Data-Driven Approaches Toward Smarter Additive Manufacturing

Chenxi Tian, Tianjiao Li, Jenniffer Bustillos, Shonak Bhattacharya, Talia Turnham, Jingjie Yeo,* and Atieh Moridi*

The latest industrial revolution, Industry 4.0, is driven by the emergence of digital manufacturing and, most notably, additive manufacturing (AM) technologies. The simultaneous material and structure forming in AM broadens the material and structural design space. This expanded design space holds a great potential in creating improved engineering materials and products that attract growing interests from both academia and industry. A major aspect of this growing interest is reflected in the increased adaptation of data-driven tools that accelerate the exploration of the vast design space in AM. Herein, the integration of data-driven tools in various aspects of AM is reviewed, from materials design in AM (i.e., homogeneous and composite material design) to structure design for AM (i.e., topology optimization). The optimization of AM tool path using machine learning for producing best-quality AM products with optimal material and structure is also discussed. Finally, the perspectives on the future development of holistically integrated frameworks of AM and data-driven methods are provided.

1. Introduction

Additive manufacturing (AM) is a process in which feedstock materials are joined layer by layer guided by a digital design. Various AM processes can be categorized into different material classes (polymer, ceramic, and metal), feedstock types (slurry, powder, wire/filament, and sheet), or energy sources (kinetic bonding and thermal fusion). Table 1 shows the most widely used AM processes.

While each process has different print resolution, print speed, and feedstock type, they share some common advantages that set AM apart from traditional manufacturing methods. Arguably, the most significant advantages are in manufacturing products with complex geometries and, in some cases, with multiple materials. The former enables lightweight designs of single components that can be topologically optimized as well as the possibility of consolidating multiple discrete parts, typically designed for conventional assembly, into a single component. The latter enables production of functionally graded parts or composites. Other complementary advantages of AM include reducing production lead time through elimination of tooling, being cost efficient for small or custom production runs, and supply chain transformation by skipping many traditional manufacturing steps. Significant advances in AM are steering AM away from merely a rapid prototyping tool and toward being a viable candidate for large-scale manufacturing of components, especially for the aerospace and biomedical industries.

The broadened material and geometrical design space of AM requires more efficient methods for discovering optimal designs. Some AM-specific challenges, such as anisotropic material properties and residual stress accumulation, require further developments in the materials used in AM as well as in fine tuning process parameters and tool path design. The complex underlying physics of AM processes also brings challenges in understanding the relationship between process parameters and quality of the product, thereby hindering industrial applications that harness the full potential of AM. However, the emergence of a diverse range of modern data-driven approaches has a great potential in enabling the AM community to overcome these challenges.

By unifying the first three scientific paradigms of theory, experiment, and computation/simulation, data-driven approaches are illuminating the way toward the fourth paradigm of science: data-driven science. Data-driven science is strongly galvanized by the enhanced ability to gather massive amounts of data, the rapid development of data infrastructures where big data are stored, as well as advances in methods of data mining that employ such data. Recently, with extensive advances in both computational hardware and algorithms, data-driven approaches are rapidly permeating many fields of science and technology. For AM in particular, data-driven approaches are increasingly being adopted. Large quantities of materials data have already been generated for, or can be utilized by, AM. Some of these data are from benchtop experiments and tests whereas a large portion of them are generated by...
In this way, data-driven methods inform the design of materials and manufacturing processes for AM. The datasets also enable high-throughput methods for material mining/screening, in which optimal materials are filtered according to a desired design attribute or technical requirement.\cite{14} In this way, data-driven approaches allow cost-effective material selection for AM. Thus, the application of such data-driven approaches in AM has the potential to address the challenges mentioned earlier.

In this Review, we highlight different crucial aspects of coupling data-driven approaches and AM. Although efficient data-driven approaches can simultaneously use various computational techniques to produce data or process data, the specifics of each computational technique will not be covered in this Review. Interested readers can learn more about these techniques from relevant reviews.\cite{19–21} We will focus on data-driven methods that are most applicable to AM of metallic products, while drawing parallels to developments in the field of polymer AM when such methods discussed are yet to be implemented in metallic AM. A detailed review is available elsewhere on the state-of-the-art techniques in data-driven AM that are more specific to polymeric materials.\cite{19} The data-driven methods discussed in our Review include various digital methods of high-throughput simulation and optimization, ML, neural networks, and genetic algorithms (GA). For a more detailed review on current developments in ML specifically for metallic AM, we refer the readers to another excellent review article.\cite{21} The structure of this Review is shown in Figure 1. Here, in the first portion, we focus on data-driven approaches that advance the development of AM in material discovery. For homogenous materials, we discuss how data-driven approaches can accelerate the discovery of materials most suited for AM’s process characteristics and intrinsic properties. For composite materials, we discuss the benefit of using data-driven tools to discover superior designs that can achieve enhanced mechanical performance and realize the full potential of AM. The second portion of our Review demonstrates how AI-aided structural design can take advantage of reduced constraints on component geometry in AM and how data-driven methods can accelerate the optimization of the tool path and process parameter design. In addition, discussions on relevant databases as fundamentals of data-driven approaches are included into each section of this Review.

\section{Homogeneous Material Design for AM}

The rapid pace of advances in manufacturing technologies across multiple industries requires significant enhancements in the performance of engineering materials, such as metals with mechanical properties that exceed those of traditional alloys.\cite{22,23} Yet, materials that are suitable for AM are limited. For instance, it is estimated that merely 0.2\% of commercial metallic alloys are printable.\cite{22,24} This statistic emphasizes the need for robust computational tools that can aid in the design of alloys that are suitable for the AM process. Efforts in computational materials design have been successful in screening promising new stable materials via density functional theory (DFT) and molecular dynamics (MD) simulations.\cite{23,25–27} Recently, along with demands from industry, materials science has seen a dramatic change in the landscape of materials design and specifically alloy design via data-driven approaches.\cite{23,25–27} Data-driven approaches based on the powerful advances in computational modeling have already demonstrated their remarkable potential for various applications in materials research. Hence, greater attention is needed from the AM community to harness the

\begin{table}[h]
\centering
\caption{List of different AM processes with their most commonly used material classes, feedstock types, and energy source.}
\begin{tabular}{|l|l|l|l|l|}
\hline
Process name & Material class & Feedstock & Energy source & Reference \\
\hline
FDM & Polymer & Filament & Resistance heating & [167–169] \\
SLA & Polymer & Liquid & Laser or light & [169,170] \\
SLS & Polymer, ceramic, or metal & Powder & Laser & [171–173] \\
PBF & Metal & Powder & Laser or electron beam & [174–177] \\
Binder jetting & Polymer, metal, or ceramic & Powder\textsuperscript{a)} & Thermal heating & [172,178] \\
DED & Metal & Powder or Wire & Laser, electron beam, or plasma arc & [179] \\
Cold spray & Metal & Powder & Kinetic energy & [180] \\
Wire arc additive mfg. (WAAM) & Metal & Wire & Plasma arc & [181] \\
Ultrasonic sheet lamination (UAM) & Metal & Sheet & Ultrasound & [182] \\
\hline
\end{tabular}
\textsuperscript{a)} Powder feedstock joint with liquid binder.
\end{table}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{Graphical illustration of the structure of the current Review.}
\end{figure}
data-driven research scheme to further expand the library of applicable materials.

2.1. DFT and High-Throughput DFT

From the standpoint of AM, first-principle atomistic modeling such as DFT and the corresponding data are highly beneficial in terms of teasing apart the complex physics underlying the manufacturing process. DFT simulates the collective behavior of atoms using quantum mechanical calculations to unravel material characteristics. These computations are vital for understanding the molecular properties that consequently affect the meso- and macroscale properties of materials. The core of atomistic modeling lies in accurate and efficient techniques for numerically solving the Schrödinger equation for complex systems that consist of multiple atoms.\[^{20,28}\] Examples of using DFT in materials science, and more specifically in materials design, include studying the formation of defects in metallic materials,\[^{29–31}\] tailoring mechanical properties of alloys,\[^{32–34}\] and polymeric materials design.\[^{20,35,36}\]

Using quantum mechanical calculation such as DFT to directly inform material research for AM is shown as the blue flow in Figure 2. For instance, in powder bed fusion (PBF) AM, the surface energies of different material phases drive the melt penetration phenomenon and strongly affect the quality of the resulting component.\[^{37}\] The surface energies are difficult to measure experimentally, but DFT calculations offer a way to determine these energies in several metallic surfaces.\[^{38,39}\] Dumontet et al. examined how the elastic properties of Ti–6Al–4V were altered by nonequilibrium structures derived from the high cooling rates in PBF-AM.\[^{40}\] Their DFT calculations showed that the α’ martensitic structure had lower Young’s modulus than the pure Ti-α phase, which was attributed to the influence of Al and V substitutional elements.\[^{40}\] In direct energy deposition (DED), the melt pool cooling rate is one of the most important factors that influence the microstructure and strength of the final product.\[^{41}\] A recent study focused on radiative cooling in DED-AM and obtained the dielectric constant of the nickel-based superalloy Inconel-718 via experiments as well as DFT calculations to determine the spectral reflectivity and emissivity.\[^{42}\] Due to the generality of the DFT calculation, these reported methods can be applied to a large variety of materials and can be readily turned into tools for materials design and quality control for AM.

Currently, DFT calculations are one of the most accepted approaches to assess material properties using their atomistic compositions and structures.\[^{14}\] However, the unprecedented boom in computational resources and rapid advances in data sciences greatly increased the awareness and ability of the material science community in storing and interpreting large quantities of data generated by these individual DFT calculations for identifying higher-level correlations and patterns for material discovery and design.\[^{14,43}\] Recently, aggregating tens and even hundreds of thousands of DFT calculations in large databases combined with data-driven methods has intensified the ability to conduct high-throughput DFT (HT DFT) simulations.\[^{44}\] One such large online database is the Open Quantum Materials Database (OQMD)\[^{17,44}\] that contains thermodynamic and structural properties of over 800 000 materials calculated using DFT and the numbers are growing steadily. OQMD was successfully used to search for the most stable set of structures in Mg-based ternary alloy systems at a given composition using grand canonical linear programming.\[^{45}\] OQMD was also used to conduct HT DFT in the search for strengthening precipitates that could be applied to a variety of alloy matrices including Fe, Al, Mg, Ni, Co, and Ti.\[^{46}\] OQMD was also used as a source for large training datasets to train an ML model in a recent study for identifying quaternary Heusler materials.\[^{47}\] Besides OQMD, there are several efforts to generate large-scale HT DFT databases, such as the Materials Project,\[^{48}\] AFLOWLIB,\[^{49}\] Computational Materials Repository,\[^{50}\] and NOMAD.\[^{51}\] Recent advances of HT DFT database developments and their applications in data-driven methods can be found in several review articles.\[^{14,26,52}\]

Figure 2. Quantum mechanical calculation (e.g., DFT) and MD simulations combined with data-driven approaches for AM materials design. Blue: using quantum mechanical calculations directly to guide materials design. Green: materials design through HT DFT. Purple: using MD simulations to directly guide materials design. Red: using machine-learned potential energies in MD simulations to guide materials design. The yellow shapes represent the output/input which characterize the boxes.
Developing these HT DFT infrastructures will enable expeditious screening of materials in general, whereas currently, to our best knowledge, their application in AM materials research is scarce. Therefore, the AM community can strongly benefit by leveraging the advances in HT-DFT for the purpose of materials discovery and design research, shown as the green flow in Figure 2.

2.2. MD Simulations and ML Potentials

In terms of AM, MD has demonstrated its strength in directly tailoring the design of the materials’ nanostructure and guiding the optimization of the manufacturing process, shown as the purple flow in Figure 2. In addition, MD simulation is known to be one of the very few methods available to provide information of the nature of phase transformation of high-entropy alloys (HEAs), which is emerging as an important material for AM, given that the transformation process is very fast.\(^{[53]}\)

Computational quantum mechanical methods are generally accurate at the molecular scale but become extremely computationally expensive at much larger spatiotemporal scales.\(^{[20]}\) These costs limit computational quantum mechanical methods to exploring phenomena that occur on the order of tens of picoseconds and a few thousand atoms. In contrast, MD methods provide an alternative for overcoming these limitations. Using the Born–Oppenheimer approximation, the quantum mechanics laws governing the interactions between atoms can be approximated with the laws of Newtonian mechanics. Due to this approximation, electron-related processes cannot be simulated, but the complexity of the calculation is significantly reduced.

The reduced complexity enables simulations of atomistic systems up to micrometers in length scale and microseconds in time scale. Furthermore, through dimensionality reduction (while preserving the conformational details, thermodynamic properties, and particle dynamics), coarse-grained MD (CGMD) methods are able to simulate systems of even larger spatiotemporal scales. Thus, MD methods are more suited for designing larger material systems.

Using MD simulations, the thermodynamic properties (size dependency of the melting point, heat of fusion, entropy of fusion, and surface energy) of Ag nanoclusters during the heating and cooling process were quantified to inform the thermal and energy management for AM.\(^{[54]}\) In another MD study of core/shell nanoparticles, the effects of the core volume fraction on the interfacial melting of the core metal and the gradual alloying of the molten shell were investigated.\(^{[55]}\) This study revealed principles of designing powder materials and implied new ways of reducing the parts of manufacturing time. Using CGMD methods such as dissipative particle dynamics (DPD) and its extended version, the many-body dissipative particle dynamics (MDPD) enables modeling and simulating systems at the meso-scale. For instance, the agglomeration of ultraviolet ink fluid and the effects of surfactants in 3D nanoinkjet printing were examined with DPD and MDPD.\(^{[56,57]}\) Similar methods were also used to study nanodroplet formation in 3D nanoinkjet printing.\(^{[58]}\) CGMD was also used to simulate the 3D polymer printing process known as continuous liquid interface production (CLIP) for identifying the factors influencing the fidelity during the manufacturing process.\(^{[59]}\) This study highlighted the effects of polymer crosslink density, thereby guiding better control over print quality.\(^{[59]}\) Such MD and CGMD simulations are important sources of raw data for constructing databases for data mining to uncover hidden properties and correlations that convention observation of such data may not reveal. For example, The Inorganic Crystal Structure Database (ICSD),\(^{[60]}\) arguably the world’s largest database of fully evaluated and published crystal structure data, was extended in 2017 and started accepting theoretical structure data, including those generated by MD simulations.\(^{[61]}\)

Taking a leaf out of these efforts, we propose that materials design and selection in AM will also benefit significantly from systematic storage, curation, and dissemination of relevant experimental and computational data. Such databases will become excellent platforms for data mining applications for materials discovery and design in AM.

Other than data mining for hidden composition–structure–function correlations from databases with MD data, data-driven approaches such as ML can also change the scheme of MD-aided materials discovery and design by introducing ML potentials. MD is particularly potent for simulating systems of relatively higher spatiotemporal scales, but it has lower accuracy compared with quantum mechanical methods. This is because the central component of MD, the potential energy surfaces (PES), is usually fitted and applied empirically, which introduces approximation errors.\(^{[20]}\) To alleviate this drawback, there is a growing interest in using ML methods to derive MD atomic potentials from large quantum mechanical datasets. This is done in two steps: first, transform the reference atomic configurations into structural descriptors and second, use ML algorithms to learn the mapping between the descriptors and the corresponding potential energy.\(^{[62]}\) This process is shown as the red flow in Figure 2. Once trained, the ML potentials provide comparable accuracy with quantum mechanical calculations, while accelerating the simulations by orders of magnitude.\(^{[63]}\) Artificial neural networks (ANNs) are extensively used to achieve this.\(^{[64,65]}\) Specifically, ANNs are increasingly used to develop interatomic potentials for a large variety of alloys.\(^{[66–69]}\) Numerous frameworks are available to learn the PES from quantum mechanical data using ANN, including SIMPLE-NN,\(^{[70]}\) SchNet,\(^{[71,72]}\) DeePMD,\(^{[73]}\) and TensorAlloy.\(^{[69]}\) Another class of ML method used for this purpose is Gaussian processes (GP), a nonparametric method with no fixed functional form. Thus, GPs are suitable for modeling highly complex potential energy landscapes. A representation of this method is the Gaussian approximation potentials (GAPs).\(^{[74–76]}\) GAP was also used to derive the atomic potential from quantum mechanical datasets to model and simulate metallic materials.\(^{[77,78]}\) In general, ANNs require greater efforts in model training compared with GP, whereas making predictions using a GP model is slower than in ANNs.\(^{[70]}\) Nevertheless, these ML potentials allow researchers to conduct atomistic simulations for materials design and discovery with higher physical accuracy and computational efficiency. The materials research community has already started exploiting the application of ML potentials in a wide range of materials research, but specifically for purposes of AM, we find that such applications are still not well documented in current literature. Hence, an intensified focus on applying ML potentials can be anticipated to firmly bolster materials research in AM.
2.3. Computational Thermodynamics

The cyclic thermal history, intrinsic to the additive process, entails a number of liquid-to-solid and solid-state phase transformations affecting the microstructure and properties of alloys. Therefore, the material design process for AM must consider a thorough understanding of the composition-dependent thermo-chemical and physical events associated with the far-from-equilibrium processing.

The calculation of phase diagram (CALPHAD) approach is widely used to predict thermodynamic and kinetic properties of new powder compositions. The CALPHAD approach is a unique tool that describes a material’s thermodynamic properties using empirical and/or theoretical data of its constituents at specific temperature ranges (room temperature—above melting temperatures). By fitting the data to the correct thermodynamic model, phase diagrams can be extracted to describe the thermo-chemical behavior of the material as a function of temperature and composition, various thermo-chemical (i.e., heat of mixing, chemical activity, vapor pressure), and thermo-physical properties (i.e., coefficient of thermal expansion, bulk modulus) based on the thermodynamic equilibrium states and chemical potentials (i.e., regular solution behavior). These properties at equilibrium can then be used to analyze closer-to-process conditions such as nonequilibrium and rapid cooling rates that are encountered during AM (Figure 3a). Readily available software packages, including ThermoCalc and FactSage, use the CALPHAD approach to simulate and predict the dynamic behavior of material.

For example, HEA is a large category of AM-specific materials, showing outstanding mechanical performance with intricate deformation mechanisms (among others), but has limited applications due to poor machinability. The uniqueness of AM makes manufacturing of HEAs feasible with reduced machining and postprocessing. This recent introduction of AM-enabled HEA applications and the lack of extensive prior studies in HEA drives attention to the use of data-driven thermodynamic tools like high-throughput CALPHAD to study the composition—structure mapping. While the data of a large amount of ternary systems required for reliable prediction in exploring design space of HEAs makes the trial-and-error approach infeasible, HT-DFT has been shown to be a viable alternative to provide enthalpy of mixing data for solid solutions and enthalpy of formation data for intermetallic phases. One particular example used high-throughput CALPHAD modeling to conduct thermodynamic data-driven design of Al-Co-Cr-Fe-Ni multiprincipal element alloy. This study successfully predicted the formation phases and microstructures of high-performing designs from 10 000 different compositions with 1% step size in the predetermined alloying element atomic percent ranges. This study illustrates the capability of CALPHAD modeling in comprehending the solidification path of complex HEAs, as well as predicting mechanical performances qualitatively from phase volume fractions.

However, accurate HEA phase diagram prediction is based on having complete data for full binary and ternary subsystems of all elements, but even the most comprehensive databases include only a limited number (about 5%) of fully assessed ternary systems. Close attention by the users is required to the quality of data, as well as its applicability to the described thermodynamic problem. For instance, many of the now-matured thermodynamic databases were established, targeting the study of conventional single-principal alloys. The aforementioned bias on single principal element alloys in many thermodynamic databases makes them sometimes inadequate for interpolation of properties for alloys with multiple principal elements, such as HEAs. The increased adoption of CALPHAD in AM-specific science.
materials design correspondingly necessitates intelligent improvement of its infrastructure and expansion of modeling to more properties of interests.[86–87] The capabilities of CALPHAD are clearly summarized in a separate review.[86] While highlighting current developments in CALPHAD, that review also points out the lack of strategies for implementing new models and data to efficiently update the databases. Furthermore, CALPHAD modeling can greatly benefit from a more sustainable data ecosystem, where individual data repositories have mechanisms to communicate and form a joint database when needed.[87]

In addition to aiding the design of AM-specific alloys, the use of CALPHAD modeling has also been extended to improve the manufacturing of powder alloy feedstock used in most AM processes. A method to quantify the uncertainty between prealloyed and target powder composition is developed using a CALPHAD-based integrated computational materials engineering (ICME) framework.[88] Uncertainty in AM part material composition, including feedstock variation between batches, localized compositional deviation from nominal feedstock, and contamination from powder recycling, negatively affects the success rate of the overall AM processes. Therefore, it is important to capture this uncertainty in manufacturing during AM materials design.[88] This study identified an optimized average composition of the high-strength low-alloy (HSLA) powder used in AM, increasing the success rate of AM builds by 44.7%.[88]

Other computational thermodynamics methods are also being integrated with CALPHAD to understand complex solidification events and phase formation in AM. An alternative diffusion module, DICTRA, enables a closer-to-practice prediction of solidification parameters by accounting for diffusion in the liquid and solid phases.[79,80,89] Furthermore, the thermodynamics and kinetics of the solidification process of the alloy can be predicted by extracting simulated melt pool conditions in laser PBF (L-PBF) via finite element analysis (FEA) (Figure 3b). These thermodynamic and kinetic computations effectively described the solidification events of commercial alloys used in AM, such as Inconel 625 and Ti–6Al–4V, by closely predicting those found experimentally.[90–93] In the case of Inconel 625, a correct cooling rate was predicted to result in the precipitation of carbides and Laves phases in the dendritically segregated material, aiding the identification of postprocessing conditions that will limit the growth of precipitates, which affect the strength of the alloy.[90] In the case of Ti–6Al–4V, the simulated diffusion-controlled phase transformations were tracked to predict the final α-phase composition after AM processing.[92] Nevertheless, CALPHAD modeling, DICTRA, and FEA simulations are excellent tools for populating the composition–structure map of AM-specific materials, but their computational cost can still be a burden when the design space is too large to exhaustively search for optimal solutions.

The power of ML algorithms lies in their capability to make informed predictions reflecting the processing–microstructure–property relationship when exhaustive search of the design space is infeasible (Figure 3d). For instance, ML may be a solution to the costly combinatorial experiments and/or computationally expensive simulations required to screen the vast space of potential metallic alloys.[23] Without explicitly modeling each composition, an adaptive search algorithm like GA could suggest alternatives that outperform known compositions based on the similarity hypothesis.[21] By evaluating the candidate’s composition through a fitness function that measures a composition-based property criteria, the algorithm can iteratively select the best-performing compositions based on the input dataset.[21,23,94] The qualifications of the alloy compositions are evaluated with a model describing the composition and process-dependent properties, and the success of the adaptive learning algorithm is based on this accelerated qualification.[94] In the case of AM alloys, these models can include computational models describing AM-specific phenomena as described throughout this section (DFT modeling, MD simulations, thermo-chemical, and thermo-physical calculations).[21,94]

In addition to GAs, metamodeling as an ML approach can swiftly screen metallic compositions based on datasets generated solely via computational modeling.[95] An example of metamodeling to design alloys with enhanced properties was demonstrated by the development of an ANN model that predicted the sigma phase fraction in a Ni superalloy.[95] The sigma phase is known to precipitate in the Ni superalloy at high operating temperatures and this phase is detrimental to the component’s structural integrity. The model consists of a complex alloy system (mole fractions of Ni, Al, Co, Cr, Mo, Re, Ta, Ti, W) and temperature as input parameters and output the target property of the sigma phase fraction. The speed of this ANN model is demonstrated by taking merely 1.7 s to calculate the target property of a 40 000-point dataset on a personal computer. Being a neural network model, this model has the advantage of producing multiple outputs simultaneously as well as the capability to be scaled up for larger design spaces.[95]

These early alloy design studies demonstrate the potential of screening a vast design space of virtual alloys with targeted composition and process-dependent properties without the need for experimentation. It is foreseen that integrating the complex AM-related thermo–physical and thermo–chemical phenomena in a single data repository will enhance the accuracy and applicability of data-driven methods, aiding the design of alloys suitable for the AM process.

3. Composite Material Design for AM

Modern structural applications strongly demand materials with high strength and excellent toughness. However, the mutual exclusivity of these properties requires undesirable compromises when designing structural components with a single homogeneous material. Composite materials allow material scientists to pick multiple sets of mechanical properties, such as the modulus, strength, ductility, and toughness, to achieve a target property that normally does not exist in any single material. Major approaches in designing composites include taking inspiration from nature or simple stacking of disparate materials.[96–98] By harnessing AM and, more specifically, multimaterial printing, material scientists have greater freedom in determining the spatial arrangement of the constituents, rather than being limited to arranging the constituents of composite material in simple laminate structures or particle-/fiber-reinforcing matrix structures. A brief overview of current progress and challenges on AM of multimaterial structures can be found in another review article.[99]
With increased design freedom, the vast design space of choosing material combinations and how they are arranged in a structure can be difficult to explore without sufficient data that usually require expensive experimentation or extensive computational simulations. This difficulty can be addressed using data-driven methods to accelerate the discovery of material—arrangement—property mapping using relatively limited information and expand the learned mapping to guide design decisions in this vast design space. As the material selection (Figure 4a) and materials' arrangement (Figure 4b,c) have intertwined effects on the properties of the composite material design (Figure 4d), this section of our Review will focus on data-driven discovery of 1) different constituent materials and 2) different constituent arrangements and their effects on composite properties. These two areas illustrate the utility of data-driven methods in all aspects of composite design for AM.

3.1. Material Selection for Composites

The emergence of easily accessible data-driven methods reduces the need for prohibitively large amounts of data in exploring the broad design space offered by AM. For example, ML was implemented to design an alumina-reinforced aluminum matrix composite with improved tribomechanical properties.[100] First, an ANN model was trained with data extracted from the literature to learn the mapping between composite properties and the composition, size, and morphology of reinforcing particles. Afterward, GA was used to conduct multiattribute optimization to design composites with desired properties. In addition, a two-step process of feature extraction and ensemble-based regression captured the elastic localization relationships in high-contrast composites (i.e., composites with large ratios between the elastic moduli of the constituents).[101] The high-contrast composite was represented in 3D voxels and the damage-inducing localized strain was captured in ML-aided prediction of the elastic response of a single voxel of interest within the whole microstructure. The relationship between this focal voxel and its neighbors was represented numerically and used as features in this study. Selected high-impact features were regressed with random forests to achieve a good balance between prediction accuracy and time consumption. This study showed the fundamental viability of using data-driven tools to make computationally efficient predictions in composition—property linkage and presented strategies to improve prediction accuracy. Similarly, a 3D convolutional neural network (CNN) was used as a data-driven tool to establish connections between the salient features of 3D microstructures and the effective material properties of interest.[102] Following this study, property prediction on high-contrast composites was also achieved using a deep learning architecture adopted from traditional CNN models.[103] Homogenization linkages were established between the macro-scale effective stiffness and the extracted higher-order neighborhood information from the 3D microstructures. This homogenization linkage allowed the arrangement—property

Figure 4. Compared with conventional methods in manufacturing, AM excels in realizing advanced composite designs due to the expanded design space from: a) an expanded selection of materials for composite constituents; b) achieving conventionally infeasible constituent arrangements; and c) introduction of structural hierarchy across length scales. Data-driven methods accelerate exploration of this vast design space through d) efficient discovery of the mapping between composite design and performance.
linkage to be directly trained on the input 3D arrangement of constituent materials, thus eliminating the need for feature engineering, and improving learning capability. The overall approach in this study was tested with generalized constituents having elastic modulus ratio of 50 to obtain high-contrast composites and therefore can be extended to study constituent pairs with other elastic modulus ratios. Similar studies on the generalized purely elastic constituent identified optimal designs for material properties such as the ratio of modulus, toughness, and volume ratio. Compared with elastic behavior of materials, the prediction of material properties with plasticity requires more complex computations due to its nonlinearity and could benefit more from data-driven methods. These benefits are demonstrated in a pioneer study, in which the data-driven framework is used to predict the constitutive model of the 2D hyperelastic composite with homogenization procedures and optimize the design of 3D inelastic composite material for toughness without costly numerical simulations. Data-driven frameworks with further development are shown to be capable of predicting stress–strain behavior beyond the elastic limit of the material with reasonable computational cost. A similar approach of predicting load-displacement curves was used for cotton fiber/PVC composites, confirming the applicability of this technique for any kind of composite systems. Other than mechanical properties, data-driven tools could predict thermal conductivities of composite materials with better accuracy than traditional analytical models.

While several of the aforementioned data-driven methods are based on data acquired from physics-based simulations, simplifications made in these simulations for reducing computational costs cannot ignore the important problems that are specific to multimaterial printing, such as the interfacial strength between constituents. A proper strategy is needed for bonding different constituents in a multimaterial print, especially functional and structural parts. Ignoring the interfaces between constituents will not only cause a potential performance decay in composite AM materials when compared with that of their individual constituents, but there may also be a potential deterioration over the material’s lifetime due to incompatible thermal expansion and cooling rates. Proper characterization of interfacial bonding strengths is essential for successfully selecting constituents for AM composite materials.

Experimental validations and computational simulations can provide greater value if these data can be stored in commonly accessible databases for different future studies, while avoiding duplicated research efforts in the field. While the development of such databases is still in its infancy, it is important to emphasize the standardization needed to increase the broader acceptance of any composite material databases. By agreeing upon common materials testing and qualification standards for each application, a group of small aircraft companies greatly reduced the cost of material certification, while simultaneously decreasing the development time of a new material with needed properties. The reduced effort for certification is built on the basis of material equivalence, that is, all users manufacturing the same material under the exact conditions recorded in the database should produce the same material property. AM composites are yet to achieve such a high level of consistency in the final product, but the benefit of a widely accepted, standardized database is clearly exemplified by the aerospace industry and is expected to be seen in future AM composite databases. While developing a standardized database will strongly supplement AM product certification, the open-access nature of a common database increases the availability of data to all stakeholders and racks up sizable savings when comparing with the cost of establishing localized databases separately.

3.2. Optimization of Constituent Arrangement in Composites

The bottom-up nature of AM processes inherently confers the freedom to arrange constituents in ways that are different from the simple stacking of layers or particle-/fiber-reinforced matrix arrangements. Taking advantage of this geometrical freedom of AM is integral to achieving further breakthroughs in composite design. Algorithm-driven optimizations of the constituent arrangements can help to unlock the full potential of geometrical freedom in AM. A pioneering study in this area iteratively arrived at near-optimized arrangements of constituents in a composite system from randomly generated starting points in a very broad design space. The choice of the starting point could potentially affect the optimized solution undesirably as an inadequate starting point that might trap the optimization at a local extremum. This potential inability to explore the whole design space was addressed using data-driven methods that identified the optimal arrangements of the constituents for a composite system to achieve high toughness and strength. Although the ML model used in this study was trained with data from less than 1% of the whole design space, it achieved high accuracy (around 90%) in predicting the resulting composite properties from geometrical arrangements of constituents. Moreover, the data used for ML model training were binary representations of the composite design properties (i.e., either good or bad). Without any information on the overall ranking, the ML model was able to accurately predict rankings of all data after training and therefore produce the best-performing arrangements. With excellent tolerance to lower order data, data-driven methods are particularly useful for extracting meaningful guidance in composite design when higher-order data are unavailable or the uncertainty in experiments compromises the quality of the data.

Another key aspect of optimizing constituent arrangement in composite material is the structural hierarchy spanning different length scales. Structural hierarchy in natural composite materials positively influenced the macro-scale property of the composite. This hierarchical complexity is difficult to mimic synthetically due to the limitations of conventional manufacturing methods.

Conventional methods are not capable of 3D spatial manipulation of the material’s internal structures. Among established methods, casting molten materials into ingots or bulk components allows addition of reinforcing material into the composite, but the reinforcement distribution has to be uniform. In addition to that, sheet lamination method introduces additional freedom to fine-tune the intra- and interlayer parameters, but the design flexibility is still limited and insufficient for production of hierarchical materials. In contrast, the geometrical freedom of AM enables the manufacturing of such hierarchy in composite material designs. Guiding the design of hierarchical composites
usually requires a mapping between the constituent arrangement and the resulting composite property. Obtaining this mapping by solving a physics-based model typically requires extensive computational power, but data-driven tools can help obtain this mapping efficiently. This capability was demonstrated in an experimentally validated identification of optimally arranged constituents of a hierarchical polymeric composite material. A CNN model was trained to learn the mapping between the hierarchical arrangements of the constituents and the toughness of the resulting composite design, thereby predicting the toughness of composite designs with arbitrary constituent arrangements. Furthermore, three unit cells were designed with the same pair of polymeric constituent materials that had different mechanical properties to obtain varying overall in-plane stiffness properties. By optimizing the arrangement of these three unit cells in a lattice-like composite system, the CNN model reduced the order of the design problem and discovered an optimal solution for a composite system with multiple levels of hierarchy. ML-aided optimization of repetitive lattice composite materials can also automatically generate structures of microscale representative volume elements for desired macroscale elastic moduli. While data-driven methods significantly advanced the AM of structurally complex polymer composites, this success is not replicated in metal composites due to the challenge of preventing the constituent materials from reacting with each other under high temperatures. Nonetheless, the success of AI-aided design in AM-enabled polymer composites can still inspire the design of AM-enabled metal composite, as shown in the optimization of the constituents’ staggered arrangements with generalized material properties that are applicable to both polymers and metals. Recent developments in particle-reinforced and functionally graded metallic composites are also very promising for improving AM metallic composite materials. Data-driven methods can aid in improving the mechanical properties of metallic composites once AM is further enhanced to solve the problem of chemical reactions between the matrix and the reinforcements. While AM-specific composite material is still an emerging field, the digital nature of AM calls for strong data infrastructures for more users to take advantage of the evolving data-driven tools. Composite material databases, due to composites’ heterogeneity and added complexity, have a shifted focus to the relationship between microscale constituent properties and macroscale composite response. Using fiber-reinforced composite material as an example, the added complexity compared with homogeneous material includes volume fraction, orientation, and spatial distribution of reinforcing fiber, matrix material property, interfacial bonding characteristic, etc. This added complexity means greater variability in the data, posing challenges on relevance and acceptance of data stored in databases. To increase the relevance of data, the completeness of the data must be emphasized. When variability in composite material property is not coupled with the data stored in database, users of such databases (i.e., material scientist and designers) have to make much more conservative design allowable based on the extracted data. It is shown that complete data on mechanical properties of a notched carbon fiber—laminate composite with statistical variance that is not included in typically reported average values for properties in handbook-based database allow engineers to use statistical tool and obtain better prediction of reliability of aerospace composite structures, proving the improved relevance of complete data. Furthermore, AM as a digital manufacturing method is also beneficial to streamlining data exchange between users and database. Despite the additional spatial freedom in composite material design brought by AM, once a complete and effective categorization scheme is put in place, it can be expected that a centralized database storing valuable experimental and simulation data can be established to greatly facilitate the use of data-driven methods in designing AM composites.

4. Structure Design for AM

4.1. Topology Optimization and Generative Design

Topology Optimization (TO) and Generative Design (GD) are two potential design tools for developing novel architectures with desirable structural and functional properties. TO mathematically optimizes the material layout in a given design domain based on a predefined set of loads and boundary conditions (mechanical, thermal, or both). GD is an iterative design process that generates multiple outputs which meet the predefined constraints. Although both methods appear to be similar and the nomenclatures are often used interchangeably, there are subtle differences between the two methods. TO is a foundational technology that aims to reduce the material of an existing object, while retaining its functional ability, whereas GD is a newer concept that builds on TO to create “imaginary designs” that satisfy the given boundary conditions. To put it simply, one could say TO is subtractive and GD is additive in nature. In both methods, the optimized outputs are often complex structures that are difficult to manufacture using subtractive methods. The geometric freedom offered by AM increased the use of these iterative tools in developing novel structures with superior performance. Based on mechanical, thermal, and fluidic boundary conditions, these tools were used in designing novel structures that were manufactured subsequently.

The next step in improving and expanding the output of these design tools is implementing data-driven approaches in creating outputs. Conventionally, iterative design tools such as TO and GD generate outputs after a series of FEA of the structures, which is very demanding computationally. The goal of using data-driven approaches in combination with these tools is to potentially replace the demanding FEA aspects that come with TO and GD. The most popular method of doing so is to apply CNN to existing designs and determine an optimized structure for the given boundary conditions. This process involves several steps. The first step is to create a dataset of previously designed structures. The program can evaluate the performance of new structures after this dataset trains the program on the performance of each existing design. The second step is to generate optimized structures based on benchmark designs, together with TO. Subsequently, a similarity criterion is used to filter similar designs from the database. This is an iterative process where the goal is to get the number of similar designs below a certain threshold. If this threshold is not achieved, the CNN is used to generate new structures based on the patterns learnt from the benchmark designs and TO-generated structures. Step 2
and 3 is repeated iteratively to filter designs until the desired threshold is achieved. Once this iterative process is complete, the next step involves evaluating the design options generated using a built-in loss function that is developed using the training dataset. The final step is a visual display of all the design options for the designer to choose from that fits the given application. Figure 5 shows the framework for one such model.[133]

Well-designed data-driven models that aid structural design optimizations can possess good predictability with high accuracies and effectively reduce the computational costs of traditional optimization techniques. For example, one such CNN model generated topologically optimized 2D structures for a given set of boundary conditions that were nearly identical to the geometries produced by FEA-based models.[134,135] Another study used a CNN-based model to generate optimized structures for automobile wheels.[133] Furthermore, data-driven TO models could also generate multiclass microstructures (i.e., structures that have more than one type of unit cell in their microstructure) which have important applications in the biomedical and aerospace industry as they can be designed to have specific properties that cannot be found naturally.[136]

Integration of data-driven approaches with design tools like TO and GD allows designers to efficiently generate novel complex structures with minimal computational cost. A crucial aspect is to construct a thorough reference dataset that can generate output designs similar to those generated by FE-based models. Unlike material databases that contain data universal to almost all users, data types like geometry or structure designs make systematic classification tremendously difficult due to the case-by-case nature of different design criteria and therefore needs to be generalized. A specific geometry design may not be useful for any other user of the database, whereas a generalized data entry can be adapted for different design applications. This generalization usually entails feature extraction from a geometry, rather than the whole geometry or structure design. One example of such a generalized database is demonstrated to be inspirational and relevant to the Design for AM (DFAM) process.[137] While the design features stored in the aforementioned database are categorized according to AM’s benefits, including functional improvement, and part consolidation requirement for assemblies, all features are generalized to be applicable to different products manufactured by various processes. To maximize the potential of AM, completely new redesigns are often required rather than incrementally improving from designs made for subtractive manufacturing methods. However, a completely new redesign usually requires designers to find optimal solutions in a much broader design space. Data-driven methods like TO and GD that are coupled with illustrative features stored in generalized databases for structure design will greatly accelerate the exploration, promoting more innovative structure designs for AM over simply improving existing designs.

4.2. Design with AM Constraints

AM-specific manufacturing constraints cannot be overlooked. Designing with AM constraints in mind is a major component of the DFAM revolution and it is essential to take advantage of the possibilities offered by AM. For detailed review on all aspects of DFAM, we direct the reader’s attention to this thorough Review.[138] Here, we focus on managing supports, overhangs, and minimum feature size in structure design.

Support structures are necessary in AM for long overhang features at a shallow angle. They not only increase material waste, but also increase postprocessing time for support removal, especially for metallic AM products. To remedy the impact of reduced manufacturability on complex AM structure designs, an iterative process is proposed to slightly modify the theoretically optimal topology and avoid introducing support structures.[139] The advantage of this method is its universal applicability as it is a digital postprocess after obtaining the optimal structure design from data-driven frameworks. This iterative process also works for all build orientations of any 2D shape and therefore promotes further optimization between structure design and proposed build orientation.[139] TO frameworks with integrated overhang constraints can also successfully create self-supporting structures that are validated experimentally in both polymer and metallic...
AM products. Data-driven methods can be incorporated into the constraint function to achieve more complex optimization output when compared with TO frameworks using fixed overhang angles and feature size constraints. With limited experimental data, the dynamic data-driven design rule can be created for determining the minimum feature sizes using overhang angles. This method allows different minimum features for varying overhang angles, as well as incorporating horizontal “bridging” features that connect two bodies in close proximity to each other.

Apart from topologically optimized structures specifically designed for AM, data-driven methods can also facilitate improvement of existing component designs to harness the benefits enabled by AM. One example of such improvement is using lattice structures. When using lattice unit cells from a predefined unit-cell library, the large-scale multivariable problem to recreate a completely new optimized structure is reduced to only two design variables: unit-cell type and size. This is vital for decreasing the computational costs in data-driven structure optimizations and it is only possible with an existing unit-cell library. This highlights the need for a more extensive database of unit-cell designs and properties. When utilizing lattice structures in AM structure design, an important constraint is that different AM methods have varying print resolutions. Thus, the minimum feature size constraint must be incorporated into the lattice structure optimization for AM to ensure manufacturability. In addition, when process imperfection causes optimal lattice design infeasible to manufacture, data-driven methods such as Bayesian ML is shown to be a robust design framework to adapt to added manufacturing constraints and provide valuable insights in lattice structures that are commonly used in AM metamaterial design.

5. Tool Path and Process Parameter Design for AM

The process of introducing, and subsequently removing support materials, lowers the efficiency of AM in industrial applications. In addition, in fusion-based AM, the materials undergo extremely rapid heating and cooling which lead to nonequilibrium phase transformation during deposition (e.g., DED, WAAM) or joining/sintering (e.g., PBF). As a result, several challenges arise, including thermally induced defects and material anisotropy in direction of the thermal gradient. Figure 6 shows the current challenges in AM, namely, support reduction, voids elimination, reducing material anisotropy, as well as reducing thermal nonuniformity and its resulting problems, such as residual stress development. Some of these challenges can be minimized by optimizing the tool path with either computational simulation or trial-and-error experimentation, but these methods are resource intensive and data-driven tools are promising to reduce this cost.

The workflow of AM consists of three major steps: preprocessing, processing, and postprocessing. Data-driven frameworks can establish the correlation between user input process parameters and product quality without solving multiphysics models that are computationally demanding. They can also guide users in adjusting the corresponding process parameters in the preprocessing stage, such as laser power and tool path. The following sections detail recent developments in using data-driven tools to aid active alteration of AM deposition strategies in the preprocessing step at the very beginning of the AM process hierarchy before irreversible fabrication takes place. In both polymer and metal AM processes, these active measures will improve the quality of the prints by achieving full density, increasing dimensional accuracy, reducing support material, and reducing material anisotropy.

5.1. Tool Path Planning to Achieve Full Density

To achieve fully dense structures without unintentional voids, several process parameters should be optimized in tool path planning. These parameters include horizontal spacing between line scans (hatch spacing), layer height, and tool path adjustments to prevent material void or pile-up at corners and intersections of line scans within a layer. A previous study showed that in FDM, the relation between the occurrence of voids induced by insufficient material and hatch spacing in each layer could be modeled with deep neural networks without requiring computationally heavy optimization of a multiattribute problem. If the geometry of the cross section and the uniformity of the deposited line is known, the tool path can be optimized to eliminate voids between lines of deposited material. This knowledge depends on understanding the dynamics of the deposition process, as well as the interactions between the deposited materials and the print surfaces or prior layers. Data-driven models can efficiently obtain such information. In an example of inkjet printing, such a model serves as a powerful predictor of the ejected droplets’ dynamics after deposition and its interactions with substrate, while requiring minimal human interventions. In this study, unsupervised deep learning was used to examine droplet evolution.
from video data of the droplet ejection. Video data, as opposed to static image data, could be used due to the lower computational resource requirement and unsupervised nature of the method, which is a strong advantage as image data does not include temporal domain information. This enhanced prediction of droplet dynamics is vital for estimating the line track geometry as well as planning for adequate hatch spacing. The benefits of requiring minimal human input were also examined in a recent study using ANN in an adaptive void-filling strategy for arbitrary-junction geometries in metal AM of rib-web structures. An optimal training set of merely 46 different junctions joining three or four edges with various angles in between was created to train the neural network.

In addition to designing the tool path trajectory, the print parameters along the tool path must be optimized to avoid defects such as keyholes or lack of fusion. High-fidelity models in different metal AM processes improved the fundamental understanding of porosity defects, temperature distribution in melt pools, and evolution across layers. However, it is infeasible to use such models to study a full component at the same resolution due to the significant computational effort involved. To facilitate more efficient identifications of print parameters, a GP-based surrogate model was constructed and trained with only 139 datasets to predict melt pool properties, resulting in low standard deviation from limited print parameters as inputs. The experimental cost of finding optimal print parameters along a tool path can be further reduced by adequately combining data mining from simple experiments or literature and using ML algorithms.

5.2. Tool Path Planning to Reduce Support Structure

Greater ease of producing geometrically complex components is a major advantage that sets AM apart from traditional manufacturing methods. However, overhang structures are sometimes inevitable in complex components. While most AM processes require adding support structures to produce overhang structures, the increase in build time and material waste as well as difficulties in removing the support material are significant problems that need to be addressed. This problem can be remedied considerably by changing the orientation of the build-platform and adjusting print parameters. Printing of free-form, thin shell parts without using support structures was achieved recently using a dynamically reorienting build-platform. An ANN-based context-dependent compensation scheme was used to apply corrections to complex tool paths. Another strategy is changing print parameters adaptively, which is more applicable especially when large-scale parts limit the capability to dynamically reorient the build-platform. With less than 50 irregularly spaced data points from characterization tests, a biharmonic spline interpolation model was used to adaptively alter print parameters along part of the tool path and achieved support-free, near-net-shape product on flat build-platforms using metal wire-arc AM (WAAM).

5.3. Tool Path Planning to Reduce Print Time and Improve Dimensional Accuracy

For designing tool path trajectories that reduce build time while achieving high-dimensional accuracy, the orientations, slicing strategies, and volume uncertainties should be considered. For example, the optimal or near-optimal build orientation, deposition direction, and slicing solution for a flat build platform were determined effectively using GA. The GA-based optimization was conducted over a vast range of orientations, while considering factors such as surface accuracy, number of turns, and start–stop positions in the tool path. Alternatively, the optimal tool path to reduce build time for a part with a predefined orientation was obtained from multiattribute optimization with evolutionary algorithms. However, having the optimal tool path for reduced support material and build time does not guarantee a perfect build, due to inevitable uncertainties in AM processes. While such uncertainties cannot be effectively reduced yet, uncertainties in material deposition and solidification were predicted using CNNs when specific process parameters and tool paths were provided.

5.4. Tool Path Planning to Address Anisotropy, Residual Stress, and Other Challenges

In addition to all factors presented earlier, the crystalline nature of most materials used for metal AM, coupled with directional heat transfer during processing, may introduce thermally induced anisotropy in printed parts. The impact of print parameters on the anisotropy of the parts can be modeled and should be considered for tool path planning in AM. Cyclic temperature profile with high peak temperatures also causes thermally induced residual stress in parts produced with AM. 

An unsupervised ML model established the strategy for identifying areas that require adaptations in tool path for peak temperature reduction. This model-clustered parts of the tool path according to their thermal history and thus this model are applicable for comparing process regimes within single components and across different components. While the tool path dependency of AM can be minimized by data-driven methods, the differences in material properties of the final product could still exist. The inevitable process uncertainty is best minimized when process-dependent property changes, such as distortional hardening in composite materials, can be predicted with data-driven methods. However, the increasing number of factors to be considered for every point in the tool path as well as the increasing number of points in the tool path to model large-scale parts are dramatically increasing the complexity of tool path planning for metal AM. By reducing the order of the model as well as establishing relationships between print parameters and print outcomes when numerical methods do not exist, ML helps users make decisions in tool path planning more accurately and efficiently.

5.5. Tool Path Design Database Development to Increase AM Accessibility

Tool path design is a complex subject with multiple parameters that are intertwined with each other, affecting the final quality of AM prints. Its complex nature and high cost of trial and errors
demand data-driven methods for efficient optimization and decision-making. Establishing a database for tool path designs is fundamental for incorporating data-driven methods and it avoids unnecessary duplicated experimental validations for the same AM methods. Different manufacturing processes have different pros and cons, limiting their capabilities to realize some designs. For example, the resolution of a typical DED system is worse than a typical PBF system, limiting its application in producing parts that have very intricate features. In contrast, the absence of a powder bed means increased versatility and scalability for DED systems, making them better candidates for building large-scale components and repairing existing components when compared with PBF systems. Certain process parameters are only applicable to a certain type of process, like powder feed rate in DED system, whereas the operating windows for certain parameters are dramatically different across processes, such as the laser/electron beam scan speed in DED or PBF systems. These variabilities described earlier necessitate a classification or organization scheme for the database that is applicable to all different processes. One study presents a viable solution to address this need, using NASA-developed materials informatics system Materials and Processing Technical Information System (MAPTIS) as an example. This system allows the user to filter data using material type, AM processes, build parameters, and so on. The connectivity of data is achieved by their classification tags. For example, when the user aims to find optimal parameters for printing Inconel 718, the build parameters for DED systems will differ from those of PBF systems, and the database therefore requires classification tags to guarantee a targeted response to the users’ query.

As long as the important criteria for tool path design databases discussed earlier are met, the opportunities of such databases are invaluable to the growing community of AM users. Despite the benefits, a mature database accessible to most companies and researchers is yet to be established, hindered by International Treaty of Arms Regulations (ITAR) restrictions and concerns about proprietary data. We still need continued effort to encourage incentives toward the establishment of a more accessible AM structure and tool path design database, that not only facilitates the increased use of data-driven methods but also accelerates the development of AM field as a whole.

6. Conclusions and Future Perspectives

Recent progress in implementing data-driven methods in designing for AM is discussed from different perspectives in this Review. Starting from designing materials specifically optimized for AM, we show examples of improvements achieved both in accuracy and in computational speed of physics-based models such as DFT and MD using data-driven methods. Coupled with the existing database of large quantum mechanical datasets, data-driven methods enable more accurate prediction of the material’s property from its composition and structure at the atomic level, as well as improve the speed of screening, accelerating the identification of optimal material suited for AM. On a larger scale, examples of data-driven methods efficiently establishing composition–structure–property mapping are presented. These high-throughput mappings bring us closer to full discovery of the expanded design space of composite material enabled by AM, thus approaching the full potential of geometrical and material freedom of AM. In addition to data-driven identification of optimal materials for AM, data-driven TO and GD further take advantage of the unique capability of AM to create complex structures. These additional design freedoms in AM have proven to be of great interest, and ML is demonstrated to be necessary for tool path optimization to achieve efficient and highly accurate production of AM-enabled designs. The challenges of solving computationally expensive multiphysics models are addressed by implementing data-driven tools in establishing parameter–property mapping, and examples of optimizing tool path to achieve full density, dimensional accuracy, and thermal stability of print in situ are discussed.

Data-driven approaches are demonstrated to drive AM toward a smarter manufacturing method. However, data-driven methods are not infallible tools that handle every problem without limitations and there is still significant room for improvements. For example, connections between different aspects of DfAM will need to be established in future works of data-driven AM to formulate the design optimization problem in a global framework where materials, structure, tool path, and process control are considered simultaneously. This article aims to showcase the state-of-the-art development of data-driven AM and inspire future work toward improved adaptation of data-driven methods in AM.

Acknowledgements

J.B gratefully acknowledges funding received by the 2020-2021 Knight @ KIC Engineering Graduate Fellowship.

Conflict of Interest

The authors declare no conflict of interest.

Keywords

additive manufacturing, artificial intelligence, machine learning, material designs, tool paths, topology optimizations

Received: January 21, 2021
Revised: June 27, 2021
Published online: July 23, 2021

[1] T. Debroy, H. L. Wei, J. S. Zuback, T. Mukherjee, J. W. Elmer, J. O. Milewski, A. M. Beese, A. Wilson-heid, A. De, W. Zhang, Prog. Mater. Sci. 2018, 92, 112.
[2] J. Y. Lee, J. An, C. K. Chua, Appl. Mater. Today 2017, 7, 120.
[3] T. D. Ngo, A. Kashani, G. Imbalzano, K. T. Q. Nguyen, D. Hui, Compos. Part B Eng. 2018, 143, 172.
[4] W. E. Frazier, J. Mater. Eng. Perform. 2014, 23, 1917.
[5] J. Liu, A. T. Gaynor, S. Chen, Z. Kang, K. Suresh, A. Takezawa, L. Li, J. Kato, J. Tang, C. C. L. Wang, L. Cheng, X. Liang, A. C. To, Struct. Multidiscip. Optim. 2018, 57, 2457.
[6] J. Schmelzle, E. V. Kline, C. J. Dickman, E. W. Reutzl, G. Jones, T. W. Simpson, J. Mech. Des. Trans. ASME 2015, 137, 1.
Chemxi Tian received his bachelor’s degree from University of California, Berkeley. Currently, he is a Ph.D. student, and he joined the Laboratory for Advanced Materials and Manufacturing (LAMM) in 2019 under the direction of Professor Atieh Moridi. His research interests focus on AM material process—property mapping, algorithm-driven optimization of AM composite material design, and manufacturing.

Tianjiao Li received his bachelor’s degree from Xi’an Jiaotong University in 2017 and master’s degree from Northwestern University in 2019. Presently, he is a Ph.D. student, and he joined the J² Lab for Engineering Living Materials in 2019 under the direction of Professor Jingjie Yeo. His research interests focus on multiscale computational modeling, as well as data- and algorithm- driven materials optimization and design.
Jennifer Bustillos received her B.S. in mechanical engineering in 2018 and M.S. in materials science in 2019 from Florida International University (FIU, Miami, FL). From 2014 to 2019, she worked as a research assistant at the Plasma Forming Laboratory (FIU, Miami, FL). Currently, she is a doctoral candidate at Cornell University (Ithaca, NY) in the Sibley School of Mechanical and Aerospace Engineering under the advisement of Moridi, joining the Laboratory for Advanced Materials and Manufacturing. Her research interests include the development of high-performance materials for additive manufacturing and evaluation of mechanical, tribological, and thermal properties of nanostructural materials.

Jingjie Yeo is an assistant professor in the Sibley School of Mechanical and Aerospace Engineering at Cornell University. He heads the J^2 Lab for Engineering Living Materials which focuses on improving the prevention and treatment of diseases associated with aging and noncommunicable diseases. The lab develops soft, adaptive, and responsive biomaterials through interdisciplinary studies of material and biological phenomena by advancing multiscale, multiphysics computational methods and simulations. Prior, he was a research scientist in the Institute of High Performance Computing, Singapore. He did his postdoctoral research in the United States at both Tufts University and Massachusetts Institute of Technology. He received his Ph.D. and his B.Eng. from Nanyang Technological University Singapore.

Atieh Moridi is an assistant professor of Mechanical and Aerospace Engineering and Director of the Laboratory for Advanced Materials and Manufacturing at Cornell University. Prior to that, she was postdoctoral fellow in the Departments of Mechanical Engineering and Materials Science and Engineering at MIT. She received her Ph.D. Cum Laude (the highest institute honor) in mechanical engineering from Politecnico di Milano, Italy. Her research interests lie in the areas of advanced materials and manufacturing. She exploits the intrinsic properties as well as flexibility of additive manufacturing to synthesize novel high-performance materials.