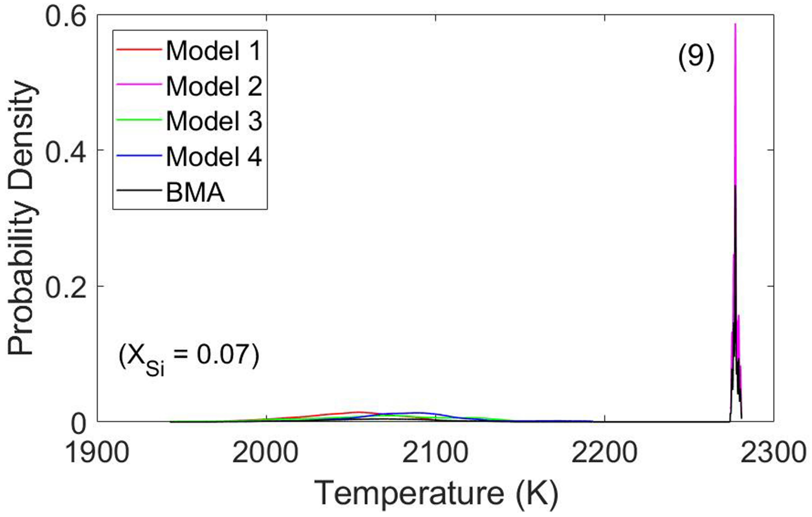

Probability Density
0.15
0.1
0.05
0
1200
1400
1600
1800
2000
2200
2400
Temperature (K)
Model 1
Model 2
Model 3
Model 4
BMA
(4)

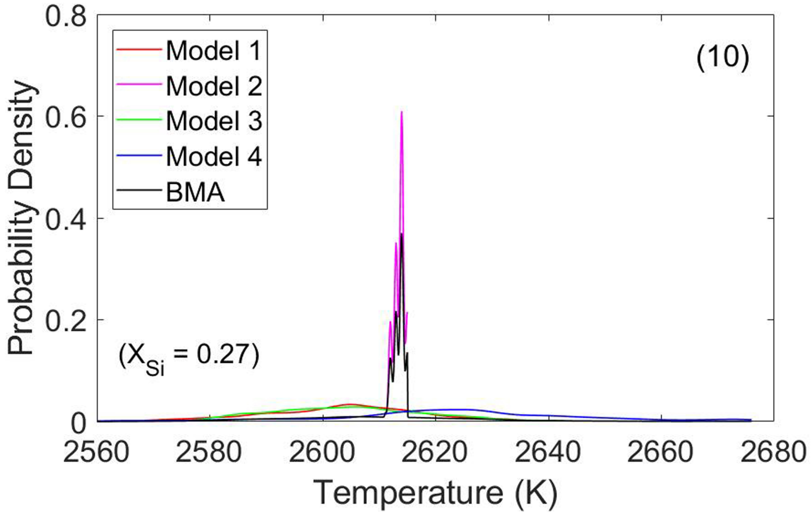

Probability Density
0.4
0.3
0.2
0.1
0
2450
2500
2550
2600
2650
2700
2750
Temperature (K)
Model 1
Model 2
Model 3
Model 4
BMA
(5)

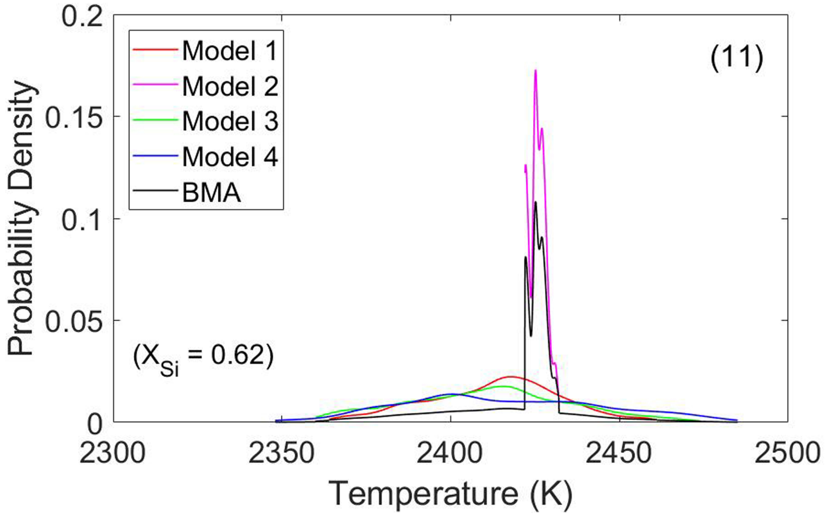

Probability Density
0.2
0.15
0.1
0.05
0
2200
2300
2400
2500
2600
2700
Temperature (K)
Model 1
Model 2
Model 3
Model 4
BMA
(6)

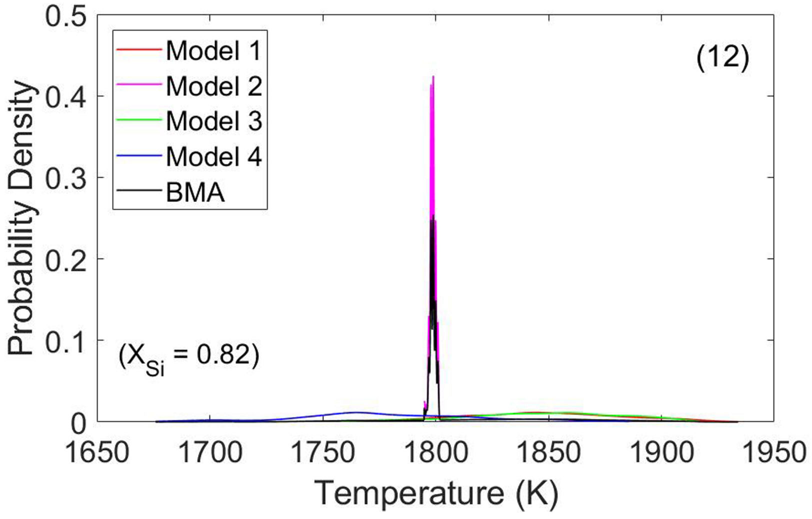

(7)
Probability Density
Temperature (K)
0
0.1
0.2
0.3
0.4
1700
1800
1900
2000
2100
2200
Model 1
Model 2
Model 3
Model 4
BMA

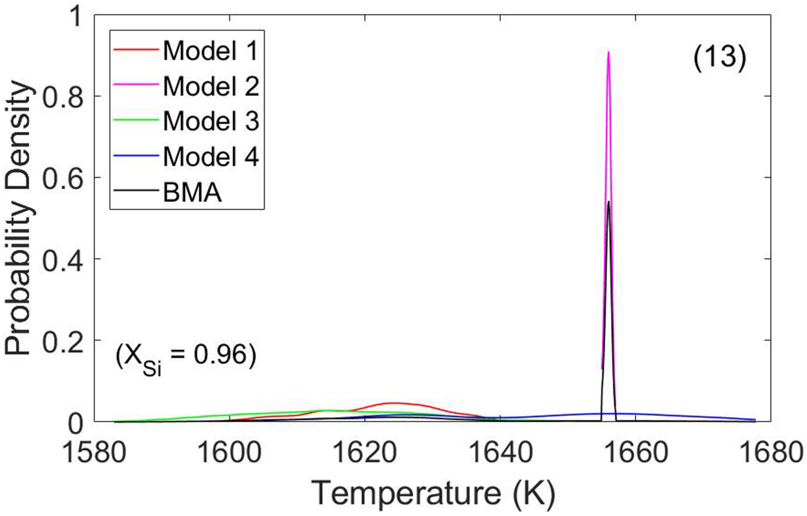

(8)
Probability Density
Temperature (K)
0
0.2
0.4
0.6
0.8
1400
1450
1500
1550
1600
1650
1700
Model 1
Model 2
Model 3
Model 4
BMA


(9)
$(X_{Si} = 0.07)$
Probability Density
Temperature (K)
0
0.2
0.4
0.6
1900
2000
2100
2200
2300
Model 1
Model 2
Model 3
Model 4
BMA


(10)
Model 1
Model 2
Model 3
Model 4
BMA
(X_{Si} = 0.27)
Probability Density
0.8
0.6
0.4
0.2
0
2560
2580
2600
2620
2640
2660
2680
Temperature (K)


(11)
Model 1
Model 2
Model 3
Model 4
BMA
(X_{Si} = 0.62)
Probability Density
0.2
0.15
0.1
0.05
0
2300
2350
2400
2450
2500
Temperature (K)


(12)
Model 1
Model 2
Model 3
Model 4
BMA
(X_{Si} = 0.82)
Probability Density
0.5
0.4
0.3
0.2
0.1
0
1650
1700
1750
1800
1850
1900
1950
Temperature (K)


Probability Density
Temperature (K)
Model 1
Model 2
Model 3
Model 4
BMA
(13)
(X_{Si} = 0.96)
0
0.2
0.4
0.6
0.8
1
1580
1600
1620
1640
1660
1680