Reliability Engineering and System Safety 164 (2017) 110–123

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

Bayesian uncertainty quantification and propagation for validation of a microstructure sensitive model for prediction of fatigue crack initiation

Saikumar R. Yeratapally[a], Michael G. Glavicic[b], Christos Argyrakis[c], Michael D. Sangid[a,*]

[a] School of Aeronautics and Astronautics, Purdue University, West Lafayette, IN 47907, USA
[b] Rolls-Royce Corporation, Indianapolis, IN 46225, USA
[c] Rolls-Royce plc, Derby DE24 8BJ, UK

ARTICLE INFO

Keywords:
Fatigue crack initiation
Global sensitivity analysis
Uncertainty quantification
Uncertainty propagation
Bayesian inference
Markov chain Monte Carlo

ABSTRACT

A microstructure and deformation mechanism based fatigue crack initiation and life prediction model, which links microstructure variability of a polycrystalline material to the scatter in fatigue life, is validated using an uncertainty quantification and propagation framework. First, global sensitivity analysis (GSA) is used to identify the set of most influential parameters in the fatigue life prediction model. Following GSA, the posterior distributions of all influential parameters are calculated using a Bayesian inference framework, which is built based on a Markov chain Monte Carlo (MCMC) algorithm. The quantified uncertainties thus obtained, are propagated through the model using Monte Carlo sampling technique to make robust predictions of fatigue life. The model is validated by comparing the predictions to experimental fatigue life data.

1. Introduction

The majority of mechanical failures can be attributed to fatigue, which is a complex problem involving many independent factors that evolve during cyclic loading. Fatigue crack initiation in polycrystalline materials can be attributed to the heterogeneous microstructure forming complex stress states resulting in strain heterogeneities and localization. Additionally, cyclic loading manifests in deformation mechanisms leading to cyclic slip irreversibilities, which ultimately increase stress concentration and thereby lead to the formation of cracks. Many empirical [1,2] and physics-based models [3–6] have been proposed to predict fatigue life in polycrystalline materials. Uncertainties exist in every model, and before such computational models are employed (to predict the life of components), careful attention must be given to understand the degree in which these uncertainties influence the predicted quantity of interest (QoI), in this case the fatigue life. Rigorous uncertainty quantification for validation purposes is a pre-requisite for such predictive models to be used in a production environment. The current work focuses on identifying, quantifying and propagating the uncertainties in a microstructure based life prediction model [6] for the purpose of validating the model. In this study, model validation is performed based on:

i) Global sensitivity analysis (GSA) to identify the set of non-influential parameters in a factor-fixing setting, which in turn helps in reducing the computational cost of the uncertainty quantification problem [7,8].
ii) Bayesian inference to quantify uncertainties in the set of influential parameters determined using GSA [8–10].
iii) Monte-Carlo sampling to propagate the quantified uncertainties to obtain distribution of predicted life, which will be used in validating the model's predictions [8,10].

Researchers in various sub-disciplines of computational materials science and engineering including computational solid (and particle) mechanics [10,11], computational fluid dynamics [12], molecular dynamics (MD) [13,14], etc., have integrated uncertainty analysis into their modeling framework. Over the past decade, uncertainty quantification has been successfully applied to fatigue crack growth models pertaining to both metals [15–19] and composites [20,21]. Zhang and Mahadevan [15] used Bayesian inference technique to quantify uncertainties via statistical distribution parameters in two competing crack growth models for metals. Cross et al. [16] used a hierarchical Bayesian inference framework to quantify uncertainties in equivalent initial flaw size and crack growth rate parameters, and hence improved the predictive capabilities of their fatigue crack growth model. Sankararaman et al. [19] used a Bayes network to propose a methodology for uncertainty quantification and model validation in fatigue crack growth analysis. Chiachio et al. [21] used a full Bayesian approach to quantify uncertainties of a set of five damage mechanics

* Corresponding author.
E-mail address: msangid@purdue.edu (M.D. Sangid).

http://dx.doi.org/10.1016/j.ress.2017.03.006
Received 12 April 2016; Received in revised form 4 March 2017; Accepted 10 March 2017
Available online 16 March 2017

models for composites and the best of the models was chosen based on an information-theoretic approach by calculating the relative probability amongst all other candidate models. The metal fatigue crack growth models discussed above are empirical in nature and are independent of the microstructure of the material, which has a great influence in crack initiation [3–6] and microstructurally small fatigue crack growth [22]. The fatigue life prediction framework which is used in the current study [6], differs from the above mentioned crack growth models in the following four ways. First, it is a microstructure based framework, where the morphological and crystallographic heterogeneities in the microstructure are considered and an attempt is made to link the variability of the microstructure with the fatigue life calculated. Second, it is not fully empirical in nature, as the model takes into consideration the physics of underpinning deformation mechanisms, which lead to cyclic slip irreversibilities during fatigue, thereby addressing fatigue at the slip system length-scale. Third, it considers complex stress states from grain-to-grain interactions. Finally, the model predicts number of cycles for fatigue crack initiation rather than calculating crack growth with number of cycles. While dealing with models that predict fatigue crack initiation, a phenomenon, which is dependent on both the local microstructure and deformation mechanisms, the number of epistemic uncertainties increases due to the complexities involving length-scale dependent deformation mechanisms. These uncertainties need to be quantified, in order to validate the model and identify an appropriate applicability regime. There is a great amount of work that needs to be done in quantifying uncertainties in complex physics based models and hence improving the predictive capabilities of such models [23].

Several micro-mechanical fatigue crack initiation models have been developed which take into consideration the heterogeneities within the microstructure and various parameters that quantify length-scale dependent deformation mechanisms [3–6]. The energy based model of Tanaka and Mura [3] takes into consideration parameters like the frictional stress, cyclic slip irreversibility and the specific fracture energy of the material. The fatigue crack initiation framework developed by Sangid et al. [4,5] takes into consideration, width of a persistent slip band (PSB), dislocation density, γ' volume fraction, grain boundary (GB) energies, extrusion height at intersection of PSB-GB, stacking fault and anti-phase boundary energies. There are uncertainties associated with all the parameters mentioned above, some of which are difficult to measure using experiments. Although these models provide great insights into understanding how certain microstructural features and competing deformation mechanisms lead to initiation of fatigue cracks, systematic sensitivity and uncertainty analysis, in an attempt to validate such physics-based models, is still lacking [23]. The current work fills this gap by using a sensitivity and uncertainty analysis framework, in order to validate a microstructure and deformation mechanism based life prediction model [6]. Although validation of the model is a driving motivation, the main contribution of the current work is the application of sensitivity and uncertainty analysis to a microstructure and deformation dependent fatigue life prediction model.

The rest of the paper is organized as follows: In Section 2, we provide a brief overview of the microstructure dependent fatigue life prediction model and an overview regarding how the sensitivity and uncertainty analysis are performed on the model. Section 3 lists all the uncertainties that prevail in the model and categorize the uncertainties. It also provides an overview of the uncertainty analysis framework used in this study. In Section 4, we show the application of GSA to identify the most influential parameters in the model, which contribute most to the uncertainty in the output. Section 5 describes Bayesian framework and the quantified uncertainties for the set of influential parameters. In Section 6, we use Monte Carlo simulations to propagate the uncertainties through the model to obtain distributions of life predictions. Section 7 discusses the dependence of various parameters on applied strain amplitude, and conclusions are presented in Section 8.

2. Overview of the microstructure based life prediction model

As uncertainty quantification and propagation are the main focus of the current work, we only present a brief overview of the PSB energy based life prediction model (or PSB model) in this section. For a detailed description of the model, please refer to Yeratapally et al. [6]. It must be noted that the PSB model takes information on state dependent variables like the resolved shear stress (τ^{α}), normal stress (σ_N^{α}), back stress (χ^{α}), critical resolved shear stress (g^{α}) and accumulated strain in a slip system (γ^{α}), output from crystal plasticity finite element (CPFE) simulations (of one-cycle loading) done on a statistically equivalent microstructure (SEM), which is sufficiently large to capture the statistics of microstructural attributes (like mean and variance of grain size distribution and percentage of twins in the microstructure) and strength properties (elastic modulus, yield strength, hardening response and reverse plasticity upon unloading) pertinent to the material of interest, RR1000, a powder processed superalloy developed by Rolls-Royce plc, is used in this study. This stress-strain information along with the GB energetics is used as input to the PSB model, to predict the potential location and number of cycles for crack initiation. Although CPFE is an integral part of the fatigue framework, the focus of the current work is to quantify the uncertainties in the PSB model itself. Quantifying all the uncertainties in CPFE framework requires information about how the stress and strain evolve relative to the microstructure with applied loading, which is beyond the scope of the current work.

Fig. 1 displays a schematic of a PSB traversing a low angle GB (or LAGB) and impinges upon a high angle GB (for example an annealing twin boundary), where the dislocations pile-up, form extrusions at the boundary plane, and thereby increase the stress concentrations at the GB, which could potentially lead to crack initiation.

With this established view of a PSB (based on experimental observations), we define the energy of a PSB as follows:

$$E_{PSB} = \sum_i \partial X_i \left(f \int_0^L \gamma_{APB} dL + (1-f) \int_0^L \gamma_{SFE} dL \right) n_{eff}^{layers} + \sum_i \partial X_i (E_{slip-GB}^{\gamma-MD} n_{ext-GB}^{dis} bh) + \sum_i \partial X_i (\sigma_{pile-up} - \Delta\tau_{CPFEM}^{\alpha} - \sigma_{hardening}) bL n^{layers}, \quad (1)$$

where ∂X_i is the incremental slip within PSB, f is the volume fraction of the γ' precipitate phase in the nickel-base superalloy, γ_{SFE} is the stacking fault energy of the γ phase, γ_{APBE} is the anti-phase boundary energy of the γ' precipitate, n_{eff}^{layers} is the number of effective layers contributing to SFE or APBE, and it decreases with a decrease in the degree of crystallinity, DC, in the PSB (see Section 3.3 for further explanation on DC), L is the length of the PSB, $E_{slip-GB}^{\gamma-MD}$ is the energy required for a dislocation to transmit across a GB, n_{ext-GB}^{dis} represents the number of dislocations forming an extrusion at the PSB-GB intersection, b is the magnitude of the Burgers vector (which represents the amount of lattice distortion due to the glide motion of a single dislocation in a crystalline lattice), h is the width of the PSB, $\Delta\tau_{CPFEM}^{\alpha}$ is applied cyclic stress on the PSB, $\sigma_{hardening}$ accounts for the hardening within the PSB, $\sigma_{pile-up}$ is the pile-up stress at the intersection of the PSB and the GB, and n^{layers} is the number of slip planes within the PSB, which is related to the PSB width, h, as $n^{layers} = \frac{h}{b}$.

Within the energy expression of PSB, the terms n_{ext-GB}^{dis}, $\sigma_{pile-up}$, σ_{stroh}, $\sigma_{hardening}$ are calculated using the following expressions:

$$n_{ext-GB}^{dis} = k \left| \frac{\tau^{\alpha} - \chi^{\alpha}}{g^{\alpha}} \right|^m \frac{L}{L_{avg}} \left(\frac{\tau^{\alpha} + \sigma_{pile-up} - \sigma_{stroh}}{\sigma_{stroh}} \right) \sqrt{n - n_{offset}}. \quad (2)$$

$$\sigma_{pile-up} = \frac{1.8 \mu \gamma^{\alpha}}{\pi(1-\nu)}. \quad (3)$$

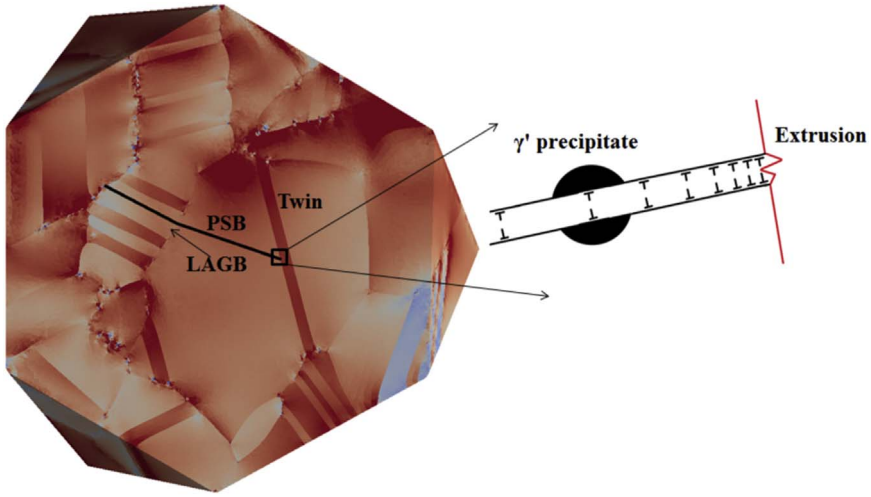

Fig. 1. Cross section view of a 3D strain state plot of a microstructure, showing the schematic of a PSB, shearing γ' precipitates, crossing an LAGB and impeded by a twin boundary, where it forms extrusions representing a preferred site for crack initiation.

$$\sigma_{stroh} = \left\{\frac{\pi E_{GB}^{trans}\mu}{2(1-\nu)L}\right\}^{\frac{1}{2}} \quad (4)$$

$$\sigma_{hardening} = 0.45\mu b\sqrt{\rho} \quad (5)$$

where k is a proportionality constant, m is the rate sensitivity parameter, τ^{α} is resolved slip-system stress, χ^{α} is back stress on a slip system and g^{α} is the critical resolved shear stress, E_{GB}^{trans} is the dislocation transmission energy for distinct GBs [4], μ is rigidity modulus, ν is the Poisson's ratio, γ^{α} is accumulated strain in the slip-system which is calculated from the CPFE simulations, n is the number of cycles, n_{offset} is the threshold number of loading cycles after which extrusions start to form at the PSB-GB intersection and ρ is the dislocation density within the PSB. It must be noted that the formulations of $\sigma_{pile-up}$, σ_{stroh} and $\sigma_{hardening}$ (shown in Eqs. (3–5)) have been adopted from the work of Schouwenaars et al. [24], Stroh [25] and Taylor [26], respectively.

Failure (e.g. fatigue crack initiation) would occur when the energy of the PSB would attain its minimum value. Mathematically speaking, cracks would initiate when the following two conditions are satisfied:

$$\frac{dE_{PSB}}{dX_i} = 0 \text{ and } \frac{d^2E_{PSB}}{d^2X_i} > 0 \quad (6)$$

The PSB model is used to link the variability in microstructure of a polycrystalline alloy to the scatter in its fatigue life, by taking into account the (i) complex stress state within the microstructure, (ii) energetics of deformation mechanisms occurring at a slip-system length scale. The variability in microstructure is simulated by generating several SEMs based on the statistics of microstructural attributes obtained from electron backscatter diffraction (EBSD) scans of the material, as described in detail in Yeratapally et al. [6]. The CPFE framework is used to solve for the complex stress and strain fields developed within the SEMs, when a cyclic load is applied. The PSB model takes the output from the CPFE simulation, in the form of the micromechanical stresses and strains, as input and deterministically calculates fatigue life on a grain-by-grain basis (assuming that a PSB exists in all the grains), and in doing so determines the weakest-link grain (the grain with the minimum number of cycles to failure) within the microstructure, which is determined to be the location where a fatigue crack is most likely to initiate. In summary, for each SEM and a set of model parameters, the PSB model links the SEM with one fatigue life (or number of cycles to failure) data point.

The PSB model predicted that fatigue crack initiation is favored to occur at the intersection of a persistent slip band (PSB) and twin boundaries, specifically when the PSB forms in a large grain [6]. Furthermore, the twins (where the PSB model predicts cracks are likely to initiate) were subjected to significant amount of normal stress (compared to the resolved shear stress), thereby implying that normal stress plays a significant role in crack initiation by acting as a mode I crack driving force unzipping the PSB. Additionally, the PSB model determined that the plastic strain accumulation (due to dislocation pileup) and elastic stress anisotropy (due to lattice mismatch at grain boundaries) act in concert to the aforementioned attributes to influence fatigue crack initiation [6]. The fatigue model delivers valuable insights regarding the location of crack initiation. The fatigue model has several parameters that possess inherent uncertainties, which need to be identified, quantified and propagated through the model, for the purpose of validating the fatigue life predictions and extending the applicability of the model in a production environment.

The current study employs GSA to identify the set of influential parameters among all the parameters in the fatigue life prediction model. Following GSA, a full Bayesian inference framework, which uses Markov chain Monte Carlo (MCMC) algorithms, is used to quantify the uncertainties in the set of influential parameters. The quantified uncertainties of the parameters are forward propagated through the model, in order to make predictions of fatigue life, using Monte Carlo sampling. This helps to quantitatively relate the input uncertainties to the output. A brief overview of the application of the sensitivity and uncertainty analysis to the fatigue model is shown in Fig. 2.

3. Uncertainties in the model

The presence of uncertainties in a computational model can be attributed to many factors including, but not limited to: (i) measurement errors, (ii) numerical errors, (iii) missing physics due to simplifying assumptions, etc. Although uncertainties arise from several sources, they are generally categorized as either aleatory or epistemic [27–29]. Aleatory uncertainty arises due to the randomness due to inherent variability in the system [27–29], for instance in polycrystalline materials such type of an uncertainty includes the randomness in size, shape and orientation attributes of nearest neighbor grains [23]. This type of uncertainty is irreducible even through the exhaustive collection of data. Epistemic uncertainty on the other hand is caused due to lack of knowledge (or data) [27–29]. This type of uncertainty can be reduced by gathering more data. For the purpose of this study, the inherent aleatory uncertainties that exist in the microstructure have not been considered in the uncertainty analysis. The uncertainties in input parameters in the current model can be reduced through experiments and physics-based computations, and hence are consid-

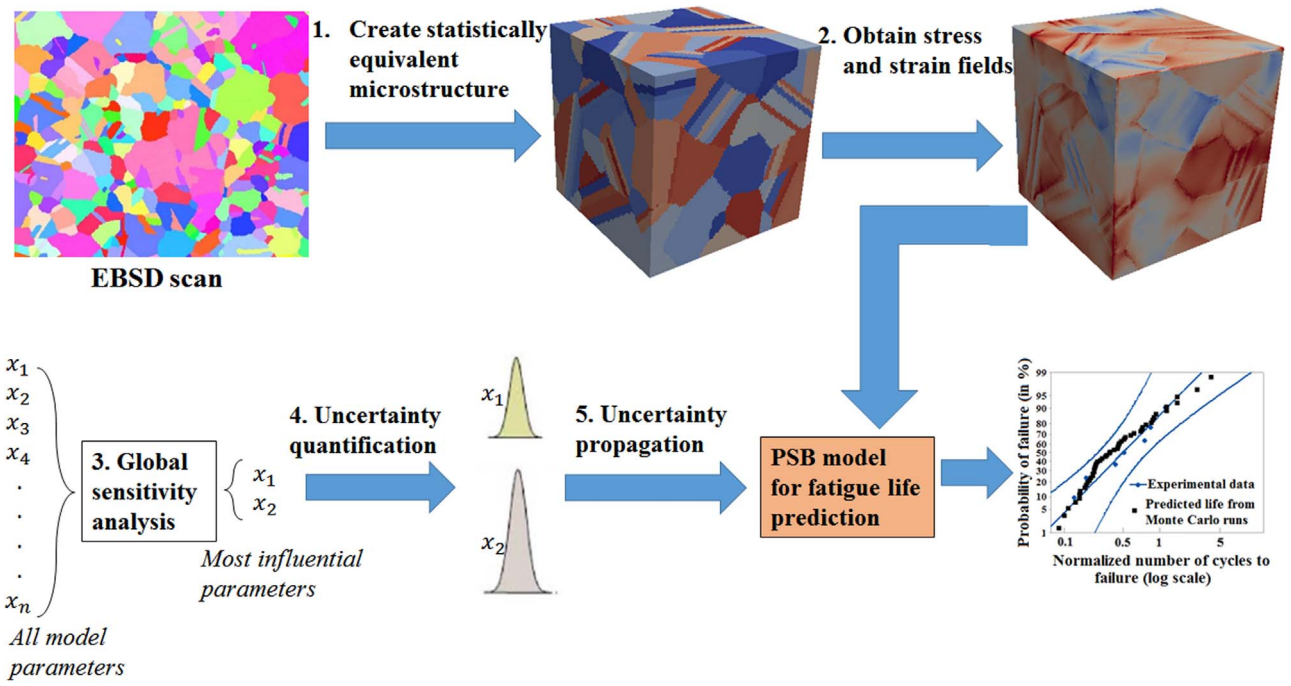

Fig. 2. Uncertainty quantification and propagation framework for the fatigue model.

ered to be epistemic in nature.

Input parameters present in the PSB model (discussed in Section 2) can be categorized into three classes based on the ease of measurability using experiments (i) physical parameters that can be calculated using experiments, (ii) physical parameters that cannot be measured using experiments, and (iii) parameters that are empirical/semi-empirical in nature, and hence cannot be measured using experiments. In addition to the aforementioned uncertainties, model discrepancy (or model bias) exists due to epistemic uncertainties present in the model.

3.1. Physical parameters that can be calculated using experiments

These physical parameters correspond to a distribution of values due to variation in the experimentally measured values. Uncertainties observed in these parameters are possibly due to measurement errors, material variability, sampling volume, or usage of various data collection techniques. The set of model parameters (within the fatigue model discussed in Section 2) that fit into this category of uncertainties are Young's modulus (E), Poisson's ratio (ν), γ' volume fraction (f), dislocation density (ρ) within a PSB, and PSB width (h). Although the measurement of E and ν are straight forward using mechanical testing equipment in a lab, the measurement of the rest of the parameters (f, ρ, h), which correspond to a lower length-scale, require advanced experimental setup, for instance the usage of neutron diffraction experiments for measuring dislocation density (ρ) or the usage transmission electron microscope to measure the width of a PSB (h). Due to this reason, a distribution was assigned to each of these input parameters based on literature review and expert opinion. While performing GSA (as discussed in Section 4), the influence of all parameters on the uncertainty of output is calculated, using sensitivity indices and global sensitivity plots. Table 1 summarizes the distribution assigned to each of the parameters, and also the source from where the data was obtained.

3.2. Physical parameters that cannot be easily measured using experiments

Some of the physical parameters in the model, like the stacking fault energy (γ_{SFE}), anti-phase boundary energy (γ_{APBE}), energy required for a dislocation to transmit across a GB ($E^{\gamma-MD}_{slip-GB}$) cannot be easily measured using experiments. Although, the intrinsic stacking fault energy can be calculated by measuring the distance between Shockley partials under a transmission electron microscope, unstable stacking fault energy strictly cannot be measured directly from experiments. Additionally, the fatigue model considered in the current study takes into account, the effect of normal stress (σ^{α}_{N}) on the energy of the PSB and hence it requires the quantitative dependence of γ_{SFE} and γ_{APBE} on the normal stress acting upon the PSB. These quantities can be obtained using MD simulations. The potentials used to measure γ_{SFE}, γ_{APBE} and $E^{\gamma-MD}_{slip-GB}$ are still empirical in nature, hence there is some degree of uncertainty involved with the calculations obtained from these simulations [13,14]. A Gaussian distribution is assigned to two

Table 1
List of model input parameters and prior distributions assigned to each of the parameter for Bayesian uncertainty quantification. Here $\widetilde{N}$ represents a normal distribution of a parameter and is defined using the mean and standard deviation, $\widetilde{LN}$ represents a lognormal distribution of a parameter and is defined using the mean and standard deviation of the associated normal distribution of the logarithm of the parameter, $\widetilde{U}$ represents a uniform distribution of a parameter defined using its upper and lower bounds.

Type	*Parameter*	*Distribution*	*Units*	*References*
Physical parameters	E	$\widetilde{N}(210,5)$	GPa	[33]
	υ	$\widetilde{N}(0.307,0.005)$	–	[33]
	f	$\widetilde{N}(0.4,0.03)$	–	[34,35]
	ρ	$\widetilde{LN}(36.3808,1)$	$\frac{1}{m^2}$	[36–38]
	h	$\widetilde{N}(200,50)$	nm	[39,40]
	γ_{SFE}	$\widetilde{N}(128,6.4)$	$\frac{mJ}{m^2}$	[4,5]
	γ_{APBE}	$\widetilde{N}(260,13)$	$\frac{mJ}{m^2}$	[33]
	$E^{\gamma-MD}_{slip-GB}$	$\widetilde{LN}(28.3634, 0.1)$	$\frac{mJ}{m^3}$	[31,32]
Empirical/ Semi-empirical parameters	DC	$\widetilde{U}(0,1)$	–	–
	k	$\widetilde{U}(0, \infty)$	–	–
Model discrepancy hyper-parameter	σ	$\widetilde{U}(0, \infty)$		

parameters, γ_{SFE} and γ_{APBE}, with the mean as the calculated value and a standard deviation equal to 5% of the calculated value.

GBs have a definitive role in localizing and accumulating strain [32], which is a precursor to crack initiation [55]. The fatigue framework employs additional information about the interactions between dislocations and specific GB characters, with the value of $E^{\gamma-MD}_{slip-GB}$. Slip-GB interactions are very complex and span an infinite parameter space, so certain simplifying assumptions are made within the model; specifically, conservative values of the energy barrier for slip-GB interaction is taken by not accounting for shear stress on the GB plane. In order to quantify the uncertainty in the values of $E^{\gamma-MD}_{slip-GB}$, we consider an example of a slip-twin boundary interaction for various twin orientations and dislocation types. Specifically, from the work of Ezaz et al. [30], MD simulations are used to quantify dislocation transmission and incorporation at a coherent twin boundary (CTB) in six different classes of slip-twin reactions. The energy barrier for dislocation transmission through CTB was observed to be proportional to the magnitude of the residual Burgers vector within the CTB after the transmission event [31]. Burgers vector increased from 0 (due to pure cross-slip of a screw dislocation) to 0.53a (with a being the lattice parameter), the energy barrier increased by approximately 24%, from 187 mJ/m^2 to 232 mJ/m^2. With the 24% deviation in $E^{\gamma-MD}_{slip-GB}$ taken as the uncertainty for all types of slip-GB interactions, a log-normal distribution is used to quantify uncertainty in the values of the GB energy barriers to slip.

The values of these parameters (γ_{SFE}, γ_{APBE} and $E^{\gamma-MD}_{slip-GB}$) are sampled from the assigned distributions to see the propagation of uncertainty through the model and its effect on fatigue life predictions. This will help in assessing the sensitivity of these parameters in fatigue life predictions.

3.3. Parameters that are empirical/semi-empirical in nature

Some parameters in the model are empirical/semi-empirical in nature. For instance, the first parameter in this category is the degree of crystallinity (DC) within a PSB. The physical meaning for this parameter is associated with the construction of the PSB energy balance, which sums the contribution of individual dislocations within the PSB. As additional dislocations are added, the entropy within the PSB increases and the degree of crystallinity decreases. Physically this parameter is associated with the latent heat of fusion for the material. This parameter decreases with increasing number of cycles as the dislocation density increases. The DC parameter can take values from 0 to 1, with 1 referring to a perfect crystal and values approaching 0 are analogous to the latent heat of fusion. Hence a uniform distribution is assigned to this parameter, with 0 and 1 as the lower and upper bounds of the distribution.

The second parameter is the proportionality constant, *k*, in Eq. (2), which is used to calculate the number of dislocations forming an extrusion at the PSB-GB intersection, n^{dis}_{ext-GB}. The PSB model shows an inverse correlation between extrusion height and fatigue life, the detailed aspect of which is discussed in Appendix A. The value of n^{dis}_{ext-GB}, and hence the extrusion height at the PSB-GB intersection scales with the value of *k*. Measurement of extrusion height within the bulk of the material is not a trivial task, and this leads to uncertainty in the parameter, which can be accounted using *k*. Hence, *k* accounts for the uncertainties arising due to missing physics and the simplifying assumptions that went into defining a stress based empirical expression for calculating n^{dis}_{ext-GB}, Eq. (2). Since n^{dis}_{ext-GB} is a strictly positive quantity, so is *k*. Therefore, a uniform distribution, $\widetilde{U}(0,\infty)$ is assigned to *k* as a prior distribution. Although such a distribution would serve as a non-informative prior while conducting Bayesian uncertainty quantification, it becomes cumbersome to sample k from $\widetilde{U}(0,\infty)$, for the purpose of the sensitivity analysis, a technique used to find the set of most influential parameters. Hence for the purpose of the sensitivity analysis, the values of k were restricted based on the following two conditions. First, for the PSB model to be able to predict crack initiation in the grains, based on Eq. (6), values of k must be restricted. In other words, if the energy required to form extrusions at the PSB-GB intersection (which can be identified by the expression, $\sum_i \partial X_i(E^{\gamma-MD}_{slip-GB} n^{dis}_{ext-GB} bh)$ in Eq. (1)) is too high, then extrusions never form and cracks never initiate. Second, the product, '$n^{dis}_{ext-GB}b$' in Eq. (1) signifies the height of the extrusion formed at the PSB-GB intersection (as shown in Fig. 1). Since the formation of extrusion at PSB-GB intersection is a fatigue specific phenomena, it is important that the aforementioned product must have physically reasonable values. For instance the height of the extrusion must be much smaller than the average grain size of the material, based on experimental observations. In an effort to satisfy the two aforementioned conditions, the values of k were restricted values between 0 and 2, using a uniform distribution $\widetilde{U}(0,2)$, for the purpose of conducting sensitivity analysis.

3.4. Model discrepancy and error in experimental data

In addition to the uncertainty caused by the aforementioned parameters, epistemic uncertainties occur in the model due to some simplifying assumptions. For instance, the PSB model assumes that PSBs exist on all active slip systems within grains (or grain clusters) within the microstructure, thereby not accounting for the number of cycles required to form PSBs. In order to address the bias created due to such assumptions, we introduce a model discrepancy (or model bias) term, δ. This can be viewed as a model error in predicting fatigue lives for a given microstructure at a given strain amplitude.

Most importantly, the scatter observed in the fatigue lives (at a particular strain amplitude) can be partly attributed to the variability in the microstructure of the material [3–6]. In addition to the inherent microstructure dependent variability, errors may occur while taking the values from experiments due to equipment alignment, data acquisition tolerances, variability in specimen machining, etc. For instance, the exact determination of when an internal fatigue crack has initiated is not possible. Due to this reason, an experimentalist relies on a percentage drop in load to define number of cycles to crack initiation within a strain controlled fatigue experiment [47]. In this context, the criteria adopted to decide crack initiation life is more phenomenological in nature and hence accounts for errors. Such measurement errors in experiments are accounted for, by introducing a measurement error term, e.

If $y(x_i)$ represents the fatigue lives obtained from experiments, and $f(x_i,\boldsymbol{\theta})$ represents the fatigue life predictions from the model, they can be related by the following simple equation:

$$y(x_i)=f(x_i, \boldsymbol{\theta})+\delta(x_i) + e \quad (7)$$

where x_i are experimental conditions (or design parameters) that can be controlled by an experimentalist. For instance, x_i can be the applied strain amplitude (Δε), temperature or R-ratio. The value of $\boldsymbol{\theta}$ represents the model parameters which cannot be directly controlled or sometimes cannot even be directly observed by the person conducting the experiment [43]. All ten parameters discussed in Sections 3.1 - 3.3 comprise the $\boldsymbol{\theta}$ vector. In the current model, $\boldsymbol{\theta}$ depends on how the material responds to an applied fatigue load based on the local microstructure features. Due to fundamental difference between x_i and $\boldsymbol{\theta}$, and the fact that $\boldsymbol{\theta}$ cannot be controlled (and in some cases cannot be measured) during experiments, we take the model discrepancy to only depend on the experimental conditions (or design parameters), x_i [8,44]. Assuming that the model discrepancies, $\delta(x_i)$, are independently and identically distributed, a zero mean Gaussian random distribution with standard deviation, σ_1, represented as $\widetilde{N}(0,\sigma_1)$, is assigned to $\delta(x_i)$. A similar argument is applied to experimental error term, e, which can also be assigned a Gaussian distribution $\widetilde{N}(0,\sigma_2)$, with a zero mean and standard deviation, σ_2. Since, the summation of two Gaussian distributions is also a Gaussian distribution, we replace the two Gaussian distributions with just one Gaussian,

$\widetilde{N}(0, \sigma)$. Hence, with the addition of σ into the parameter estimation problem, we now have a new augmented parameter set, $\{\boldsymbol{\theta}, \sigma\}$ or $\{E, \nu, f, \rho, h, \gamma_{SFE}, \gamma_{APBE}, E_{slip-GB}^{\gamma-MD}, DC, k, \sigma\}$. Since there is no information available about the standard deviation, σ, it is an unknown hyper-parameter that needs to be inferred using the Bayesian framework. Due to lack of any information for this parameter and based on the fact that can only take positive values, a uniform prior distribution, $\widetilde{U}(0,\infty)$, is assigned to this parameter.

4. Parameter selection using global sensitivity analysis (GSA)

Generally, in computational models, different input parameters will have varying degree of influence on the uncertainty of the model output. Moreover, a higher parameter dimension will increase the computational cost of uncertainty analysis. Hence, parameter selection technique is applied to isolate the most influential parameters in the model [7]. This is done using GSA, the objective of which is to ascertain how uncertainty in model outputs can be apportioned to uncertainties in model inputs, when considered over the entire range of input values [7,8]. The GSA focuses on the model parameters, $\boldsymbol{\theta}$. In other words, the hyper parameter, σ, is not considered, because given the extreme ranges in σ values; it can overshadow the sensitivities in the model parameters. This consideration is in agreement with the work of Chiachío et al. [21]. We consider σ in the uncertainty quantification problem along with other influential model parameters (determined using GSA). In other words, while doing sensitivity analysis, we ignore the combination of discrepancy and error terms (δ and e), in order to rank, just the model parameters with respect to their influence on the uncertainty of the output. In this current work, we use various existing graphical tools and techniques [41–43] to select the set of influential model parameters based on how they influence the fatigue life calculations. Scatter plots give us a rough idea of how an input parameter affects the output, they cannot be used to easily assess the importance of one parameter over the other. If the number of parameters that we are trying to analyze increase, scatter plots become cumbersome to analyze. Furthermore, although the well-established variance based sensitivity indices [8] also provide information on how various input parameters contribute to the uncertainty in the calculated output, graphical tools used in this work provide more information than the variance based sensitivity indices. For instance, once the most important input has been detected, variance based sensitivity indices do not provide any information on how much reduction in the range of uncertainty in the influential parameters is required to achieve a target reduction of the output variance [42]. Graphical tools maintain a rich set of information in addition to the sensitivity indices, as they estimate the contribution of a particular range of input parameter values on the sample mean [41] or sample variance [42] of the output quantity of interest. Before using any of the graphical tools for GSA, Monte Carlo simulations are run, in which all input parameters are varied simultaneously, over their entire range, by sampling from distributions listed in Table 1, except for the model parameter k, which was sampled from a uniform distribution $\widetilde{U}(0,2)$. For the purpose of this study, 50,000 Monte Carlo iterations were run. For each Monte Carlo iteration, each of the model parameter is randomly sampled from its respective distribution, and the fatigue life is calculated as the output, as described in Section 2.

In the current work, we use two such graphical plots [41–43] called, contribution to sample mean (CSM) plot [41] and contribution to sample variance (CSV) plot [42], which are used to assess the influence of an input parameter on the sample mean or sample variance of the output, respectively. Based on the deviation of the CSM (or CSV) curve of a parameter X_j from the diagonal (e.g. straight line with slope of 1), its influence on the sample mean (or sample variance) of the output can be determined. The CSM and CSV plots for all variables in model parameters in the fatigue model are shown in Fig. 3a and b, respectively. It should be noted that the two graphical tools (CSM and CSV), are independent of the type of model (additive/non-additive/linear/non-linear) being used and just need the data generated using thousands of Monte Carlo simulations.

In order to ensure that stability exists in the results shown by the graphical tools, the deviation (measured as the maximum perpendicular distance) of the CSM and CSV curves from the diagonal, and the evolution of the deviations is plotted against the number of iterations, and hence the sample size. The stability plots for both CSM and CSV curves for all the parameters are shown in Fig. 4. From Fig. 4, it is clear that the deviations stabilize after around 10,000 iterations for all the three curves.

It can be inferred from the global sensitivity plots shown in Fig. 3, that the parameters $E_{slip-GB}^{\gamma-MD}$, k, ρ and h, significantly influence the output (fatigue life prediction), as these curves deviate from the diagonal (for CSM and CSV plots in Fig. 3a and b, respectively). The curves of the non-influential parameters coincide (or are in close proximity) with the diagonal (for CSM and CSV plots), thus demonstrating little impact on the overall model output. Additionally, Fig. 4 clearly shows that the deviation between the two curves and their respective diagonal lines, for the parameters $E_{slip-GB}^{\gamma-MD}$, k, ρ and h, are significantly larger and stable over a range of sample sizes. The efforts of the Bayesian inference will be directed towards quantifying the uncertainties in the set of influential parameters.

5. Uncertainty quantification using Bayesian inference

We represent the most influential model parameters (along with the hyper-parameter σ) using a vector, $\boldsymbol{\alpha}=\{E_{slip-GB}^{\gamma-MD}, k, \rho, h, \sigma\}$, for use within the uncertainty quantification framework. The parameters that are considered relatively non-influential are still needed as the input to the model, and are assigned to their mean values of their respective prior distributions (shown in Table 1).

5.1. Bayesian method

Bayesian inference technique is used in updating probabilities, and more generally, our current state of knowledge of parameters, $\boldsymbol{\alpha}$, using observed (or experimental) data, D. The updated probability distribution (posterior distribution), $\pi(\boldsymbol{\alpha}|D)$, can be obtained by applying Bayes theorem as follows:

$$\pi(\boldsymbol{\alpha}|D) = \frac{\pi(D|\boldsymbol{\alpha})\pi_0(\boldsymbol{\alpha})}{\pi(D)} \tag{8}$$

where $\pi_0(\boldsymbol{\alpha})$ represent our current state of knowledge or prior beliefs on parameter set $\boldsymbol{\alpha}$, $\pi(D|\boldsymbol{\alpha})$ represents the likelihood of observing the data D, given parameter realizations $\boldsymbol{\alpha}$ and $\pi(D)$ is the marginal density obtained by integrating the joint density $\pi(\boldsymbol{\alpha},D)$ over all possible values of $\boldsymbol{\alpha}$. It can be treated as a normalization factor and this allows us to proportionally relate the prior, posterior and likelihood distributions as follows:

$$\pi(\boldsymbol{\alpha}|D) \propto \pi(D|\boldsymbol{\alpha})\pi_0(\boldsymbol{\alpha}) \tag{9}$$

The prior distributions of all parameters, $\pi_0(\boldsymbol{\alpha})$, are listed in Table 1. Using a statistical model shown in Eq. (7), in which the model errors are assumed independently and identically distributed, the likelihood $\pi(D|\boldsymbol{\alpha})$ of observing the data follows a normal distribution [8,10]:

$$\pi(D|\boldsymbol{\alpha}) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} e^{\sum_{i=1}^{n}\left(\frac{-(y_i-g(\boldsymbol{\alpha}))^2}{2\sigma^2}\right)} \tag{10}$$

where σ is a hyper parameter in the likelihood distribution, which can be inferred using the fatigue life data $\{y_1, y_2,....,y_n\}$ collected independently from testing n different specimens, and $g(\boldsymbol{\alpha})$ represents the prediction made by the model. Although the non-influential parameters are also used in the model, for brevity we just show the set $\boldsymbol{\alpha}$ (set of influential parameters) as it is those parameters that we are

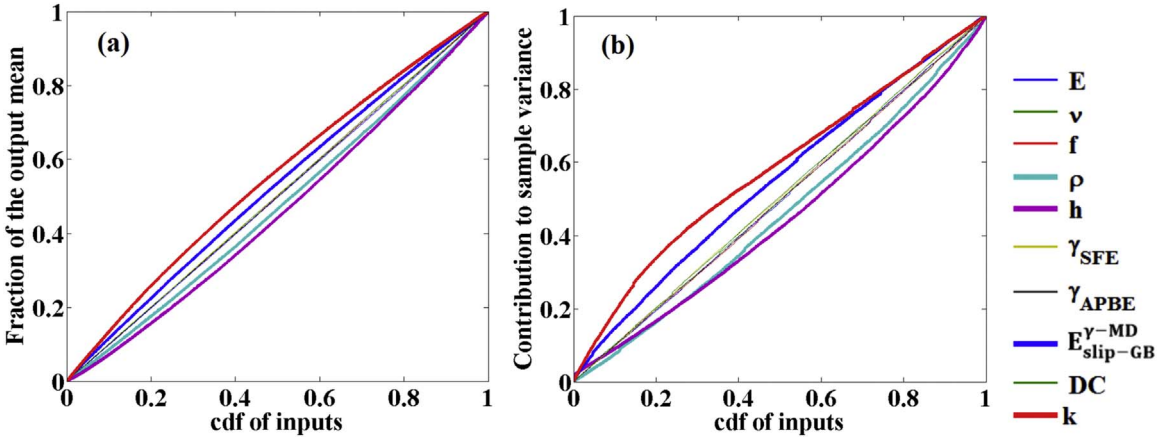

Fig. 3. Graphical tools to qualitatively understand the influence of various parameters on the mean and variance of output. (a) CSM plot and (b) CSV plot. The data for generating these plots was acquired from running 50,000 Monte Carlo simulation runs.

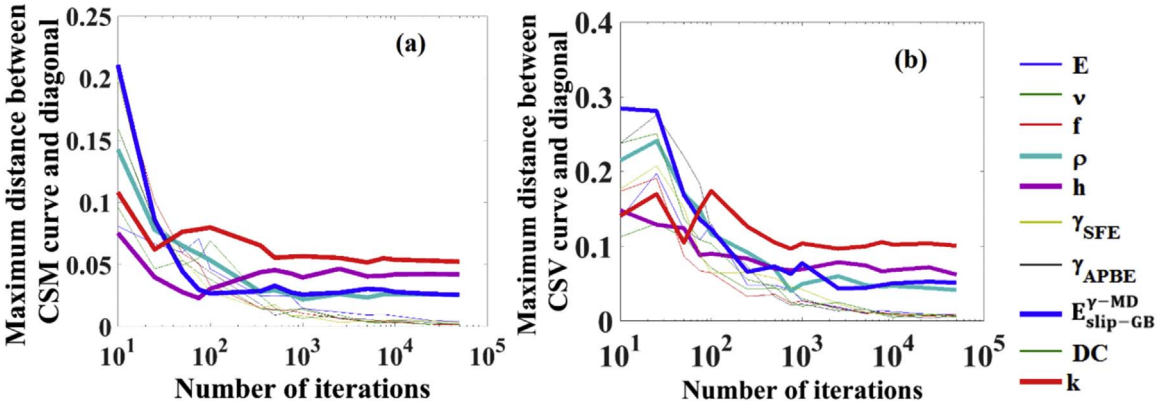

Fig. 4. (a) Evolution of maximum deviation (or distance) measured between CSM curves of all the parameters and the diagonal, (b) evolution of maximum deviation (or distance) measured between CSV curves of all the parameters and the diagonal.

trying to estimate. The parameter estimation (or inverse Bayesian uncertainty quantification) problem in this study is to infer distributions of all parameters (in the set parameter $\boldsymbol{\alpha}$), which makes predicted fatigue lives to be close to experimental fatigue life data, by using MCMC algorithm [45].

5.2. Markov chain Monte Carlo method for evaluating posterior densities of model parameters

MCMC uses the attributes of posterior densities (which are in turn dependent on the likelihood and prior) to specify parameter values that adequately explore the geometry of the distribution [8]. It is based on a simple idea of comparing the posterior densities of a candidate point (or a newly proposed point), $\boldsymbol{\alpha}^*$, with a current location in the state space $\boldsymbol{\alpha}$. If the candidate point yields a posterior density greater than the posterior density at the current location, then the proposed point is accepted with a probability of one, otherwise the candidate point is accepted with a probability of r, less than one.

Mathematically, the ratio (r) between the posterior densities between the candidate point ($\boldsymbol{\alpha}^*$) and the current point ($\boldsymbol{\alpha}$) can be represented as:

$$r=\frac{\pi(\boldsymbol{\alpha}^*|D)}{\pi(\boldsymbol{\alpha}|D)}=\frac{\pi(D|\boldsymbol{\alpha}^*)\pi_0(\boldsymbol{\alpha}^*)\pi(\boldsymbol{\alpha}\mid\boldsymbol{\alpha}^*)}{\pi(D|\boldsymbol{\alpha})\pi_0(\boldsymbol{\alpha})\pi(\boldsymbol{\alpha}^*\mid\boldsymbol{\alpha})} \tag{11}$$

Symmetrical forms of proposal distributions, $\pi(\boldsymbol{\alpha}^*\mid\boldsymbol{\alpha})$ (used to propose a new candidate point, $\boldsymbol{\alpha}^*$), are preferred to construct posterior densities, which most likely will have a symmetrical geometry. But in general, posterior densities of some model parameters can have highly asymmetrical shapes, in which case using a symmetrical proposal distribution could lead to long convergence times. Therefore, in this work we use non-symmetrical proposal distributions which are a more generic case of sampling from any type of posterior densities and is the basis of the Metropolis-Hastings (m-H) algorithm [46]. A high level overview of the component-wise sampling using M-H algorithm is provided in the Appendix B.

The convergence of the Markov chain is monitored to ensure a stationary posterior distribution of parameters is obtained. We use a convergence test for which multiple Markov chains were run in parallel with different initial values of the parameters [17,48]. If the variances of a parameter (p) between n chains and within chains are represented as B_p and W_p, respectively, then an estimate of variance of p, V_p, can be represented as:

$$V_p=\frac{n-1}{n}W_p+\frac{1}{n}B_p \tag{12}$$

A convergence test statistic, R_p, can be calculated as:

Table 2
Mean, variance, and convergence statistic of the posterior distributions of parameters.

Parameter (units)	Mean	Standard deviation	Convergence test statistic (R_p)
$E^{\gamma-MD}_{slip-GB}$ $(\frac{mJ}{m^3})$	2.2×10^{12}	0.13×10^{12}	1.003
k	1.67×10^{0}	1.85×10^{-1}	1.021
ρ $(\frac{1}{m^2})$	8.68×10^{15}	1.48×10^{15}	1.010
h (nm)	1.91×10^{2}	2.41×10^{1}	1.040
σ	6.53×10^{3}	2.06×10^{3}	1.006

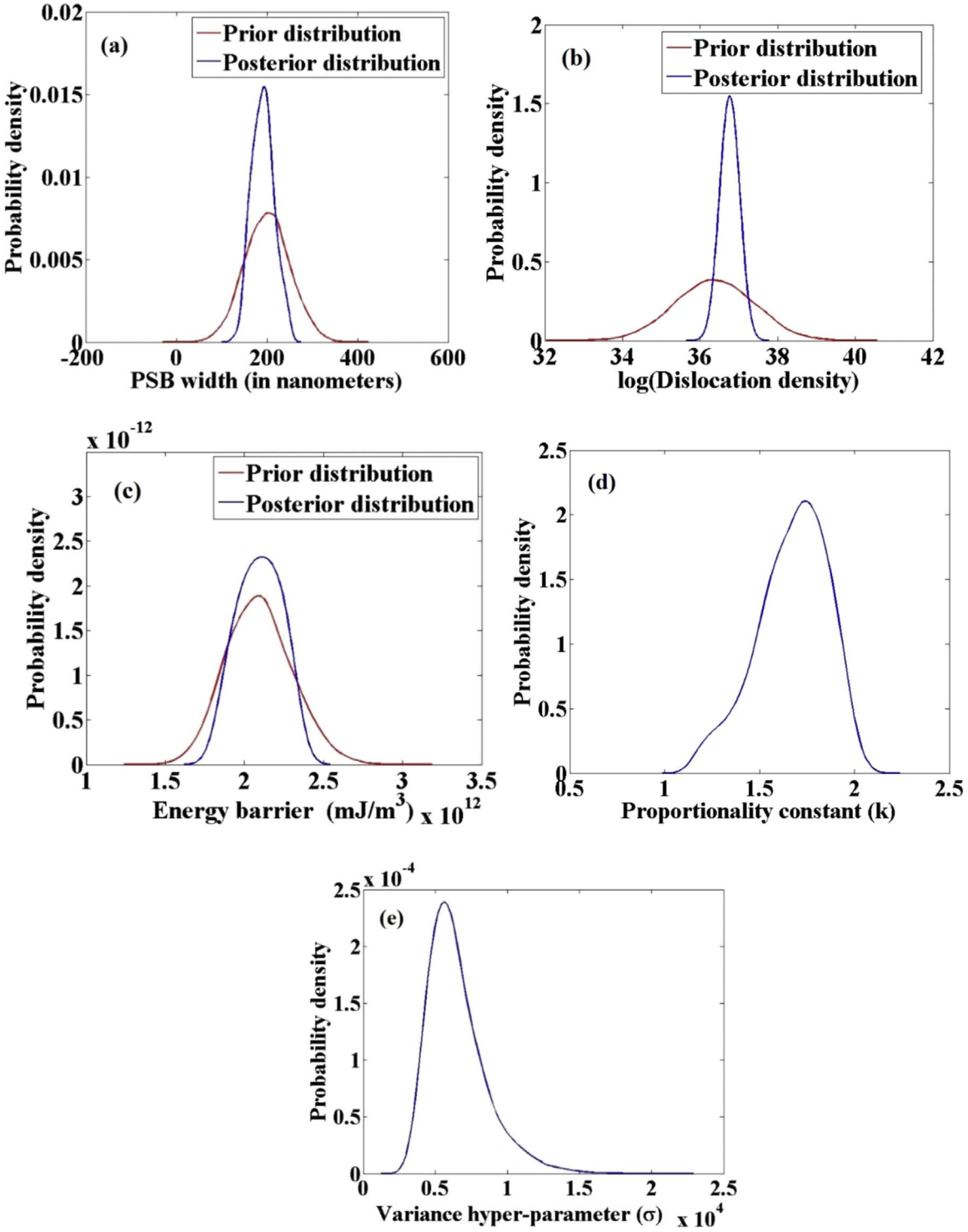

Fig. 5. Prior and posterior densities of (a) PSB width (h), (b) dislocation density, ρ, in log scale, (c) Energy barrier, $E_{slip-GB}^{\gamma-MD}$, (d) proportionality constant, k and (e) variance hyper-parameter (σ). It must be noted that the prior density for k and σ is, $\bar{U}(0,\infty)$, and hence is coincident with the x-axis.

$$R_p = \sqrt{\frac{V_p}{W_p}} \tag{13}$$

The value of R_p when calculated for each parameter should be close to unity to quantitatively ensure convergence. For the full form expressions of B_p and W_p, please refer to Cross et al. [17].

5.3. Marginal posterior densities of the parameters

Multiple Markov chains are run with different initial guesses of the parameters, and posterior distributions are extracted after checking for convergence in each of the chains. The mean, variance, and the convergence test statistic (R_p) for the posterior distributions of all parameters are shown in Table 2. The posterior densities for the set of influential parameters (along with parameter σ) are shown in Fig. 5.

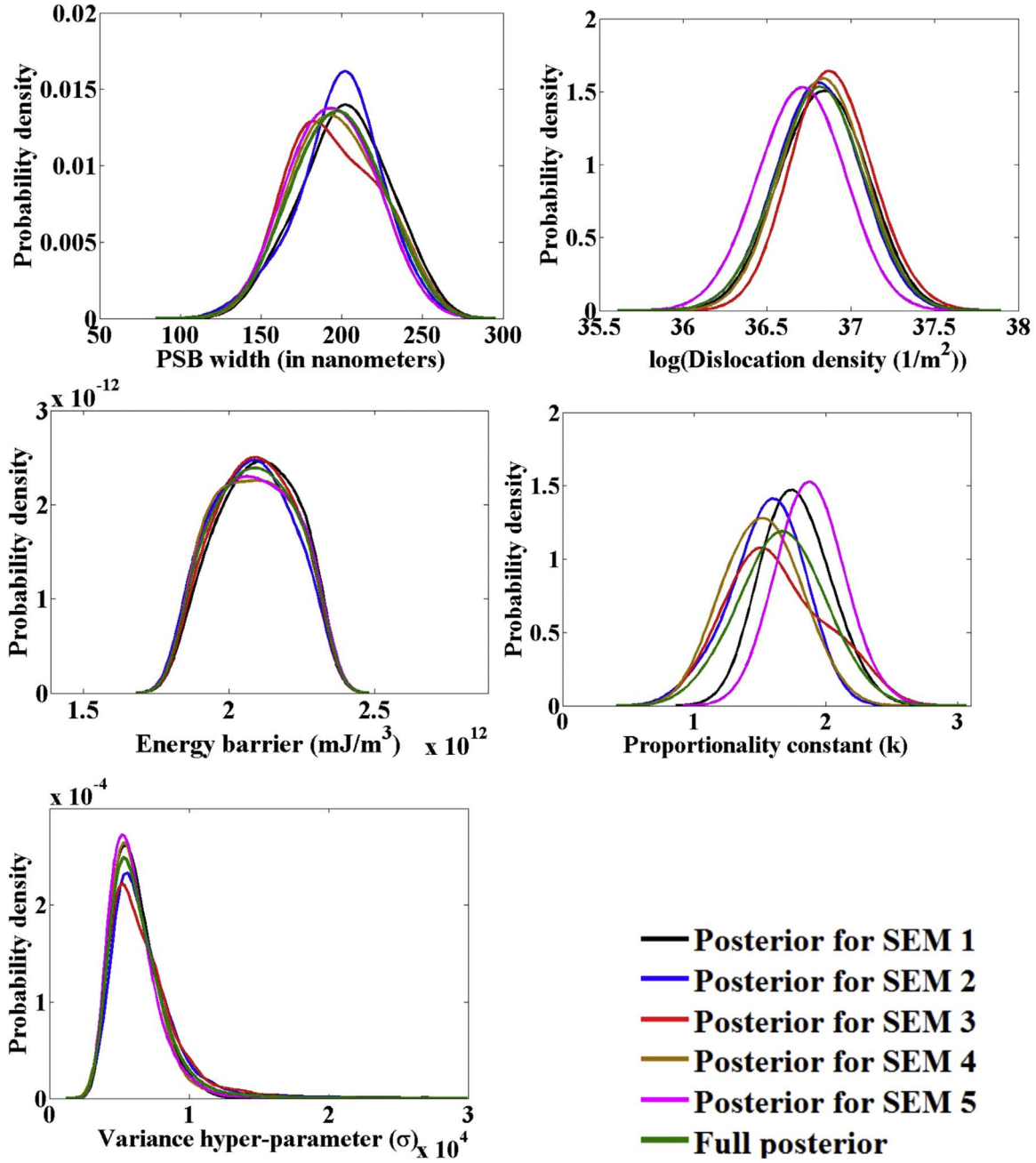

Fig. 6. Plots showing the overlay of sub-posterior distributions of all the parameters obtained using five different SEMs, and also the full posterior distribution obtained from the sub-posterior distributions.

All the posterior distributions of the physical and semi-empirical parameters (like the transmission energy barrier, PSB width, dislocation density) are showing physically reasonable values as indicated in the literature. For instance, the PSB width (h) is close to the prior estimate and the typically observed widths of the PSBs are on the order of hundreds of nanometers [39,40]. Recall, a non-informative flat prior (which has minimal influence on the posterior distribution of parameters) was assigned to the parameters k and σ. The MCMC algorithm quantified the uncertainties pertaining to those parameters, which is evident from the unimodal distribution shown in Fig. 5d and e, respectively.

5.4. Constructing full posterior distributions for all parameters

Instead of using a single large representative volume element (RVE), which makes it computationally prohibitive to conduct CPFE simulations, we consider several SEMs, each of which encompasses a small volume compared to the RVE, but still are sufficiently large to capture the statistics of the microstructural attributes and the strength properties (elastic modulus, yield strength, strain hardening behavior and reverse plasticity upon unloading). In accordance with Niezgoda et al. [49], we treat an RVE as an ensemble of several SEMs, and construct the fatigue life distribution by pooling the distributions obtained by probing the fatigue model through individual SEMs. Similarly, we construct the full posterior distributions (representing the ensemble of all SEMs) for parameters, by pooling together the corresponding posterior densities obtained from the individual SEMs. Full posterior distribution for each parameter is constructed by combining multiple posterior distributions, as discussed by Miroshnikov et al. [50] and Neiswanger et al. [51]. Fig. 6 shows five different posteriors (for each parameter, obtained by using five different SEMs) and also a full posterior distribution for each para-

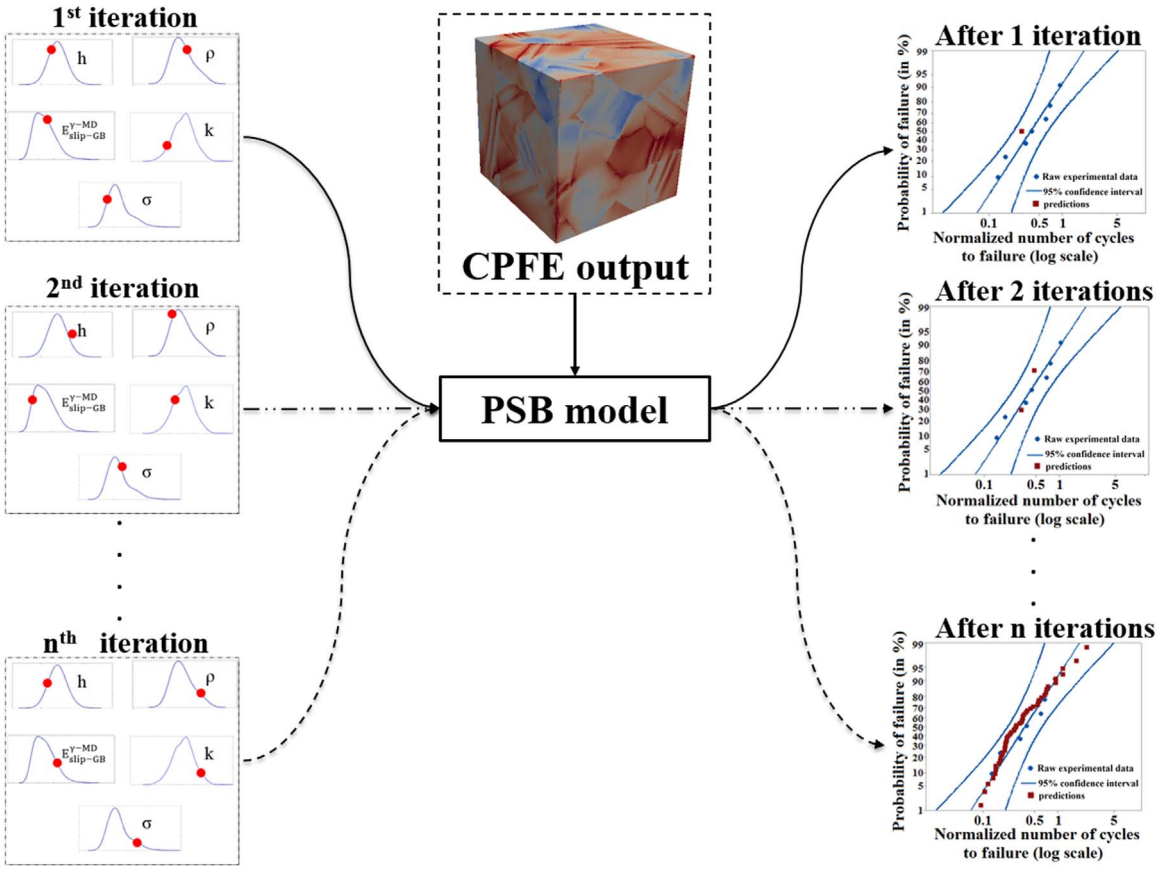

Fig. 7. Schematic of uncertainty propagation using Monte Carlo sampling. In every Monte Carlo iteration each of the influential parameters is assigned a value (depicted using a solid red circle) by sampling from their respective posterior distributions and fed into the PSB model along with the microstructural attributes, stress, strain and other necessary state dependent variable data obtained from the CPFE simulations. For a given microstructure, each Monte Carlo iteration generates one fatigue life data point. The schematic also shows that the CPFE output is kept the same for all the Monte Carlo iterations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

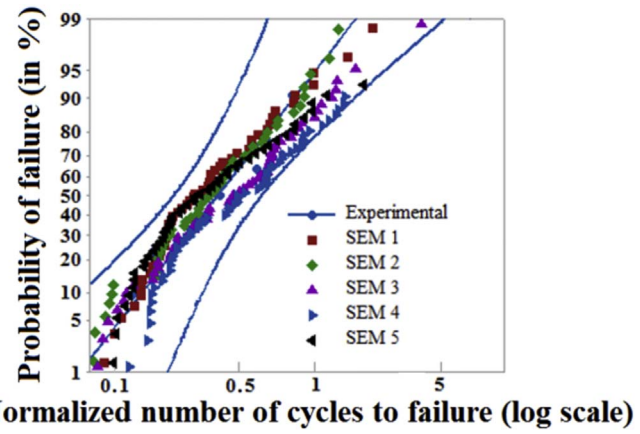

Fig. 8. Comparisons of fatigue life predictions obtained for five different SEMs (by propagating the uncertainties through the PSB model individually for each of the SEM) with the 95% CI bounds generated from experimental data. The number of cycles to failure (log scale) is normalized by the maximum experimental value.

meter.

6. Uncertainty propagation

The posterior distributions of uncertainties calculated using Bayesian inference are propagated through the model, in order to make robust fatigue life predictions. There are many techniques available to propagate uncertainties through a model, such as, i) sampling techniques, ii) perturbation methods, iii) spectral methods, etc., [8]. In this work we use Monte Carlo based sampling techniques due to their simplicity in implementation and the fact that the efficiency of these sampling techniques is independent of the number of parameters within the model. In each Monte Carlo iteration, the set of influential model parameters are sampled from their respective full posterior distributions (shown in Section 5.3) and are fed into the PSB model to obtain fatigue life distributions for a population of SEMs. Although within each Monte Carlo simulation, the PSB model calculates fatigue life on a grain-by-grain basis for the entire SEM, we only consider the fatigue life of the weakest link grain (the one with least number of cycles to failure) and hence each Monte Carlo simulation is associated with one fatigue life data point. Further, it must be noted that, while propagating uncertainties through the model for each SEM, the stress and strain attributes derived from CPFE simulations (that go into the PSB model) are kept the same and only the most influential parameters are changed by sampling from the full posterior distributions. For each SEM, the fatigue life distribution generated through several Monte Carlo simulations is pooled together and for validation purposes, this data was compared to the 95% confidence interval (CI) plots generated using experimental fatigue life data. The schematic of the uncertainty propagation methodology is shown in Fig. 7.

For the purpose of this study, five different SEMs were chosen and life predictions obtained for each of the five SEMs by propagating the uncertainties through the PSB model for each of the five microstructures. Based on the convergence of mean and variance for the output QoI (fatigue life), it was determined that fifty Monte Carlo iterations were sufficient to obtain life predictions with equivalent mean and variance of a much larger sample of fatigue life predictions. Hence, fifty

Table 3
Normalized mean, variance and corresponding 95% CI bounds of experimental data and PSB model calculations using various microstructures.

Fatigue life data source	Normalized mean($\frac{\mu}{\mu_{experiment}}$)	Normalized 95% CI bounds for mean ($\frac{\mu_{lower\ lound}}{\mu_{experiment}}, \frac{\mu_{upper\ lound}}{\mu_{experiment}}$)	Normalized variance ($\frac{\sigma}{\sigma_{experiment}}$)	Normalized 95% CI bounds for variance ($\frac{\sigma_{lower\ lound}}{\sigma_{experiment}}, \frac{\sigma_{upper\ lound}}{\sigma_{experiment}}$)
Experiment	1.0000	(0.5676, 1.4324)	1.0000	(0.5967, 2.3277)
SEM 1	0.9133	(0.7759, 1.0507)	1.0332	(0.8511, 1.3077)
SEM 2	0.9412	(0.7956, 1.0868)	1.0726	(0.9203, 1.3032)
SEM 3	1.1198	(0.9619, 1.2777)	1.2731	(1.0823, 1.5510)
SEM 4	1.2399	(1.0810, 1.3988)	1.3041	(1.0929, 1.6095)
SEM 5	0.9975	(0.8185, 1.1766)	1.3477	(1.0911, 1.7325)

Table 4
Outcome of the two-sample KS test conducted (at 5% significance level) to test the equality between predicted fatigue life distributions (from each of the five microstructure) and the experimental fatigue life distribution. The KS test decision was made by comparing the p-value to the significance level.

Microstructure	KS test statistic	p-value	KS test decision
SEM 1	0.2543	0.7573	Accept null hypothesis
SEM 2	0.1935	0.9572	Accept null hypothesis
SEM 3	0.2431	0.7968	Accept null hypothesis
SEM 4	0.2712	0.6729	Accept null hypothesis
SEM 5	0.2343	0.8381	Accept null hypothesis

Monte Carlo iterations were run for each of the five SEMs and fatigue life distributions were obtained for each SEM, individually. It can be seen from the log-log plot shown in Fig. 8 that the fatigue life predictions obtained using various SEMs lie within the 95% CI bounds of the experimental fatigue life data. In addition to visual comparison for the predictions the 95% CI generated from the experimental fatigue life data (shown in Fig. 8), Table 3 gives a quantitative overview of the normalized mean, variance and corresponding 95% CI bounds for experimental fatigue life data and PSB model predictions using the five microstructures. The 95% CI bounds (for both mean and variance) of the fatigue life predictions from the model (obtained for various microstructures) are overlapping with that of the corresponding 95% CI bounds obtained from the experimental data. Given the quantified uncertainties, SEMs representing the microstructure of the material, along with the heterogeneous stress and strain data obtained from CPFE as inputs, the model is able to predict the life quite well. The calculated fatigue life predictions obtained by propagating the quantified uncertainties through the PSB model shown in Fig. 8 is for a specific fatigue condition, e.g. a single applied strain range, an intermediate elevated temperature, R-ratio, frequency, etc.

Furthermore, two-sample Kolmogorov-Smirnov (KS) test was conducted (at 5% significance level) to test the null hypothesis which states that there is an equality between the predicted fatigue life distributions (obtained by propagating uncertainties through the PSB model) and the experimental fatigue life distribution. The outcome of the two-sample KS test conducted for the fatigue life distributions obtained by using five different microstructures are shown in Table 4. KS test accepts the null hypothesis (as the p-value comes out to be greater than the significance level which is 5%), further affirming that there is good agreement between model predictions and experimental data.

7. Dependency of model parameters on applied strain

Certain parameters involved in the life prediction model are not just material dependent, but also depend on the applied strain. Specifically, the model takes into consideration cyclic slip irreversibilities (please refer to Appendix A for more details), which are dependent on applied strain [52,53]. Cyclic slip irreversibilities manifest as extrusions when PSBs intersect with the surface or GBs. Experimental studies [54,55] quantified the dependence of the extrusion heights (and heights of slip steps formed on the surface) with the applied macroscopic strain. Mughrabi [53] provided a systematic review of studies done to quantify the cyclic slip irreversibilities (in both single crystals and polycrystals) and concluded that an inverse correlation exists between accumulated cyclic slip irreversibilities and fatigue lives. In the current model, the factor which scales the height of extrusion formed at PSB-GB intersection is k, which is also an influential parameter in the model (as shown in Section 4). Hence, the dependence of the parameter k, on the applied strain is characterized, thus linking the extrusion height to the applied macroscopic strain.

For this purpose, the uncertainty in k is quantified at three different strain amplitudes of interest. Uncertainty quantification was done using the same procedure described in Section 5, by utilizing the experimental fatigue life data at the three strains. Only a single SEM is subjected to three different strain amplitudes, solely for the purpose of

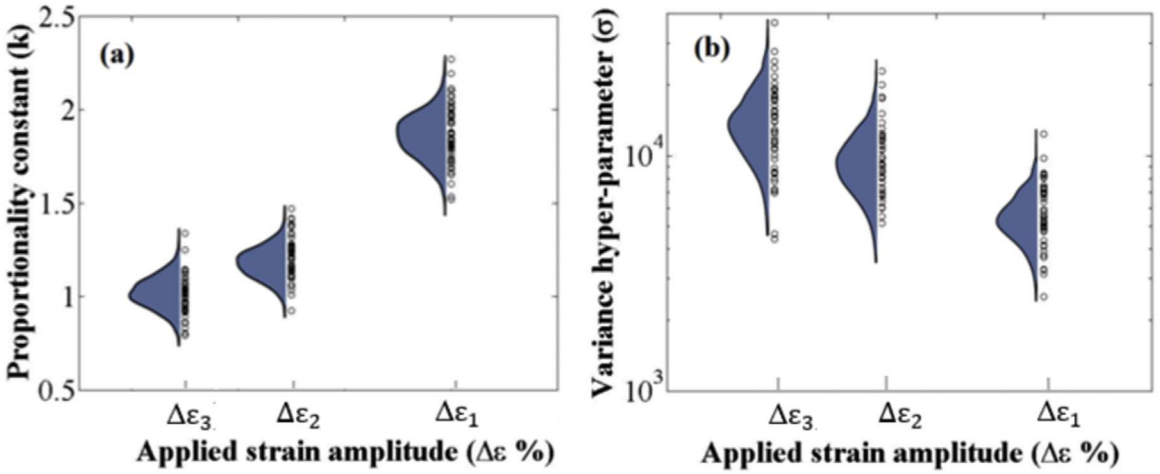

Fig. 9. (a) The variation of proportionality constant term (k) and (b) hyper-parameter (σ) with applied strain.

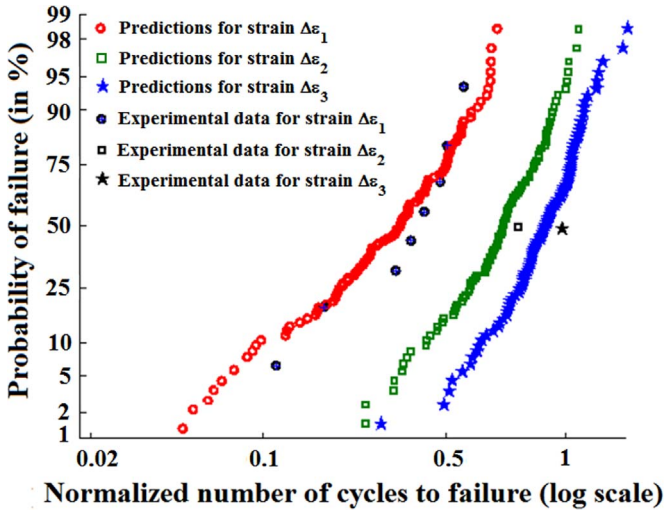

Fig. 10. Strain-life plots at three different strain amplitudes ($\Delta\varepsilon_1 > \Delta\varepsilon_2 > \Delta\varepsilon_3$). The number of cycles to failure (log scale) is normalized by the maximum experimental value.

establishing the dependence of k on the applied strain. Fig. 9a shows the posterior densities of k at three different strain amplitudes. It is evident from Fig. 9a that as the applied strain amplitude increases, the mean value of the posterior distribution of the scaling factor, k, also increases, indicating that the extrusions grow in size when the strain amplitude increases. This makes the model consistent with experimental observations [54,55].

In experiments conducted to characterize fatigue life at an applied strain amplitude, a distinct percentage of load drop is used as a measure to define number of cycles for crack initiation [47], as it is not possible to precisely determine when a fatigue crack initiates, especially when the crack initiates within the bulk of the material. This manifests as experimental error, e, due to a fundamental difference between the predictions from the PSB model (number of cycles to fatigue crack initiation) and the experimental fatigue life data at hand (which is essentially a combination of number of cycles to crack initiation and additional cycles required for an initiated crack to incubate corresponding to a specific percentage of load drop). Since the crack driving forces increase with applied strain amplitude, we postulate that, e (and hence σ), is inherently dependent on the applied strain amplitude (Δε). Hence, in order to establish a relation between the hyper parameter, σ, and applied strain amplitude (Δε), we obtain posterior densities for σ at three different strain amplitudes of interest (as shown in Fig. 9b). The posterior density (of σ) corresponding to the highest applied strain amplitude has the lowest mean value, which can be attributed to the fact that the load drop percentage is achieved at a faster rate within the experiments conducted, due to the presence of larger crack driving forces. Similarly, it takes relatively more cycles to observe a distinct load drop in experiments conducted at lower applied strain amplitudes, and hence the posterior densities of σ, at lower strain amplitudes, have higher mean values.

Further, the quantified uncertainties in k and σ (which are dependent on the applied strain amplitudes) are propagated through the PSB model at three different strain amplitudes ($\Delta\varepsilon_1$, $\Delta\varepsilon_2$ and $\Delta\varepsilon_3$), and fatigue life distributions are obtained at each of the three different strain amplitudes (using the uncertainty propagation methodology described in Section 6). For the purpose of comparing fatigue life distributions at the three different strain amplitudes, we only choose a single SEM and implement uncertainty propagation at the each of the three strain amplitudes of interest. Fatigue life predictions obtained at the three strain amplitudes are shown in Fig. 10. The life predictions obtained (by the PSB model) at the three strain amplitudes are in agreement with the limited experimental fatigue life data available at the three strain amplitudes, overlaid on the fatigue predictions shown in Fig. 10. For the strain amplitudes ($\Delta\varepsilon_2$ and $\Delta\varepsilon_3$), since only one experimental data point is available, a 50% probability of failure was assigned to that data point.

8. Conclusions

A microstructure and deformation mechanism based fatigue life prediction model, which uses the stability of a persistent slip band (PSB) as a criterion for fatigue crack initiation, is validated by using rigorous sensitivity and uncertainty analysis. Various types of uncertainties were identified in the model parameters based on the ease of their measurability using experiments (i) physical parameters that can be measured using experiments, (ii) physical parameters that cannot be easily measured using experiments, (iii) parameters that are empirical/semi-empirical in nature, and hence cannot be measured using experiments. Following parameter identification, global sensitivity analysis (GSA) was used to identify the set of most influential parameters in the model, thereby reducing the dimensionality of the Bayesian uncertainty quantification framework. By using a component-wise Markov chain Monte Carlo (MCMC) algorithm, which takes into consideration the experimental fatigue life data and the prior beliefs of the parameters, posterior densities of the uncertain parameters were obtained using multiple statistically equivalent microstructures (SEMs). Uncertainty propagation was applied using a Monte Carlo framework, in which the uncertainties of the parameters from the full

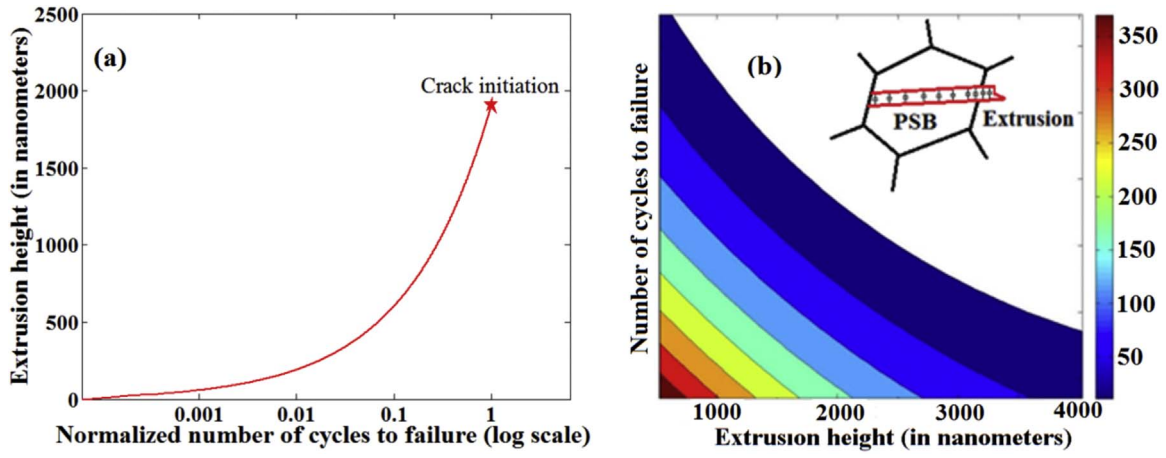

Fig. 11. (a) Evolution of extrusion height with number of cycles, until crack initiation in a specific grain. (b) Influence of extrusion height on fatigue life. The colored contour regions encompass certain number of data points, which can be inferred from the color bar. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

posterior distributions were propagated through the model, thereby calculating the fatigue life prediction in the presence of uncertainties. Fatigue life predictions obtained by using five different SEMs were overlaid on the 95% confidence interval plots of the experimental fatigue data and a good quantitative agreement was observed. Further, two sample Kolmogorov-Smirnov test showed a good agreement between fatigue life distributions calculated by the model and those obtained from the fatigue experiments.

Acknowledgements

Support for this work was provided primarily by Rolls-Royce Corporation and partially by the National Science Foundation (CMMI 13-34664). We thank Jeff Stillinger, Dr. John Matlik, Dr. Mark Hardy, Patrick E. Leser, Dr. James E. Warner, Dr. Robert Kenley, and Andrea Rovinelli for thought provoking discussions on uncertainty quantification.

Appendix A

The PSB model (discussed in Section 2) considers the formation and evolution of extrusions at the PSB-GB intersection (using Eq. (2)). As an example, the cyclic evolution of an extrusion (until crack initiation) at one of the PSB-GB intersections within an SEM is plotted in Fig. 11a. Similarly, fatigue lives and extrusion heights (at failure) were extracted for around 1500 grains in 10 different SEMs. From this data, a contour plot (shown in Fig. 11b) is generated between extrusion height (at PSB-GB intersection) and fatigue life, with the contours encompassing various number of data points. It can be inferred from Fig. 11b that the fatigue life is inversely proportional to the height of extrusions, which is also in agreement with observations in the literature [52–55].

Appendix B

The M-H algorithm can be implemented either by proposing a new state for all the parameters ($\boldsymbol{\alpha}$) at once in a 'block-wise' way (by choosing a proposal distribution which has number of dimensions equal to the number of parameters ($\boldsymbol{\alpha}$) we are trying to estimate) or by proposing each parameter α_i individually in a 'component-wise' way (by using a corresponding univariate proposal distribution assigned for each parameter). For block-wise sampling, depending on the number of dimensions and the type of parameters, an ideal n-dimensional proposal distribution, which takes care of all parameters at once, is difficult to identify. If the proposal distribution is not appropriate, a large number of the proposed sample will be rejected and hence takes a long time for convergence to be achieved. Due to the aforementioned reasons, we use component-wise sampling in the current work.

A high level overview of the component-wise sampling using M-H algorithm is given below.

i) Set loop counter i=0,
ii) Assign initial values to all the individual parameter in $\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, \ldots \alpha_n\}$ by randomly sampling from their respective prior distributions, and
iii) Repeat the following steps until i= M (desired number of iterations).

Increment i by 1,

Repeat the following for each parameter, α_j, in parameter set $\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, \ldots \alpha_n\}$,

Generate a new candidate $\alpha_j^* \sim \pi(\alpha_j^* \mid \alpha_j^{(t-1)})$,

Calculate probability $p = \min\left(1, \frac{\pi(D|\alpha_j^*)\pi_0(\alpha_j^*)\pi(\alpha_j^{(t-1)} \mid \alpha_j^*)}{\pi(D|\alpha_j^{(t-1)})\pi_0(\alpha_j^{(t-1)})\pi(\alpha_j^* \mid \alpha_j^{(t-1)})}\right)$,

Sample a random number u from $\tilde{U}(0, 1)$,

if $u \leq p$, accept proposed state α_j^* and set $\alpha_j^t = \alpha_j^*$,

else, set $\alpha_j^t = \alpha_j^{(t-1)}$.

References

[1] Coffin LF. A study of the effects of cyclic thermal stresses on a ductile metal. Trans Am Soc Mech Eng 1954;76:931–50.
[2] Manson SS. Behavior of materials under conditions of thermal stress. National Advisory Commission on Aeronautics: Report1170 Cleveland: Lewis Flight Propulsion Laboratory; 1954.
[3] Tanaka K, Mura T. A dislocation model for fatigue crack initiation. J Appl Mech 1981;48:97–103.
[4] Sangid MD, Maier MD, Sehitoglu H. A physically based fatigue model for prediction of crack initiation from persistent slip bands in polycrystals. Acta Mater 2011;59:328–41.
[5] Sangid MD, Maier HJ, Sehitoglu H. The role of grain boundaries on fatigue crack initiation – an energy approach. Int J Plast 2011;27:801–21.
[6] Yeratapally SR, Glavicic MG, Hardy M, Sangid MD. Microstructure based fatigue life prediction framework for polycrystalline nickel-base superalloys with emphasis on the role played by twin boundaries in crack initiation. Acta Mater 2016;107:152–67.
[7] Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D. Global Sensitivity Analysis: the Primer. Wiley Interscience; 2008.
[8] Smith RC. Uncertainty quantification: Theory, Implementation, and Applications. SIAM Computational Science & Engineering Series: Philadelphia, PA, USA; 2014.
[9] Kennedy MC, O'Hagan A. Bayesian calibration of computer models. J R Stat Soc: Ser B 2011;63:425–64.
[10] Hadjidoukas PE, Angelikopoulos P, Rossinelli D, Alexeev D, Papadimitriou C, Koumoutsakos P. Bayesian uncertainty quantification and propagation for discrete element simulations of granular materials. Comput Methods Appl Mech Engrg 2014;282:218–38.
[11] Madireddy S, Sista B, Vemaganti K. A Bayesian approach to selecting hyperelastic constitutive models of soft tissue Comput. Methods Appl Mech Engrg 2015;291:102–22.
[12] Cheung SH, Oliver TA, Prudencio EE, Prudhomme S, Moser RD. Bayesian uncertainty analysis with applications to turbulence modeling. Reliab Eng Syst Saf 2011;96:1137–49.
[13] Koslowski M, Strachan A Uncertainty propagation in a multiscale model of nanocrystalline plasticity. Reliab. Engrg. Syst. Saf. 96 (1011) 1161.
[14] Angelikopoulos P, Papadimitriou C, Koumoutsakos P. Bayesian uncertainty quantification and propagation in molecular dynamics simulations: a high performance computing framework. J Chem Phys 2012;137:144103.
[15] Zhang R, Mahadevan S. Model uncertainty and Bayesian updating in reliability-based inspection. Struct Saf 2000;22:145–60.
[16] Makeev A, Nikishkov Y, Armanios E. A concept for quantifying equivalent initial flaw size distributions in fracture mechanics based life prediction models. Int J Fatigue 2007;29:141–5.
[17] Cross R, Makeev A, Armanios E. Simultaneous uncertainty quantification of fracture mechanics based life prediction model parameters. Int J Fatigue 2007;29:1510–5.
[18] Sankararaman S, Ling Y, Shantz C, Mahadevan S. Uncertainty quantification in

fatigue crack growth prognosis. Int J Progn Health Manag 2011;2:1–15.
[19] Sankararaman S, Ling Y, Mahadevan S. Uncertainty quantification and model validation of fatigue crack growth prediction. Eng Fract Mech 2011;78:1487–504.
[20] Chiachío M, Chiachío J, Rus G, Beck JL. Predicting fatigue damage in composites: a Bayesian framework. Struct Saf 2014;51:57–68.
[21] Chiachío J, Chiachío M, Saxena A, Sankararaman S, Rus G, Goebel K. Bayesian model selection and parameter estimation for fatigue damage progression models in composites. Int J Fatigue 2015;70:361–73.
[22] Rovinelli A, Lebensohn RA, Sangid MD. Influence of microstructure variability on short crack behavior through postulated micromechanical short crack driving force metrics. Eng Fract Mech 2015;138:265–88.
[23] Panchal JH, Kalidindi SR, McDowell DL. Key computational modeling issues in Integrated Computational Materials Engineering. Comput Aided Des 2013;45:4–25.
[24] Schouwenaars R, Seefeldt M, Van Houtte P. The stress field of an array of parallel dislocation pile-ups: implications for grain boundary hardening and excess dislocation distributions. Acta Mater 2010;58:4344–53.
[25] Stroh AN. A theory of the fracture of metals. Adv Phys 1957;6:418–65.
[26] Taylor GI. The mechanism of plastic deformation of crystals. Proc Roy Soc 1934;145:362–87.
[27] Kiureghian AD, Ditlevsen O. Aleatory Or Epistem* Does it Matter* Struct Saf 2009;31:105–12.
[28] Flage R, Baraldi P, Zio E, Aven T. Probability and Possibility-Based Representations of Uncertainty in Fault Tree Analysis. Risk Anal 2013;33:121–33.
[29] Aven T, Zio E. Some considerations on the treatment of uncertainties in risk assessment for practical decision making. Reliab Eng Syst Saf 2011;96:64–74.
[30] Ezaz T, Sangid MD, Sehitoglu H. Energy barriers associated with slip–twin interactions. Philos Mag 2011;91:1464–88.
[31] Sangid MD, Ezaz T, Sehitoglu H. Energetics of residual dislocations associated with slip–twin and slip–GBs interactions. Mat Sci Eng A 2012;542:21–30.
[32] Abuzaid WZ, Sangid MD, Carroll JD, Sehitoglu H, Lambros J. Slip transfer and plastic strain accumulation across grain boundaries in Hastelloy X. J Mech Phy Solids 2012;60:1201–20.
[33] Collins DM, Stone HJ. A modelling approach to yield strength optimisation in a nickel-base superalloy. Int J Plast 2014;54:96–112.
[34] Li HY, Sun JF, Hardy MC, Evans HE, Williams SJ, Doel TJA, Bowen P. Effects of microstructure on high temperature dwell fatigue crack growth in a coarse grain PM nickel based superalloy. Acta Mater 2015;90:355–69.
[35] Hardy MC, Zirbel B, Shen G, Shankar R. Green KA, editor. Superalloys. The Minerals, Metals & Materials Society; 2004. p. 83.
[36] Huang EW, Barabash RI, Yandong W, Bj C, Li L, Liaw PK, Ice GE, Yang R, Hahn C, Pike LM, Klarstrom DL. Plastic behavior of a nickel-based alloy under monotonic tension and low-cycle-fatigue loading. Int J Plast 2008;24:1440–56.
[37] Huang M, Zhao L, Tong J. Discrete dislocation dynamics modelling of mechanical deformation of nickel-based single crystal superalloys. Int J Plast 2012;28:141–58.
[38] Jiang J, Yang J, Zhang T, Dunne FPE, Britton TB. On the mechanistic basis of fatigue crack nucleation in Ni superalloy containing inclusions using high resolution electron backscatter diffraction. Acta Mater 2015;97:367–79.
[39] Petrenec M, Obrtlík K, Polák J. Inhomogeneous dislocation structure in fatigued INCONEL 713 LC superalloy at room and elevated temperatures. Mat Sci Eng A 2005;400–401:485–8.
[40] Ho HS, Risbet M, Feaugas X. On the unified view of the contribution of plastic strain to cyclic crack initiation: impact of the progressive transformation of shear bands to persistent slip bands. Acta Mat 2015;85:155–67.
[41] Bolado-Lavin R, Castaings W, Tarantola S. Contribution to the sample mean plot for graphical and numerical sensitivity analysis. Reliab Eng Syst Saf 2009;94:1041–9.
[42] Tarantola S, Kopustinskas V, Bolado-Lavin R, Kaliatka A, Uspuras E, Vaisnoras M. Sensitivity analysis using contribution to sample variance plot: application to a water hammer model. Reliab Eng Syst Saf 2012;99:62–73.
[43] Plischke E. An adaptive correlation ratio method using the cumulative sum of the reordered output. Reliab Eng Syst Saf 2012;107:149–56.
[44] Arendt PD, Apley DW, Chen W. Quantification of model uncertainty: calibration, model discrepancy, and identifiability. J Mech Des 2012;134(2012):1–12.
[45] Metropolis N, Rosenbluth A, Rosenbluth M, Teller A, Teller E. Equation of state calculations by fast computing machines. J Chem Phys 1953;21:1087–92.
[46] Hastings WK. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 1970;57:97–109.
[47] ASTM standard, E606/E606M-12, Standard test method for strain-controlled fatigue testing.
[48] Gelman A, Carlin JB, Stern HS, Rubin DB. Bayesian data analysis. New York: Chapman & Hall; 2004. p. 289–98.
[49] Niezgoda SR, Turner DM, Fullwood DT, Kalidindi SR. Optimized structure based representative volume element sets reflecting the ensemble-averaged 2-point statistics. Acta Mater 2010;58:4432–45.
[50] Miroshnikov A, Conlon EM. Parallel MCMC combine: an R package for bayesian methods for big data and analytics. PLoS ONE 2014;9(9):e108425.
[51] Neiswanger W, Wang C, Xing EP. Asymptotically exact, embarrassingly parallel MCMC. In Proceedings of the International Conference on Uncertainty in Artificial Intelligence; 2014.
[52] Mughrabi H. Microstructural fatigue mechanisms: cyclic slip irreversibility, crack initiation, non-linear elastic damage analysis. Int J Fatigue 2013;57:2–8.
[53] Mughrabi H. Fatigue, an everlasting materials problem - still en vogue. Procedia Eng 2010;2:3–26.
[54] Risbet M, Feaugas X, Guillemer-Neel C, Clavel M. Use of atomic force microscopy to quantify slip irreversibility in a nickel-base superalloy. Scr Mater 2003;49:533–8.
[55] Differt K, Essmann U, Mughrabi H. A model of extrusions and intrusions in fatigued metals. Part II: surface roughening by random irreversible slip. Philos Mag A 1986;54:237–58.