Autonomous Development of a Machine-Learning Model for the Plastic Response of Two-Phase Composites from Micromechanical Finite Element Models

ANDREW MARSHALL[1] and SURYA R. KALIDINDI[1,2]

1.—Georgia Institute of Technology, Atlanta, GA, USA. 2.—e-mail: surya.kalidindi@me.gatech.edu

We demonstrate a novel strategy for the autonomous development of a machine-learning model for predicting the equivalent stress–equivalent plastic strain response of a two-phase composite calibrated to micromechanical finite element models. A unique feature of the model is that it takes a user-defined three-dimensional, two-phase microstructure along with user-defined hardening laws for each constituent phase, and outputs the equivalent stress–plastic strain response of the microstructure modeled using J_2-based isotropic plasticity theory for each constituent phase. Previously, this task was addressed using linear regression approaches on a large training dataset. In this work, it is demonstrated that the use of Gaussian process regression together with a Bayesian sequential design of experiments can lead to autonomous protocols for optimal generation of the training dataset and the development of the model. It is shown that this strategy dramatically reduces the time and effort expended in generating the training set.

INTRODUCTION

Reduced-order, low-computational cost, process-structure-property (PSP) linkages play a central role in the efforts related to the design of materials for targeted combinations of effective properties or performance characteristics.[1–5] This is because the materials design space (comprising a broad range of potential chemistries and available options in processing) is typically too large to explore with the computationally expensive physics-based materials modeling toolsets (e.g., micromechanical finite element models, phase-field models). The formulation and use of high fidelity, reduced-order, surrogate PSP linkages offer one of the most promising research avenues for realizing the goals of accelerated materials innovation targeted by the US Materials Genome Initiative.[6] In addition to facilitating rapid exploration of the vast materials design spaces, the low computational cost PSP linkages play an important role in the objective calibration of

the multiscale physics-based models to the available, often limited, experimental data.[7–9] This is because the calibration of parameters present in multiscale models is inherently an inverse problem that also demands many executions of the forward models.

Establishing high-fidelity surrogate PSP linkages requires a suitable framework for the rigorous quantification of the material internal structure (i.e., the quantification of 'S' in the PSP linkages).[10] This is because the material internal structure often spans a hierarchy of material length scales, from the subatomic to the macroscale, with each scale exhibiting a rich variety of defects and/or disorder. Consequently, the rigorous quantification of the material internal structure needs a versatile framework that can address the diverse features observed in the material structure at the different length scales in the different material systems. A second challenge comes from the fact that any rigorous quantification of the material internal structure is likely to produce a large number of statistical measures of the microstructure. The number of microstructure measures constitutes the dimensionality of the material structure representation for the

present discussion (i.e., each microstructure measure corresponds to one dimension). Our goal of formulating low computational cost PSP linkages demands that we seek high-value, but low-dimensional, representations of the material structure. In recent years, our research group has advanced a systematic framework called Materials Knowledge Systems (MKS)[2] for extracting data-driven reduced-order PSP linkages from materials datasets aggregated from either multiscale simulations or multiscale experiments. At the core of MKS lies an approach for high-value, low-dimensional representation of the material internal structure by bringing together the concepts of n-point spatial correlations[11–13] and principal component analysis (PCA).[14,15] The efficacy of the MKS framework in producing high-fidelity reduced-order PSP linkages has been demonstrated in prior work.[16–18]

There exist many opportunities for the further development of the MKS framework. Of particular relevance to this work are the approaches used to build the reduced-order models after the feature engineering step (i.e., after the low-dimensional representations of the material structure are obtained employing n-point spatial correlations and PCA). In most of the prior work,[16,17,19,20] surrogate models were built using standard regression techniques. These techniques are indeed ideally suited when there is a reasonably large dataset available, and there is already evidence (or physical insights) suggesting a specific model form for capturing the governing physics of the problem. In many practical applications of the MKS framework for a broad range of multiscale materials phenomena, one or both of these requirements are not met. There is therefore a critical need to expand beyond the standard regression techniques. In the present work, we specifically explore the utility and benefits of employing modern machine learning (ML) tools, specifically Gaussian process regression (GPR),[21] in establishing the desired surrogate PSP linkages.

Several benefits are anticipated from the use of GPR in the MKS framework for extracting PSP linkages from materials data. Unlike standard regression, GPR does not demand the aggregation of a sufficiently large dataset before building the surrogate model. In other words, one can start with a very small number of data points and gradually build the desired PSP linkages as more data become available. Furthermore, since GPR provides a rigorous assessment of the uncertainty (i.e., variance) in the predictions made from the surrogate model, it can provide objective guidance on where one should generate new training data points. This is particularly important when the data are being generated using expensive computations or experiments. In the basic regression techniques, one usually aims to spread the data points uniformly in the input domain. Such a simple strategy is unlikely to be optimal in terms of maximizing the predictive accuracy of the extracted surrogate model. Using GPR, one can select inputs for generating new training data based on systematically minimizing the uncertainty of the model predictions over the entire input domain of interest. One can therefore implement a dynamic selection process that leads to autonomous (without the need for human intervention) and economical workflows for the extraction of robust surrogate models for any selected physics-based multiscale materials model. A second major advantage of using GPR is that that there is no need for a priori selection of a specific parametric model form. GPR employs specialized kernels[22,23] on the training data that serve as efficient interpolators. These kernels have hyperparameters that are optimized to produce the most reliable surrogate models from the available training data.

GPR has already been used successfully to address a limited set of materials problems. Bélisle et al.[24] presented a scalable GPR approach for predictions of martensite start temperature and electrical conductivity. Tapia et al.[25] used GPR to predict porosity in stainless steel parts fabricated in an additive manufacturing process, and to predict the melt pool depth.[26] GPR was used by Hoang et al.[27] to predict the compressive strength of concrete samples based on composition and process variables. Yabansu et al.[18] used GPR models to predict the microstructural evolution of Ni-based superalloys in a thermal aging process. Tallman et al.[28] used a GPR-based approach to drive data collection for a homogenization model in order to study deformation of α-Ti. Hashemi et al.[29] used an autoregressive GPR model to predict the time evolution of structure in FCC metals during static recrystallization processes. This work is specifically aimed at using GPR for the autonomous extraction of PSP linkages.

The main goal of this study is to develop and demonstrate GPR-based MKS protocols for autonomous and economical extraction of reduced-order PSP models for the bulk (effective) plastic response of multiphase materials. Specifically, our focus will be on two-phase composites, where each phase exhibits an isotropic plastic response. This specific case study was selected because the viability of the standard MKS regression techniques for this problem has already been established.[19] Therefore, it offers an opportunity to evaluate critically the computational cost of training the new GPR-based MKS models and their predictive accuracy compared to the previously established models based on standard regression methods.

METHODS

Mean-Field Theory Approximations for Plastic Response of Composites

Mean-field theories approximate the effective response of composites by assuming uniform stress and/or strain fields within the constituents.[30,31] The specific approach employed in this work builds on

Autonomous Development of a Machine-Learning Model for the Plastic Response of Two-Phase Composites from Micromechanical Finite Element Models

2087

the approach described initially by Stringfellow et al.,[32,33] and more recently by Latypov et al.[19] The problem of interest here is the homogenized equivalent stress–equivalent plastic strain response of a composite material made of two isotropic constituents (i.e., thermodynamic phases), both following a J_2-based associated flow rule[34] with distinct hardening responses. The composite microstructure is assumed to be adequately captured in a voxelized 3-D representative volume element (RVE).[34,35] The main state variable of interest in these models is the resistance to plastic deformation, henceforth referred to as strength, and denoted by g. During any imposed deformation, the evolution of strength (due to strain hardening) at any material point in the RVE can be expressed as:

$$\dot{g} = h\dot{\epsilon}, \tag{1}$$

where $\dot{\epsilon} = \sqrt{\frac{2}{3}\mathbf{D}\cdot\mathbf{D}}$ is the equivalent plastic strain rate, and h represents a hardening law expressed typically as a nonlinear function of the strength, i.e., $h = h(g)$. In the present application, we expect the two phases present in the composite to exhibit two distinct hardening laws denoted as $h_1(g)$ and $h_2(g)$, where the subscripts index the different phases present in the composite.

Prior work employing mean-field approaches[32,33] has demonstrated the remarkable utility of simplified approaches for predicting the overall response of the composite by tracking only the volume averaged values of the plastic strain rates and strengths for each phase. Following the prior work of Stringfellow and Parks[33], we assume here that Eq. 1 can be applied to phase–volume averaged quantities as (no summation is implied on the repeated index):

$$\dot{g}_k = h_k(g)\dot{\epsilon}_k \tag{2}$$

where g_k and $\dot{\epsilon}_k$ now denote the volume averaged strength and the volume averaged equivalent plastic strain rate, respectively, for the kth phase. With this simplifying assumption, the problem of predicting the overall response of the composite can be addressed by partitioning the overall imposed deformation into the phase–volume averaged quantities, $\dot{\epsilon}_k$. The phase–volume averaged strengths in each phase, g_k, are updated using Eq. 2, and the effective (composite) strength, $\overline{g}$, is computed. Stringfellow and Parks[33] addressed these challenges using the self-consistent approach based on the Eshelby solution for a spherical inclusion in a homogeneous effective medium. Although the approach yields a simple model, it only incorporates highly simplified information on the material microstructure, i.e., volume fractions of the constituents.

In the recent work of Latypov et al., [19] the approach of Stringfellow and Parks[33] was extended to include the higher-order details of the material microstructure. This was accomplished by extending the MKS framework[2] as shown schematically in Fig. 1. In this new strategy, calibrated MKS models are established to predict the strain rate partitioning ratios for the soft and hard phases (denoted as χ_1 and χ_2, respectively). The ratios are defined through $\dot{\epsilon}_k = \chi_k\dot{\overline{\epsilon}}$, where $\dot{\overline{\epsilon}}$ denotes the macroscopically imposed equivalent plastic strain rate on the RVE. The inputs for these models are the microstructure's reduced-order representation (PC scores) and strength contrast, $\eta = g_2/g_1$. A key underlying assumption in this approach is that the changes in morphology of the microstructure are minimal for the small plastic strains imposed; therefore, the PC scores are kept constant for each prediction of the stress–plastic strain response. Once the volume-averaged strain rates $\dot{\epsilon}_k$ are computed, the strain hardening in each phase is computed using Eq. 2, and the strengths of each phase are updated as $g_k(\tau) = g_k(t) + \dot{g}_k\Delta t$, where $\Delta t = (\tau - t)$ is the time increment. The value of the strength contrast η is then updated and used in a different MKS model developed specifically to predict the effective yield strength of the composite ($\overline{g}$). This MKS model takes the PC scores of the microstructure and the strength contrast, η, as inputs. Repeating the computations described above, one can then predict the equivalent stress–equivalent plastic strain curve for the RVE with high fidelity and very low computational cost. The PC scores used as inputs to the effective yield strength MKS model are the same PC scores used as inputs by the other two MKS models described above. The only difference is that the value of η used as input to the yield strength MKS model is the updated value after the prediction of the strain rate partitioning ratios. The viability of this approach was demonstrated in prior work.[19]

Full Field Numerical Evaluation of Micromechanical Responses

Datasets are needed to train and subsequently test the GPR models built in this study. These were generated using finite element (FE) micromechanical simulations on digitally generated 3-D RVEs.[19,36,37] The results from the micromechanical FE simulations are treated as the ground-truth for building the GPR models. The protocols used for generating the datasets are essentially the same as those used in prior work,[19] and are briefly described next.

Commercial FE software Abaqus[38] was used for generating the datasets. Each voxel of the 3-D RVE was converted into a C3D8 element in the FE model. A J_2 rate-independent plasticity model based on the von Mises yield criterion was selected for this study. Each RVE was subjected to a velocity gradient tensor $(\overline{L})$, defined as $\overline{L}_{11} = 10^{-3}\,\text{s}^{-1}$ and $\overline{L}_{22} = \overline{L}_{33} = -\overline{L}_{11}/2$, with all off-diagonal components set to zero. These are imposed as periodic velocity boundary conditions on the RVE. To minimize the effects of elasticity, the elastic modulus of
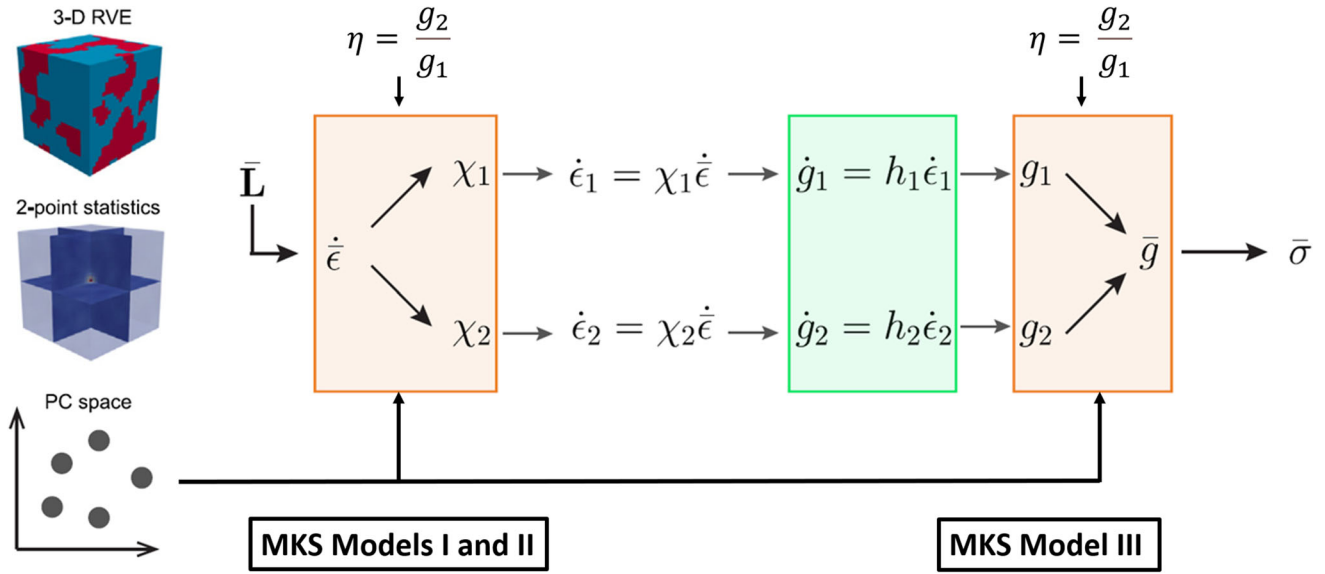
Fig. 1. The MKS plasticity framework applies a velocity gradient tensor to the RVE. Predictions are then made for the strain partitioning ratios (via MKS Models I and II) and then for effective yield strength (via MKS Model III). The phase-averaged strain rates are utilized in updating the yield strengths of each constituent phase. Adapted from.[19]

each phase was selected to be 10^6 times larger than its yield strength.

The MKS models identified in Fig. 1 take only the microstructure and the strength contrast as inputs and output the strain rate partitioning ratios and the effective strength. The data needed to train these models only require FE simulations performed to small strain levels. Specifically, digitally generated microstructures were deformed to an equivalent strain of 0.002. The constituent phases were assigned strength values without hardening to enforce the desired strength contrast ratio (one of the inputs to the MKS models to be built).

Test datasets were also generated the critical evaluation of the GPR-based framework depicted in Fig. 1. For this purpose, FE simulations were conducted on selected digitally created microstructures to an equivalent strain of 0.1. The hardening laws for both constituent phases were described by a power law as:

$$\dot{g}_k = h_k^0 \left(1 - \frac{g_k}{g_k^\infty}\right)^{a_k} \dot{\epsilon}_k \qquad (3)$$

where h_k^0, g_k^∞, and a_k denote hardening parameters for the phase k. These parameters were selected to correspond to literature values reported for ferrite and martensite,[39] and are summarized in Table I.

Following the Hill–Mandel principle,[40,41] the macroscopic stress was computed as the volume-averaged stress tensor over the RVE. The equivalent stress was then computed using the von Mises definition, which is the same as effective strength in the rate-independent plastic theory used in this study. The equivalent plastic strain rate is computed at each integration point in the FE mesh, and the strain rate partitioning ratios, χ_k, are obtained

Table I. Values of quantities used for strain hardening law calculations.

Quantity	Ferrite (k = 1)	Martensite (k = 2)
$g_k(0)$	285 MPa	974.4 MPa
g_k^∞	1236 MPa	2330.4 MPa
h_k^0	3.0 GPa	1351.2 GPa
a_k	2.25	2.25

These values were obtained from a prior study by Tasan et al. [39]

by dividing the average equivalent strain rate in the kth phase by the macroscopically imposed equivalent strain rate.

Generation of Three-Dimensional (3-D) Digital Microstructure Volumes

Voxelized 3-D RVEs reflecting a diverse range of morphological features were generated to serve as the training and test data, where each voxel is assigned to either the soft phase or the hard phase. The RVE size was selected to be $27 \times 27 \times 27$, which produced good results in previous work.[19] Two specific methods were used in this study for the generation of a large ensemble of diverse RVEs for the present study. The first technique generated RVEs through the use of a Gaussian filter,[42] where each voxel is assigned a value drawn randomly from a uniform distribution over the range (0,1). A 3-D Gaussian filter with diagonal covariance matrix, Σ, is applied to the random field via circular convolution to obtain a smooth (locally averaged) field.

Thresholding is then applied to assign label 1 or 2 (denoting a soft or hard phase, respectively) to each voxel; this process results in a two-phase 3-D RVE. The thresholding value determines the phase volume fractions, while the diagonal values of Σ control the average phase size and shape. By suitably changing the diagonal values of Σ, one can generate a diverse range of morphologies ranging between equiaxed morphologies to plate-like morphologies to rod-like morphologies, with volume fractions in the range (0,1). A total of 512 RVEs were generated using this strategy. The top row of Fig. 2 shows example RVEs generated using the Gaussian filter approach.

The Universal Microstructure Generation Framework (UMGF) developed recently[43] was used as a second approach for generating the digital RVEs used in the present study. This approach utilizes a user-specified cost function to identify the optimal placement of objects in the microstructure. Objects to be used in the microstructure may be called from a library or defined by the user, allowing for tremendous morphological flexibility. This method is also computationally efficient, because it uses convolution filters for carrying out many of the computations involved. The UMGF approach used in this work systematically controls object overlap; this feature allows it to generate morphologies that are unobtainable using the Gaussian filter approach. A total of 286 RVEs were generated using the UMGF approach. Examples are presented in the bottom row of Fig. 2, where D represents the sizes of objects used in the generation.

For each RVE, 10 different values of η, corresponding to integers between 1 and 10, were assigned. As a result, the total number of distinct inputs generated for the models was 7980.

Microstructure Quantification

Voxelized microstructures can be conveniently represented by the array, m_s^k, which reflects the volume fraction of local state, k, in voxels, s.[10] Since each voxel is assigned entirely to a single phase, the values of m_s^k are either 0 or 1 for the RVEs studied. These microstructures can be featurized using two-point spatial correlations and PCA.[14,15] For two-phase microstructures, the autocorrelation of a single local state provides an adequate set of spatial correlations.[44,45] These are mathematically defined as:

$$f_t^{kk} = \frac{1}{|S|} \sum_s^S m_s^k m_{s+t}^k \qquad (4)$$

where f_t^{kk} is the probability of finding local state, k, in spatial bins separated by a specified discretized vector, t.[46] For this study, $k = 1$ was selected; this has no consequence on the results.[47] Fast Fourier
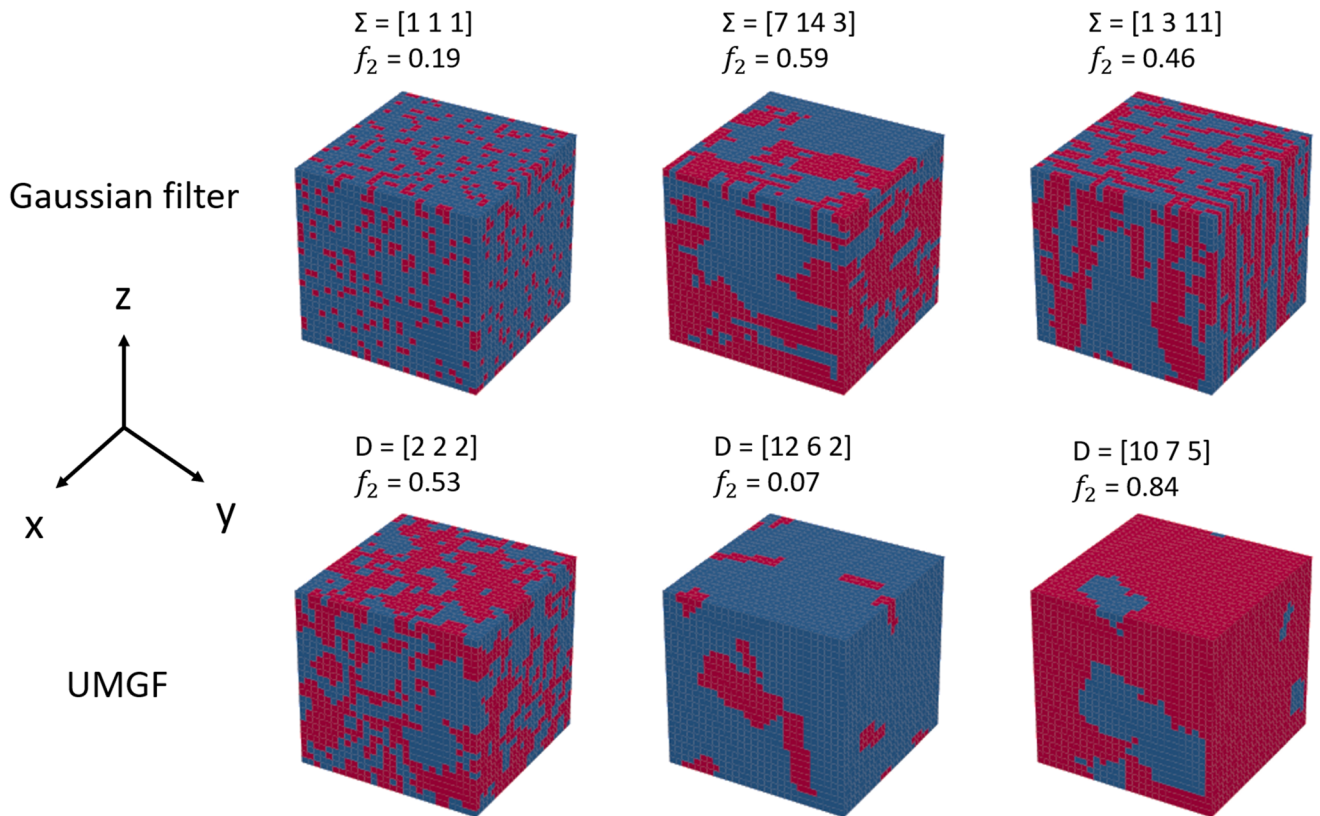

Fig. 2.. (Top) Example RVEs generated with the Gaussian filter method. (Bottom) Example RVEs generated with the UMGF method. For both methods, the ensemble was generated to target different hard phase volume fractions (f2) and morphologies.

transform algorithms allow highly efficient computation of the autocorrelations.[46] Figure 3 shows four example RVEs exhibiting different phase volume fractions and morphologies. The figure also presents three orthogonal sections of the corresponding 3-D autocorrelations. The value at the origin of the 3-D autocorrelations reflects the phase volume fraction, and the averaged inclusion size and shape is reflected in the contours surrounding the central peak. The first set of 3-D autocorrelations in Fig. 3 shows a very small central peak because its corresponding RVE has very small inclusions. The second set of 3-D autocorrelations in Fig. 3 clearly show a vertical fiber as the central peak, reflecting the dominant morphology of its RVE. Furthermore, the spacings between the vertical bands in this set of autocorrelations reflect the distributions of spacings between vertical fibers in the RVE. The third and fourth sets of autocorrelations reflect coarse equiaxed features in their corresponding RVEs.

Although two-point correlations provide rigorous statistical descriptions for the microstructure, they result in high dimensional representations (each value of f_t^{kk} is treated as a dimension). In the present application, the dimensionality of the autocorrelations is 19,683, which is the number of voxels in the generated RVEs. As a result, dimensionality reduction becomes essential for the development of ML models. In the MKS framework, dimensionality reduction is achieved performing PCA on the full ensemble of microstructures. The autocorrelations of the nth microstructure (the superscripts on the two-point statistics have been dropped since we selected $k = 1$) can then be expressed as:

$$f_t^{(n)} \approx \sum_{j=1}^{R^*} \alpha_j^{(n)} \varphi_{jt} + \overline{f_t} \qquad (5)$$

where $\alpha_j^{(n)}$ denote PC scores, φ_{jt} describes the basis vectors of the PC space, $\overline{f_t}$ denotes the ensemble mean, and R^* denotes the truncation level (i.e., the reduced dimensionality of the microstructure representation). Several prior applications of MKS have shown that a small number of PCs (typically 3–8) are adequate for capturing PSP linkages with high fidelity.[17,20,36,37]

The PC space representation of the 798 digital microstructures generated for this study is presented in Fig. 4 in multiple subspaces. It is noted that the two microstructure generation methods produced microstructures occupying different regions, even in the 3-D PC subspace. Each PC score reflects the strength of the corresponding PC basis (φ_{jt} in Eq. 5). Each PC basis essentially reflects a normalized and weighted set of autocorrelations. Figure 4 also presents the basis maps for the first two PC scores, along with example RVEs corresponding to selected locations, from which it is clear that an increase in the value of PC1 results in an increase in the phase volume fraction, while a decrease in the value of PC2 results in the phase regions being elongated in the y-axis while also becoming coarser. While some of the microstructural features represented in the PC scores may be readily interpreted from their PC basis maps, there will be many other features in the high-dimensional (19,683 for the present case) PC basis that are not easily interpreted.

A general practice is to use a truncated set of PC scores in the model building effort. Percent variance capture is a useful measure for selecting the number of PC scores, since it quantifies the differences between microstructures. However, this is not necessarily the only criterion: some of the PC scores contributing to only a small fraction of variance can exhibit a strong correlation to the effective property
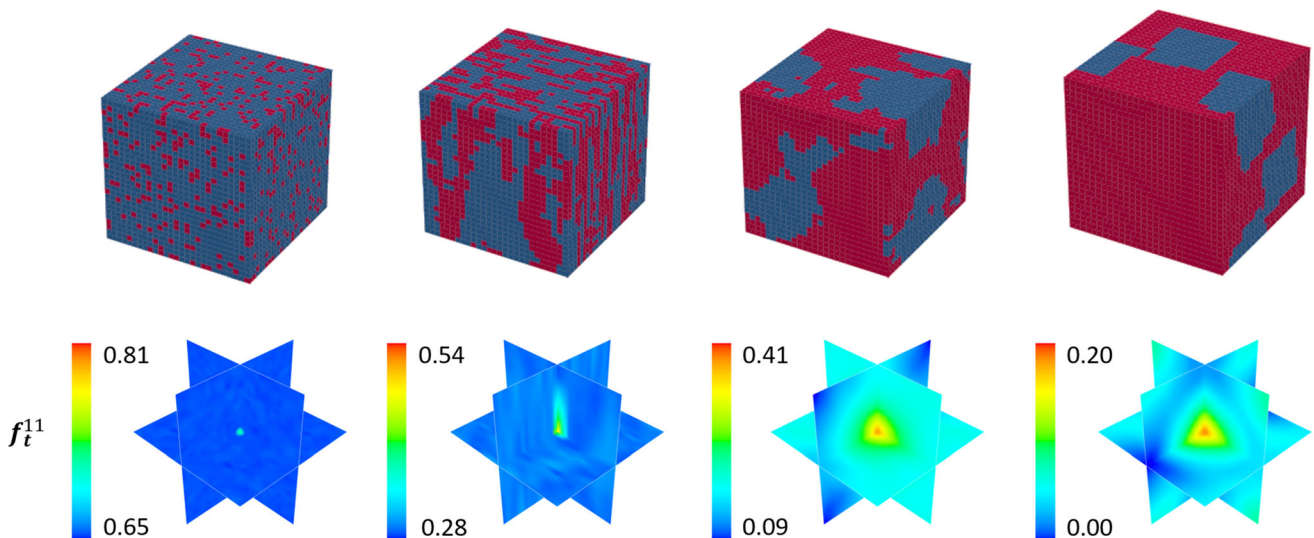

Fig. 3. Example RVEs (top row) and their autocorrelations of the soft phase (bottom row). Only three orthogonal sections of the 3-D autocorrelations are shown. The autocorrelations contain information such as the hard phase volume fraction (value at the origin), morphology, and spacing.

Autonomous Development of a Machine-Learning Model for the Plastic Response of Two-Phase Composites from Micromechanical Finite Element Models
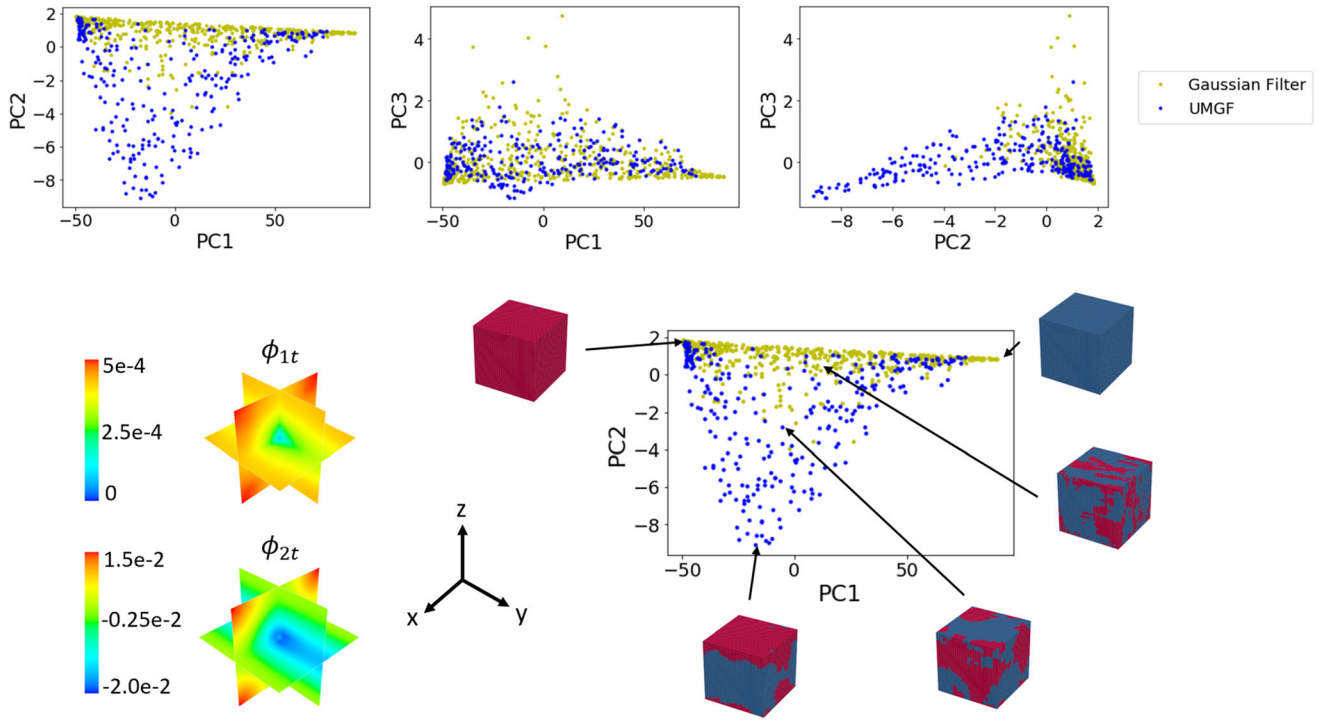
2091

Fig. 4. Projections of the microstructure ensemble in PC space: (Top left) PC1-PC2 projection, (Top Middle) PC1-PC3 projection, (Top right) PC2-PC3 projection. (Bottom left) 3-D basis vectors for first two PC scores. (Bottom right) Illustration of RVEs on the PC1-PC2 projection of the PC space. It may be observed that PC1 strongly correlates with soft phase volume fraction, while PC2 strongly correlates with elongation of soft phase domains along the *y*-axis.

of interest. Based on the prior study,[19] we have decided to use the first 10 PC scores in this study. These 10 PCs account for over 99.9% of the variance between the elements of the microstructure ensemble utilized for this study.

Gaussian Process Regression (GPR)

GPR is a nonparametric ML approach to building surrogate models. A Gaussian process[21] is completely defined by specifying its mean and covariance. Therefore, the target, y, can be expressed as

$$y = \hat{y} + \epsilon, \hat{y} \sim N(\mu(\boldsymbol{x}), \boldsymbol{K}(\boldsymbol{x}, \boldsymbol{x}')) \quad (6)$$

where $\hat{y}$ denotes the predicted value, ϵ is the residual, $\boldsymbol{x}$ is the vector of inputs, and $\mathcal{N}(\mu, \boldsymbol{K})$ is a multivariate Gaussian distribution with a mean μ and a covariance $\boldsymbol{K}$. Generally, the mean of a Gaussian process is assumed to be zero. The covariance is generally estimated using a kernel function designed for efficient interpolation between the training data points. An automatic relevance determination squared exponential (ARD-SE) kernel[21, 22] is used in this study, and is expressed as:

$$k(\boldsymbol{x}, \boldsymbol{x}') = \sigma_f^2 \exp\left(-\frac{1}{2} \sum_{d=1}^{D} \frac{(x_d - x_d')^2}{l_d^2}\right) + \sigma_n^2 \delta_{xx'} \quad (7)$$

where σ_f, l_d, and σ_n serve as hyperparameters controlling the accuracy of the predictions; σ_f scales the variance associated with the predictions, with

increasing the value of σ_f increasing the probability of capturing outlier values, but also increasing the prediction noise; l_d is an interpolation hyperparameter, where low values of l_d imply short-range interpolations (and noisier predictions); and σ_n controls the homoscedastic Gaussian noise applied to predictions. The central feature of the ARD-SE kernel is that it allows different interpolation parameters for the different input variables. This feature allows the model to capture the different levels of contributions of the different inputs to the accuracy of the prediction, improving its fidelity and interpretability. The main task in building a reliable GPR model reduces to the estimation of the values of the hyperparameters in Eq. 7. This task is generally approached as an optimization problem based on the maximization of the joint log marginal likelihood of the hyperparameters, given the training dataset. In this work, a quasi-Newton method, known as the limited-memory Broyden–Fletcher–Goldfarb–Shanno algorithm[48–50] is used to optimize the hyperparameters in Eq. 7.

Once the optimized hyperparameters have been established, the mean ($\boldsymbol{\mu}^*$) and covariance ($\boldsymbol{\Sigma}^*$) of the Gaussian predictive distribution for test inputs $\boldsymbol{x}^*$ are computed as:

$$\boldsymbol{\mu}^* = \boldsymbol{k}^{*T} \boldsymbol{K}^{-1} y \quad (8)$$

$$\boldsymbol{\Sigma}^* = K^* - \boldsymbol{k}^{*T} \boldsymbol{K}^{-1} \boldsymbol{k}^* \quad (9)$$

where k^* is the covariance between the training inputs x and the test inputs x^*, K is the covariance within the training inputs, and y denotes the target responses for the training inputs. Superscript T denotes the transpose operation.

Dynamic Selection Methods

As already discussed, our main focus is on optimal generation of the training data for establishing surrogate PSP models from expensive physics-based micromechanical FE models. This will be accomplished through dynamic sequential selection of the next input to the physics-based simulations. Specifically, we will employ a relatively simple approach based on the highest predictive variance, known as the mean squared error (MSE) approach.[51–54] In this approach, a single sample point is added in each selection cycle in a sequential manner,[52] and may be interpreted as seeking the maximization of information added to the surrogate model.

AUTONOMOUS PROTOCOLS FOR FORMULATION OF THE MKS MODELS

The present application requires the establishment of three different surrogate models, identified in Fig. 1 as MKS Models I, II, and III. In prior work,[19,36] these were calibrated using all of the available RVEs and their corresponding micromechanical responses, without considering the order in which new data should be added to the training set. By contrast, this work seeks to develop protocols for the autonomous generation of training data that optimally improve the surrogate models being built. This is accomplished using a variant of the MSE approach described in "Dynamic Selection Methods" section, where we identify one new set of inputs for each model being built. In other words, in each cycle of training data generation, we have identified three different sets of inputs: each set is expected to maximize the information gain for each of the MKS Models I, II, and III, respectively. Additionally, when a candidate input is selected, it is removed from the candidate set so that it cannot be selected again in future data generation cycles. Note that the input set for all three MKS models comprises the same set of variables, which include the first 10 PC scores of the microstructure (see "Microstructure Quantification" section) and the strength contrast value, η. A total of 7980 distinct inputs were established in "Generation of Three-Dimensional (3-D) Digital Microstructure Volumes" section for this purpose. Before applying GPR, each input variable was scaled to exhibit a unit standard deviation over the complete set of inputs; this is to improve the optimization of the hyperparameters in Eq. 7. A subset of 500 inputs from the total set of 7980 inputs were set aside as the test set. The test inputs were selected such that they uniformly covered the eleven-dimensional input space. Therefore, we were left with 7480 inputs as the candidate training set for the present work.

Figure 5 provides a schematic of the autonomous workflow developed and implemented in this work to build the desired surrogate models. This workflow was comprised of the following steps:

1. An initial training dataset was created by selecting 10 inputs from the set of 7480 candidate inputs. FE simulations were conducted for these, and the results were used to train the initial GPR-based MKS Models I, II, and III.
2. The single-point MSE approach was used to select three new inputs (each aimed at maximizing the fidelity of each MKS model) from the remaining candidate set.
3. FE simulations were performed for the three new inputs, and the results were added to the training dataset.
4. All three GPR models were updated using the enhanced training set.
5. Steps 2–4 were repeated until the model exhibited the desired predictive accuracy on the test set (not included in the training set).

Figure 6 demonstrates the improvement in the GPR-based MKS models with each model update, as measured by the errors on the test set. It is seen that the proposed workflow is quite efficient in systematically improving the MKS models. For comparison, Fig. 6 also shows the model improvement with random selection of training points. It can be seen that the single-point MSE approach used in this work significantly outperforms the random selections (repeated 10 times). In terms of error, there are only a very few times in which the random selection outperforms the single-point MSE selection; even in these rare cases, the advantage is not seen in all three models simultaneously, and the benefit is often lost in the next few updates of the model. There is a clear benefit to the sequential design approach in the selection of training points. The models appear to learn quickly, displaying a roughly asymptotic behavior. The training in the present work was stopped after using 190 training points. For all models, the dynamic selection outperformed the random selection after this number of training points.

The final model errors were found to be 3.3%, 7.5%, and 6.1% for $\overline{g}$, χ_1, and χ_2, respectively. These errors are higher than those obtained in prior work,[19] where errors of 0.6%, 5.0%, and 4.7%, respectively, were reported after using 1140 training points and 380 test points. However, the RVEs used in the test set for that work covered a much smaller range in PC space compared to the test set used for the present work. Even with much fewer training points and a test set covering a larger range PC space, the χ_1 and χ_2 errors are only slightly higher for this work compared to those
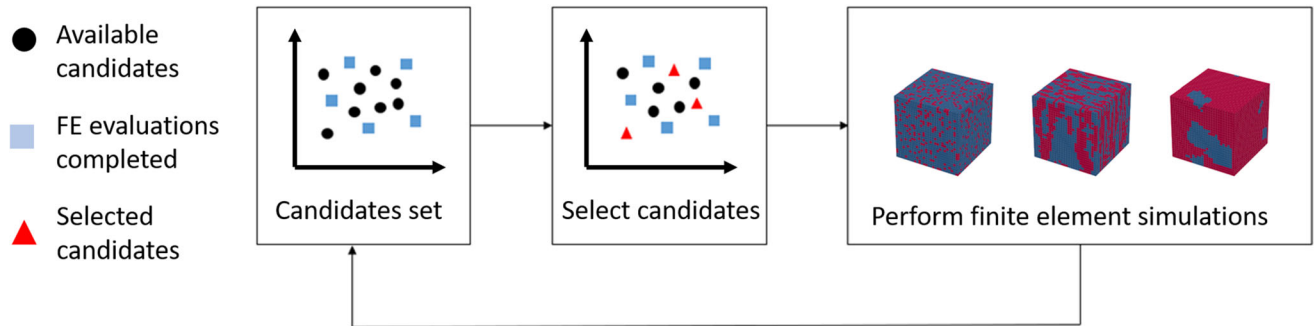
Fig. 5. Schematic describing the autonomous workflow for the development of GPR-based MKS Models identified in Fig. 1. The workflow involves three main steps: (1) establishing the candidate set of inputs that may be selected for FE evaluation, (2) identifying the next set of inputs to be evaluated using FE models, and (3) performing the FE simulations and adding them as new training data. This cycle of steps is repeated as many times as needed until the models exhibit the desired predictive accuracies.
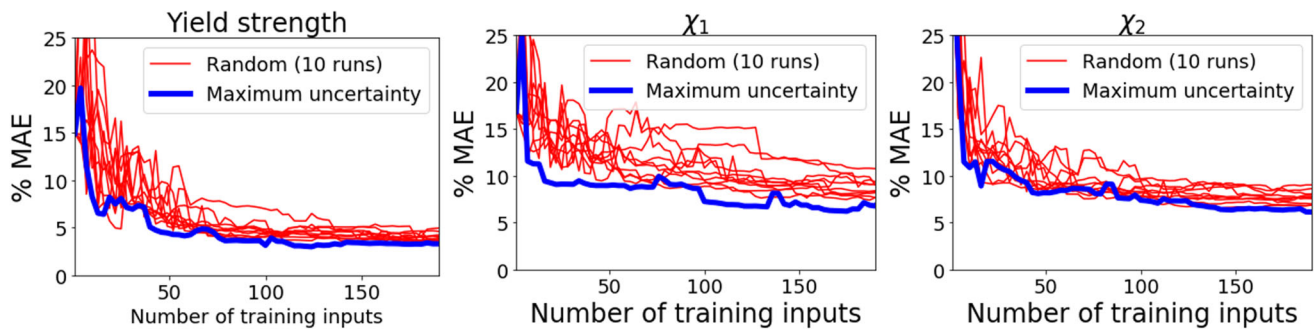

Fig. 6. Model learning efficiency of max uncertainty versus random selection approaches for a test set of 500 inputs. The max uncertainty selection scheme shows faster error convergence compared to random selection for all models generated in this study.

obtained previously. Clearly, the dynamic selection approach utilized in this work produced significant economy in the training effort (only 190 training points were used in the present work). Figure 6 also shows that the model error does not always decrease in subsequent cycles of model improvement. This is mainly because of the simplicity of the single-point MSE approach. Other approaches based on maximization of information gain are likely to improve on this aspect.

If one aims to improve the model accuracy, one might try to include more PC scores (10 PC scores were used in this work) or explore other feature engineering approaches (PCA on two-point statistics was used in this work) combined with other surrogate model building approaches (GPR was used in this work).

PREDICTION OF THE STRESS–STRAIN RESPONSES WITH UNCERTAINTY QUANTIFICATION

Prior work,[19] based on simple polynomial-based regression approaches for establishing the three MKS models, made predictions only for the expected stress–strain responses of new microstructures, without quantifying their prediction uncertainty. Replacing the polynomial regression models with GP-based models not only allows for a richer representation of the underlying model form but also allows for the quantification of the prediction uncertainty.

The overall workflow for the prediction of the equivalent stress–equivalent plastic strain response follows the workflow shown in Fig. 1, and was summarized in "Mean-Field Theory Approximations for Plastic Response of Composites" section. The stress–strain response is computed by performing 200 increments, with each increment corresponding to an increase of 5×10^{-3} in the equivalent plastic strain. For each strain step, the expected value of the equivalent stress is obtained using the GPR-based MKS models I, II, and III following the workflow described earlier. The main changes in the workflow are related to the computation of the prediction uncertainty and its propagation throughout the sequential strain increments in the predicted stress–strain response.

In the workflow employed in this work, the prediction uncertainty is propagated to the stress–strain response by sampling the variables of interest. At the beginning of each strain step, variables χ_k are sampled from their respective distributions and evolved to the end of the time step, providing a distribution of stress values at the end of the time step. Standard deviation in the predicted stress value is then computed from the sampled values. In the first time step, the phase strengths are assumed

to be known with zero variance. For all subsequent strain steps, the GPR models provide the information on the variances for both χ_k, which are used to compute the phase strengths. It was found adequate to represent each distribution at each strain step using 10 samples (i.e., a higher number of samples did not significantly influence the computed variance on the predicted stress value). Sampling from the normal distributions was performed using built-in functions in the Numpy module[55] of Python. A total of 1000 samples are obtained from each of the $10\overline{g}$ values computed in each step, yielding a population of $10,000\overline{g}$ samples. The mean and variance of this population are computed to establish the prediction mean and its uncertainty.

The accuracy of the predicted stress-strain responses was validated using multiple RVEs that were not in the training set. Three examples of these RVEs are shown in Fig. 7, exhibiting hard phase volume fractions of 0.25, 0.35, and 0.75. The phases were assigned the hardening parameters in Table 1, corresponding to dual-phase steels. Note that the initial strength contrast for these examples is 3.42 (= 974.4/285), which is distinct from the values used in the training set. Figure 7 shows the predicted equivalent stress–equivalent plastic strain responses using the protocols developed in this work. These predictions not only include the expected values but also their associated uncertainties (depicted using ± 2 standard deviations). The predictions are compared with corresponding results from FE simulations (considered as ground-truth).

The GPR-MKS framework developed in this work shows good agreement with the FE-predicted stress–strain responses for the very different microstructures shown in Fig. 7. Note that the modeling framework developed here is applicable to virtually any two-phase microstructure 3D RVE with arbitrary assignment of hardening laws for each phase. Most importantly, the GPR-MKS framework is able to generate the full stress-strain curve prediction in under 3 s on a personal laptop with an Intel i5 processor, compared to approximately 4 h required for the FE simulation, executed on the PACE computing cluster at Georgia Tech. Therefore, it is clear that the proposed GPR-MKS framework provides remarkable savings in computational cost, while retaining a high degree of accuracy in its predictions.

SUMMARY

This work has demonstrated the benefits of GPR approaches for the development of low computational cost MKS surrogate models for the prediction of the equivalent stress–equivalent plastic strain responses of arbitrary (user-specified) 3D RVEs of two (isotropic)-phase microstructures. It has been shown that the use of GPR approaches allows for optimal training of the surrogate models with a minimal number of training data points. This is particularly useful for building surrogate models calibrated to computationally expensive physics-based simulation data. In this work, this was accomplished using a simple single-point maximum variance-based approach. It has been shown that the dynamic selection strategy significantly outperforms the random selection of the training data, especially when large-dimensional input domains are involved.
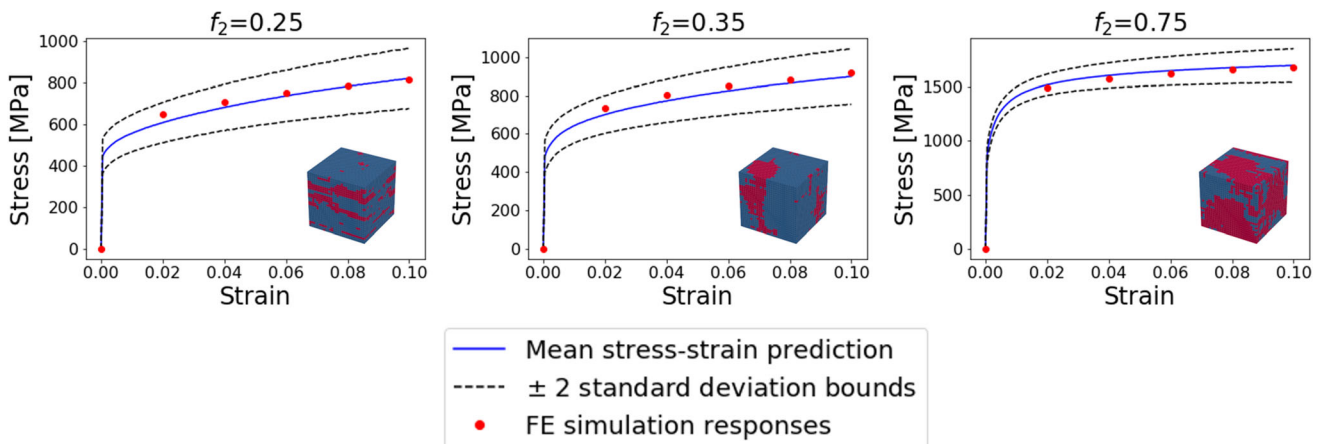

Fig. 7. Comparison of the predicted equivalent stress–equivalent plastic strain responses for three different microstructures. The mean prediction is shown by the center line, the ± 2 SD (4 standard deviations in total) showing prediction uncertainty are shown by the outer lines, and the FE simulation results (i.e., ground-truth) are shown as circles.

Autonomous Development of a Machine-Learning Model for the Plastic Response of Two-Phase Composites from Micromechanical Finite Element Models

2095

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

1. S. Ganapathysubramanian, and N. Zabaras, *Comput. Method Appl. M* 193, 5017. (2004).
2. S.R. Kalidindi, *Hierarchical Materials Informatics: Novel Analytics for Materials Data* (Elsevier, Amsterdam, 2015).
3. McDowell, D.L. and G.B. Olson, Concurrent design of hierarchical materials and structures, in *Scientific Modeling and Simulations*. (Berlin, Springer, 2008). p. 207.
4. D.L. McDowell, J. Panchal, H.-J. Choi, C. Seepersad, J. Allen, and F. Mistree, *Integrated Design of Multiscale, Multifunctional Materials and Products* (Butterworth-Heinemann, Oxford, 2009).
5. G.B. Olson, *J. Comput.-Aided Mater. Des.* 4, 143. (1998).
6. Materials Genome Initiative for Global Competitiveness., N.S.a.T. Council, Editor. 2011.
7. B. Aashranth, S. Kumar, D. Samantaray, and U. Borah, *JOM* 71, 2705. (2019).
8. A.R. Castillo, and S.R. Kalidindi, *Front. Mater.* 6, 136. (2019).
9. K. Pierson, A. Rahman, and A.D. Spear, *JOM* 71, 2680. (2019).
10. S.R. Niezgoda, Y.C. Yabansu, and S.R. Kalidindi, *Acta Mater.* 59, 6387. (2011).
11. Adams, B.L., P. Etinghof, and D.D. Sam. Coordinate free tensorial representation of n-point correlation functions for microstructure by harmonic polynomials. in Mater. Sci. Forum. 1994. Trans Tech Publ.
12. S. Torquato, *Random heterogeneous materials: microstructure and macroscopic properties* (Springer, Berlin, 2013).
13. B.L. Adams, S.R. Kalidindi, and D.T. Fullwood, *Microstructure Sensitive Design for Performance Optimization* (Elsevier Science, Oxford, 2012).
14. I. Jolliffe, *Principal Component Analysis* (Springer, Berlin, 2011).
15. S. Wold, K. Esbensen, and P. Geladi, *Chemom. Intell. Lab. Syst.* 2, 37. (1987).
16. N.H. Paulson, M.W. Priddy, D.L. McDowell, and S.R. Kalidindi, *Acta Mater.* 129, 428. (2017).
17. E. Popova, T.M. Rodgers, X. Gong, A. Cecen, J.D. Madison, and S.R. Kalidindi, *Integr. Mater. Manuf. Innov.* 6, 54. (2017).
18. Y.C. Yabansu, A. Iskakov, A. Kapustina, S. Rajagopalan, and S.R. Kalidindi, *Acta Mater.* 178, 45. (2019).
19. M.I. Latypov, L.S. Toth, and S.R. Kalidindi, *Comput. Method Appl. M* 346, 180. (2019).
20. Y.C. Yabansu, P. Steinmetz, J. Hötzer, S.R. Kalidindi, and B. Nestler, *Acta Mater.* 124, 182. (2017).
21. Rasmussen, C.E. Gaussian processes in machine learning. in Summer School on Machine Learning (Berlin: Springer, 2003).
22. C.M. Bishop, *Pattern Recognition and Machine Learning* (Springer, Berlin, 2006).
23. Wilson, A. and R. Adams. Gaussian process kernels for pattern discovery and extrapolation. in *International Conference on Machine Learning*. 2013.
24. Bélisle, E., Z. Huang, and A. Gheribi. Scalable gaussian process regression for prediction of material properties. in *Australasian Database Conference*. 2014. Springer.
25. G. Tapia, A. Elwany, and H. Sang, *Addit. Manuf.* 12, 282. (2016).
26. G. Tapia, S. Khairallah, M. Matthews, W.E. King, and A. Elwany, *Int. J. Adv. Manuf. Tech.* 94, 3591. (2018).
27. N.-D. Hoang, A.-D. Pham, Q.-L. Nguyen, and Q.-N. Pham, *Adv. Civ. Eng.* 2016, 2861380. (2016).
28. A.E. Tallman, K.S. Stopka, L.P. Swiler, Y. Wang, S.R. Kalidindi, and D.L. McDowell, *JOM* 71, 2646. (2019).
29. S. Hashemi, and S.R. Kalidindi, *Comput. Mater. Sci.* 188, 110132. (2021).
30. J.D. Eshelby, *Proc. R. Soc. Lond. Ser. A* 241, 376. (1957).
31. J.W. Hutchinson, *Proc. R. Soc. Lond. Ser. A* 348, 101. (1976).
32. R. Stringfellow, D. Parks, and G.B. Olson, *Acta Metall. Mater.* 40, 1703. (1992).
33. R.G. Stringfellow, and D.M. Parks, *Int. J. Plast.* 7, 529. (1991).
34. S. Bargmann, B. Klusemann, J. Markmann, J.E. Schnabel, K. Schneider, C. Soyarslan, and J. Wilmers, *Prog. Mater Sci.* 96, 322. (2018).
35. S.R. Kalidindi, *Int. Mater. Rev.* 60, 150. (2015).
36. M.I. Latypov, and S.R. Kalidindi, *J. Comput. Phys.* 346, 242. (2017).
37. A. Gupta, A. Cecen, S. Goyal, A.K. Singh, and S.R. Kalidindi, *Acta Mater.* 91, 239. (2015).
38. Hibbett, Karlsson, and Sorensen, ABAQUS/standard: User's Manual. (Providence, RI: Hibbitt, Karlsson & Sorensen, 1998).
39. C.C. Tasan, M. Diehl, D. Yan, C. Zambaldi, P. Shanthraj, F. Roters, and D. Raabe, *Acta Mater.* 81, 386. (2014).
40. R. Hill, *J. Mech. Phys. Solids* 11, 357. (1963).
41. Mandel, J., Contribution théorique à l'étude de l'écrouissage et des lois de l'écoulement plastique, in Applied Mechanics. (Berlin: Springer, 1966). p. 502.
42. S.R. Niezgoda, A.K. Kanjarla, and S.R. Kalidindi, *Integr. Mater. Manuf. Innov.* 2, 54. (2013).
43. A. Cecen, *Calculation, utilization, and inference of spatial statistics in practical spatio-temporal data* (Georgia Institute of Technology, Atlanta, 2017).
44. D.T. Fullwood, S.R. Niezgoda, B.L. Adams, and S.R. Kalidindi, *Prog. Mater Sci.* 55, 477. (2010).
45. A. Gokhale, A. Tewari, and H. Garmestani, *Scr. Mater.* 53, 989. (2005).
46. A. Cecen, T. Fast, and S.R. Kalidindi, *Integr. Mater. Manuf. Innov.* 5, 1. (2016).
47. S. Niezgoda, D. Fullwood, and S. Kalidindi, *Acta Mater.* 56, 5285. (2008).
48. Andrew, G. and J. Gao. Scalable training of L 1-regularized log-linear models. in *Proceedings of the 24th International Conference on Machine Learning*. 2007. ACM.
49. J. Nocedal, *Math. Comput.* 35, 773. (1980).
50. A. Skajaa, *Limited Memory BFGS for Nonsmooth Optimization* (New York University, New York, 2010).
51. Burnaev, E. and M. Panov. Adaptive design of experiments based on gaussian processes. in *International Symposium on Statistical Learning and Data Sciences*. (Berlin: Springer, 2015).
52. Jin, R., W. Chen, and A. Sudjianto. On sequential sampling for global metamodeling in engineering design. in *ASME 2002 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. 2002. American Society of Mechanical Engineers Digital Collection.
53. J. Sacks, W.J. Welch, T.J. Mitchell, and H.P. Wynn, *Stat. Sci.* 4, 409. (1989).
54. M.C. Shewry, and H.P. Wynn, *J. Appl. Stat.* 14, 165. (1987).
55. C.R. Harris, K.J. Millman, S.J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, and N.J. Smith, *Nature* 585, 357. (2020).