Proceedings of the ASME 2016 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference
IDETC/CIE 2016
August 21-24, 2016, Charlotte, North Carolina, USA

DETC2016-60506

INVESTIGATING PREDICTIVE METAMODELING FOR ADDITIVE MANUFACTURING

Zhuo Yang, Douglas Eddy, Sundar Krishnamurty, Ian Grosse
University of Massachusetts Amherst
Department of Mechanical and Industrial Engineering
Amherst, MA 01003
Email: [zhuoyang, dceddy]@engin.umass.edu, [skrishna, grosse]@ecs.umass.edu

Peter Denno, Felipe Lopez
National Institute of Standards and Technology
Engineering Laboratory
Gaithersburg, MD 20899
Email: [peter.denno, felipe.lopez]@nist.gov

ABSTRACT
Additive manufacturing (AM) is a new and disruptive technology that comes with a set of unique challenges. One of them is the lack of understanding of the complex relationships between the numerous physical phenomena occurring in these processes. Metamodels can be used to provide a simplified mathematical framework for capturing the behavior of such complex systems. At the same time, they offer a reusable and composable paradigm to study, analyze, diagnose, forecast, and design AM parts and process plans. Training a metamodel requires a large number of experiments and even more so in AM due to the various process parameters involved. To address this challenge, this work analyzes and prescribes metamodeling techniques to select optimal sample points, construct and update metamodels, and test them for specific and isolated physical phenomena. A simplified case study of two different laser welding process experiments is presented to illustrate the potential use of these concepts. We conclude with a discussion on potential future directions, such as data and model integration while also accounting for sources of uncertainty.

Keywords: Metamodels, additive manufacturing, space filling sampling, DOE, model updating

1. INTRODUCTION

Additive Manufacturing (AM) processes are more complex, variable, and difficult to understand than subtractive manufacturing [1, 2]. Typical AM processes implement material patterning, energy patterning, new layer creation, and support from previous layers [3] to realize shape, material, and hierarchical complexities [4].

Material properties of AM-produced parts often depend upon the process parameters. For example, platform temperature, building direction, and post heat treatment influence the part microstructure that determines fatigue properties of selective laser melting parts [5]. Further, variations in layer thickness and hatching distance settings have affected material porosity along with hardness and density [6].

Various models have been developed in recent years to describe complex AM process-structure-property relationships. In spite of advances in model accuracy, the enormous computational cost of complex, high-fidelity physics-based simulations of AM makes these models impractical to adopt in industry [7, 8]. A more preferable strategy is to utilize surrogates, or metamodels, as they provide a "model of the model" to replace the expensive simulation model in design and optimization

Certain commercial equipment, instruments, or materials are identified in this paper are not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the materials or equipment are necessarily the best available for the purpose.

processes [9]. Metamodeling has been used successfully as an alternative to computationally expensive simulations in aerospace and other advanced manufacturing domains. [7, 10].

Currently, varieties of metamodeling techniques are applied in engineering design. Several comparative studies present the performance of these various techniques under different modeling criteria [7, 11]. Generally, different modeling methods show both advantages and disadvantages for different types of problems. These disadvantages include orders of nonlinearity and problem scales [7]. To simplify explanation in this paper, we mainly focus on the Polynomial Regression and Kriging Method for metamodel construction to illustrate selection of the most applicable metamodeling techniques for these specific cases.

In spite of the benefits envisioned through the use of such metamodeling techniques, very little research has been done in this area. Some notable exceptions include a polynomial regression model of density, hardness, and porosity of a carbon steel selective laser sintering process [6], porosity predictions in selective laser melting [12] and an energy density model of CoCrMo powder material [13]. These approaches are limited to experimental designs for a specific portion of an AM process. There is a need for a complete AM metamodeling methodology to construct and integrate local metamodels [14] for robust prediction of AM process results. Challenges for the AM situation include cost of experimentation [15], accuracy of simulation capabilities [16], and complex interactions of different physical phenomena during the AM process [17].

This study aims to investigate metamodeling as a means to generate accurate predictive models compatible with a composable multilevel structure, defined as made up of highly reusable models that can be used together and mirror the general AM process model [14]. Such a metamodeling methodology will be able to address the challenges in AM processes such as high system complexity, uncertainty, and limitations of legacy data conducted by design of experiments (DOE) that designers may need to rely on due to the expense of producing experimental sample parts [14, 15, 18].

The following section covers the necessary background to set the stage for Section 3, which introduces methodical approaches to construct and test individual metamodels. A pair of case studies in Section 4 illustrates the potential effectiveness of these approaches. Section 5 discusses this work and potential future work.

2. OVERVIEW OF METAMODELING TECHNIQUES

This section provides a brief background summary of basic metamodeling approaches. These traditional metamodeling techniques consist of several steps. First, data sets are used to construct metamodels. The composition of these data sets depend upon the experimental design used to represent and sample that design space.

Traditional sampling methods, such as full factorial, fractional factorial, and central composite, etc., have been frequently used for AM process modeling [13, 19]. Typically, variables could be either discrete or continuous [20]. Since some or most of the input variables are continuous in this case, it would be impossible to generate all possible combinations in a data set. These traditional experimental designs discretize the continuous variables to limit the number of experimental trials necessary.

The full factorial design selects all possible combinations of input variables at specific locations to maximize the amount of information and data accuracy for a prescribed sample size. Fractional factorial design symmetrically selects a fraction of full factorial samples. Central Composite Design (CCD) includes center points that can estimate curvature of the response function [21]. CCD and other classical DOE methods tend to reduce the experimental cost and cover an entire design space by placing the samples at or near the boundary of the space. However, it leaves the interior of the design space unexplored [22]. In this study the data sources used to construct the metamodels consist of DOE data sets that represent manufacturing processes empirically. The goal is to develop a method that may be universally applied to all of these main types of classical DOE sampling methods.

To capture the important characteristics of unknown systems, it is preferred to collect data that represents the entire design space to include information about the most critical regions [21, 23]. This is necessary to overcome the limitations of classical DOE approaches. Unlike the random errors in results often exhibited by physical experiments, computer experiments are often deterministic [24]. This work focuses on metamodel construction from legacy empirical data rather than computer experiments to address the issues of difficulty with obtaining accurate simulations of AM processes and the high cost of producing sample parts for new experimental designs.

Thus, any approach needs flexibility to accommodate larger DOE sample sizes and combinations of different DOE data sets, which may not conform to the classical DOE sample locations. For example, the welding test by Khan et al. was based on full factorial DOE strategy that included 18 data points [25]. A similar experiment operated by Balasubramanian et al. with the same response was based on a fractional factorial design that included 15 data points [26]. Aforementioned experiments cannot conform to one uniform DOE method due to their different design space and levels of value.

Space Filling Sampling (SFS) methods have been developed to address these various limitations of the classical DOE approaches [22]. Many SFS methods have been widely used in simulation-based metamodels, such as Grid Sampling, Lattice Design, Audze-Eglais, Orthogonal Array, and Latin Hypercube Design (LHD) [23, 27-30]. The number of sample points of LHD is the same as the number of discrete cells defined by the level, or grid spacing, of each input variable. For example, the Latin square that contains four sample points in two dimensions appears only once in each row and each column [21]. LHD fills each cell location with one sample point, but the location within each cell is randomized. LHDs are usually at least as accurate as random sampling and stratified sampling techniques [31]. Thus, Latin Hypercube can be a most suitable candidate for situations that involve nonflexible and non-uniform data locations in a design space [32].

2.1 Traditional Metamodel Construction Techniques

Since SFS is often not adequate to generate desired model accuracy, reliability, predictability, or robustness, Sequential Infilling Sampling (SIS) techniques are widely used. Unlike SFS methods that distribute sample points into a design space at one single stage, SIS methods assign sample points sequentially at "particular" locations [21]. SIS methods are more efficient for an unknown system [33] by providing options to a designer for determining when to stop the data collection process as sufficient information has been gathered [21]. The combination of SFS and SIS can significantly improve the results of metamodel construction [34]. The most popular construction techniques include polynomial, kriging, splines, artificial neural network, and hybrid methods [35].

Response surface modeling (RSM) techniques were originally developed to analyze the results of physical experiments and create empirical models of the observed response values [36]. The typical form of RSM is:

$$y(\tilde{x}) = f(\tilde{x}) + \varepsilon \tag{1}$$

where $y(\tilde{x})$ represents the unknown function, $f(\tilde{x})$ is a known polynomial function of $\tilde{x}$ derived statistically, and ε is random error assumed to be normally distributed. $\tilde{x}$ is the set of the system's independent input variables. A second order quadratic polynomial function would have the form of:

$$\hat{y} = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_i \sum_j \beta_{ij} x_i x_j \tag{2}$$

where β_0, β_i, and β_{ij} are regression coefficients, and k is the number of design variables.

RSM provides the advantage of generating a mathematical function that can easily compute the data location predicted by that equation. However, since RSM deploys curve-fitting techniques between the data points, it can tend to smooth out such regions without data to lessen predictive accuracy for highly nonlinear responses of systems [21].

Kriging methods were initially developed for analyses of random processes and have been known to outperform RSM and other metamodel construction techniques, especially when dimensionality of a system increases [37]. The fundamental assumption of predicting in the kriging method is modeled through the variogram, or spatial correlation functions, which describes spatial correlation of observed data [38, 39]. The general form of a kriging model is:

$$\tilde{y} = f(\tilde{x}; \tilde{\beta}) + \varepsilon(\tilde{x}) \tag{3}$$

where $f(\tilde{x}; \tilde{\beta})$ is a least squares fit regression to the global trend in the observed data, $\varepsilon(\tilde{x})$ is the correlated prediction error, which is assumed to be the realization of a stochastic/Gaussian process with zero mean, non-zero variance, and covariance. Although defined as statistical error, this term represents an imposed spatial correlation in the parameter space that makes observations positively correlated when close. Kriging can potentially provide improved accuracy and reliability over other techniques available to AM metamodeling.

3 METAMODEL CONSTRUCTION METHOD

This section introduces a predictive metamodeling approach (Figure 1) to address some of the unique challenges particular to metamodel construction to represent the various sub-processes in AM. The following subsection explains the rationale of this methodical approach based on the challenges identified previously in this paper related to predictive metamodeling for AM processes.

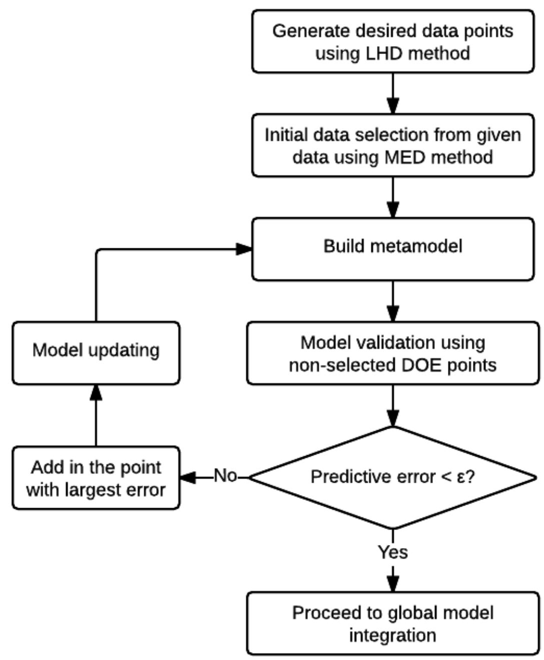

Figure 1. Maximum Predictive Error Updating (MPEU) Method

3.1 Rationale of Approach

One of the major objectives for developing a reusable metamodeling methodology is to give engineers the opportunity to use historical AM data to construct their own design, which can potentially save the cost of collecting DOE data. Thus, this work strives for compatibility with most of the AM processes and different empirical data sets. Existing data that conforms to classical DOE does not provide guidance towards selecting sample points. Further, existing data may not account for the same set of conditions (design space) provided by a metamodel. Variability of the classical DOE set-ups further complicates the quest for a method compatible with existing DOE data. The variability of different DOE methods deployed in data collection includes different numbers of levels, fractional factorials, or input variables relative to the problem non-uniformity. To overcome these disadvantages, the following subsection introduces the Minimum Euclidean Distance (MED) method for selecting a limited number of data points for metamodel construction from nonflexible given data.

3.2 Minimum Euclidean Distance (MED) Method

The coauthors' prior work addressed a similar situation of metamodel construction in non-ideal data locations for design space filling [32][Eddy 2015]. This work identified LHD as a potential approach for the reasons given in Section 2. Given the inability to choose the points at the exact locations identified by an LHD sample set generation, this work proposed a method to find the minimum Euclidean distance between each identified data location and the data point identified in the data set [32].

The procedural steps begin with the generation of the desired amount of LHD points from a given data set. Next the Euclidean distance is calculated between each DOE data point to the generated LHD points by the Maximin method. Those points closest to the desired LHD points are selected for constructing the metamodel.

Since the selected DOE points depend upon the LHD points, selection of the initial points from these LHD results is critical to improvement of metamodel construction. Such methods as Maximin LHD [40], orthogonal array-based LHD [30], and optimal Audze-Eglais uniform LHD [41] can generate optimal LHD sampling points.

3.3 Maximum Predictive Error Updating (MPEU) Method

As mentioned in Section 2, the initial step of space filling is often not adequate to obtain the desired accuracy and predictability of a metamodel. In such cases, SIS becomes a necessary next step.

Utilizing the concepts from SIS, Shao and Krishnamurty developed a surrogate model based design optimization (SMBDO) method to sequentially update a surrogate model by capturing the critical features of an unknown system in a simulation-based experiment [42]. Similary, a comprehensive adaptive sampling methodology is presented in Sandia's Dakota framework [43] to enable selection of successive sample points based on the maximum distance from existing points or the uncertainty of model prediction. Based on these methods, the initial LHD sample points are used to construct the initial surrogate model. During each updating step, potential optimal locations predicted by the current model are then validated. Those points that exhibit high predictive error are then added into the current model iteratively until the desired model accuracy is obtained. However, the clustering based multilocation search procedure of SMBDO and the adaptive sampling method both rely upon an ample supply of data points from efficient computer simulations, which is simply not realistic in this case. Thus, there is a need to develop a model updating method for the limited data sizes inherent with using historical data for metamodel construction.

To address these challenges and limitations, we introduce the Maximum Predictive Error Updating (MPEU) method to gradually improve model accuracy. Figure 1 outlines the general framework of the MPEU method using the MED method to select the most appropriate sampling points from original DOE data. The method begins by generating LHD points using the maximin method in the design space given by the data. The kriging method with a Gaussian correlation function was employed to build the surrogate models. Then, the points are selected to construct the initial metamodel by using the MED method. A validation procedure next determines whether the current surrogate model needs further improvement. Maximum relative error magnitude (MREM) and average relative error magnitude (AREM) are calculated using the following equations to test the metamodel for predictability at each iterative stage of model updating:

$$MREM = \max\left(\frac{|y_i - \hat{y}_i|}{y_i}\right) \tag{4}$$

$$AREM = \frac{1}{m}\left(\frac{\sum_{i=1}^{m}|y_i - \hat{y}_i|}{y_i}\right) \tag{5}$$

where y_i is the observed value from given data, $\hat{y}$ is the value predicted by the metamodel of the DOE points that were not selected to construct the metamodel, and m is the number of data points.

If either error calculation exceeds a preset threshold for robustness, the point presenting the largest predictive error is added into the initial sample pool. A new model is created based on the new sample set. The model creating-validating procedure will iteratively proceed until both MREM and AREM satisfy the preset threshold value for robustness. This is the MPEU method that sequentially infills the sample set by updating the model to improve the resulting metamodel construction.

Verification and validation techniques must test the metamodel at each stage [21]. Verification tests the internal consistency of constructed metamodels and validation tests reliability with external data [44]. To validate the newly built metamodel, a model validation criterion is established based on the prediction accuracies [21] of all non-selected DOE points. If the MREM and AREM are both greater than a specified preset threshold value for robustness, the data point with the lowest prediction accuracy (or highest MREM value) would be added into the current sample pool and the metamodel is updated accordingly. Subsequently, the newly built metamodel will be validated again with the same process iteratively until convergence to within the threshold MREM and AREM values. Thus, effective model construction can be achieved efficiently by combining predictive metamodel construction simultaneously with validation to robustness requirements.

The preset threshold values of MREM and AREM are based on design requirements such as penetration depth and melt pool width. Both average and maximum error are involved in the validation process since they represent general and distinguished model performance. A designer would need to decide on what model accuracy and predictability are necessary or acceptable before model construction [32]. An unnecessarily low threshold value may significantly increase the computational cost. Conversely, an excessively high threshold value may reduce model predictability and utility. The following section demonstrates the potential use of these proposed space filling and sequential infilling techniques in a pair of case study examples.

4 CASE STUDY: PREDICTIVE METAMODELS IN LASER WELDING PROCESSES

The laser welding process is used in the case studies reported in this section due to its similarities to directed energy deposition processes. In both applications, a heat source fuses metal as it is being deposited. The processes share similar process parameters and their quality is determined by similar metrics (dimensional accuracy, surface finish, residual stresses and mechanical properties, all of which can be traced back to the geometry of the melt pool). With that said, data is more readily available for laser welding, making the process a good candidate for demonstrating proof-of-concept.

The following two case studies illustrate the potential applicability of the proposed MPEU method for different DOE data. Both cases focus on the same response of the penetration depth (P). The cases have similar experimental methods but different DOE strategies. These two simple and somewhat similar experiments help to illustrate various results that can be expected from different data sets. This section shows the potential to deploy methods to construct and test various individual AM metamodels by use of the method introduced in the prior section.

4.1 Full Factorial DOE with Different Levels of Value for Input Variables

In the first case by Kahn, et al. [25], laser power (LP), welding speed (WS), and fiber diameter (FD) are the input variables. Among those three variables, LP and WS ranged from 800–1100 W and 4.5–7.5 m/min by three linear levels, with midpoint 950 and 6.0 respectively. The third variable of FD has only two levels at values of 300 µm and 400 µm FD [25]. The full factorial DOE consists of eighteen total data points for penetration depth, measured in micrometers after a standard washing procedure and with no special heating treatment.

The first step is generating an LHD sample set in the design space. In this case the LHD set consists of five points in order to give the initial model enough options for future updating. Fewer start points may not adequately cover the design space. Using the MED method described in the previous section, the Euclidean distance between each LHD point and DOE point are calculated. Table 1 lists the initial data points selected by the MED method. The first column represents the standard order number of each point in the original DOE. The initial metamodel would be constructed from these five points.

Table 1. Initial data points generated by MED method

	Input variables			Observed value
Data point number	**LP (W)**	**WS (m/min)**	**FD (µm)**	**P (µm)**
1	800	4.5	300	960
5	950	6	300	950
6	1100	6	300	1180
14	950	6	400	727
17	950	7.5	400	580

From the collected data, the initial metamodel is built using a standard kriging method. Kriging has built-in verification of internal consistency to prevent the error that can occur when RSM is used. The remaining thirteen data points next validate the metamodel by calculation of MREM and AREM as explained in the prior section. Model updating is next done iteratively by applying the MPEU method, as described in the prior section, to the preset thresholds for robustness of $\varepsilon_{MREM} \leq 10\%$, and $\varepsilon_{AREM} \leq 5\%$ in this case.

Table 2. MREM and AREM at each stage

Iteration	Data point number	Input variables			Observed value	Predictive value	MREM	AREM
		LP (W)	**WS (m/min)**	**FD (µm)**	**P (µm)**			
stage 1	3	1100	4.5	300	1610	1108	31.12%	28.37%
	7	800	7.5	300	560	891	59.18%	
	12	1100	4.5	400	1307	875	33.02%	
	13	800	6.0	400	577	818	41.86%	
	16	800	7.5	400	492	759	54.43%	
stage 2	3	1100	4.5	300	1610	1339	16.79%	11.10%
	9	1100	7.5	300	880	1019	15.76%	
	11	950	4.5	400	1043	899	13.82%	
	12	1100	4.5	400	1307	1094	16.31%	
	16	800	7.5	400	492	385	21.70%	
stage 3	2	950	4.5	300	1290	1107	14.20%	14.26%
	3	1100	4.5	300	1610	1244	22.71%	
	11	950	4.5	400	1043	845	18.93%	
	12	1100	4.5	400	1307	882	32.53%	
	15	1100	6.0	400	920	806	12.37%	

stage 4	2	950	4.5	300	1290	1241	3.81%	3.71%
	3	1100	4.5	300	1610	1472	8.57%	
	8	950	7.5	300	730	702	3.87%	
	13	800	6.0	400	577	539	6.65%	
	15	1100	6.0	400	920	963	4.62%	

Table 2 lists the results for this example of the first four iterations of the MPEU method. Note that each sequential iteration represents the validation results calculated by the current updated metamodel. Only the points showing the most significant error are included in Table 2. At each iteration, the point with the greatest MREM is marked in grey in Table 2. At stage 1, the point at standard order 7 is selected by adding it into the initial sample pool since it shows the highest MREM (59.18%). As a result, after the third iteration both the MREM and the AREM values satisfy the preset threshold value. According to the MPEU method, the updating process converged to construct the final metamodel with eight DOE data points at 8.57% MREM and 3.35% AREM. Figure 2 shows the error values at each iteration. Note that the MREM and AREM values did not always decrease monotonically prior to the final stage, as one would expect in the early stages in any numerical iterative approach, but shows monotonicity and convergence towards the end. Similar trends were observed in the application of SMBDO to several classical simulation-based model updating case studies [25].

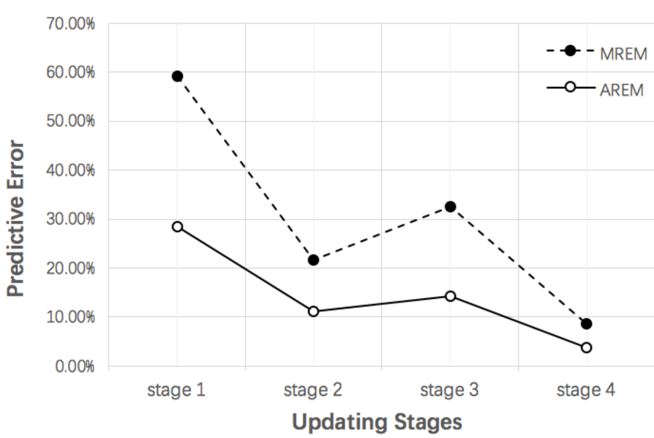

Figure 2. Error at iterations

In this first case study a two level, three factor DOE strategy becomes the only choice if one prefers to create the model by classical DOE techniques without collecting new experimental data. Beyond the proposed MPEU method and a two level DOE method, another compatible sampling method is a random search method. However, this method is not recommended here due to its uncontrolled behavior. The comparison results of AREM and MREM between random search method, two level full factorial DEO method, and MPEU method with different threshold values are listed in Table 3.

As shown in the table, both AREM and MREM of the model built by the MPEU method are significantly lower than the random search and DOE methods when the sample size is the same. When gradually reducing the threshold values of AREM and MREM, the MPEU method typically incorporates a few more points to improve the model accuracy to the new convergence requirements.

With the MPEU method, the model is iteratively improved by updating sample points. However, only the initial sample points can evenly distribute across the given design space by use of the MED method. Newly updated points are selected based on the validation results from previous iterations without considering their location in a design space. As shown in Table 4 the metamodel that is constructed by sequential infilling reduces prediction errors at comparable sample sizes. Thus, despite the possibility of the initial five data points not adequately filling the design space, the MPEU method shows potential to generate a more accurate model through the updating strategy. The following subsection examines the results of applying this same method to a situation that provides fewer data points in a data set.

Table 3. AREM and MREM results of random search method, two levels full factorial DOE method, and MPEU method

	Random	**DOE**	**MPEU**		
Sample size	n=8	n=8	n=8	n=10	n=12
AREM	22.76%	7.76%	3.35%	2.44%	1.94%
MREM	59.16%	12.99%	8.57%	4.74%	5.18%

Table 4. Comparison of different point selection strategies

	Single stage sampling	**MPEU method**	**Improvement**
Sample size	n=8	n=8	
AREM	7.60%	3.35%	55%
MREM	12.66%	8.57%	32%

4.2 Fractional Factorial DOE with Same Levels of Value

The second case study of laser welding DOE data is based on a three factor, three level Box-Behnken design with full replication [26]. "Beam angle" (BA) in this experiment replaced the input of "fiber diameter" from the first case study. The experimental design generated fifteen data points. A mean value of the data set's three replicate points reduces the size of the data set from fifteen to thirteen.

Table 5 lists the MREM and AREM values at each stage for those points having significant predictive error. As shown, the

MREM started with five sample points from 82.00% and gradually decreased to 4.80% after five updates, or six stages. The error at the start could have a significant effect on the number of iterations required. It is notable that the error at the first stage is 39% higher than the amount shown for the first stage in Table 2 for the first case study. It is also notable that this second case study is covering more levels with less data than the first case study.

After applying the MPEU method, accuracies of the final model satisfied the threshold values (ε_{MREM}≤10%, ε_{AREM}≤5%). The error convergence progression is shown in Figure 3. Both MREM and AREM increased slightly during the middle stages as new points were added into the previous sample pool.

Table 5. MREM and AREM at each stage

iteration	Std. order	MREM	AREM
stage 1	3	45.89%	34.93%
	4	82.00%	
	6	52.50%	
	8	54.79%	
stage 2	1	15.46%	11.09%
	6	22.16%	
	8	21.71%	
	13	4.79%	
stage 3	1	41.51%	15.71%
	3	33.93%	
	8	0.85%	
	13	16.64%	
stage 4	3	4.48%	4.60%
	8	1.30%	
	9	2.71%	
	13	13.32%	
stage 5	3	17.07%	6.16%
	8	3.69%	
	9	3.22%	
	12	0.65%	
stage 6	9	4.80%	3.72%

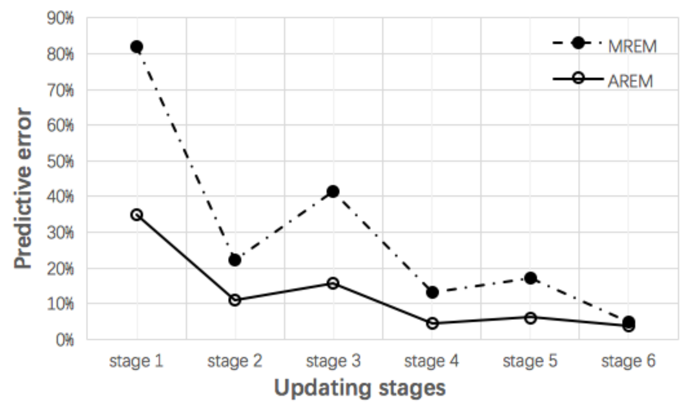

Figure 3. Error at iterations

5 DISCUSSION AND FUTURE WORK

The objective of this work was to explore a metamodeling methodology tailored for AM and adaptable to different types of empirical data. To address the challenges identified, this work introduces an MED method to select usable sample points from different types of given data and an MPEU method with an updating procedure to create predictive metamodels to predetermined robustness requirements from limited data sets. The proposed MED method can select usable initial sample points from various types of DOE data since its foundation is based on the LHD sampling method, which is adaptable for most any design space. Though the generated LHD sample locations may not be occupied by given DOE data, the MED method can improve selection of more appropriate existing points over other methods.

The MPEU method allows model developers to balance the tradeoff between model accuracy and computational cost by adjusting the threshold values of MREM and AREM to achieve specified levels of robustness. As shown in Table 3, with the same number of sample points, the MPEU method, which also utilizes the MED method, provides a more accurate model than the random search and eight DOE data points for the example that was tested. The MPEU method also provides an option if one intends to improve the model at the expense of slightly higher computational costs. In the first case study, two added new points can significantly reduce the MREM from 8.57% to 4.74%. Furthermore, the updating strategy of the proposed method can contribute more to capture the critical features of an unknown system than simply picking up points from the given data set. As shown in Table 4, MPEU method significantly reduces both MREM and AREM. In other words, the proposed method focuses more on capturing the critical system features rather than the point locations.

Despite the advantages in model construction with the MED and MPEU methods, there are some limitations. Such disadvantages can potentially limit the application of proposed methods. For example, since the MED method selects the initial points through randomly generated LHD samples, each time the MPEU method may produce different models to the same convergence criteria. It cannot guarantee that the generated LHD sampling set is optimal in a given DOE design space. Rather, the updating procedure depends highly upon the initial MED selected points. Without a confirmed starting point, the overall performance of the MPEU method may decrease. For example, in Section 4.2, the final model required ten points for construction but left only three points for validation. Thus, another limitation relates to model validation. Unlike metamodels generated by computer simulations, historical DOE data is often not reproducible. One can only rely on the existing data since it is impossible to gather additional information. In this second case study, the start point accuracy and resulting number of points remaining to validate the model were not as acceptable as found in the first case study. This second case study also had less data than the first case study. While not conclusive, this supports the assertion that the amount of data or information can have a significant effect on the results of using metamodeling methods. Methods such as Grey System Theory that work with little data or information may be introduced along with this current proposed method [45]. Nonetheless, future work could potentially improve the MPEU method by adding a check and adjustment process based on the error at the first stage.

Two laser welding case studies in the prior section show that the proposed MPEU method is compatible with different DOE data sets in these cases. The two data sets have similar experimental conditions such as the same laser source, common input variables of laser power and welding speed, the same response of penetration depth. However, one must use these two metamodels separately due to their different ranges of data locations in the design space. To overcome such a shortcoming or data limitation, a future goal is to build towards a global metamodel by combing two local data sets. Such a development may more efficiently utilize different historical data sources to know more about a process and also raises the issue of uncertainty between data sets with different sources.

Another aspect to explore in future studies is the uncertainty in data points, which may be measurement error for experimental data or prediction uncertainty for computational predictions. Experimental variability is significant in additive manufacturing processes due to the random interactions between powder particles and melt pool, variations between different machines and models, and different manufacturing practices followed in different shops. Additive manufacturing models have a large degree of variability as well because of the different modeling assumptions that may be taken in the development of computational models. Merging of data sets of different sources requires knowledge of the amount of uncertainty in each source, to ensure that the metamodels stay closer to more accurate points and that an appropriate metric is adopted to determine the adequacy of the metamodels. The methods presented in this paper originate from the metamodeling literature, which has traditionally dealt with deterministic data. Extensions to stochastic data sets are under development.

The MPEU method lays the foundation for a predictive metamodeling methodology to use in AM. Future work could investigate development of a hybrid metamodeling method through the application of clustering techniques [34] and multi-surrogate approximation (MSA) methods [46] to build the global model by combining data sets with different input variables, process conditions, or material parameters.

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation (NSF) under Grant No. 1439683, the National Institute of Standards and Technology (NIST) under Cooperative Agreement number NIST 70NANB15H320, and industry members of the NSF Center for e-Design.

REFERENCES

[1] Gibson, I., Rosen, D.W., and Stucker, B., 2010, "Additive manufacturing technologies," Springer, .

[2] Kim, D. B., Witherell, P., Lipman, R., 2015, "Streamlining the Additive Manufacturing Digital Spectrum: A Systems Approach," Additive Manufacturing, **5**pp. 20-30.

[3] Williams, C. B., Mistree, F., and Rosen, D. W., 2011, "A Functional Classification Framework for the Conceptual Design of Additive Manufacturing Technologies," Journal of Mechanical Design, **133**(12) pp. 121002.

[4] Rosen, D.W., 2007. Design for additive manufacturing: a method to explore unexplored regions of the design space. In Eighteenth Annual Solid Freeform Fabrication Symposium pp. 402-415.

[5] Brandl, E., Heckenberger, U., Holzinger, V., 2012, "Additive Manufactured AlSi10Mg Samples using Selective Laser Melting (SLM): Microstructure, High Cycle Fatigue, and Fracture Behavior," Materials & Design, **34**pp. 159-169.

[6] Chatterjee, A., Kumar, S., Saha, P., 2003, "An Experimental Design Approach to Selective Laser Sintering of Low Carbon Steel," Journal of Materials Processing Technology, **136**(1) pp. 151-157.

[7] Jin, R., Chen, W., and Simpson, T. W., 2001, "Comparative Studies of Metamodelling Techniques Under Multiple Modelling Criteria," Structural and Multidisciplinary Optimization, **23**(1) pp. 1-13.

[8] Khairallah, S. A., Anderson, A. T., Rubenchik, A., 2016, "Laser Powder-Bed Fusion Additive Manufacturing: Physics of Complex Melt Flow and Formation Mechanisms of Pores, Spatter, and Denudation Zones," Acta Materialia, **108**pp. 36-45.

[9] Kleijnen, J.P., 1986, "Statistical tools for simulation practitioners," Marcel Dekker, Inc., .

[10] Kleijnen, J. P., and Sargent, R. G., 2000, "A Methodology for Fitting and Validating Metamodels in Simulation," European Journal of Operational Research, **120**(1) pp. 14-29.

[11] Varadarajan, S., Chen, W. and Pelka, C.J., 2000, “Robust concept exploration of propulsion systems with enhanced model approximation capabilities,” Engineering Optimization A35, 32(3), pp. 309-334.

[12] Tapia, G., and Elwany, A., 2015, “Prediction of porosity in SLM parts using a MARS statistical model and Bayesian inference,” Proceedings of the 2015 Annual International Solid Freeform Fabrication Symposium, pp. 1205-1219.

[13] Ciurana, J., Hernandez, L., and Delgado, J., 2013, "Energy Density Analysis on Single Tracks Formed by Selective Laser Melting with CoCrMo Powder Material," The International Journal of Advanced Manufacturing Technology, **68**(5-8) pp. 1103-1110.

[14] Witherell, P., Feng, S., Simpson, T. W., 2014, "Toward Metamodels for Composable and Reusable Additive Manufacturing Process Models," Journal of Manufacturing Science and Engineering, **136**(6) pp. 061025.

[15] Thomas, D. S., and Gilbert, S. W., 2014, "Costs and Cost Effectiveness of Additive Manufacturing," .

[16] Ding, J., Colegrove, P., Mehnen, J., 2014, "A Computationally Efficient Finite Element Model of Wire and Arc Additive Manufacture," The International Journal of Advanced Manufacturing Technology, **70**(1-4) pp. 227-236.

[17] Michopoulos, J. G., Lambrakos, S., and Iliopoulos, A., 2014, "Multiphysics challenges for controlling layered manufacturing processes targeting thermomechanical performance," ASME 2014 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Anonymous American Society of Mechanical Engineers, pp. V01AT02A050-V01AT02A050.

[18] Van Elsen, M., Al-Bender, F., and Kruth, J., 2008, "Application of Dimensional Analysis to Selective Laser Melting," Rapid Prototyping Journal, **14**(1) pp. 15-22.

[19] Spierings, A., Levy, G., and Wegener, K., 2014, "Designing material properties locally with additive manufacturing technology SLM," ETH-Zürich, .

[20] Montgomery, D.C., 2008, "Design and analysis of experiments," John Wiley & Sons, .

[21] Shao, T., 2007, “Toward a structured approach to simulation-based engineering design under uncertainty,” Ph.D. Thesis, University of Massachusetts Amherst, Amherst, MA.

[22] Giunta, A. A., Wojtkiewicz, S. F., and Eldred, M. S., 2003, "Overview of modern design of experiments methods for computational simulations," Proceedings of the 41st AIAA aerospace sciences meeting and exhibit, AIAA-2003-0649, Anonymous .

[23] Husslage, B. G., Rennen, G., van Dam, E. R., 2011, "Space-Filling Latin Hypercube Designs for Computer Experiments," Optimization and Engineering, **12**(4) pp. 611-630.

[24] Simpson, T. W., Booker, A. J., Ghosh, D., 2004, "Approximation Methods in Multidisciplinary Analysis and Optimization: A Panel Discussion," Structural and Multidisciplinary Optimization, **27**(5) pp. 302-313.

[25] Khan, M., Romoli, L., Fiaschi, M., 2010, "Experimental Investigation on Laser Beam Welding of Martensitic Stainless Steels in a Constrained Overlap Joint Configuration," Journal of Materials Processing Technology, **210**(10) pp. 1340-1353.

[26] Balasubramanian, K., Siva Shanmugam, N., Buvanashekaran, G., 2008, "Numerical and Experimental Investigation of Laser Beam Welding of AISI 304 Stainless Steel Sheet," Adv.Produc.Engineer.Manag, **3**(2) pp. 93-105.

[27] Cole, R., Healy, T., Wood, M., 2001, "Statistical Analysis of Spatial Pattern: A Comparison of Grid and Hierarchical Sampling Approaches," Environmental Monitoring and Assessment, **69**(1) pp. 85-99.

[28] Wang, Y., and Hickernell, F.J., 2002, "An historical overview of lattice point sets," Springer, .

[29] Audze, P., and Eglais, V., 1977, "New Approach for Planning Out of Experiments," Problems of Dynamics and Strengths, **35**pp. 104-107.

[30] Tang, B., 1993, "Orthogonal Array-Based Latin Hypercubes," Journal of the American Statistical Association, **88**(424) pp. 1392-1397.

[31] McKay, M. D., Beckman, R. J., and Conover, W. J., 2000, "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code," Technometrics, **42**(1) pp. 55-61.

[32] Eddy, D. C., Krishnamurty, S., Grosse, I. R., 2015, "A Predictive Modelling-Based Material Selection Method for

Sustainable Product Design," Journal of Engineering Design, **26**(10-12) pp. 365-390.

[33] Sacks, J., Welch, W. J., Mitchell, T. J., 1989, "Design and Analysis of Computer Experiments," Statistical Science, pp. 409-423.

[34] Shao, T., and Krishnamurty, S., 2008, "A Clustering-Based Surrogate Model Updating Approach to Simulation-Based Engineering Design," Journal of Mechanical Design, **130**(4) pp. 041101.

[35] Wang, G. G., and Shan, S., 2007, "Review of Metamodeling Techniques in Support of Engineering Design Optimization," Journal of Mechanical Design, **129**(4) pp. 370-380.

[36] Box, G.E., and Draper, N.R., 1987, "Empirical model-building and response surfaces," Wiley New York, .

[37] Simpson, T. W., 1998, "Comparison of Response Surface and Kriging Models in the Multidisciplinary Design of an Aerospike Nozzle," .

[38] Shan, S., and Wang, G. G., 2010, "Survey of Modeling and Optimization Strategies to Solve High-Dimensional Design Problems with Computationally-Expensive Black-Box Functions," Structural and Multidisciplinary Optimization, **41**(2) pp. 219-241.

[39] Cressie, N., 2015, "Statistics for spatial data," John Wiley & Sons, .

[40] Van Dam, E. R., Husslage, B., Den Hertog, D., 2007, "Maximin Latin Hypercube Designs in Two Dimensions," Operations Research, **55**(1) pp. 158-169.

[41] Bates, S. J., Sienz, J., and Langley, D. S., 2003, "Formulation of the Audze–Eglais Uniform Latin Hypercube Design of Experiments," Advances in Engineering Software, **34**(8) pp. 493-506.

[42] Shao, T., and Krishnamurty, S., 2009, "A Preference-Performance Hybrid Method for Surrogate Model Updating in Engineering Design Optimisation," International Journal of Product Development, **9**(1-3) pp. 218-264.

[43] Eldred, M.S., Giunta, A.A., van Bloemen Waanders, Bart G, 2007, "DAKOTA, a multilevel parallel object-oriented framework for design optimization, parameter estimation, uncertainty quantification, and sensitivity analysis: Version 4.1 reference manual," Citeseer, .

[44] Sargent, R. G., 2005, "Verification and validation of simulation models," Proceedings of the 37th conference on Winter simulation, Anonymous winter simulation conference, pp. 130-143.

[45] Deng, J., 1989, "Grey Information Space," The Journal of Grey System, **1**(1) pp. 103-117.

[46] Zhao, D., and Xue, D., 2011, "A Multi-Surrogate Approximation Method for Metamodeling," Engineering with Computers, **27**(2) pp. 139-153.