

TOPICAL REVIEW • **OPEN ACCESS**

Modeling process–structure–property relationships in metal additive manufacturing: a review on physics-driven versus data-driven approaches

To cite this article: Nadia Kouraytem *et al* 2021 *J. Phys. Mater.* **4** 032002

View the [article online](#) for updates and enhancements.



240th ECS Meeting ORLANDO, FL

Orange County Convention Center **Oct 10-14, 2021**

Abstract submission deadline extended: April 23rd

SUBMIT NOW



TOPICAL REVIEW

OPEN ACCESS

RECEIVED
1 July 2020REVISED
25 August 2020ACCEPTED FOR PUBLICATION
13 November 2020PUBLISHED
14 April 2021

Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



Modeling process–structure–property relationships in metal additive manufacturing: a review on physics-driven versus data-driven approaches

Nadia Kouraytem¹ , Xuxiao Li², Wenda Tan², Branden Kappes^{3,4} and Ashley D Spear² ¹ Department of Mechanical and Aerospace Engineering, Utah State University, Logan, UT, United States of America² Department of Mechanical Engineering, University of Utah, Salt Lake City, UT, United States of America³ KMMD, Castle Rock, CO, United States of America⁴ Department of Mechanical Engineering, Colorado School of Mines, Golden, CO 80401, United States of AmericaE-mail: nadia.kouraytem@usu.edu**Keywords:** physics-driven, data-driven, process–structure–property relationships, metal additive manufacturing, digital twin

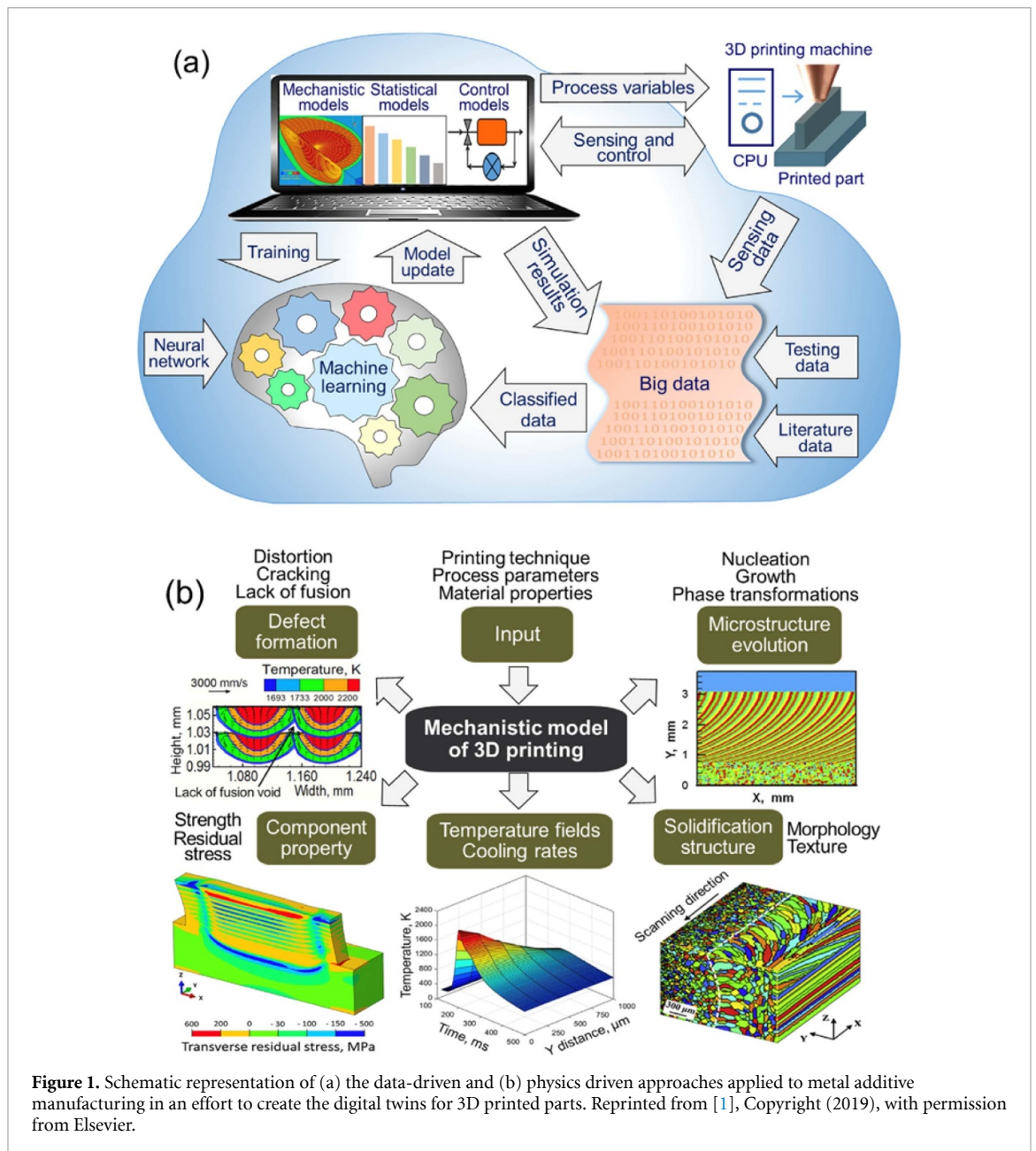
Abstract

Metal additive manufacturing (AM) presents advantages such as increased complexity for a lower part cost and part consolidation compared to traditional manufacturing. The multiscale, multiphase AM processes have been shown to produce parts with non-homogeneous microstructures, leading to variability in the mechanical properties based on complex process–structure–property (p-s-p) relationships. However, the wide range of processing parameters in additive machines presents a challenge in solely experimentally understanding these relationships and calls for the use of digital twins that allow to survey a larger set of parameters using physics-driven methods. Even though physics-driven methods advance the understanding of the p-s-p relationships, they still face challenges of high computing cost and the need for calibration of input parameters. Therefore, data-driven methods have emerged as a new paradigm in the exploration of the p-s-p relationships in metal AM. Data-driven methods are capable of predicting complex phenomena without the need for traditional calibration but also present drawbacks of lack of interpretability and complicated validation. This review article presents a collection of physics- and data-driven methods and examples of their application for understanding the linkages in the p-s-p relationships (in any of the links) in widely used metal AM techniques. The review also contains a discussion of the advantages and disadvantages of the use of each type of model, as well as a vision for the future role of both physics-driven and data-driven models in metal AM.

1. Introduction

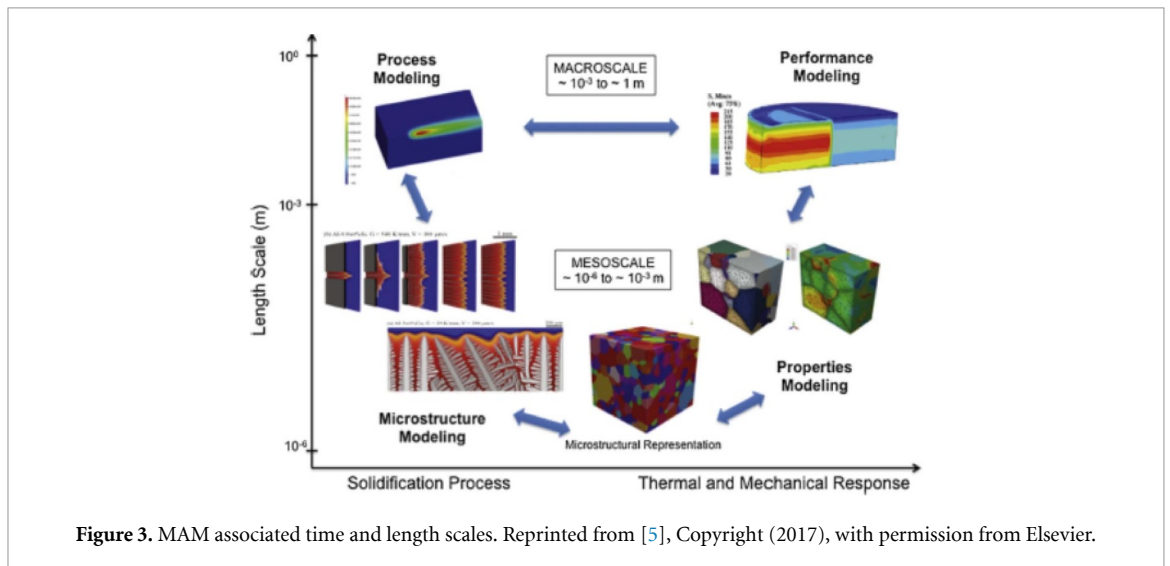
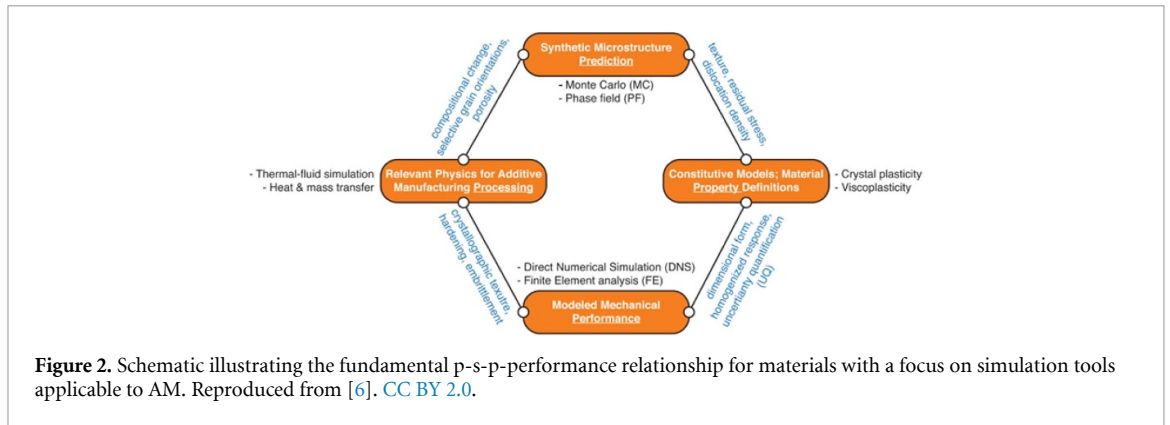
Additive manufacturing (AM), colloquially known as three-dimensional (3D) printing, of metal parts is the process of manufacturing a metallic bulk part in a layer-by-layer fashion following a computer aided design (CAD). AM techniques provide a number of advantages compared to traditional manufacturing, including the ability to consolidate parts, providing the possibility to manufacture complex geometries with less material waste and a lower cost per part. AM techniques include a realm of processes that commonly include a source of energy such as laser or electron beam. Then, based on the type of substrate used, AM processes are categorized into powder bed or directed deposition. Examples of AM processes include, but are not limited to, laser powder bed fusion (LPBF), directed energy deposition (DED), electron beam melting (EBM), etc. Most popular metallic materials used in AM include titanium, nickel-based, steel, and aluminum alloys.

Depending on the material in use, AM processing parameters are manipulated to yield specific material properties for a given application. AM processes allow the alteration of processing parameters, including the source power, scanning speed, scan strategy, layer thickness, powder size and sieve number (when in use). The large number of processing parameters available and their wide range of options or values make it unrealistic to test for optimal processing parameters experimentally in a trial-and-error process. During the



process, each alteration would also result in complicated multiphase and multiscale interaction that would result in different structures (defects and microstructure). The complex AM processes have been shown to result in non-homogeneous microstructures, which are expected to result in variability in the material's mechanical properties. Hence, it is naive to evaluate the process–structure–property (p-s-p) relationships in AM solely by trial and error. Instead, numerical approaches provide a powerful tool in evaluating the p-s-p relationships of a digital twin of an additively manufactured metallic part.

Numerical modeling of the AM processes, resulting structures, and subsequent properties supports experimental studies in the quest for achieving optimal printed parts with tailored properties. Numerical modeling of the p-s-p relationship encompasses two types of approaches: physics-driven and data-driven, as shown in Figure 1. In the physics-driven approach, a more traditional approach to simulate manufacturing processes, the p-s-p relationships are simulated using constitutive equations and mathematical relationships that describe the underlying physics of the problem. Whereas, in the data-driven approach, which has been gaining interest recently as the fourth science paradigm [2–4], available data (from experiments, simulations, or a combination of the two) are used to identify correlations between the selected inputs and outputs needed to model a given phenomenon. A number of studies that have used both approaches to investigate the p-s-p relationship in metal AM (MAM) are reviewed herein. The review article is divided into physics-driven (section 2) and data-driven (section 3) approaches. Under each section, each branch of the p-s-p relationships will be presented separately, followed by the three-way p-s-p relationship. The review is



followed by a discussion of the advantages and disadvantages of each approach and recommendations for the MAM community for both approaches moving forward.

2. Review on available physics-driven p-s-p literature

Modeling of the MAM p-s-p relationships requires the integration of phenomena that occur at different associated time and length scales, as envisioned by Francois *et al* [5] and shown in figure 3. Process modeling investigates phenomena on the mesoscale and macroscale; whereas, structure modeling focuses on the mesoscale and microscale, and performance modeling generally on the macroscale. In terms of techniques used in the physics-driven methods to explore the p-s-p relationships, multiple fundamental relationships for materials are usually implemented, as detailed by Rodgers *et al* [6] and depicted in figure 2. In this section of the review, physics-driven models are further split by specific links that are made between process and structure (section 2.1), structure and property (section 2.2), and p-s-p (section 2.3) (in which some have modeled directly the link between process and property).

2.1. Process-structure relationship

Modeling the process-structure (p-s) relationship in MAM involves computational techniques that were established to model traditionally manufactured metals, welding processes that translate to MAM, and, in some cases, specific MAM processes. In process modeling, the physical variables are predicted as a function of space and time during the MAM process. Due to the fact that the MAM process is a multi-scale process, process models can be subcategorized into mesoscale models and macroscale models [7].

Mesoscale models resolve the geometry of individual particles and predict the temperature, pressure, and fluid flow (velocity) field when the source interacts with the particles [8–11]. The calculation of these physical variables is achieved using computational fluid dynamics (CFD) simulations, which numerically solve a set of partial differential equations (PDEs). In order to distinguish the geometry of particles and molten pool, the interface between the metal phase and the gaseous phase needs to be captured. That has

been achieved by a variety of interface capturing techniques for multi-phase CFD simulations. Thermo-fluid physics in the building process, such as surface tension, Marangoni effects, laser absorption, and metal evaporation, can be incorporated into the model through special treatment of the boundary conditions and source terms when solving the PDEs. These multiphysics considerations enable the model to predict key physical events, e.g., balling [9–11], keyholing [12, 13] and spattering [14], that lead to mesoscale defects such as lack-of-fusion and porosity [15]. Therefore, mesoscale process models can play a crucial role in quantitatively understanding the formation of these defects.

Macroscale models implement simplifying assumptions of the physics to relieve the computational costs and allow to simulate the building process of an entire part. A major simplification is to treat the metal phase as a continuum without resolving individual particles or the fluid flow in the molten pool [7]. The temperature and stress/strain fields in the part during its building process can be predicted with a thermo-mechanical approach [16]. The macroscale models have shown to predict with considerable accuracy the residual thermal stress/strain after the part is finished [17, 18]. With this predictive power, the macroscale models have been used to optimize process parameters [19]. Compared with mesoscale models, macroscale models are more relevant to solve practical problems but have reduced physical fidelity. The hybrid macroscale-mesoscale modeling approach is a prominent way to balance physical fidelity and practicality [20].

Efforts of structure modeling have been primarily focused on two length scales, mesoscale and microscale. Mesoscale modeling predicts the grain structure resulting from the solidification of the metal, while microscale modeling predicts sub-grain features such as dendrite morphology and phase distribution.

Mesoscale models are used to predict the grain structure in MAM based on either the cellular automata (CA) [21–23] or the Monte Carlo method [24, 25]. The temperature field is a necessary input for the model, and the temperature reduction during solidification is the thermodynamic driving force for nucleation and grain growth to occur. The CA and Monte Carlo methods use different algorithms to predict grain structure evolution based on the temperature information. Each method has its own advantages and disadvantages, and they appear to have comparable popularity in literature. The mesoscale models can reproduce typical grain structures observed in experiments and have shown to capture several key mechanisms of grain structure formation, including the epitaxial nucleation and growth, bulk nucleation, competitive growth, and columnar-to-equiaxed transition. This modeling capability may potentially facilitate the process design to achieve site-specific control of grain structures [26].

Microscale models primarily simulate the evolution of sub-grain microstructure morphologies (e.g., dendrite) and/or phase transformation during solidification based on given temperature history. Phase field is arguably the most widely used method to simulate the sub-grain microstructure evolution for AM processes [27–31]. The method is composed of a set of thermodynamically based PDEs. It is applicable to fairly complex metal alloy systems if the thermodynamic descriptions of the systems are available, usually from the CALPHAD (calculation of phase diagram) database. The computational costs for the phase field are rather demanding, which limits the simulations to relatively small spatial and temporal domains.

2.2. Structure-property relationship

Structure-property (s-p) modeling is based on the use of solid-mechanics models to link structure to properties in MAM. In this case, constitutive equations from continuum mechanics are used to describe the evolution of plasticity and, in some cases, damage. The constitutive equations are solved over the simulation domain using numerical methods like the finite-element method (FEM) or fast Fourier transform (FFT) method. In MAM, efforts have been broadly divided into three types of predictions. In the first one, conventional continuum-based modeling is used to predict mechanical properties of AM structures, disregarding any types of AM defect structures. In the second modeling approach, defect structures (e.g., gas porosity, lack-of-fusion) in MAM are incorporated. In the third modeling approach, the underlying crystal structure of the MAM part is incorporated.

Many researchers have performed s-p predictions for AM parts by neglecting the process modeling and excluding any defects and crystal structure resulting from the AM process, thereby treating the part as a conventional, continuum domain absent voids. In an application to lattice components, a number of investigations simulated the mechanical response of AM lattices [32–34] by modeling the geometry of the component, treating each strut as a material continuum. Specifically, Hedayati *et al* [33] investigated the fatigue behavior of an open-cell lattice component with AM Ti-6Al-4V mechanical properties using FEM. Similarly, work by Andani *et al* [34] compared experimentally measured and simulated elastic moduli of NiTi open-cell lattice components with a user material subroutine in Abaqus. Another recent effort in which researchers modeled the mechanical response of an AM component treated as a continuum is highlighted in the Third Sandia Fracture Challenge [35–44]. In the challenge scenario, participants were tasked with predicting ductile fracture in an AM 316L part geometry containing multiple internal features, including

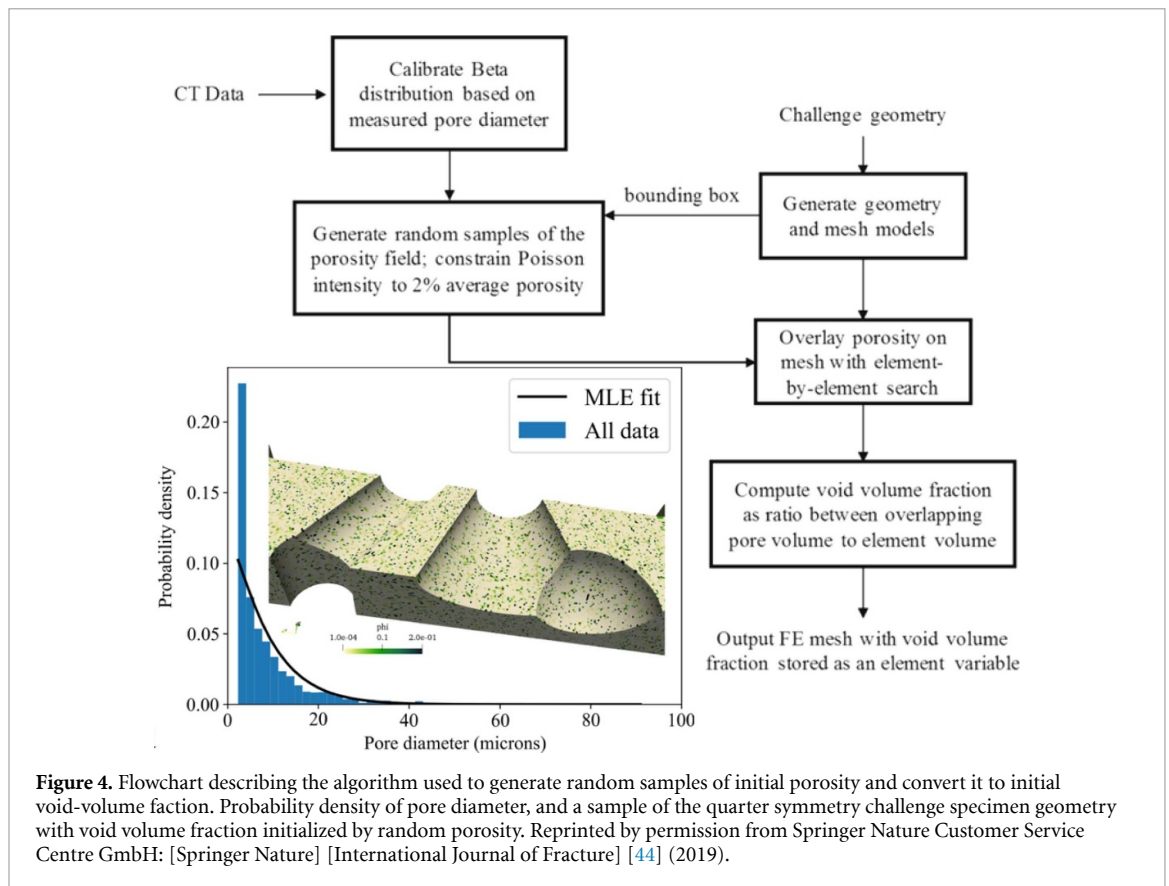


Figure 4. Flowchart describing the algorithm used to generate random samples of initial porosity and convert it to initial void-volume fraction. Probability density of pore diameter, and a sample of the quarter symmetry challenge specimen geometry with void volume fraction initialized by random porosity. Reprinted by permission from Springer Nature Customer Service Centre GmbH: [Springer Nature] [International Journal of Fracture] [44] (2019).

angled channels, a hole, and a large void. Most of the participants modeled the AM material as a continuum, neglecting to explicitly include details like pore structure, surface roughness, and grain structure, which were provided. Because most of the mechanical response of the AM part was governed by the large-scale geometrical features of the part, many of the teams were able to perform reasonably well in the blind challenge, despite neglecting to model the p - s - p relationships explicitly.

In the second type of modeling approach to investigate the s - p link, multiple efforts have incorporated defect structures in the MAM structure. A number of studies simulate the stress distribution over synthetic grain and/or pore structures under certain loading conditions and predict the effective mechanical properties of the simulated materials [45–47]. In the context of the Sandia Fracture Challenge, one team [44] accounted for AM-induced pore structure in their solid–mechanics simulations. The team leveraged x-ray computed tomography data (which was provided to all participants) to initialize an element-wise value of a damage parameter based on the local pore volume to element volume by juxtaposing the tomography data with a finite-element mesh as shown in figure 4. Another effort [42] adopted a similar constitutive model to incorporate the effects of AM-induced void defects in their predictions with more accurate prediction results. In another effort, Erickson *et al* [48] established a new, unique descriptor of pore structures that shows a promising relationship with a variety of mechanical properties. The new descriptor is based on a radial distribution function that takes into account the sizes, clustering, and locations of pores within a specific part. The maximum value of the radial distribution function (or the location of the maximum value) was shown to have a stronger correlation with ultimate tensile strength, elongation, toughness modulus, and fracture location than any of the pore metrics previously reported in the literature.

Finally, there have been recent efforts in physics-driven, s - p modeling to incorporate the crystal structure of the AM material into predictions of mechanical properties. Work by Ozturk and Rollett [49] provides an example of modeling the s - p relationships using statistically representative volumes of AM Ti-6Al-4V microstructures. The microstructural instantiations were created in DREAM.3D [50] using grain-scale statistics for AM Ti-6Al-4V from the literature. Subsequently, an elasto-viscoplastic FFT code was used to simulate the full-field micromechanical response and corresponding homogenized stress–strain curve of each volume. In another approach, Ahmadi *et al* [45] simulated an AM microstructure by repeating a melt-pool geometry with a grain structure generated via Voronoi tessellation. They predicted the effective stress–strain response of the polycrystalline domain using finite-element modeling with cohesive elements along the melt-pool interfaces and compared the results to experimental observations. Likewise,

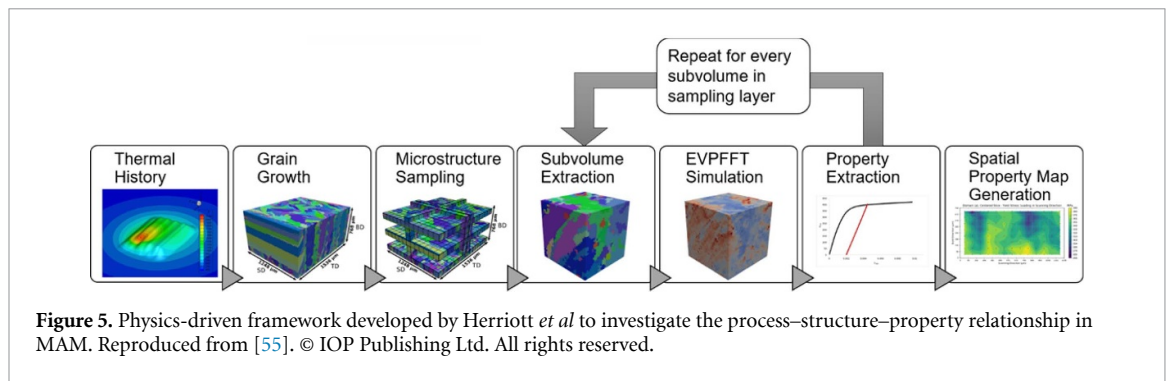


Figure 5. Physics-driven framework developed by Herriott *et al* to investigate the process–structure–property relationship in MAM. Reproduced from [55]. © IOP Publishing Ltd. All rights reserved.

Andani *et al* [46] used a similar grain-structure model but investigated the response of representative volume elements of interest within the simulated microstructure. While the above works consider a simulated AM microstructure in the mechanical simulation, none used physics-based process modeling to predict the microstructure. The accuracy of all s-p predictions described in this section hinged upon rigorous selection and calibration of the material constitutive model(s) used to represent the mechanical behavior of the AM material, which is further discussed later in the manuscript.

2.3. Process-structure-property relationship

In MAM, the investigation of the p-s-p relationships is not limited to the three-way relationship paradigm. A number of efforts bypassed the (micro)structure prediction, and the physics-based model was used to predict, for example, the thermal history during the process followed immediately by a prediction of the residual stress field and part distortions [51–54]. In the work by Johnson *et al* [52], the transient temperature field was first solved over a finite-element mesh for a 304L stainless steel tube manufactured using laser engineered net shaping (LENS). Output from the thermal model was then used to predict the thermo-mechanical strain as a function of temperature-dependent material properties and, finally, the residual-stress state was predicted using an assumed constitutive model.

While the above modeling efforts focus on modeling either the first link, last link, or bypass the structure link in the p-s-p lifecycle, there are modeling approaches that capture the complete lifecycle from process to mechanical properties for MAM. In this area, Yan *et al* [47] implemented a framework that, given the processing parameters, predicts using FEM a number of mechanical properties (e.g., yield strength and fatigue) of the simulated microstructure. The process (powder bed flow, heat transfer, and fluid flow) and structure (grains and voids) are modeled and incorporated into the framework. Rodgers *et al* [6], used direct numerical simulations to simulate tensile loading on multiple hollow cylindrical structures with synthetic microstructures ranging from equiaxed to columnar grains (by changing the processing parameters, namely the scan speed). The 3D, synthetic voxelized microstructures were generated using the kinetic Monte Carlo method. The resultant microstructures were mapped onto a highly refined conformal FE mesh of a part geometry. A grain-scale anisotropic crystal elasticity model was then used to represent the constitutive response of each grain. In a similar framework, detailed in figure 5, Herriott *et al* [55] implemented a parallelized elasto-viscoplastic FFT model to predict the spatial variability of the yield strength and elastic modulus throughout different MAM build domains, exemplified for SS316L DED build volumes. To achieve that, four microstructural volumes, ranging from fully equiaxed to columnar, were predicted using process modeling (*viz.*, the CA method) to mimic different combinations of processing parameters in the DED process. Then, the full-field mechanical response of microstructural subvolumes extracted from multiple layers of interest within the simulated build volumes was predicted using EVPFFT. Later, the full-field response of each subvolume was homogenized, and stress–strain curves of each subvolume were extracted and analyzed for values of the elastic modulus and yield strength. Finally, the homogenized mechanical properties were plotted using contour maps to investigate their location-specific variability with respect to the laser scan tracks.

More recently, Shi *et al* [56] established a physics-driven framework to establish the p-s-p relationships and tested it on a Ti-6Al-4V single-track LPBF system. The framework consists of a high-fidelity, powder-scale, 3D simulation of transient heat transfer and melt flow dynamics (figure 6(B)), a CA simulation of solidification grain structure and texture, phase field modeling of precipitation and dissolution of second-phase precipitate during repeated thermal cycles as presented in figure 6(A), and microstructure-based micro- and mesoscopic elastic response calculation (figure 6(D)). The framework was capable of simulating concurrent nucleation, growth, and coarsening of the equilibrium α phase in a biccristalline β matrix (figure 6(C)).

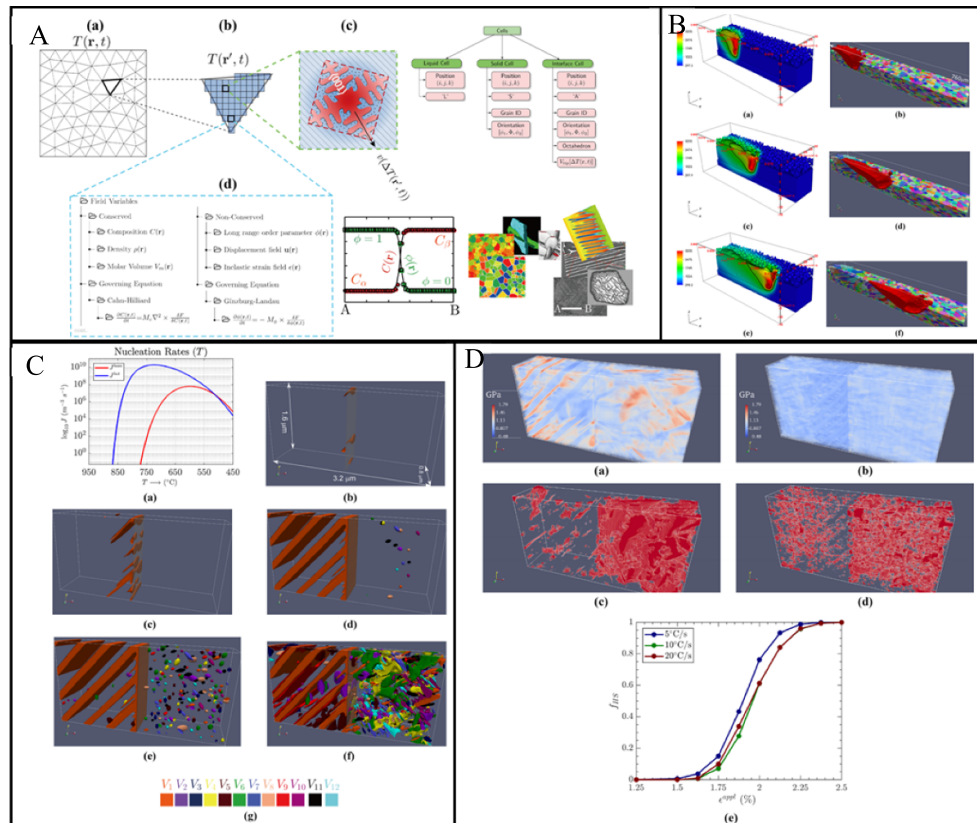


Figure 6. Outputs from sections of the p-s-p framework implemented by Shi *et al* [56]. (A) Microstructural prediction based on the interaction among (a) an arbitrary Lagrangian–Eulerian-based powder-scale model to predict spatiotemporal evolution of thermal history, (b) CA model to simulate solidification grain structure evolution (size as well as orientation), and (d) a quantitative PF model to predict the subgrain/intragranular solid-state precipitate microstructure. (B) Time snapshots showing the evolution of thermal profile (left column) and grain structure (right column) at $t = 327 \mu\text{s}$, $t = 647 \mu\text{s}$, $t = 967 \mu\text{s}$. The laser scan speed is 500 mm s^{-1} , moving to the right with a power of 75 W . (C) (a) Rates for both homogeneous nucleation and heterogeneous nucleation of α phase as function of temperature; (b)–(f) α precipitation upon cooling from $950 \text{ }^\circ\text{C}$ to $700 \text{ }^\circ\text{C}$ at a rate of $5 \text{ }^\circ\text{C s}^{-1}$; (b) $860 \text{ }^\circ\text{C}$, (c) $855 \text{ }^\circ\text{C}$, (d) $770 \text{ }^\circ\text{C}$, (e) $765 \text{ }^\circ\text{C}$, (f) $740 \text{ }^\circ\text{C}$; (g) Color scheme for all 12 α variants. (D) Microstructure-based elastic response calculations: computed von Mises stress profiles σ_v in Ti-64 two-phase microstructures under $\epsilon_{11}^{\text{appl}}$ for cooling rates of (a) $5 \text{ }^\circ\text{C s}^{-1}$ and (b) $20 \text{ }^\circ\text{C s}^{-1}$; spatial distribution of the identified stress hotspots (i.e., $\sigma_v >$ yield strength) for the microstructures formed at the cooling rates of (c) $5 \text{ }^\circ\text{C s}^{-1}$ and (d) $20 \text{ }^\circ\text{C s}^{-1}$, and (e) evolution of hotspot volume fractions (f_{HS}) with increasing $\epsilon_{11}^{\text{appl}}$. Reprinted by permission from Springer Nature Customer Service Centre GmbH: [Springer Nature] [JOM Journal of the Minerals, Metals and Materials Society] [56] (2019).

Some other approaches have combined both physics-driven methods combined with experimental support, which are not discussed in the current review. An extensive review of those efforts are detailed by Smith *et al* and can be found in [57].

Physics-driven approaches provide a number of advantages and disadvantages that are discussed in detail in section 4. Another approach to model the p-s-p relationship in MAM is data-driven models, which predict the p-s-p relationships of MAM by correlating the expected properties to build conditions (e.g., laser processing parameters, location of a part on the build plate, gas flow in the printer, etc.) that uniquely define the manufacturing process. Such models utilize specific data-driven techniques such as machine learning or data-mining algorithms, for example, to establish the p-s-p relationships and are capable of predicting the behavior of metal components that are manufactured using specific processing parameters without the need to explicitly identify the constitutive model governing its behavior.

3. Review on available data-driven p-s-p literature

With the availability of various types of materials-related data (mostly big data from different sources), there has been a natural advancement in data-driven modeling techniques as a major area of research in materials science. Data-driven modeling in materials science encompasses the overlap of the domain knowledge in materials science, mathematics and statistics, and computer science and technology (figure 7). The generation of big data used in these techniques is characterized by the ‘five V’s’ (volume, velocity, variety, veracity, and value), as detailed in figure 7. Data-driven techniques in the investigation of the p-s-p relationships in MAM are a very attractive approach.

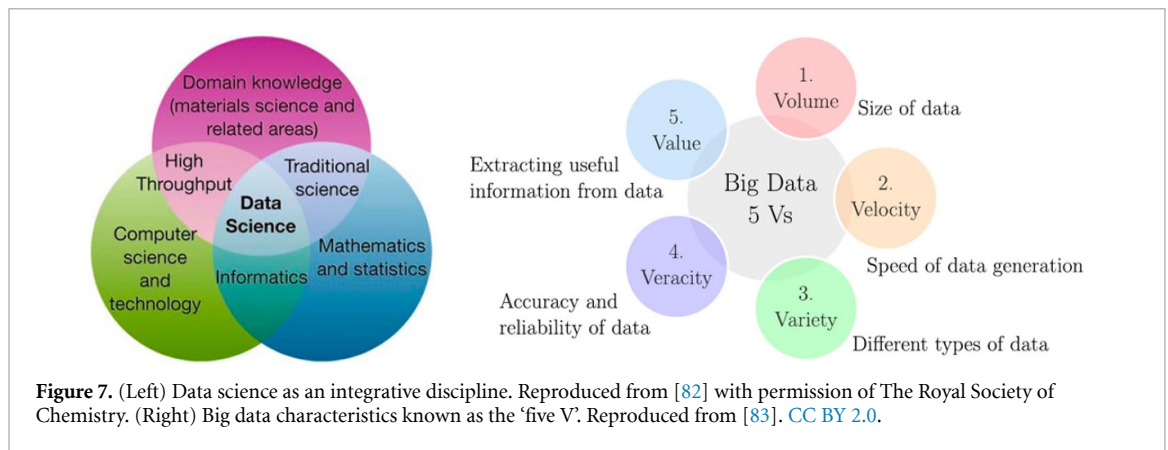


Figure 7. (Left) Data science as an integrative discipline. Reproduced from [82] with permission of The Royal Society of Chemistry. (Right) Big data characteristics known as the ‘five V’. Reproduced from [83]. CC BY 2.0.

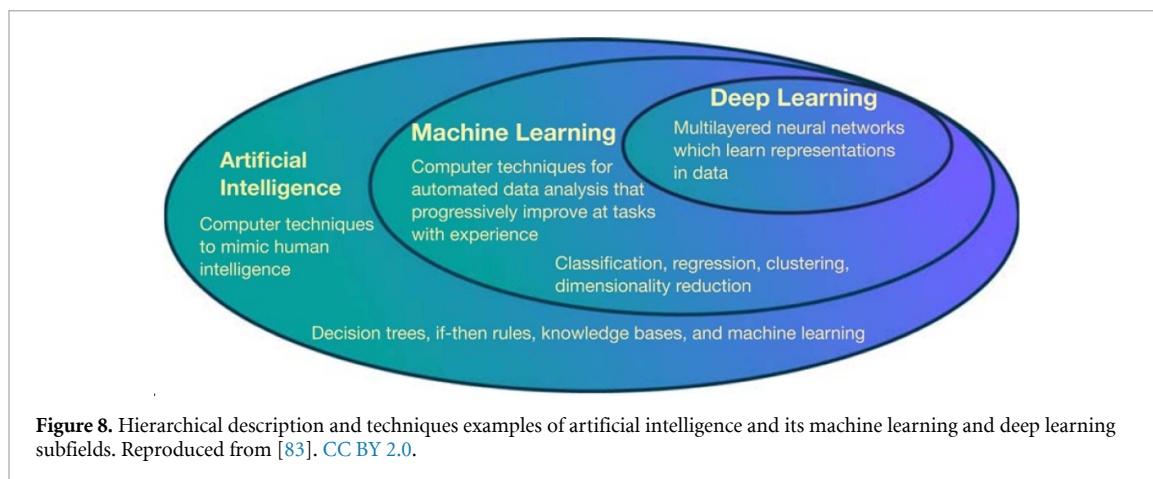
Data-driven materials modeling addresses the need to extract high-value information from many and diverse data sources—from modeling-derived data (e.g., from density functional theory (DFT), phase field, CALPHAD, and a range of continuum models), to experimentally derived data (e.g., mechanical, structural, chemical, or thermal analysis) [58]. Algorithmic optimization of materials, using heuristic algorithms—such as genetic algorithms [59] or particle swarm optimization [60]—or using machine learning algorithms—such as Gaussian process regression (GPR) [61], artificial neural networks (ANNs) [62], or kernel ridge regression [63], require materials data to be *featurized*. Featurization is the process of extracting parameters (*features*) from experiment [64, 65], literature [66], or modeling/simulation data [67–69]. In the context of MAM, data-driven approaches have been used in the investigation of the p-s-p relationships to support the integrated computational materials engineering (ICME) effort of the scientific community in understanding the links between each of the branches for MAM digital twins.

In the terminology of machine learning, *structure*, in the paradigmatic *process–structure–property* relationship of materials science, represents a latent space. A stable, correlative relationship between process and properties is all that is required in many practical applications. Although the complexity of the process–property models are often better served by the explicit inclusion of structural information through modeling or experiment, such as CALPHAD, phase field, DFT, microscopy, or x-ray diffraction, data-driven algorithms provide a mathematical and modeling framework to identify process–property correlations that implicitly account for structure. The microstructure of AM parts is complicated by the thermal history of the additive process. As a result, only the most basic parts have a well-characterized microstructure [26], but AM is amenable to exploration of a broad process–property space through high-throughput sample fabrication and mechanical testing.

Data-driven methods that are used in the investigation of the p-s-p relationships in MAM include an array of techniques that mimic human intelligence, also known as artificial intelligence. The most widely used techniques under artificial intelligence are machine learning models that include classification, regression, clustering, and dimensionality reduction. Another subcategory of machine learning falls under deep learning, which encapsulates neural networks, as represented in figure 8. In the subsequent section, recent efforts that focus on any of the branches of the p-s-p relationships using data-driven techniques will be reviewed. As in the previous section 2 on physics-driven models, the links in this section are divided into the relationships between process and structure (section 3.1), structure and property (section 3.2), and p-s-p (section 3.3).

3.1. Process-structure relationship

The prediction of the p-s link in MAM using data-driven approaches is still at its infancy due to the complexity of the problem and the multiphase, multiscale processes involved. Data-driven methods have been used to uncover the p-s relationship by predicting phase formation in the microstructures of traditionally manufactured metals. In fact, phase formation and microstructures are predicted by mapping thermal processing using continuous cooling transformation (CCT) and time–temperature–transformation (TTT) diagrams. Primarily, artificial intelligence techniques have been employed to predict CCT curves [70–72]. Initial efforts by Wang *et al* [72] trained an ANN to using 151 CCT diagrams to predict the diagrams of class Fe-xC-0.4Si-0.8Mn-1.0-Cr-0.003P-0.002 S (x within 0.1 through 0.6) steels. Similarly, Dobrzański *et al* [71] trained a neural network using 400 CCT diagrams made for structural and machinable steels. The model takes as inputs the chemical composition and austenitising temperature and predicted the temperature of the beginning and the end of transformation as a function of cooling rate, the chemical composition, and the hardness of steel cooled from austenitising temperature with a fixed rate which then



allow to predict the CCT diagram of the material. Using a similar approach, Trzaska *et al* [70] trained a neural network and combined it with regression methods to predict the CCT curve for steel with a known chemical composition. The method allowed the analysis of the influence of particular elements on the characteristic points and transformation curves of the supercooled austenite using a set of 300 experimental data sets that are available in the literature. The CCT and TTT diagram of AM metals, however, have been shown to be different from their traditionally manufactured counterparts. Moving forward, the aforementioned approaches could be adapted to MAM processes and used in the prediction of the p-s relationship in MAM.

In the context of p-s linkage for MAM, Popova *et al* [73] presented a workflow, reproduced in figure 9(A), that can be applied to augment computational materials science using data science approaches. The authors also suggested a general workflow template that could be used to establish the p-s link, as shown in figure 9(B). The blue boxes describe the general steps to achieve that (data pre-processing, microstructure quantification, dimensionality reduction, and extraction/validation of the p-s relationship), while the white boxes present examples of methods that can be used to achieve each. In a specific application of the framework to MAM, the authors established a microstructural descriptor (chord length) on a synthetic microstructure and established the p-s link using distribution calculation, principal component analysis, and multivariate polynomial regression methods. In a different approach, Wang *et al* [74] successfully trained a two-level surrogate model based on 280 thermal simulations and 150 grain growth simulations to predict temperature profile and materials microstructures. The authors quantified the sensitivity of the material microstructure to specific heat capacity and grain growth activation energies and its uncertainty quantification and presented optimal combination of process parameters for EBM Ti-6Al-4V.

Data-driven approaches have been also used to identify microstructural features and defects in MAM structures using *in situ* data. Combining acoustic emission (recorded using a fiber Bragg grating sensor) and training a neural network, Shevchik *et al* [75, 76] performed *in situ* quality monitoring in MAM to identify the formation of porosity during the build, as presented in figure 10. The inputs to the neural network were the relative energies of the narrow frequency bands of the wavelet packet transform, and the outputs of the models provided a classification of the build quality of the layer of interest with accuracy around 83%–89%. In a different approach for the classification of *in situ* defect formation, other studies focused on the use of visual data and image processing techniques. For example, Scime and Beuth [77, 78] and Caggiano *et al* [79] employed classification techniques in identifying defects such as keyhole-induced pores and balling from *in situ* process images. Both investigations implemented convolutional neural networks to identify and classify the types of defects in the printed layer. More details of the employed techniques can be found in the flowcharts in figures 11(A) [79] and (B) [78]. Other classification efforts focused on the identification of microstructural features [80] and identifying feedstock powder source and type for AM metals [81] based on images of the microstructure.

After printing, MAM parts have been shown to exhibit geometrical deviations post solidification compared to the intended design (CAD model). Such deviations in the geometry from the initial design may limit the certification and qualification of MAM parts, especially in applications where qualification requirements are stringent. To control the geometrical fidelity of MAM parts, Koeppel *et al* [84] trained a neural network based on FEM simulation results of MAM cell structures to predict load-displacement curves and maximum stress, which were validated using experimental results. The neural network was then used to reproduce the loading history of the FEM simulation and later as a substitute for the physics-driven simulations.

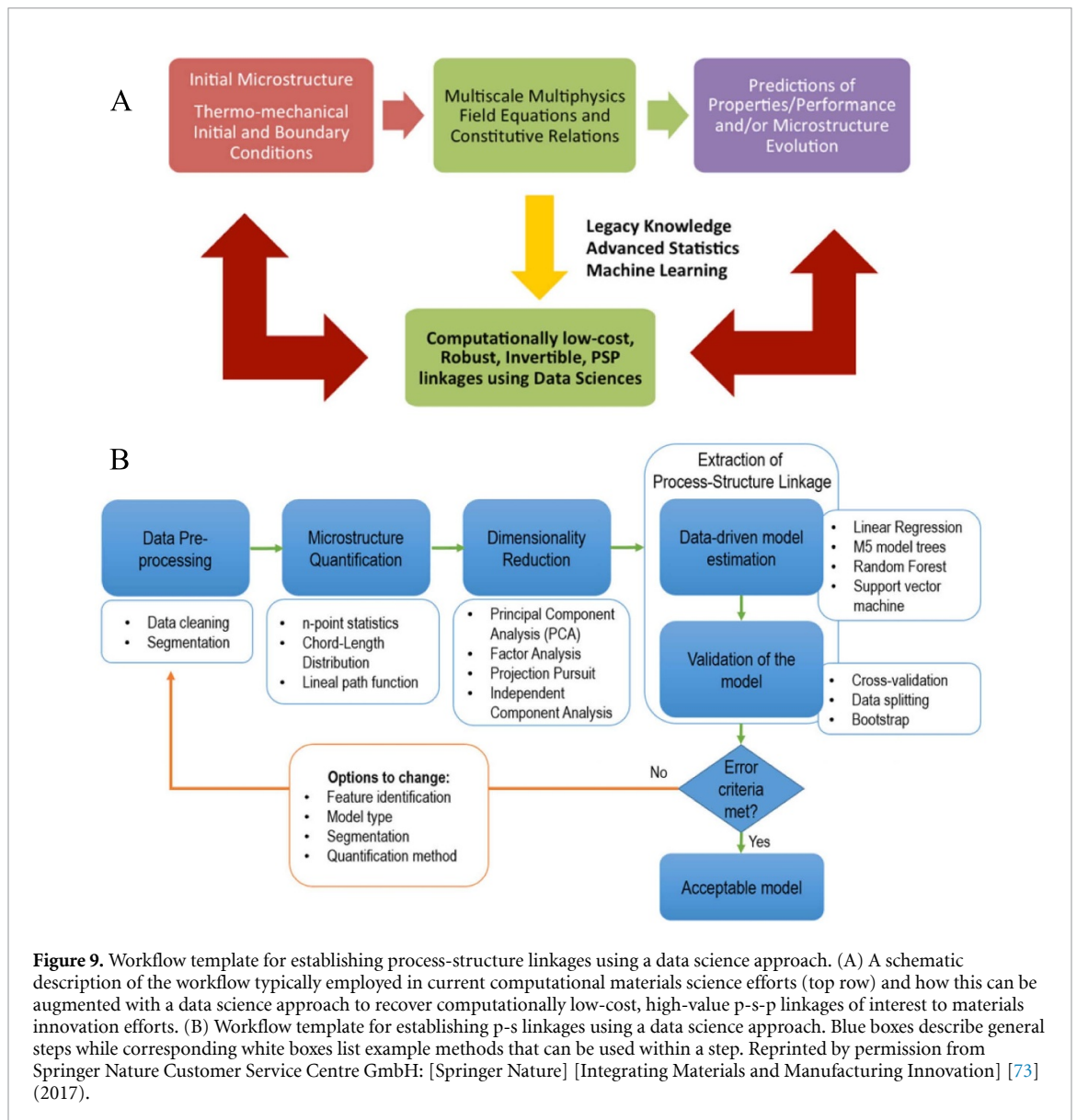
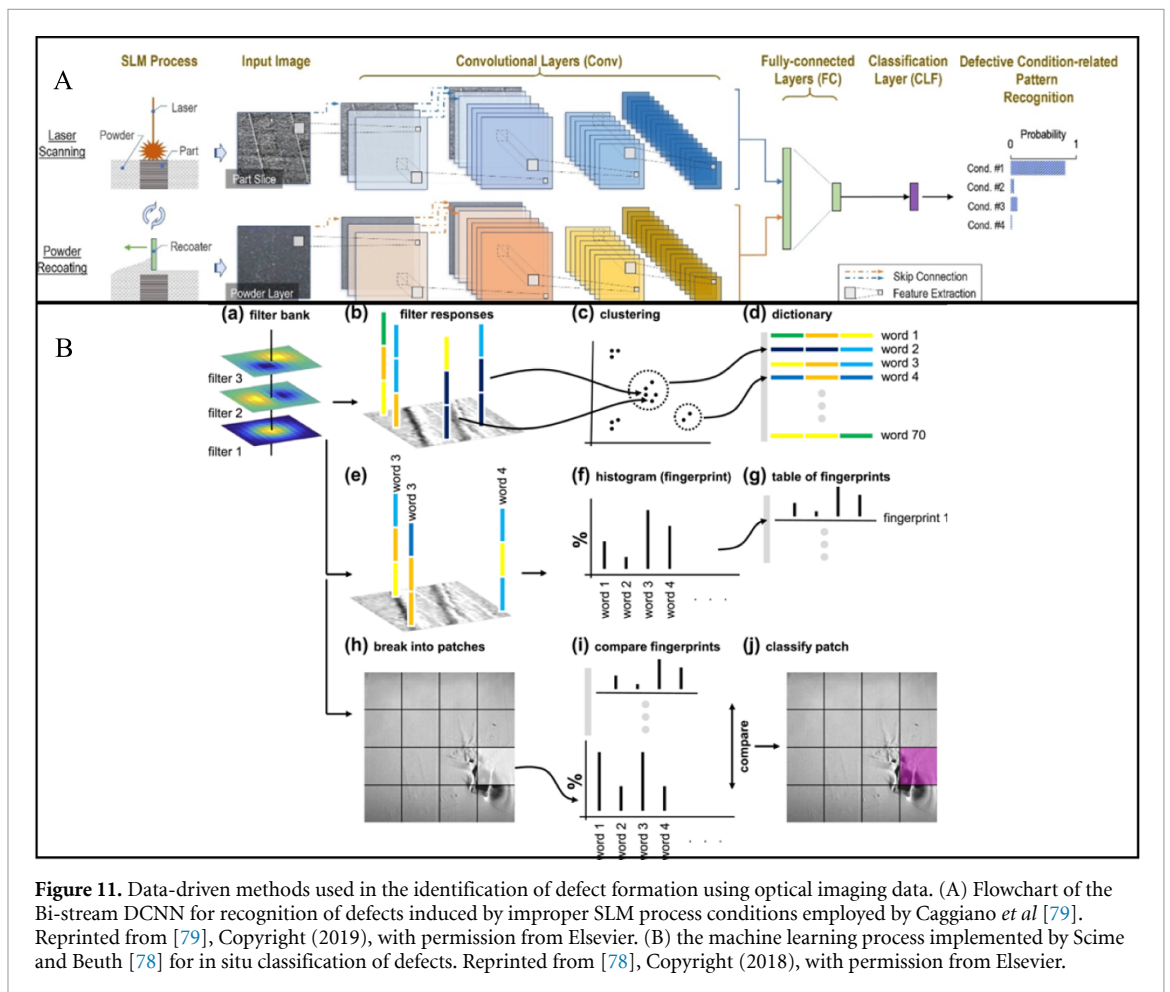
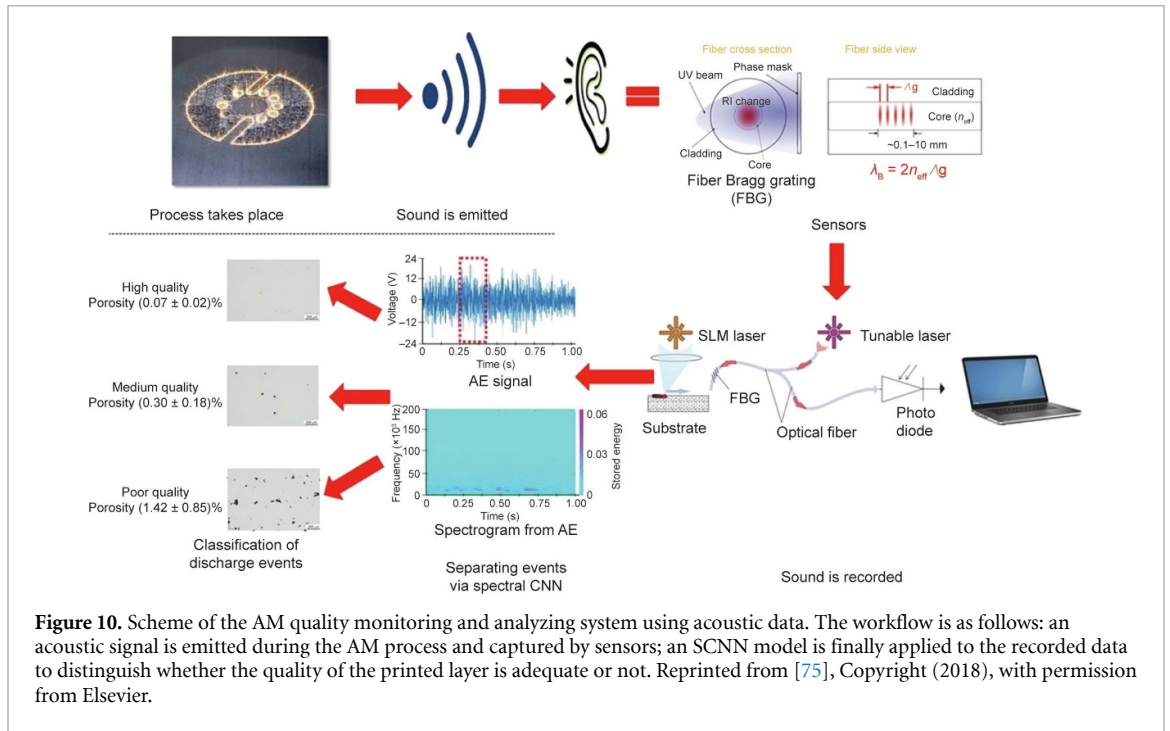


Figure 9. Workflow template for establishing process-structure linkages using a data science approach. (A) A schematic description of the workflow typically employed in current computational materials science efforts (top row) and how this can be augmented with a data science approach to recover computationally low-cost, high-value p-s-p linkages of interest to materials innovation efforts. (B) Workflow template for establishing p-s linkages using a data science approach. Blue boxes describe general steps while corresponding white boxes list example methods that can be used within a step. Reprinted by permission from Springer Nature Customer Service Centre GmbH: [Springer Nature] [Integrating Materials and Manufacturing Innovation] [73] (2017).

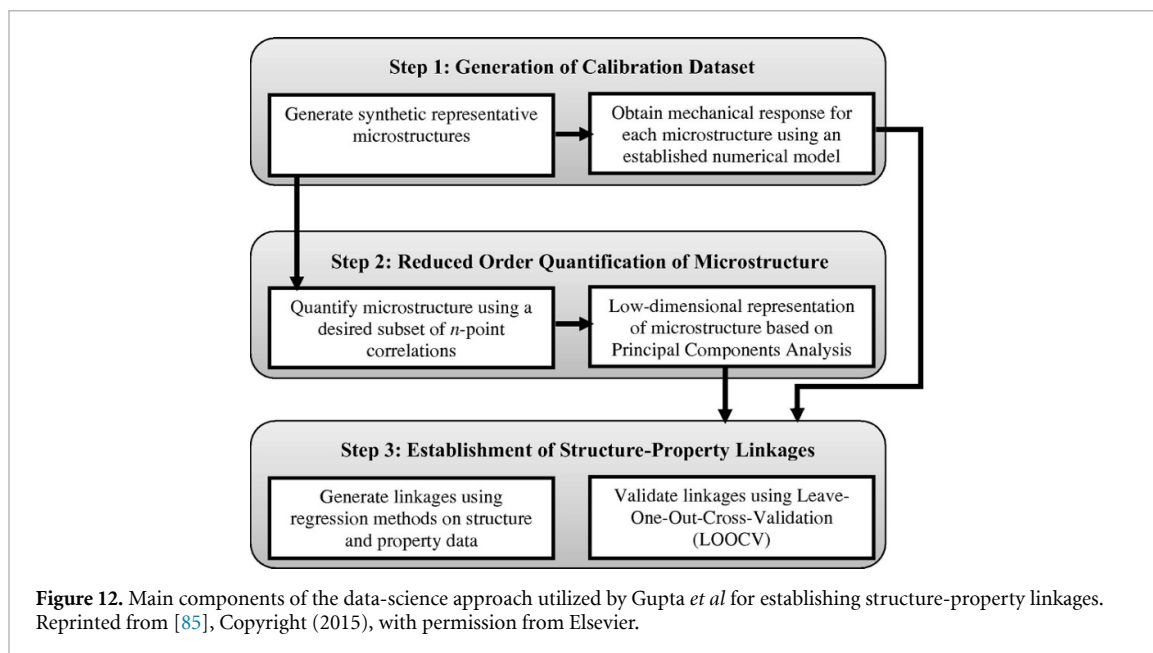
3.2. Structure-property relationship

Fewer efforts have been invested in the exploration of the s-p relationship in MAM using data-driven approaches as compared to the physics-driven approaches. In an effort to comprehend the s-p link, Gupta *et al* [85] proposed a general data-driven framework to investigate the aforementioned branch as detailed in figure 12. The framework consists of three steps that start with the generation of a calibration dataset consisting of representative microstructures (either from measurements or generated synthetically) and their subsequent mechanical response (from an established physics-based numerical model, or experimental results). The second step consists of extracting objective, reduced-order (based on principal component analysis), quantitative measure(s) of the microstructure. In the third and final step, the s-p linkages are established using regression methods and then validated using leave-one-out-cross-validation. The authors evaluated the framework on non-metallic inclusions/a steel composite system and compared it to results from conventional linkage techniques. The results demonstrated high accuracy of the surrogate model that was based on the data-driven framework compared to the conventional approaches. Moving forward, the presented framework could also be applied to s-p data that is available in the MAM literature to establish those linkages.

In a different approach, recent efforts applied machine learning techniques to predict the mechanical properties of given metallic microstructures [86–88]. Even though these efforts have been presented for non-additively manufactured materials, the approaches present powerful tools for potential applications in MAM. In the first effort, Jung *et al* [86] employed the GPR method to predict the full-field response of two-phase microstructures generated with periodic boundary conditions using DREAM.3D. Over 1100



different synthetic microstructures were used to predict the uniform elongation, ultimate tensile strength, and strain localization index defined in the soft phase of each microstructure and later validated using another 1000 microstructural instantiations. The accuracy of the surrogate model was calculated as relative error from finite element simulation results and showed that as little as 40 training points were enough to



decrease the error between the surrogate and finite element models to less than 1%. The authors also employed the Efficient Global Optimization method, a Bayesian optimization technique that utilizes an expected improvement algorithm to maximize an unknown black-box objective function to predict the optimal microstructure in the dual phase steel to result in the targeted material properties. The results were shown to predict with high accuracy using a relatively low number of inputs. In a different approach, Mangal and Holm predicted the formation of stress hotspots in face centered cubic materials [87] and hexagonal close packed materials [88] under uniaxial tensile deformation by integrating full-field crystal plasticity-based deformation models and machine learning techniques as summarized in figure 13. In both studies, synthetic microstructural data were generated using DREAM.3D [50]. The local micro-mechanical fields under uniaxial tensile deformation were then simulated using a 3D full-field, image-based, FFT technique, and stress hotspots were defined as grains having stress values above the 90th percentile of the stress distribution. Then, hotspots were characterized using metrics that reflect local crystallography, geometry, and connectivity. The data were used to create input feature vectors to train a random forest learning algorithm, which predicted the grains that will become stress hotspots. Other work, tailored specifically towards MAM, have also investigated hotspots in build structures with respect to the processing parameters.

In a similar approach tailored towards MAM, Herriott and Spear [89] utilized the results from their p-s-p framework discussed in section 2.3 to establish a data-driven model that predicts the s-p link in four AM SS316L build domains exhibiting vastly different microstructures. The main purpose of the data-driven method is to omit the high cost and computational time associated with the crystal-plasticity physics-driven method. To predict the s-p relationship, machine-learning (Ridge regression and XGBoost) and deep-learning (convolutional neural networks) models were trained using microstructural inputs as features and the effective mechanical properties (calculated from the full-field response resulting from the crystal plasticity FFT method) as targets. Output predictions of the effective yield strength in a certain loading direction were then tested for build layers that were omitted from the training set of different MAM build domains. The results were compared to a reference property map (generated using the physics-driven model). An example of the output from the machine- and deep-learning models are presented in figure 14. Among all of the data-driven models tested, the deep-learning models that use crystal orientation as the primary input significantly outperform all other models considered, achieving an R^2 value upwards of 0.86 in blind predictions of the holdout layers for two different MAM build domains. Furthermore, the deep-learning models offer the benefit of requiring very little pre-processing and feature extraction, instead allowing the model, itself, to extract relevant higher-level features from the microstructural image data. In all cases, the data-driven approach was capable of predicting the properties in a matter of seconds once trained.

In another application that is based on the use of structure data, Kantzos *et al* [90] trained a convolutional neural network to predict stress concentration of rough surfaces, which are detrimental to fatigue life of MAM parts. The model was trained using the elasto-viscoplastic FFT stress field results of a

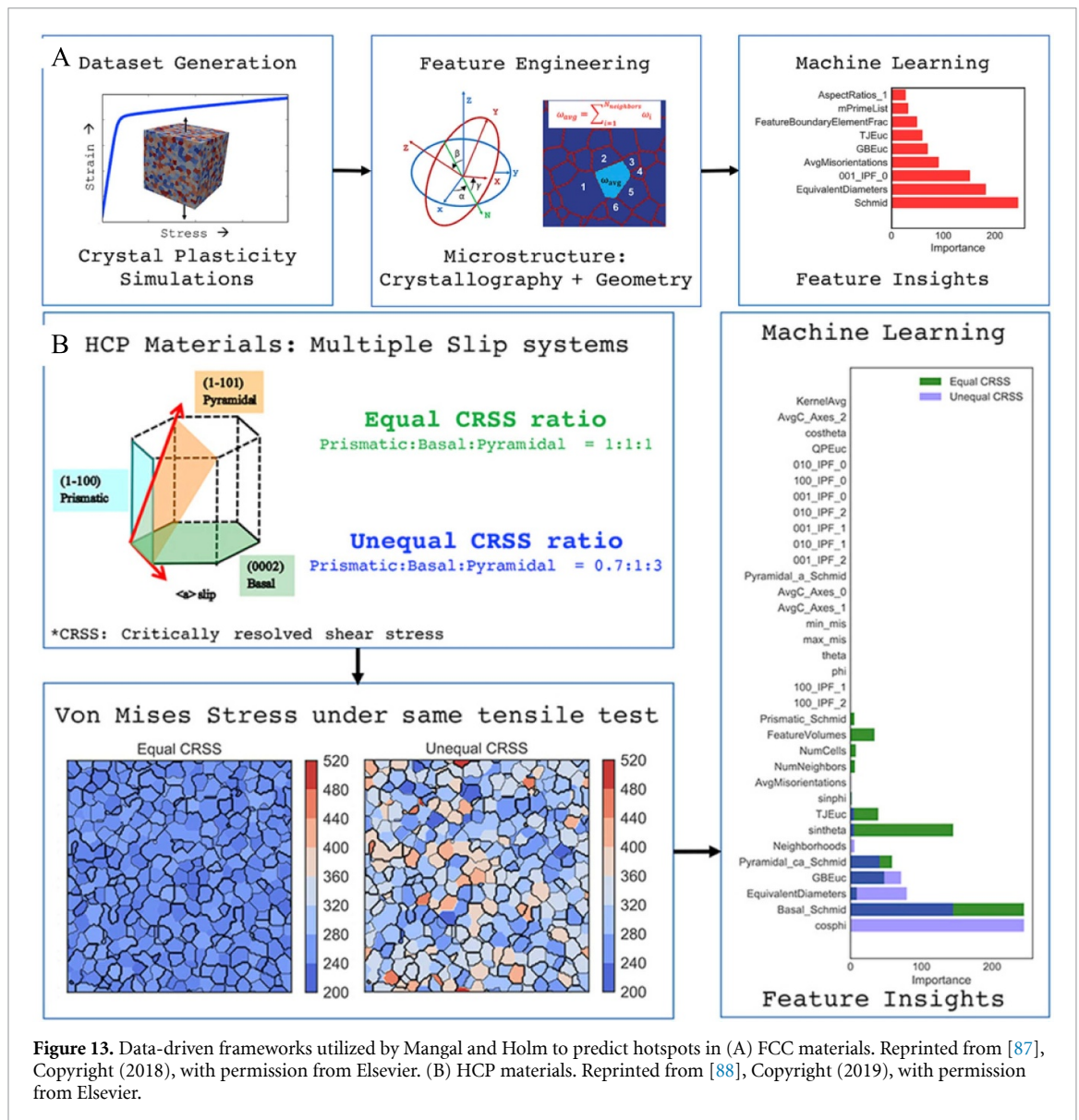


Figure 13. Data-driven frameworks utilized by Mangal and Holm to predict hotspots in (A) FCC materials. Reprinted from [87], Copyright (2018), with permission from Elsevier. (B) HCP materials. Reprinted from [88], Copyright (2019), with permission from Elsevier.

database of 512 synthetically generated surface height maps (figures 15(A) and (B)) that mimic the surface roughness of AM metals and predicted the strain energy metric for each surface. The images were augmented to 1536 data points by performing transformations on the data (rotations and mirroring), and the results showed that the convolutional neural network was able to achieve predictions with an R^2 value around 0.75 over surfaces that were unknown to the trained model (figure 15(A)). One interesting contribution of this work is the integration of a ‘viewpoint’ within the convolutional neural network for better interpretability by looking at the activation of the second convolution layer, as shown in figures 15(B)(d).

In a review of the efforts undertaken by GE, Aggour *et al* [91] detailed GE’s current efforts to implement artificial intelligence and machine learning methods to manufacturing techniques (both traditional and additive) to support design, processing, and inspection. In terms of MAM, GE presented a framework utilizing probabilistic machine learning and intelligent sampling and optimization protocols, coupled with high-throughput printing, testing, and characterization tools to accelerate process parameter development for LPBF MAM [92]. The process optimization framework was tested on the crack-susceptible Nickel-based alloys used in high temperature locations of aircraft engines. The protocol is based on a Bayesian hybrid model [93, 94], which utilizes a Gaussian process to develop a predictive, data-driven modeling technique, and the GE intelligent design and analysis of computer experiments, as shown in figure 16. The initial model is built with key inputs that are dependent on the processing parameters in use (e.g., laser power, scan speed, hatch spacing, and beam spot size), along with initial constraints and experimental results. In addition to predicting mean values of the model response, the data-driven model also provides uncertainty in the model

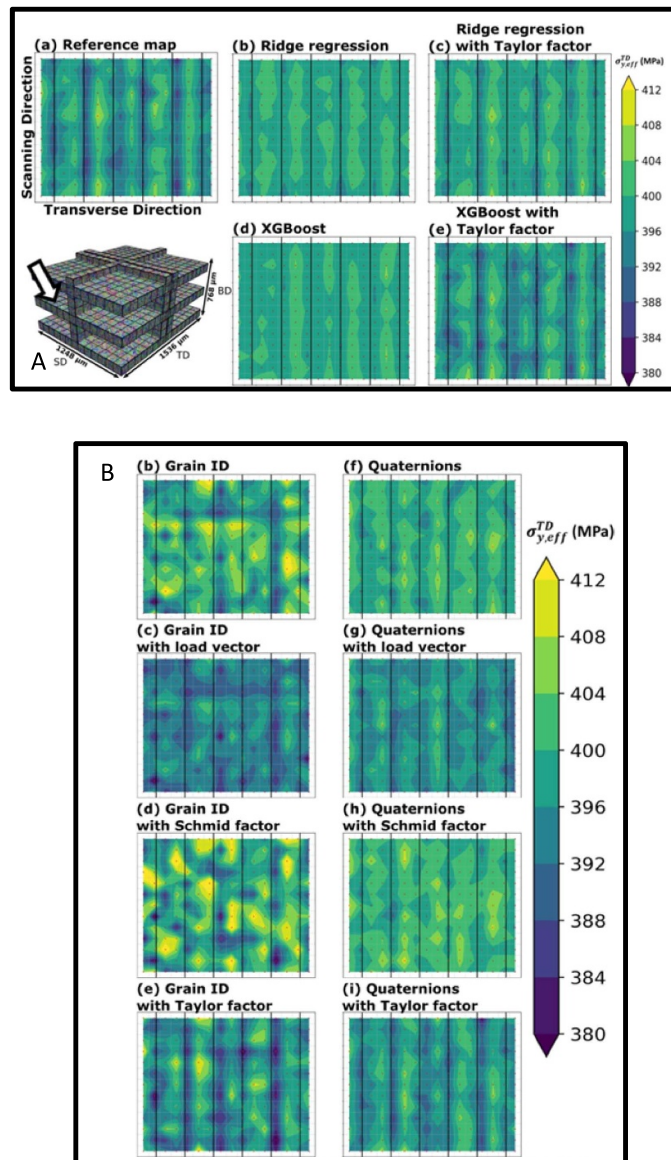


Figure 14. Comparison of property maps (viz., effective yield strength in the transverse direction) models for the holdout layer indicated in Build Domain D, predicted by: (A) different machine-learning models trained with 33–35 microstructural descriptors: (a) for reference, property map generated using crystal-plasticity modeling; (b) ridge regression, excluding M^{micro} from training set; (c) ridge regression, including M^{micro} in training set; (d) XGBoost, excluding M^{micro} from training set; (e) XGBoost, including M^{micro} in training set. Both models are able to recover the banded pattern present in (a). (B) Deep-learning models: A–(a) for reference, property map generated using crystal-plasticity modeling; (b)–(e) CNNs with grain ID as the main feature in the training data; (f)–(i) CNNs with quaternion as the main feature in the training data. Auxiliary features used to train each model include: (b),(f) none; (c),(g) one-hot vector representing the global loading direction; (d),(h) Schmid factor; (e),(i) micromechanical Taylor factor. Vertical black lines correspond to laser scan tracks. Reprinted from [89], Copyright (2020), with permission from Elsevier.

response. The uncertainty data is then used to iteratively identify new process parameter combinations for test coupons and builds. The framework has been used to minimize the model uncertainty or optimize quantities of interest such as porosity and mechanical properties and was shown to improve with increasing number of iterations. Using these tools, LPBF AM process parameter development cycle time has been reduced by about four times, as shown in figure 16(B) (from [95]), leading to significant savings in material cost, labor, and machine resources compared to traditional design-of-experiment methods. In a different application of part repair using AM, GE used machine learning to define toolpaths (physical path through which the spray nozzle moves when applying material) during a cold-spray process to improve yield and reduce manufacturing time and cost. To perform the repair, GE incorporated artificial intelligence to guide the 12 degree-of-freedom robotic arm movements and allow the robotic arm to adjust *in situ* to deliver high quality builds of parts such as aviation airfoils. Furthermore, GE shares their experience where the cold-spray deposition process may be improved by incorporating *in situ* 3D scanning of the deposition surface to

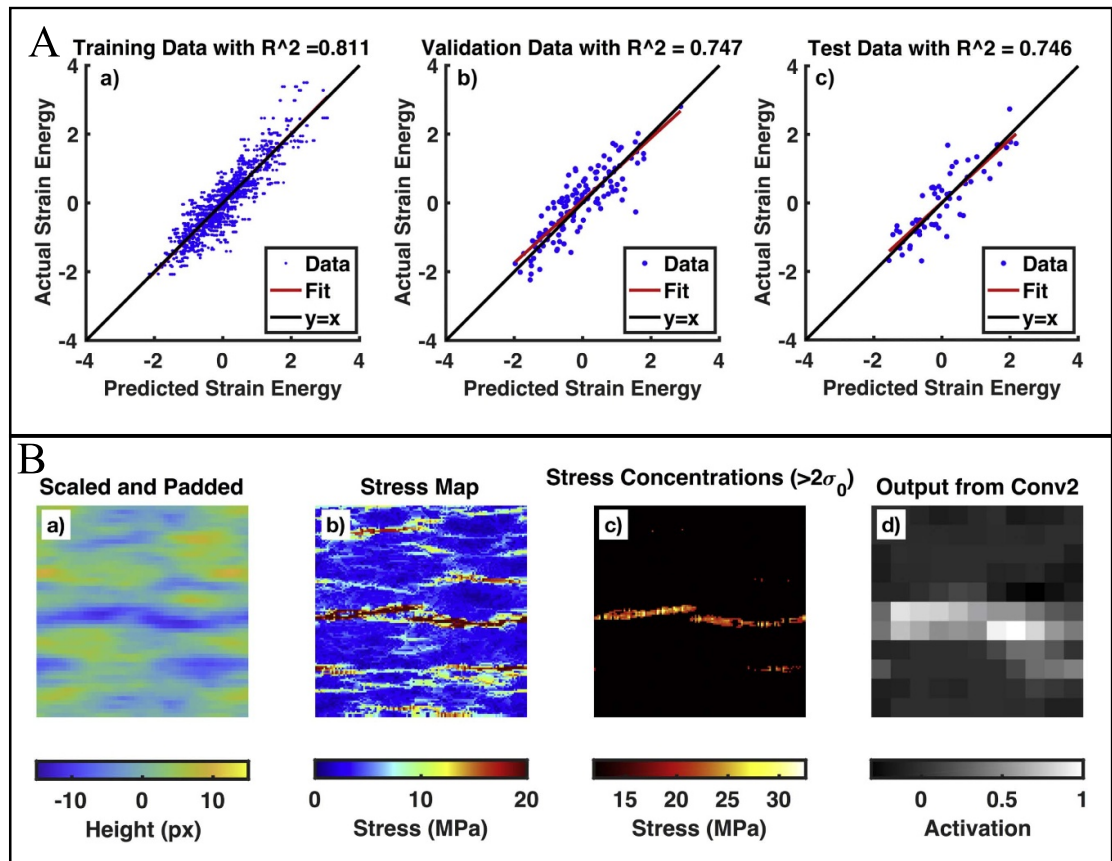


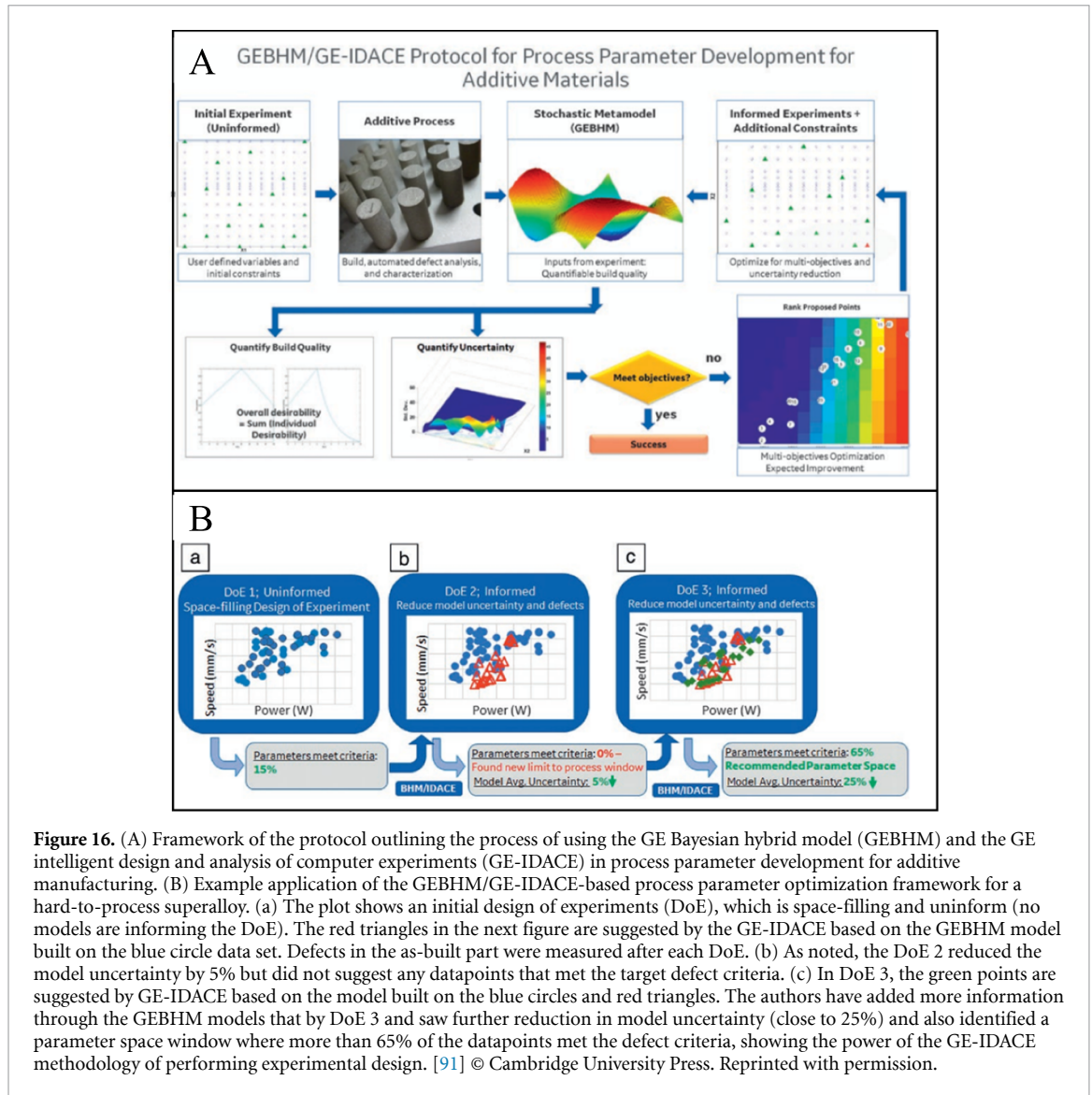
Figure 15. (A) Graphs showing the prediction of the strain energy metric for the 1536 training images in (a), the 128 reserved validation data, and the 64 test images. The R^2 value for each image is shown in the title of the graph. The high R^2 value for the test images (0.75) suggests that the CNN is capable of predicting the stress concentration metric for surfaces it has never seen. (B) Example of a single image. Section (a) shows the scaled and padded height map, and the stress field is shown in section (b) with the corresponding stress concentrations shown in (c). Section (d) shows the activation of the second convolution layer (Conv2), where the bright areas correspond to the stress concentrations shown in (c). Reprinted from [90], Copyright (2019), with permission from Elsevier.

intelligently modify the toolpaths on the fly, completing the deposition process more quickly using less material.

3.3. Process-structure-property relationship

To establish the p-s-p relationships for MAM, Yan *et al* [96] proposed a vision for the use of a data-driven multi-scale multi-physics framework, detailed in figure 17. The framework aims at quantitatively predicting the working performance of the final MAM products from given manufacturing parameters for a range of AM technologies (EBM, LPBF, and LENS) by combining process modeling, mechanical modeling, and data mining. The inputs to the model are derived from physics-driven models for all three links (process, structure, and property). Data mining is then used in a cycle of design–predict–optimize, and the preliminary results were validated experimentally. In the future, real-time monitoring is envisioned by Yan *et al* as a means to validate the numerical models.

In another effort, Wang *et al* [97] employed a physics-informed data-driven surrogate model applicable for MAM EBM technique. Surrogate models for each of the process, structure, and property links were trained using data from physics-driven results. Namely, the melt pool was simulated using FEM (280 thermal simulations), grain growth was modeled using the phase-field method (150 grain growth simulations), and the properties were predicted using the FFT crystal plasticity method (150 elasto-viscoplastic FFT simulations) as shown in figure 18(A). The authors then trained a Gaussian process model to serve as a surrogate model for each of the levels (process, structure, property) and tested the model quality using leave-one-out cross-validation, figure 18(B). Even though the authors communicated the difficulty of generating the thermal surrogate model that is a high-dimensional response, results were shown to capture the temperature fields with reasonable accuracy (figure 18(C)). The final step of Wang's framework [97] consisted of uncertainty quantification using a brute-force Monte Carlo approach, from process to property



estimates. Sensitivity analysis was applied to a test case of EBM Ti-6Al-4V alloy, as summarized in figure 18(D). More details about the uncertainty quantification can be found in [97].

As it is clear from the previously presented works, the scientific literature provides a wealth of modeling, simulation, and experimental data that captures the processing, structure, and properties of materials. Mining this information from the literature requires the development of a number of tools [98]. Using these tools, data-driven techniques provide a powerful means to learn from ‘prior experience’. Extracting information from the literature, Liu *et al* [99] built a GPR model to correlate process conditions (i.e., laser power, laser speed, hatch spacing, layer thickness, and powder size) across seven laser powder bed machines to density and microhardness of the finished part. Then, using this model, they predicted the build conditions necessary to achieve target densities and microhardness on a build platform that was not part of the original training data. This is one of many studies that leverage existing information in the scientific literature to map process–property relationships directly [99–101].

It is worth noting that another category of data-driven models, models not captured in this review, and not feasible to implement using physics-based models but worth mentioning due to their implementation using data-driven approaches, are property–property models. That is, models that use measurable properties, such as yield strength or springback, to model and predict qualitative properties (e.g., formability and rollability). An example of these efforts, which have not yet been implemented for AM materials, have looked at formability or rollability [102–105].

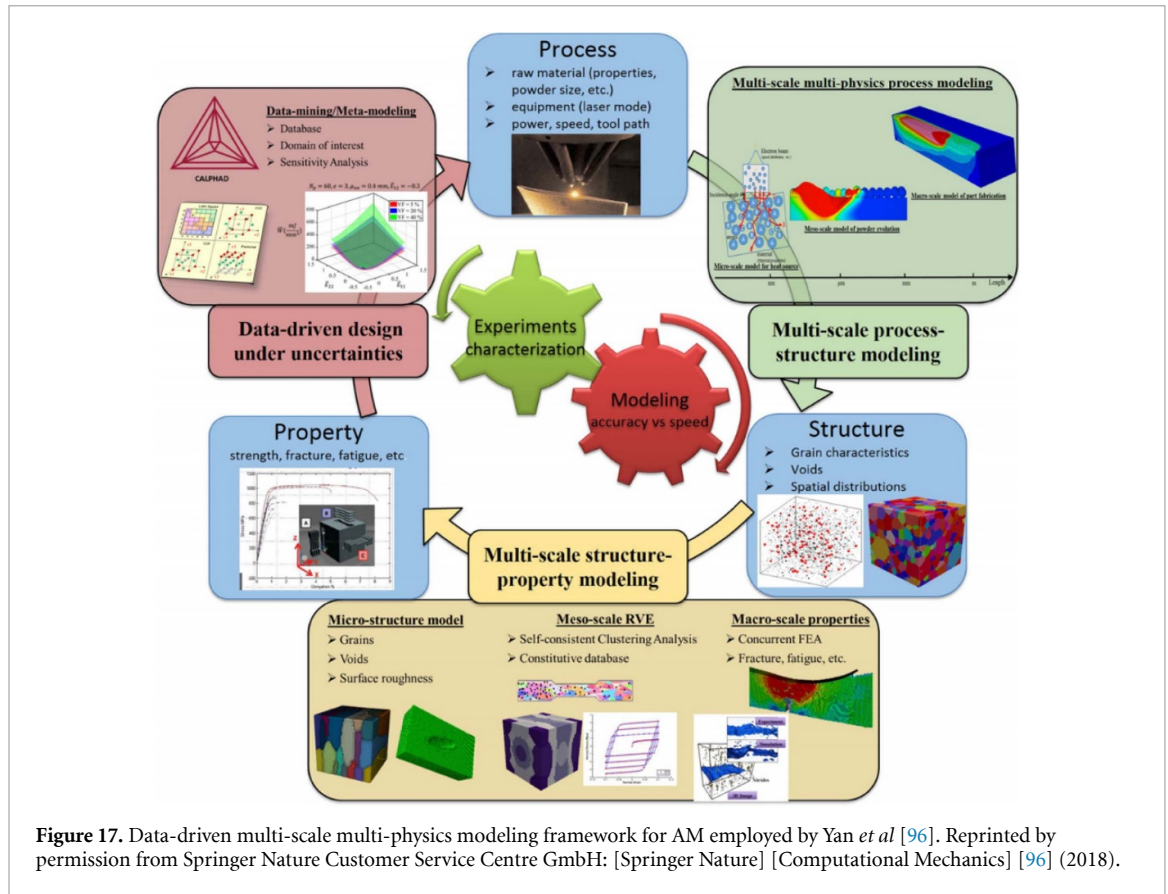


Figure 17. Data-driven multi-scale multi-physics modeling framework for AM employed by Yan *et al* [96]. Reprinted by permission from Springer Nature Customer Service Centre GmbH: [Springer Nature] [Computational Mechanics] [96] (2018).

4. Discussion

At the highest level, physics-driven and data-driven models in MAM share a similar objective: to predict, for a given set of AM process conditions, the resulting material microstructure and/or mechanical response of the material. However, the manner in which this is achieved by each type of model is fundamentally different, leading to advantages and disadvantages for each of the modeling approaches. The advantages and disadvantages of both physics-driven and data-driven models used in predicting the complex p-s-p relationships in MAM are summarized in table 1 and detailed in the following discussion.

Despite their common aim, fundamental differences exist between physics-driven and data-driven approaches. Physics-driven modeling approaches require a set of governing constitutive equations that represent the physical phenomena underpinning the process, structure, and property relationships. Numerical solvers are then used to solve the governing equations over some finite domain, which can be expensive and therefore preclude exploration of a high-dimensional MAM design space. As discussed in depth by Lindgren and co-workers [51, 106], the proper choice of constitutive model (and corresponding material parameters used within each constitutive model) is imperative to making accurate predictions for both welding and, by extension, MAM. On the other hand, the training algorithms underpinning many data-driven models, including machine-learning models, are designed to handle problems for which the governing equations relating inputs to outputs are typically unknown *a priori* but rather inferred through the correlative relationship between the control (input) variables and the response (output) variables.

Both types of modeling approaches require that the parameters embedded within the models are calibrated; however, the calibration procedure is different between the two approaches. For physics-based models, parameter calibration is generally performed in either an ad hoc manner or by an estimation algorithm (for further discussion, see work by An *et al* [107]). Some parameter values can be obtained from (pre-calibrated) values in the literature; while others might be calibrated through either trial-and-error or optimization processes until satisfactory comparison with some experimental or ground-truth measure. On the other hand, the parameters embedded within a data-driven model (e.g., the weights and biases within a machine-learning model) are calibrated using a formalized process known as ‘training’, through which the parameters are iteratively adjusted. For example, neural networks are typically trained/optimized with some form of a gradient-descent algorithm.

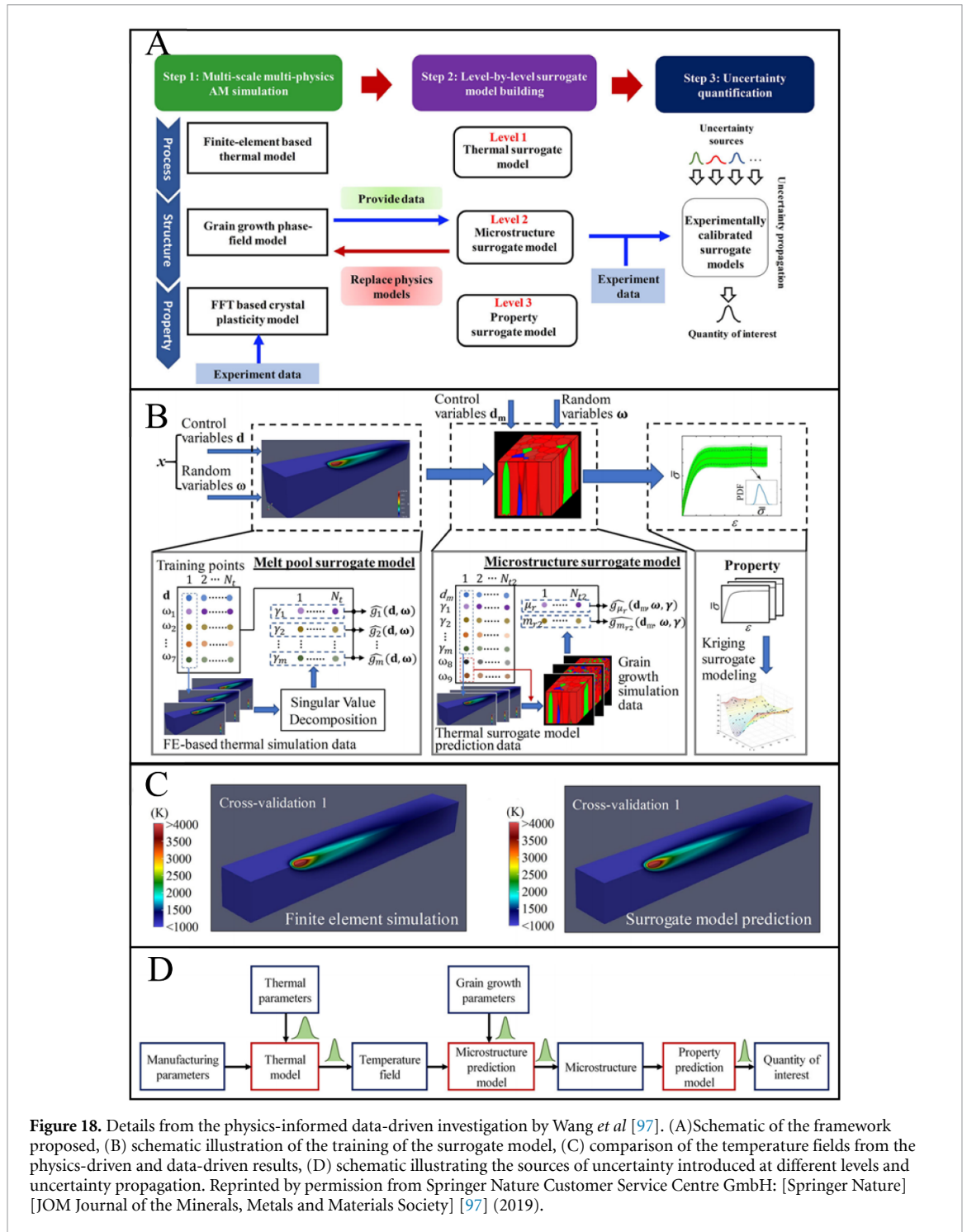


Figure 18. Details from the physics-informed data-driven investigation by Wang *et al* [97]. (A) Schematic of the framework proposed, (B) schematic illustration of the training of the surrogate model, (C) comparison of the temperature fields from the physics-driven and data-driven results, (D) schematic illustrating the sources of uncertainty introduced at different levels and uncertainty propagation. Reprinted by permission from Springer Nature Customer Service Centre GmbH: [Springer Nature] [JOM Journal of the Minerals, Metals and Materials Society] [97] (2019).

Another difference between physics-driven and data-driven modeling approaches is in their treatment of probabilistics and uncertainty. In physics-driven models, the resulting solution to a set of governing constitutive equations for a single simulation typically provides a deterministic prediction, whether of a microstructural volume, a residual-stress field, or a full-field mechanical response. Of course, owing to the inherent randomness that can be injected at different stages (e.g., the assignment of random crystal orientations to grain-nucleation points during prediction of the phase-transformation process), one could, in theory, run many deterministic simulations to populate a statistical model. In practice, this is generally intractable due to the computational cost to perform even a single multi-scale, multi-physics numerical simulation. On the other hand, data-driven models rely on the correlative relationship between the independent variables (inputs) and the dependent variables (outputs), using statistical tools to solve for this correlative relationship, tools built on point estimation, hypothesis testing, and statistical inference. Hence,

Table 1. Advantages and disadvantages of physics-driven and data-driven approaches in the prediction of the p-s-p relationships in MAM.

	Physics-driven models	Data-driven models
Advantages	<ul style="list-style-type: none"> • Simulate complex multiphase phenomena • High-fidelity representation of the underlying physics • Results are generally interpretable 	<ul style="list-style-type: none"> • Low computational cost enables the exploration of high-dimensional design space • Predictions of complex phenomena when physics are implicitly embedded within training data • Formalized calibration via model training
Disadvantages	<ul style="list-style-type: none"> • Need for input parameter calibration • Need for simplifications • High computational cost precludes exploration of the complete design space • Valid within the limitations and context of their original formulations 	<ul style="list-style-type: none"> • Lack of interpretability • Complicated validation • Potential for model bias and brittleness (easy to fool) • Valid within the limitations and context of their original formulations

many types of data-driven models (e.g., Bayesian networks) account inherently for probabilistic predictions and various sources of uncertainty.

Even though physics-driven models face a high computational cost that could be avoided using data-driven models, the latter methods are susceptible to bias and brittleness, depending on the type and amount of data used in training. A biased data-driven model is one that fails to capture the trends present in the data set used in training and as a result may lead to inaccurate predictions. On the other hand, a brittle model (in analogy to metalworking) is one that is less malleable to a wide range of data besides the assumptions made over the training data. Therefore, a brittle model may be easily fooled if those assumptions are not stated, hence not met by the input data. As a result, while using data-driven models, one should maintain a good judgment of the validity and limitations of the model in terms of its original formulation.

While physics-driven models have a more explicit and obvious representation of the governing physics, data-driven models—if trained using an appropriate data set—might encode the physics through a more implicit and hidden representation. For example, if a data-driven model is trained using experimental data for which the training input and output have some fundamental physical relationship among them, then it is possible that the data-driven model could learn to predict the output accurately⁵, even though it has no explicit knowledge of the governing constitutive relationships. Because data-driven models do not include an obvious representation of the physics, the lack of interpretability of their predictions is perhaps the biggest source of criticism of data-driven approaches, and they tend to be regarded as ‘black-box’ models. A recent article by Holm entitled, ‘In defense of the black box’ [108], presents a rational explanation for the use of such models in specific applications. For example, the black box method can be used where the cost of a wrong answer is low relative to the value of a correct answer, where a machine can perform a monotonous task, or where the black box can actually produce similar or even better results than experiential judgment.

Both data-driven and physics-based models will play key roles in realizing the ICME approach for design of MAM. Despite the criticism of data-driven models as black boxes, the use of data-driven approaches will likely play a key role in linking process to structure to property for MAM, given the high dimensionality of the MAM design space, the inherent uncertainty associated with MAM, and the computational expense of integrating p-s-p linkages using purely physics-based approaches. The rapid-prediction capabilities of data-driven models make them especially attractive for high-throughput predictions, AM-parameter optimization, and real-time AM-process monitoring [77, 109]. As Francois *et al* [5] point out in their article highlighting challenges and opportunities for modeling p-s-p-performance in AM, ‘To be useful, the predictions must be able to be run in a reasonable amount of time while retaining sufficient physics fidelity so as to yield trustworthy results’. Similarly, the development and use of physics-based models will continue to play a key role in MAM by providing improved scientific understanding of the fundamental relationships

⁵ Data-driven models are inherently unable to account for bias that might exist in the training data set.

between AM processes, structures, and properties and by providing ‘ground-truth’ predictions to assess the predictions made by data-driven models.

4.1. Potential for future work

Both physics-driven and data-driven modeling approaches result in the generation of big/rich data. With the wealth of data in the scientific community comes a growing challenge of data storage, management, and sharing, especially for data relevant to AM. Materials informatics, which accelerate materials, products, and manufacturing innovations, is the way forward in facilitating materials innovation, as detailed in the discussion by Ramakrishna *et al* [110]. With the availability of the large datasets, a number of databases (such as Citrination [111], the AM Materials Database developed by NIST [112], and the Materials Commons repository by PRISMS [113]) are available to the scientific community to store and share scientific materials data. This list is not exhaustive but is intended to offer examples of how the community is handling data storage and curation. More efforts are needed in unifying and sharing data via established materials databases.

A new paradigm of theory-guided data science was introduced by Karpatne *et al* [114], in which physics-driven knowledge and data science are integrated to leverage the potential of data science in the advancement of materials discovery. The strength of ICME is the ability to incorporate physics-based modeling data into data-driven models [67, 115–119]. In the specific application of predicting p-s-p relationships for MAM design, output from physics-driven model predictions for MAM digital twins can be used as training data for a data-driven model; in which case, the latter could be trained to predict the relevant output, despite not having explicit access to the governing equations. In turn, the forward predictions from trained data-driven models could be used to inform subsequent iterations of physics-driven models. In this way, data-driven and physics-driven models are not mutually exclusive but can be used to inform each other.

Combining both approaches will allow the physics-driven approaches to support the interpretability of results from the data-driven techniques, revealing the contents of the black box or even making it slightly less opaque. Outputs from the physics-driven models used as inputs to the data-driven models may provide larger sets of input data for training based on governing physics. In turn, the interpretability of the data-driven modeling results may support in formulating more sophisticated governing equations to be used in the physics-driven models in support of the digital twin of MAM parts. The understanding of the p-s-p relationships in the digital twins of MAM parts is ultimately indispensable to the exploration of the MAM design space, which is not feasible solely through experimentation. The exploration of the process design space of MAM will enable the process optimization and the production of parts with specific material properties for a given application to enable qualification and certification of MAM parts.

Acknowledgments

N.K., B.K., and A.D.S. would like to acknowledge support from the Department of Defense—Office of Economic Adjustment under Award No. ST1605-19-03. X.L. and W.T. wish to gratefully acknowledge the partial financial support provided by the National Science Foundation under Grant No. CMMI-1752218.

ORCID iDs

Nadia Kouraytem  <https://orcid.org/0000-0002-7183-2774>

Ashley D Spear  <https://orcid.org/0000-0002-3933-3131>

References

- [1] Mukherjee T and DebRoy T 2019 A digital twin for rapid qualification of 3D printed metallic components *Appl. Mater. Today* **14** 59–65
- [2] Tolle K M, Tansley D S W and Hey A J 2011 The fourth paradigm: data-intensive scientific discovery [point of view] *Proc. IEEE* **99** 1334–7
- [3] Hey T, Tansley S and Tolle K 2009 *The Fourth Paradigm: Data-Intensive Scientific Discovery* (USA: Microsoft Research)
- [4] Agrawal A and Choudhary A 2016 Perspective: materials informatics and big data: realization of the ‘fourth paradigm’ of science in materials science *APL Mater.* **4** 053208
- [5] Francois M M *et al* 2017 Modeling of additive manufacturing processes for metals: challenges and opportunities *Curr. Opin. Solid State Mater. Sci.* **21** 198–206
- [6] Rodgers T M, Bishop J E and Madison J D 2018 Direct numerical simulation of mechanical response in synthetic additively manufactured microstructures *Model. Simul. Mater. Sci. Eng.* **26** 055010
- [7] King W E, Anderson A T, Ferencz R M, Hodge N E, Kamath C, Khairallah S A and Rubenchik A M 2015 Laser powder bed fusion additive manufacturing of metals; physics, computational, and materials challenges *Appl. Phys. Rev.* **2** 41304
- [8] Khairallah S A, Anderson A T, Rubenchik A and King W E 2016 Laser powder-bed fusion additive manufacturing: physics of complex melt flow and formation mechanisms of pores, spatter, and denudation zones *Acta Mater.* **108** 36–45

- [9] Yan W, Ge W, Qian Y, Lin S, Zhou B, Liu W K, Lin F and Wagner G J 2017 Multi-physics modeling of single/multiple-track defect mechanisms in electron beam selective melting *Acta Mater.* **134** 324–33
- [10] Panwisawas C, Qiu C, Anderson M J, Sovani Y, Turner R P, Attallah M M, Brooks J W and Basoalto H C 2017 Mesoscale modelling of selective laser melting: thermal fluid dynamics and microstructural evolution *Comput. Mater. Sci.* **126** 479–90
- [11] Körner C, Attar E and Heinel P 2011 Mesoscopic simulation of selective beam melting processes *J. Mater. Process. Technol.* **211** 978–87
- [12] Khairallah S A, Anderson A T, Rubenchik A M and King W E 2017 Laser powder-bed fusion additive manufacturing: physics of complex melt flow and formation mechanisms of pores, spatter, and denudation zones *Additive Manufacturing Handbook: Product Development for the Defense Industry* vol 108 pp 613–28
- [13] Kouraytem N, Li X, Cunningham R, Zhao C, Parab N, Sun T, Rollett A D, Spear A D and Tan W 2019 Effect of laser-matter interaction on molten pool flow and keyhole dynamics *Phys. Rev. Appl.* **11** 1–16
- [14] Ly S, Rubenchik A M, Khairallah S A, Guss G and Matthews M J 2017 Metal vapor micro-jet controls material redistribution in laser powder bed fusion additive manufacturing *Sci. Rep.* **7** 1–12
- [15] Zheng M, Wei L, Chen J, Zhang Q, Zhong C, Lin X and Huang W 2019 A novel method for the molten pool and porosity formation modelling in selective laser melting *Int. J. Heat Mass Transf.* **140** 1091–105
- [16] Hodge N E, Ferencz R M and Solberg J M 2014 Implementation of a thermomechanical model for the simulation of selective laser melting *Comput. Mech.* **54** 33–51
- [17] Heigel J C, Michaleris P and Reutzel E W 2015 Thermo-mechanical model development and validation of directed energy deposition additive manufacturing of Ti-6Al-4V *Addit. Manuf.* **5** 9–19
- [18] Denlinger E R, Heigel J C and Michaleris P 2015 Residual stress and distortion modeling of electron beam direct manufacturing Ti-6Al-4V *Proc. Inst. Mech. Eng. B* **229** 1803–13
- [19] Parry L, Ashcroft I A and Wildman R D 2016 Understanding the effect of laser scan strategy on residual stress in selective laser melting through thermo-mechanical simulation *Addit. Manuf.* **12** 1–15
- [20] Luo Z and Zhao Y 2018 A survey of finite element analysis of temperature and thermal stress fields in powder bed fusion additive manufacturing *Addit. Manuf.* **21** 318–32
- [21] Li X and Tan W 2018 Numerical investigation of effects of nucleation mechanisms on grain structure in metal additive manufacturing *Comput. Mater. Sci.* **153** 159–69
- [22] Koepf J A, Gotterbarm M R, Markl M and Körner C 2018 3D multi-layer grain structure simulation of powder bed fusion additive manufacturing *Acta Mater.* **152** 119–26
- [23] Lian Y, Gan Z, Yu C, Kats D, Liu W K and Wagner G J 2019 A cellular automaton finite volume method for microstructure evolution during additive manufacturing *Mater. Des.* **169** 107672
- [24] Wei H L, Knapp G L, Mukherjee T and DebRoy T 2019 Three-dimensional grain growth during multi-layer printing of a nickel-based alloy Inconel 718 *Addit. Manuf.* **25** 448–59
- [25] Rodgers T M, Madison J D and Tikare V 2017 Simulation of metal additive manufacturing microstructures using kinetic Monte Carlo *Comput. Mater. Sci.* **135** 78–89
- [26] Dehoff R R, Kirka M, Sames W J, Bilheux H, Tremsin A S, Lowe L E and Babu S S 2015 Site specific control of crystallographic grain orientation through electron beam additive manufacturing *Mater. Sci. Technol.* **31** 931–8
- [27] Nie P, Ojo O A and Li Z 2014 Numerical modeling of microstructure evolution during laser additive manufacturing of a nickel-based superalloy *Acta Mater.* **77** 85–95
- [28] Sahoo S and Chou K 2016 Phase-field simulation of microstructure evolution of Ti-6Al-4V in electron beam additive manufacturing process *Addit. Manuf.* **9** 14–24
- [29] Keller T et al 2017 Application of finite element, phase-field, and CALPHAD-based methods to additive manufacturing of Ni-based superalloys *Acta Mater.* **139** 244–53
- [30] Wu L and Zhang J 2018 Phase field simulation of dendritic solidification of Ti-6Al-4V during additive manufacturing process *JOM* **70** 2392–9
- [31] Wang X, Liu P W, Ji Y, Liu Y, Horstemeyer M H and Chen L 2019 Investigation on microsegregation of IN718 alloy during additive manufacturing via integrated phase-field and finite-element modeling *J. Mater. Eng. Perform.* **28** 657–65
- [32] Smith M, Guan Z and Cantwell W J 2013 Finite element modelling of the compressive response of lattice structures manufactured using the selective laser melting technique *Int. J. Mech. Sci.* **67** 28–41
- [33] Hedayati R, Hosseini-Toudeshky H, Sadighi M, Mohammadi-Aghdam M and Zadpoor A A 2016 Computational prediction of the fatigue behavior of additively manufactured porous metallic biomaterials *Int. J. Fatigue* **84** 67–79
- [34] Taheri Andani M et al 2016 Achieving biocompatible stiffness in NiTi through additive manufacturing *J. Intell. Mater. Syst. Struct.* **27** 2661–71
- [35] Kramer S L B et al 2019 The third Sandia Fracture Challenge: predictions of ductile fracture in additively manufactured metal *Int. J. Fract.* **218** 5–61
- [36] Brinckmann S 2019 A framework for material calibration and deformation predictions applied to additive manufacturing of metals *Int. J. Fract.* **218** 85–95
- [37] Behzadinasab M and Foster J T 2019 The third Sandia Fracture Challenge: peridynamic blind prediction of ductile fracture characterization in additively manufactured metal *Int. J. Fract.* **218** 97–109
- [38] Neilsen M K 2019 Predicting ductile tearing of additively manufactured 316L stainless steel *Int. J. Fract.* **218** 195–207
- [39] Keim V, Cerrone A and Nonn A 2019 Using local damage models to predict fracture in additively manufactured specimens *Int. J. Fract.* **218** 135–47
- [40] Spear A D et al 2019 The third Sandia Fracture Challenge: from theory to practice in a classroom setting *Int. J. Fract.* **218** 171–94
- [41] Sobotka J C, McFarland J M and Stein J 2019 Application of uncertainty quantification techniques to ductile damage predictions in the third Sandia Fracture Challenge *Int. J. Fract.* **218** 111–33
- [42] Karlson K N, Alleman C, Foulk J W, Manktelow K L, Ostien J T, Stender M E, Stershic A J and Veilleux M G 2019 Sandia Fracture Challenge 3: detailing the Sandia Team Q failure prediction strategy *Int. J. Fract.* **218** 149–70
- [43] Tancogne-Dejean T, Gorji M B, Pack K and Roth C C 2019 The third Sandia Fracture Challenge: deterministic and probabilistic modeling of ductile fracture of additively-manufactured material *Int. J. Fract.* **218** 209–29
- [44] Johnson K L, Emery J M, Hammetter C I, Brown J A, Grange S J, Ford K R and Bishop J E 2019 Predicting the reliability of an additively-manufactured metal part for the third Sandia Fracture Challenge by accounting for random material defects *Int. J. Fract.* **218** 231–43

- [45] Ahmadi A, Mirzaeifar R, Moghaddam N S, Turabi A S, Karaca H E and Elahinia M 2016 Effect of manufacturing parameters on mechanical properties of 316L stainless steel parts fabricated by selective laser melting: a computational framework *Mater. Des.* **112** 328–38
- [46] Taheri Andani M, Karamooz-Ravari M R, Mirzaeifar R and Ni J 2018 Micromechanics modeling of metallic alloys 3D printed by selective laser melting *Mater. Des.* **137** 204–13
- [47] Yan W, Lian Y, Yu C, Kafka O L, Liu Z, Liu W K and Wagner G J 2018 An integrated process–structure–property modeling framework for additive manufacturing *Comput. Methods Appl. Mech. Eng.* **339** 184–204
- [48] Erickson J, Rahman A and Spear A 2020 A void descriptor function to uniquely characterize pore networks and predict ductile-metal failure properties *Int. J. Fract.* **225** 47–67
- [49] Ozturk T and Rollett A D 2018 Effect of microstructure on the elasto-viscoplastic deformation of dual phase titanium structures *Comput. Mech.* **61** 55–70
- [50] Groeber M A and Jackson M A 2014 DREAM.3D: a digital representation environment for the analysis of microstructure in 3D *Integr. Mater. Manuf. Innov.* **3** 56–72
- [51] Lindgren L E, Lundbäck A, Fisk M, Pederson R and Andersson J 2016 Simulation of additive manufacturing using coupled constitutive and microstructure models *Addit. Manuf.* **12** 144–58
- [52] Johnson K L, Rodgers T M, Underwood O D, Madison J D, Ford K R, Whetten S R, Dagle D J and Bishop J E 2018 Simulation and experimental comparison of the thermo-mechanical history and 3D microstructure evolution of 304L stainless steel tubes manufactured using LENS *Comput. Mech.* **61** 559–74
- [53] Li C, Fu C H, Guo Y B and Fang F Z 2015 A multiscale modeling approach for fast prediction of part distortion in selective laser melting *J. Mater. Process. Technol.* **229** 703–12
- [54] Denlinger E R, Gouge M, Irwin J and Michaleris P 2017 Thermomechanical model development and in situ experimental validation of the Laser Powder-Bed Fusion process *Addit. Manuf.* **16** 73–80
- [55] Herriott C, Li X, Kouraytem N, Tari V, Tan W, Anglin B, Rollett A D and Spear A D 2019 A multi-scale, multi-physics modeling framework to predict spatial variation of properties in additive-manufactured metals *Model. Simul. Mater. Sci. Eng.* **27** 025009
- [56] Shi R, Khairallah S, Heo T W, Rolchigo M, McKeown J T and Matthews M J 2019 Integrated simulation framework for additively manufactured Ti-6Al-4V: melt pool dynamics, microstructure, solid-state phase transformation, and microelastic response *JOM* **71** 3640–55
- [57] Smith J, Xiong W, Yan W, Lin S, Cheng P, Kafka O L, Wagner G J, Cao J and Liu W K 2016 Linking process, structure, property, and performance for metal-based additive manufacturing: computational approaches with experimental support *Comput. Mech.* **57** 583–610
- [58] Hill J, Mulholland G, Persson K, Seshadri R, Wolverson C and Meredig B 2016 Materials science with large-scale data and informatics: unlocking new opportunities *MRS Bull.* **41** 399–409
- [59] Whitley D 1993 A genetic algorithm tutorial *Stat. Comput.* **4** 65–85
- [60] Kennedy J and Eberhart R 1995 Particle swarm optimization *Proc. ICNN'95—Int. Conf. Neural Networks (IEEE) 1942–1948*
- [61] Quinonero-Candela J and Rasmussen C E 2005 A unifying view of sparse approximate Gaussian process regression *J. Mach. Learn. Res.* **6** 1939–59
- [62] Fine T L 2005 Fundamentals of artificial neural networks [book reviews] *IEEE Trans. Inf. Theory* **42** 1322
- [63] Vovk V 2013 Kernel ridge regression *Empirical Inference* (Berlin, Heidelberg: Springer) pp 105–16
- [64] Fullwood D T, Niezgodka S R and Kalidindi S R 2008 Microstructure reconstructions from 2-point statistics using phase-recovery algorithms *Acta Mater.* **56** 942–8
- [65] Niezgodka S R, Fullwood D T and Kalidindi S R 2008 Delineation of the space of 2-point correlations in a composite material system *Acta Mater.* **56** 5285–92
- [66] Honnibal M and Johnson M 2015 An improved non-monotonic transition system for dependency parsing *Conf. Proc.—EMNLP 2015 Conf. Empir. Methods Nat. Lang. Process* pp 1373–8
- [67] Tancret F 2013 Computational thermodynamics, Gaussian processes and genetic algorithms: combined tools to design new alloys *Model. Simul. Mater. Sci. Eng.* **21** 045013
- [68] Smith D H, Bicknell J, Jorgensen L, Patterson B M, Cordes N L, Tsukrov I and Knezevic M 2016 Microstructure and mechanical behavior of direct metal laser sintered Inconel alloy 718 *Mater. Charact.* **113** 1–9
- [69] Kappes B B and Ciobanu C V 2015 Materials screening through GPU accelerated topological mapping *Mater. Manuf. Process.* **30** 529–37
- [70] Trzaska J and Dobrzański L A 2007 Modelling of CCT diagrams for engineering and constructional steels *J. Mater. Process. Technol.* **192–193** 504–10
- [71] Dobrzański L A and Trzaska J 2004 Application of neural networks for the prediction of continuous cooling transformation diagrams *Comput. Mater. Sci.* **30** 251–9
- [72] Wang J, van der Wolk P J and van der Zwaag S 1999 Effects of carbon concentration and cooling rate on continuous cooling transformations predicted by artificial neural network *ISIJ Int.* **39** 38–46
- [73] Popova E, Rodgers T M, Gong X, Cecen A, Madison J D and Kalidindi S R 2017 Process-structure linkages using a data science approach: application to simulated additive manufacturing data *Integr. Mater. Manuf. Innov.* **6** 54–68
- [74] Wang Z, Liu P, Xiao Y, Cui X, Hu Z and Chen L 2019 A data-driven approach for process optimization of metallic additive manufacturing under uncertainty *Trans. ASME, J. Manuf. Sci. Eng.* **141** 081004
- [75] Shevchik S A, Kenel C, Leinenbach C and Wasmer K 2018 Acoustic emission for in situ quality monitoring in additive manufacturing using spectral convolutional neural networks *Addit. Manuf.* **21** 598–604
- [76] Wasmer K, Le-Quang T, Meylan B and Shevchik S A 2019 In situ quality monitoring in am using acoustic emission: a reinforcement learning approach *J. Mater. Eng. Perform.* **28** 666–72
- [77] Scime L and Beuth J 2019 Using machine learning to identify in-situ melt pool signatures indicative of flaw formation in a laser powder bed fusion additive manufacturing process *Addit. Manuf.* **25** 151–65
- [78] Scime L and Beuth J 2018 Anomaly detection and classification in a laser powder bed additive manufacturing process using a trained computer vision algorithm *Addit. Manuf.* **19** 114–26
- [79] Caggiano A, Zhang J, Alfieri V, Caiazza F, Gao R and Teti R 2019 Machine learning-based image processing for on-line defect recognition in additive manufacturing *CIRP Ann.* **68** 451–4
- [80] Azimi S M, Britz D, Engstler M, Fritz M and Mücklich F 2018 Advanced steel microstructural classification by deep learning methods *Sci. Rep.* **8** 1–14

- [81] DeCost B L, Jain H, Rollett A D and Holm E A 2017 Computer vision and machine learning for autonomous characterization of AM powder feedstocks *JOM* **69** 456–65
- [82] Sun B and Barnard A S 2019 Visualising multi-dimensional structure/property relationships with machine learning *J. Phys. Mater.* **2** 034003
- [83] Schleder G R, Padilha A C M, Acosta C M, Costa M and Fazzio A 2019 From DFT to machine learning: recent approaches to materials science—a review *J. Phys. Mater.* **2** 032001
- [84] Koeppel A, Hernandez Padilla C A, Voshage M, Schleifenbaum J H and Markert B 2018 Efficient numerical modeling of 3D-printed lattice-cell structures using neural networks *Manuf. Lett.* **15** 147–50
- [85] Gupta A, Cecen A, Goyal S, Singh A K and Kalidindi S R 2015 Structure-property linkages using a data science approach: application to a non-metallic inclusion/steel composite system *Acta Mater.* **91** 239–54
- [86] Jung J, Yoon J I, Park H K, Kim J Y and Kim H S 2019 An efficient machine learning approach to establish structure-property linkages *Comput. Mater. Sci.* **156** 17–25
- [87] Mangal A and Holm E A 2018 Applied machine learning to predict stress hotspots I: face centered cubic materials *Int. J. Plast.* **111** 122–34
- [88] Mangal A and Holm E A 2019 Applied machine learning to predict stress hotspots II: hexagonal close packed materials *Int. J. Plast.* **114** 1–14
- [89] Herriott C and Spear A D 2020 Predicting microstructure-dependent mechanical properties in additively manufactured metals with machine- and deep-learning methods *Comput. Mater. Sci.* **175** 109599
- [90] Kantzos C, Lao J and Rollett A 2019 Design of an interpretable Convolutional Neural Network for stress concentration prediction in rough surfaces *Mater. Charact.* **158** 109961
- [91] Aggour K S et al 2019 Artificial intelligence/machine learning in manufacturing and inspection: a GE perspective *MRS Bull.* **44** 545–58
- [92] Dheeradhada V, Chennimalai Kumar N, Gupta V K, Dial L, Vinciguerra J and Hanlon T 2020 Machine Learning Assisted Development in Additive Manufacturing
- [93] Chennimalai Kumar N, Subramaniyan A K and Wang L 2012 Improving high-dimensional physics models through Bayesian calibration with uncertain data *ASME Turbo Expo 2012 Turbine Tech. Conf. Expo.*
- [94] Chennimalai Kumar N, Subramaniyan A K, Wang L and Wiggs G 2013 Calibrating transient models with multiple responses using Bayesian inverse techniques *ASME Turbo Expo 2013 Turbine Tech. Conf. Expo.*
- [95] Kristensen J, Ling Y, Asher I and Wang L 2016 Expected-improvement-based methods for adaptive sampling in multi-objective optimization problems *ASME 2016 Int. Des. Eng. Tech. Conf. Comput. Inf. Eng. Conf.*
- [96] Yan W et al 2018 Data-driven multi-scale multi-physics models to derive process–structure–property relationships for additive manufacturing *Comput. Mech.* **61** 521–41
- [97] Wang Z, Liu P, Ji Y, Mahadevan S, Horstemeyer M F, Hu Z, Chen L and Chen L Q 2019 Uncertainty quantification in metallic additive manufacturing through physics-informed data-driven modeling *JOM* **71** 2625–34
- [98] Hawizy L, Jessop D M, Adams N and Murray-Rust P 2011 ChemicalTagger: A tool for semantic text-mining in chemistry *J. Cheminf.* **3** 1–13
- [99] Liu S, Kappes B B, Stebner A P and Zhang X A Data-mining assisted learning framework to simultaneously optimize multiple properties in metals additive manufacturing
- [100] Menou E, Ramstein G, Bertrand E and Tancret F 2016 Multi-objective constrained design of nickel-base superalloys using data mining- and thermodynamics-driven genetic algorithms *Model. Simul. Mater. Sci. Eng.* **24** 055001
- [101] Kim E, Huang K, Saunders A, McCallum A, Ceder G and Olivetti E 2017 Materials synthesis insights from scientific literature via text extraction and machine learning *Chem. Mater.* **29** 9436–44
- [102] Forcellese A, Gabrielli F and Ruffini R 1998 Effect of the training set size on springback control by neural network in an air bending process *J. Mater. Process. Technol.* **80–81** 493–500
- [103] Forcellese A, Gabrielli F and Simoncini M 2011 Prediction of flow curves and forming limit curves of Mg alloy thin sheets using ANN-based models *Comput. Mater. Sci.* **50** 3184–97
- [104] Toros S and Ozturk F 2011 Flow curve prediction of Al–Mg alloys under warm forming conditions at various strain rates by ANN *Appl. Soft Comput. J.* **11** 1891–8
- [105] Liu J, Chang H, Hsu T Y and Ruan X 2000 Prediction of the flow stress of high-speed steel during hot deformation using a BP artificial neural network *J. Mater. Process. Technol.* **103** 200–5
- [106] Lindgren L E 2006 Numerical modelling of welding *Comput. Methods Appl. Mech. Eng.* **195** 6710–36
- [107] An D, Kim N H and Choi J H 2015 Practical options for selecting data-driven or physics-based prognostics algorithms with reviews *Reliab. Eng. Syst. Saf.* **133** 223–36
- [108] Holm E A 2019 In defense of the black box **364** 26–7
- [109] Everton S K, Hirsch M, Stavroulakis P I, Leach R K and Clare A T 2016 Review of in-situ process monitoring and in-situ metrology for metal additive manufacturing *Mater. Des.* **95** 431–45
- [110] Ramakrishna S, Zhang T Y, Lu W C, Qian Q, Low J S C, Yune J H R, Tan D Z L, Bressan S, Sanvito S and Kalidindi S R 2019 Materials informatics *J. Intell. Manuf.* **30** 2307–26
- [111] O'Mara J, Meredig B and Michel K 2016 Materials data infrastructure: a case study of the citrination platform to examine data import, storage, and access *JOM* **68** 2031–4
- [112] Lu Y, Witherell P and Donmez A 2017 A collaborative data management system for additive manufacturing *ASME 2017 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* vol 1 p [IDETC201768457](#)
- [113] Puchala B, Tarcea G, Marquis E A, Hedstrom M, Jagadish H V and Allison J E 2016 The materials commons: a collaboration platform and information repository for the global materials community *JOM* **68** 2035–44
- [114] Karpatne A, Atluri G, Faghmous J H, Steinbach M, Banerjee A, Ganguly A, Shekhar S, Samatova N and Kumar V 2017 Theory-guided data science: a new paradigm for scientific discovery from data *IEEE Trans. Knowl. Data Eng.* **29** 2318–31
- [115] Rettig R, Ritter N C, Helmer H E, Neumeier S and Singer R F 2015 Single-crystal nickel-based superalloys developed by numerical multi-criteria optimization techniques: design based on thermodynamic calculations and experimental validation *Model. Simul. Mater. Sci. Eng.* **23** 035004
- [116] Reed R C, Tao T and Warnken N 2009 Alloys-by-design: application to nickel-based single crystal superalloys *Acta Mater.* **57** 5898–913

- [117] Das P, Mukherjee S, Ganguly S, Bhattacharyay B K and Datta S 2009 Genetic algorithm based optimization for multi-physical properties of HSLA steel through hybridization of neural network and desirability function *Comput. Mater. Sci.* **45** 104–10
- [118] Jha D, Ward L, Paul A, Liao W K, Choudhary A, Wolverson C and Agrawal A 2018 ElemNet: deep learning the chemistry of materials from only elemental composition *Sci. Rep.* **8** 1–13
- [119] Egorov-Yegorov I S and Dulikravin G 2005 Chemical composition design of superalloys for maximum stress, temperature, and time-to-rupture using self-adapting response surface optimization *Mater. Manuf. Process.* **20** 569–90