MACHINE LEARNING AND PROBABILISTIC DESIGN FRAMEWORK FOR LASER POWDER BED FUSION PROCESS

by

Lingbin Meng

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Department of Mechanical and Energy Engineering at IUPUI

Indianapolis, Indiana

May 2020

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Jing Zhang, Co-Chair

Department of Mechanical and Energy Engineering

Dr. Bexin Wu, Co-Chair

School of Mechanical Engineering

Dr. Xiaoping Du

Department of Mechanical and Energy Engineering

Dr. Mohammad Al Hasan

School of Computer & Information Science

Shengfeng Yang

Department of Mechanical and Energy Engineering

Approved by:

Dr. Jie Chen

For my wife Lijie Song and son Deer Meng.

ACKNOWLEDGMENTS

First and foremost, I would like to express my sincere appreciation for my academic advisors, Professor Jing Zhang, and Professor Benxin Wu, for their expert advice, valuable guidance, and kind support during my study and research at Purdue University. I would also like to thank my committee members Professor Xiaoping Du, Professor Mohammad Al Hasan, and Professor Shengfeng Yang, for serving as my committee members, and their great help in my studies.

The work is conducted under CCDC Army Research Laboratory Cooperative Research and Development Agreement (CRADA) 19-013-001. I would like to thank Dr. Brandon McWilliams from Army Research Laboratory for suggestions on machine learning model. Special thanks to Prof. Xiaoping Du for his valuable guidance in the probabilistic engineering design and the first order reliability method, Dr. Tao Sun from Argonne National Laboratory for providing the high-speed synchrotron X-ray imaging data, and Prof. Yeon-Gil Jung from Changwon National University for 3D printing process.

I also would like to thank my lab mates Dr. Linmin Wu, Dr. Yi Zhang, Xuehui Yang, Jian Zhang and many others, who gave me help in many aspects.

Last but not the least, I would like to thank my parents, Xianghai Meng and Lei Feng, and my wife, Lijie Song, for their love and support during my doctoral studies and in my whole life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

There has been increasing demand for 3D printed metals from aerospace & defense and automotive end-use industries, due to their low manufacturing cost, and reduction in lead times. Although the significant advancement in metal 3D printing promises to revolutionize industry, it is constrained by a widespread problem: the cracks and other defects in the metal 3D printed parts. In this work, two major causes of defects in the laser power bed fusion (L-PBF) process are focused: keyhole mode and spattering phenomena. Both defect sources are highly dependent to the processing parameters. Although extensive efforts have been made on experiments and computational models to improve the quality of printed parts, the high experimental costs and large computational intensity still limit their effect on the optimization of the processing parameters. In addition, the uncertainty in the design process further limits the accuracy of these models.

The aim of this thesis is to develop a probabilistic design framework for reliability-based design in the L-PBF process. The modeling framework spans physical models, machine learning models, and probabilistic models. First, the keyhole mode and spattering phenomena are simulated by physical models including computational fluid dynamics (CFD) and smoothed particle hydrodynamics (SPH) methods, respectively. Then, the data acquired by the physical models serve as the training data for machine learning models, which are used as surrogates to alleviate the high computational cost of physical models. Finally, the feasible design region is computed by probabilistic models such as Monte Carlo simulation (MCS) and the first order reliability method (FORM). The feasible design region can be used assessing a satisfactory reliability not lower than the specified reliability level.

The developed Gaussian process (GP) based machine learning model is capable of predicting the remelted depth of single tracks, as a function of combined laser power and laser scan speed in the L-PBF process. The GP model is trained by both simulation and experimental data from the literature. The mean absolute prediction error magnified by the GP model is only 0.6 μm for a powder bed with layer thickness of 30 μm, suggesting the adequacy of the GP model. Then, the process design maps of two metals, 316L and 17-4 PH stainless steel, are developed using the trained model. The normalized enthalpy criterion of identifying keyhole mode is

evaluated for both stainless steels. For 316L, the result suggests that the $\frac{\Delta H}{h_s} \geq 30$ criterion should be related to the powder layer thickness. For 17-4 PH, the criterion should be revised to $\frac{\Delta H}{h_s} \geq 25$.

Moreover, a new and efficient probabilistic method for the reliability analysis is developed in this work. It provides a solution to address quality inconsistency due to uncertainty in the L-PBF process. The method determines a feasible region of the design space for given design requirements at specified reliability levels. If a design point falls into the feasible region, the design requirement will be satisfied with a probability higher or equal to the specified reliability. Since the problem involves the inverse reliability analysis that requires calling the direct reliability analysis repeatedly, directly using MCS is computationally intractable, especially for a high reliability requirement. In this work, a new algorithm is developed to integrate MCS and FORM. The algorithm finds the initial feasible region quickly by FORM and then updates it with higher accuracy by MCS. The method is applied to several case studies, where the normalized enthalpy criterion is used as a design requirement. The feasible regions of the normalized enthalpy criterion are obtained as contours with respect to the laser power and laser scan speed at different reliability levels, accounting for uncertainty in seven processing and material parameters. The results show that the proposed method dramatically alleviates the computational cost while maintaining high accuracy. This work provides a guidance for the process design with required reliability.

The developed SPH model is used to simulate the spattering phenomenon in the L-PBF process, to overcome the limitation of commercial CFD packages, including their incapability of phase change and particle sticking phenomena, which are however commonly seen in the spattering process. The SPH model is capable to couple heat transfer, particle motion and phase change. The sticking phenomenon observed in the experiment is successfully reproduced by the SPH model using a similar scenario.

In summary, the modeling framework developed in this thesis can serve as a comprehensive tool for reliability-based design in the L-PBF process. The work is helpful for applying machine learning models in the additive manufacturing field.

1. INTRODUCTION

1.1 Background

Additive manufacturing (AM) is defined as a group of layer-upon-layer three-dimensional fabrication processes and previously applied for rapid prototyping. In the last decade, AM technology has rapidly evolved into a decent alternative for direct fabrication of tools and functional parts. Powder bed fusion (PBF) is an AM technique that uses a high-energy power source to selectively melt or sinter a metallic powder bed [1]. Due to the capability of producing parts with improved density, resolution, and surface finish, PBF tends to be the dominant AM technique for metallic materials [2]. Figure 1.1 shows the schematic of the L-PBF equipment [1]. In a typical PBF process, for each new layer which is usually in the range of 20-50 μm thickness, a recoating blade pushes a layer of fresh powder from the powder tank to the top of the previously built surface or the substrate. Then, a high-energy power source is fired along the designed path controlled by a computer-aided design (CAD) model. Metal powders are selectively sintered by the high-energy power source. This process is repeated in a layer-upon-layer manner until the product is printed. Finally, the unfused powders will be wiped off from the powder bed for recycle and the printed product will be ready for use or post-processing.

Based on the type of the energy source, PBF can be further divided into two major techniques [3]: (1) laser powder bed fusion (L-PBF), also known as selective laser sintering (SLS), selective laser melting (SLM) and direct metal laser sintering (DMLS), which uses a high-intensity laser beam as the power source, and (2) electron beam melting (EBM), which uses an electron beam. This thesis focuses on the L-PBF process.

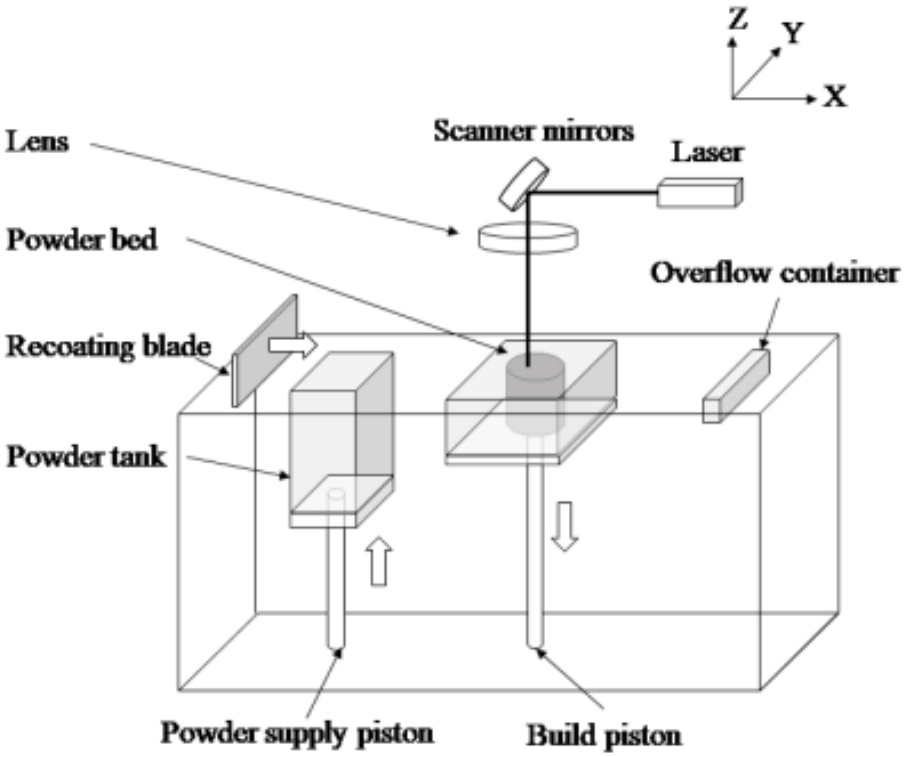


Figure 1.1. Schematic of the L-PBF equipment [1]

PBF technique has many advantages compared with conventional methods through machining. Firstly, take the advantage of its layer-by-layer nature, PBF allows fabricating complex metallic parts with internal and external features, which is not achievable in conventional manufacturing methods. Secondly, the material utilization rate in PBF process is nearly 100%, which is impossible for conventional subtraction-based methods. Thirdly, the time and labor cost will be greatly reduced when fabricating products that do not need mass production, e.g. customized products and rapid prototyping. With these advantages, PBF becomes a strong alternative method in fabricating metallic products.

1.2 Challenges in the L-PBF Process

Despite of the advantages, a common issue in the L-PBF process is the defect formation, which may result in the lack of robustness and stability of metal AM processes and limits the widespread applications of L-PBF in many industries [4], such as aerospace and medical. Figure 1.2 [5] and Figure 1.3 [6] illustrate two typical types of defects: pores and cracks. In L-PBF process,

the occurrence of the defects is highly dependent on numerous processing parameters [4], including laser power, laser scan speed, laser spot size, powder size, layer thickness, external pressure and material absorptivity, especially those parameters related to energy input. These process-structure-property (PSP) relationships have been discussed in many review articles [1, 6, 7].

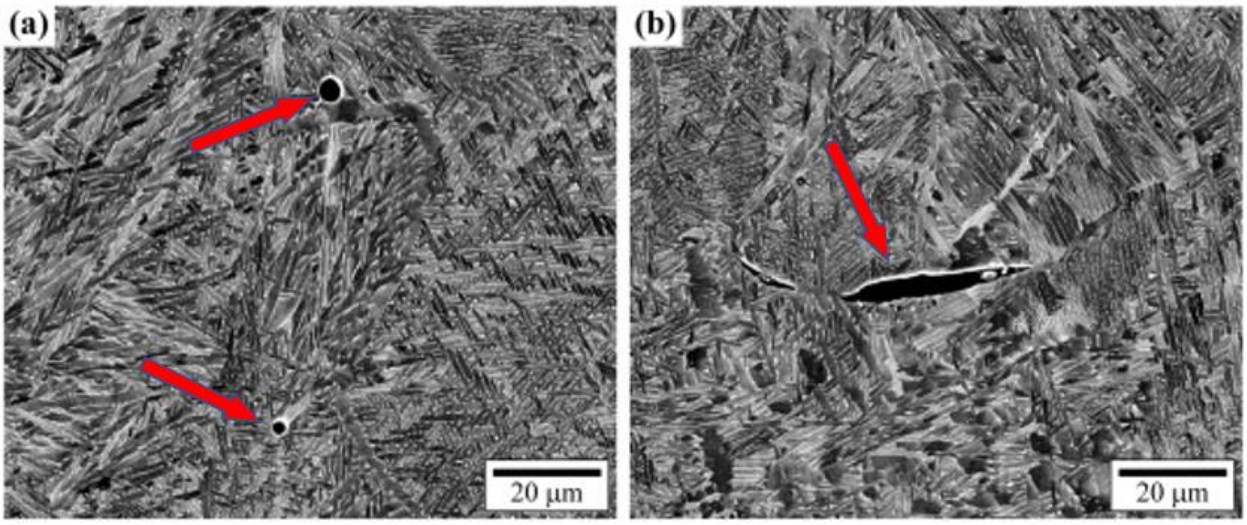


Figure 1.2. Examples of typical pores in PBF process: (a) two circular pores and (b) an acicular pore [5]

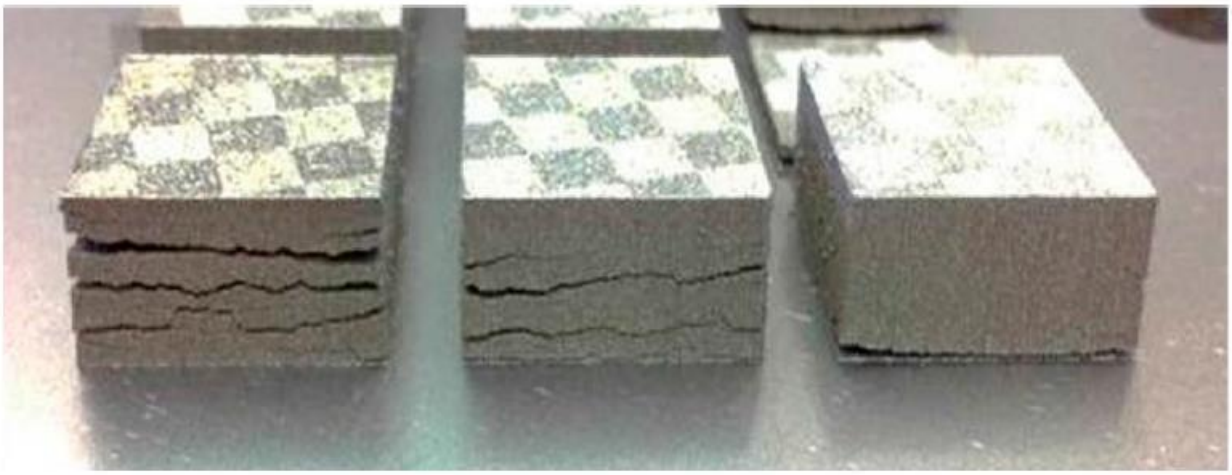


Figure 1.3. Examples of severe cracks in PBF process [6]

Forerunners have made lots of efforts in process optimization by experimental [8-13] and computational study [14-17]. Figure 1.4 [18] shows an image from an experimental study of L-

PBF process using high-speed x-ray *in situ* monitoring system. The formation of pore and spatters can be observed from this image. Figure 1.5 [19] shows sample result from a computation study of L-PBF process. The melt pool dynamics is simulated by this study. Through these experimental and computational studies, the two major causes of defects in the L-PBF process, keyhole mode and spattering, have been well exposed and understood. They are discussed detailedly in Section 1.2.1 and 1.2.2, respectively. Both defect sources are highly dependent to many processing parameters. Thus, the optimization of processing parameters is critical to avoid these phenomena and ensure the quality of the printed parts.

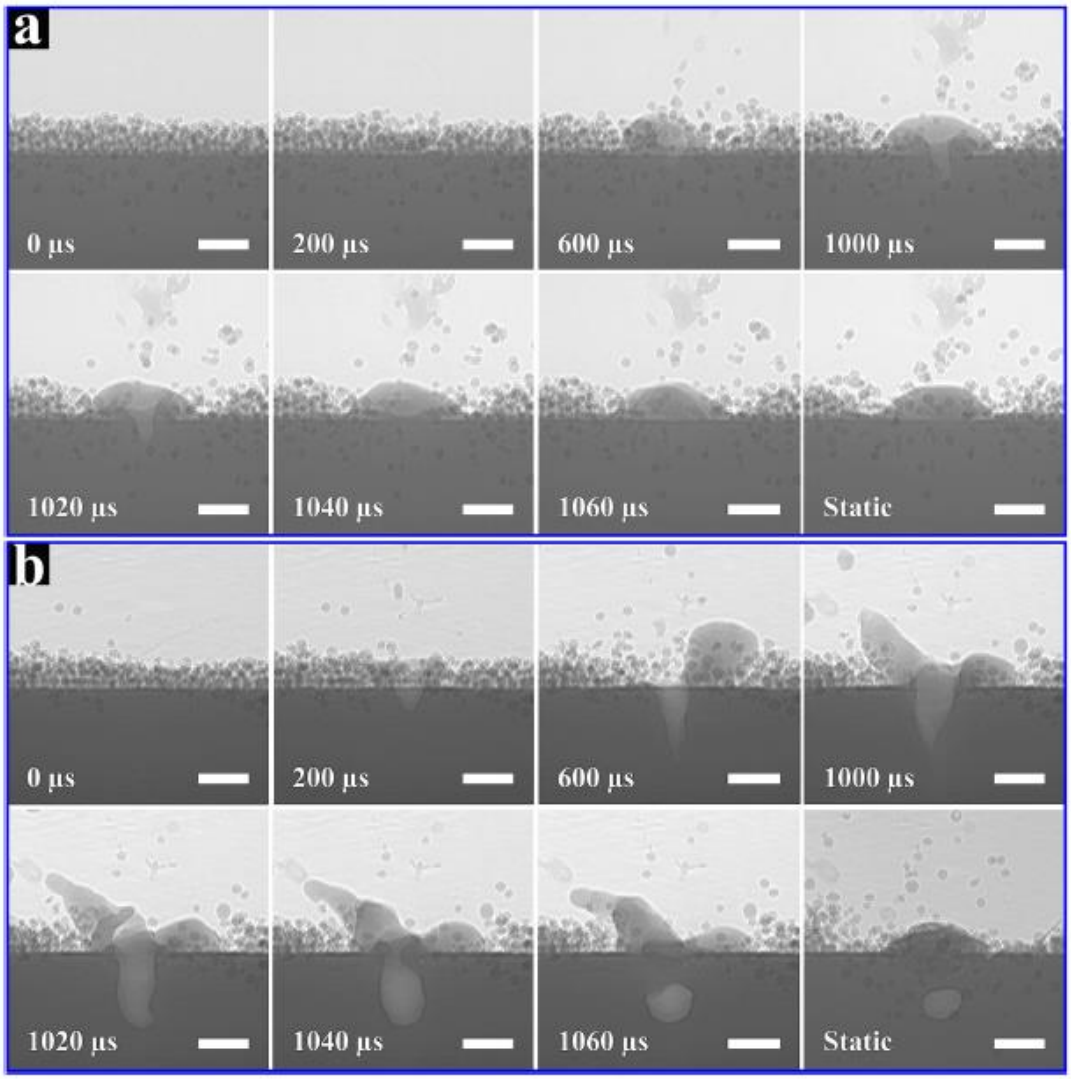


Figure 1.4. Image taken from high-speed x-ray *in situ* monitoring system during L-PBF process: (a) laser energy is 340 W and (b) laser energy is 520 W. Formation of spatters and pore can be observed [20]

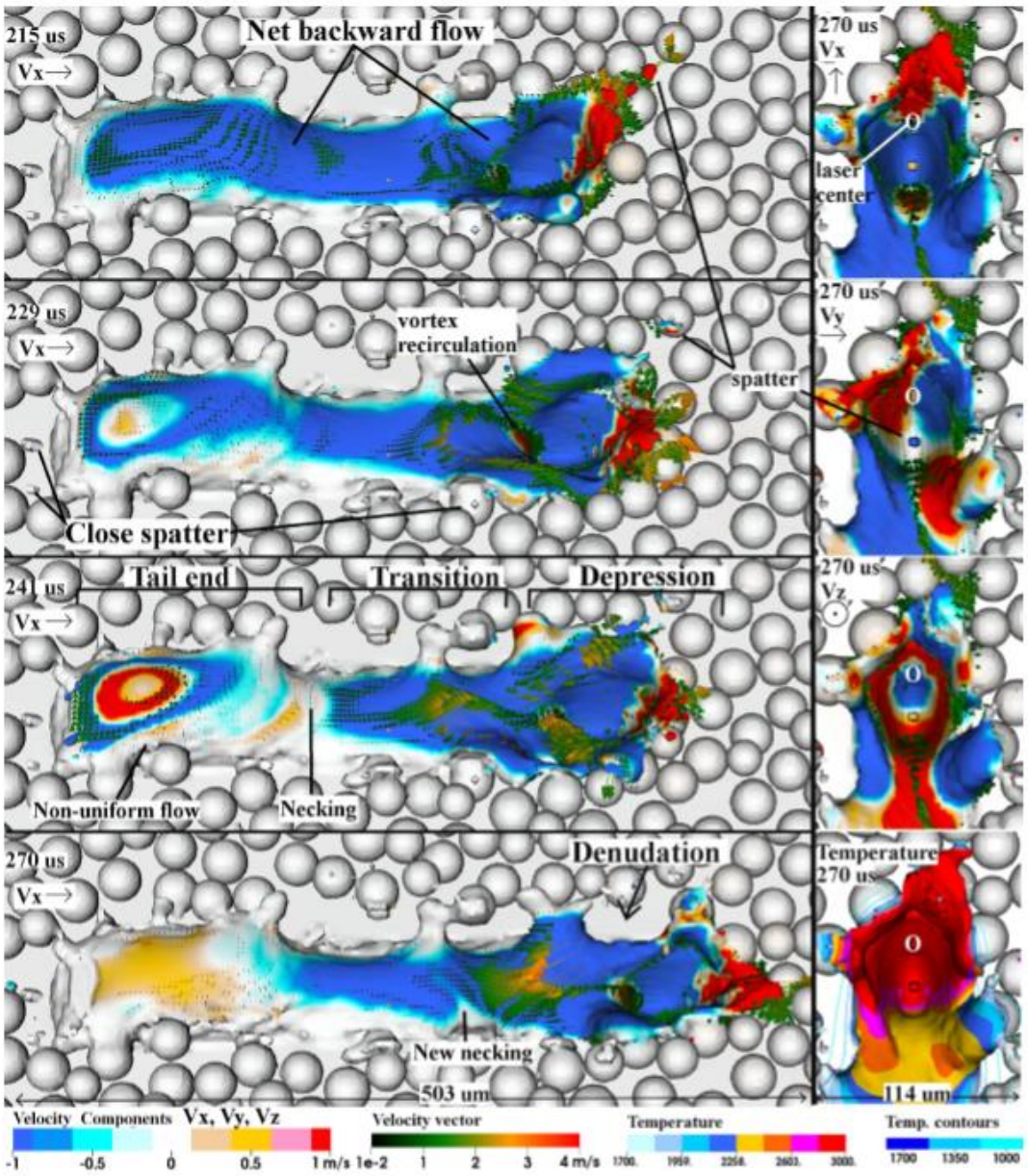


Figure 1.5. Sample result from a computational study of the melt pool dynamics of L-PBF process [19]

1.2.1 Keyhole Mode Identification

Keyhole mode is a typical phenomenon in L-PBF process. In laser processing field, there are two modes of heat transfer mechanisms related to defect formation: (1) conduction mode, in which the melting is controlled by heat conduction, and (2) keyhole mode, in which the input power density of the laser beam is sufficient to vaporize the metal and drill a much deeper cavity

than in conduction mode [21]. Since the collapse of the cavities in keyhole mode may leave voids in the printed parts [22], the conduction mode is preferred in laser additive manufacturing.

Since keyhole mode is a major cause of pore formation, identifying keyhole mode is critical to provide guidance on how to avoid its occurrence. According to previous experimental and computational studies [3, 10, 23, 24], the basic criterion of identifying the keyhole mode is

$$\frac{2d}{w} \geq 1 \tag{1.1}$$

where d is the remelted depth, w is the melt pool width, and the quantity, $\frac{d}{w}$, is the normalized depth. Besides this criterion, King *et al.* [10] presented another criterion for 316L stainless steel to identify the keyhole mode using normalized enthalpy:

$$\frac{\Delta H}{h_s} = \frac{AP}{\rho h_s \sqrt{\pi D v \left(\frac{\sigma}{4}\right)^3}} \geq 30 \tag{1.2}$$

where ΔH is the specific enthalpy, h_s is the enthalpy at melting, A is the material absorptivity of the laser power, P is the laser power, ρ is density, D is thermal diffusivity, v is the laser scan speed, σ is the laser spot size. This criterion enables designers to identify the preferred conduction mode in L-PBF conveniently. However, this is an empirical method proposed by Hann *et al.* [25]. By experimental results, they first found that the normalized depth is proportional to the product of power density, $P_d = \frac{P}{\pi\sigma^2}$, and square root of the laser interaction time, $\tau = \frac{2\sigma}{v}$. In other words,

$$\frac{d}{w} \propto P_d \sqrt{\tau}. \tag{1.3}$$

Then, the slope was determined by dimension analysis. Since this criterion is an empirical method, it is crucial to test it to other materials before applying it.

1.2.2 Spattering Phenomenon

Spattering is a physical phenomenon in laser powder bed fusion (L-PBF) process observed from experiments [26, 27] and is found to be the major cause of the structural defects in the printed products [28, 29]. Forerunners have made unremitting efforts in experiments to investigate the spattering phenomenon. Zhao *et al*. [18] applied high-speed synchrotron X-ray to monitor the L-PBF process of Ti-6Al-4V *in situ*, and found that some spatters tend to merge together and form larger particles. Andani *et al*. [30] used SEM images to quantify the spatters sizes and their distribution during PBF process with multi-laser technology. Through the comparison between

multi-laser and single laser scanning, they found that a higher number of working laser beams induces greater recoil pressure above the melting pools and thus more spatters are ejected from the molten pool. In addition, the spatter particles size is much larger than the raw powders. Large spatters tend to create an inclusion or remain as a non-melted region into the printed part, and thus degrade its quality. Guo *et al*. [31] used in-situ high-speed high-energy x-ray imaging to detailedly study the spattering mechanism and quantification as a function of time, environment pressure, and location in the L-PBF process. Under stationary laser impulse, as time goes by, the vaporization occurs after melting and generates intensive vapor jet which ejects the metal powders, and then argon gas flow forms and forces the particles surrounding the molten pool to move toward the molten pool. If the pressure of the surrounding argon gas is decreased, the intensity of the argon gas flow will also decrease, and thus powders are ejected with a large divergence angle as the vapor can expand freely. Figure 1.6 [31] Shows the schematic of the pressure and time dependent spattering mechanism. Regarding the effect of location, the experiment was conducted under moving laser beam at normal pressure. They found that particles behind the laser beam tend to move towards laser beam entrained by argon gas flow, particles around the laser beam tend to be ejected by the evaporation recoil pressure, and particles ahead of the laser beam tend to incline towards the laser beam. Yin *et al*. [32] applied a laser confocal microscope to characterize the morphological features of Inconel 718 sample. The correlation between the laser power and the ejection angle was revealed by their study. Overall, the spattering mechanisms and its effect are well exposed by experiments.

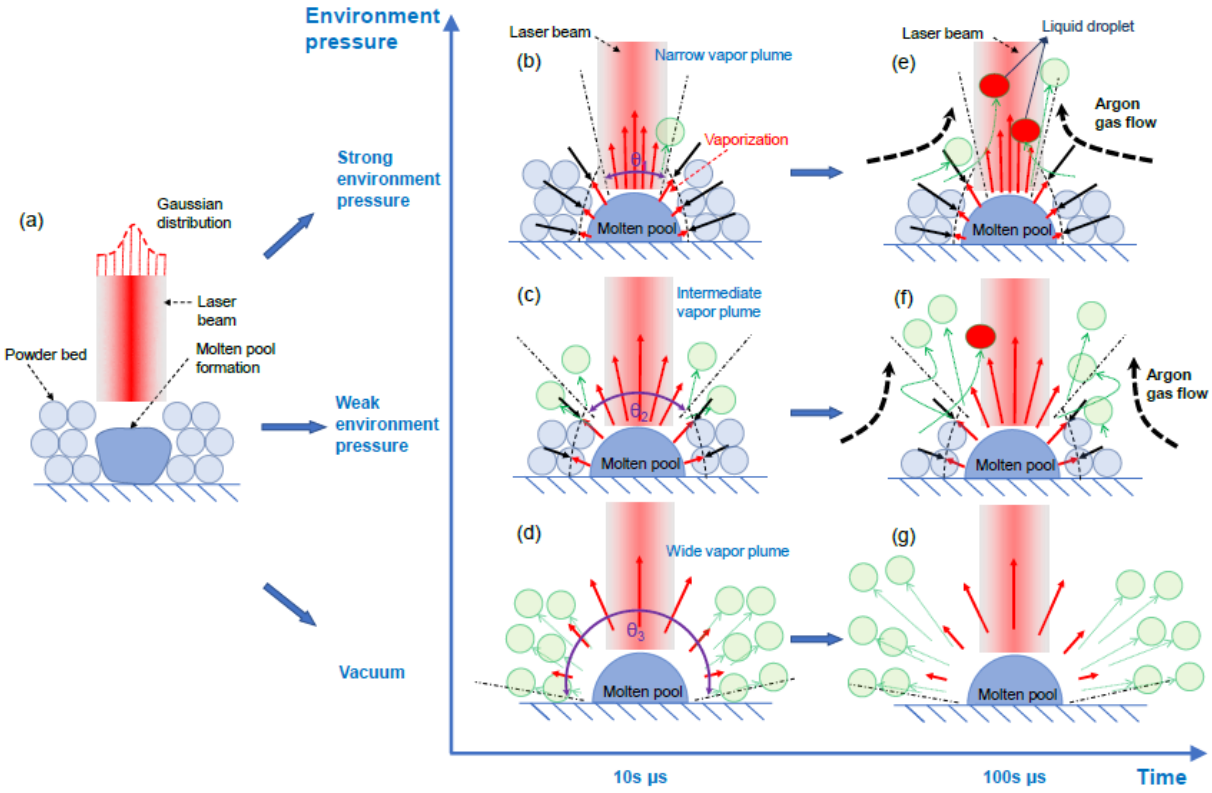


Figure 1.6. Schematic of the pressure and time dependent spattering mechanism [31]

Although the spattering mechanism is well exposed from experimental studies, simulation and modelling of the spattering phenomenon is still a challenge due to its multi-phase nature and complexity in physics. Pioneers have made efforts to simulate the L-PBF process using numerous computational models. These models range from nanoscopic model such as molecular dynamics (MD), and mesoscopic models such as computational fluid dynamics (CFD).

Molecular dynamics is a numerical method for studying the atomic motion. The governing equation in MD model is the equations of motion and the Newton's second law. As long as the atomic force between any two atoms with respect to their distance is well defined by the pair-wise potential function, the atomic motion can be solved by MD model without any other information such as the thermal conductivity or specific heat. MD model is not a popular method for L-PBF simulation due to its nanoscopic nature. As typical powder size for L-PBF is dozens of microns, each metal particle contains ~10^{22} atoms, which is beyond the computational limitation of MD model. Researchers have utilized it to investigate the sintering behavior for nanoparticles [33-36]. For L-PBF, to achieve simulation using MD, metal particles are assumed to be scaled down to

nanoscale. Yi Zhang *et al*. [17] applied different constant uniform temperature field to represent different laser power in their MD simulation of the sintering process of 41 well stacked nickel particles, and concluded that higher laser power will lead to higher resultant tensile strength. Yue Zhang *et al*. [1] applied non-translational kinetic energy to a moving laser spot using the "fix heat" method in LAMMPS to represent the laser energy in their MD simulation of $Cu_{50}Zr_{50}$ metallic glass in a simulated powder bed of 5*5 particles with a fix power density of 15.4 J/mm^2. The multi-layer sintering process was successfully simulated by their model. Despite of these efforts by forerunners, even under the scaling assumption, the scope of spattering phenomenon is still beyond the simulation capability of MD, due to its physical complexity.

CFD is a class of numerical methods aiming at solving problems involve fluid flows. As the L-PBF process involves molten pool dynamics, CFD models are frequently applied to simulate the sintering process in L-PBF. The basic governing equation in conventional CFD models is the Navier–Stokes (NS) equations, which is usually computationally intensive. Lattice Boltzmann method (LBM), on this other hand, solve the discrete Boltzmann equation instead, and can easily consume on the order of 10,000 CPU hours taking the advantage of its massively parallel architectures. Lee *et al*. [37] used the NS equations to simulate the L-PBF process to investigate the balling formation, but evaporation induced recoil pressure is not included in their simulation. The Marangoni effect and the recoil pressure is included in Yi Zhang's model [3] using the NS equations, but spattering phenomenon did not occur in his result. As typical simulation using the NS equation for a 1 mm laser scan take more than 20 hours [3], researchers tend to use the LBM in recent study [38, 39]. Khairallah *et al*. [40] developed a numerical simulation software tool, ALE3D, based on LBM and the arbitrary Lagrangian-Eulerian (ALE) techniques, and applied it to simulate the spattering phenomenon in L-PBF process [19, 41, 42]. Though liquid-phase spatters driven by the recoil pressure can be observed from their simulation results, solid particles are stationary and cannot be ejected in CFD models. In addition, the recoil pressure in all the simulations mentioned above is simulated by adding a boundary condition that reflects its effect, instead of simulating the vapor expansion. In this regard, Bidare *et al*. [43] simulated the argon gas flow induced by the metal vapor using a 2D finite element analysis (FEA) model, but powder bed and molten pool dynamics were not included in their model, as it's not a CFD based model. Overall, though CFD models are suitable to simulate the molten pool dynamics, the fact that the

solid metal powders can only be stationary in CFD model limits its application in simulating the spattering behavior in L-PBF process.

1.2.3 Surrogate Modeling

So far, the two major causes of defects formation in the L-PBF process and their current research status are reviewed. However, these experimental and computational studies are not sufficient for design and optimization purpose, due to their high time cost. Thus, a surrogate model is desirable to alleviate the needs of experiments and simulations.

For designers, on one hand, experiments and testing are helpful to determine the appropriate range of the processing parameters but are also time consuming and expensive. On the other hand, the modeling and simulation tools enable designers to make predictions of the printed products of L-PBF so as to optimize the processing parameters and decrease the need of conducting experiments. In addition, those mesoscopic simulation models are extremely helpful in understanding the physical meaning of the process. Despite these advantages, models with high accuracy are usually computationally intensive. It is impractical to optimize the processing parameters using these simulation models, especially when the number of involved processing parameters is large.

An alternative solution of this dilemma is the application machine learning (ML) models, a subset of artificial intelligence (AI). With a reliable training dataset, the ML models learn knowledge from the training set and make inference based on the knowledge. On one hand, the trained machine learning models can make predictions as a surrogate of experiments and simulations and determine the optimal processing parameters in an efficient way. On the other hand, it can also deal with *in situ* data for defect detection in real time. In general, the ML applications can be regarded as the art of data manipulation. This capability makes ML a key aspect of Industry 4.0 [44].

In general, machine learning tasks can be divided into three groups, supervised learning, unsupervised learning, and reinforcement learning. In AM field, the type of applied machine learning techniques is usually the supervised learning, in which each input is labeled with an output, and predictions are made from an unobserved input to an output based on example input-output pairs. The input is a vector contains all involved processing parameters, and the output can be a classification such as quality assessment, or a post-processing parameter such as porosity or

penetration depth. As the number of observation data points in PBF field is limited by experimental cost or computational intensity, a key challenge in the application of machine learning models in PBF field is how to achieve accurate predictions using limited training set.

Though machine learning models are powerful in making predictions and optimizing the processing parameters, a training dataset from either experiments or high-fidelity simulations is still essential. Therefore, developing high-fidelity simulation modelling that can provide physical explanation to the phenomena occurred during the L-PBF process and make predictions to the printed products is still critical. Thus, a modeling framework combining the physical models and machine learning models is desirable to address the design problem.

1.2.4 Probabilistic Process Design

So far, all the above discussions are based on an uncertainty free environment. However, uncertainty generally exists in the L-PBF process, and is the major cause of the process inconsistency, unpredictability and unrepeatability. For example, in a process that the laser power is set to be 100W, the outcome of the laser power may be in a range of 90-110W due to natural variability and printer inaccuracy. As a result, the quality of this process may not be accurately predicted if 100W is used as the input of laser power to the physical models and machine learning models. Instead, the distribution of the laser power should be used as input. In fact, the uncertainty of all input parameters should be taken into account, and the output quantity such as quality or mechanical properties of the products will be in a distribution form. This is the general task in probabilistic process design.

To mitigate the effect of uncertainty in engineering field, there are generally two methods: (1) robust design and (2) reliability-based design. A robust design focuses on diminishing the sensitivity of the output quantity to the uncertainty, thus making the process consistent. However, this sensitivity mainly depends on the process itself. Some fabricating methods may be insensitive to uncertainty. Since we only focus on the L-PBF process, the robust design method cannot be applied. A reliability-based design focuses on reliability, defined as the probability of the output quantity satisfying the design requirements, or factor of safety greater than 1. In this regard, a reliability contour map in terms of the design variables is desirable for designers to look up. Monte Carlo simulation (MCS), which is a numerical algorithm that relies on repeated random sampling,

is easy to acquire such a reliability contour map with high accuracy but extremely inefficient. A more efficient method for reliability-based design is in urgent need.

1.3 Problem Statement

A longstanding barrier in the L-PBF process is the gap between knowledge and design. To date, the mechanisms of laser melting process have been extensively studied from experiments and simulations, which provide the guidance for selecting processing parameters in P-S-P relation chain. However, for the well understood keyhole mode, the basic criterion in Eq.(1.1) is a post identification method, which can only be used in a trial-and-error manner. Even with the convenient normalized enthalpy criterion in Eq. (1.2) which can make predictions, it still needs verification for materials other than 316L stainless steel. In addition, with uncertainty in the processing parameters, the criterion becomes inaccurate. With sufficient knowledge, the probabilistic design problem is still in need of a comprehensive modeling framework to address.

In this thesis, a modeling framework that span physical models, machine learning models, and probabilistic models will be developed to address the probabilistic design problem in the L-PBF process, shown in Figure 1.7. The keyhole mode and spattering phenomena are simulated by physical models including CFD and SPH. The data acquired by the physical models are then used as training data for machine learning models. A guide for machine learning applications in the L-PBF field including taxonomy, model selection, model assessment, as well as methods to monitor and avoid overfitting is provided. A Gaussian process-based machine learning model is developed for demonstration of an application in the L-PBF field in calibration and verification of the normalized enthalpy criterion for keyhole model identification. Finally, an efficient probabilistic model integrating the MCS and the first order reliability method (FORM) is proposed to mitigate the effect of uncertainty in the reliability-based design. With a design requirement generated by machine learning models, the model can provide the feasible design region, within which the design requirement will be satisfied with a probability higher or equal to the specified reliability.

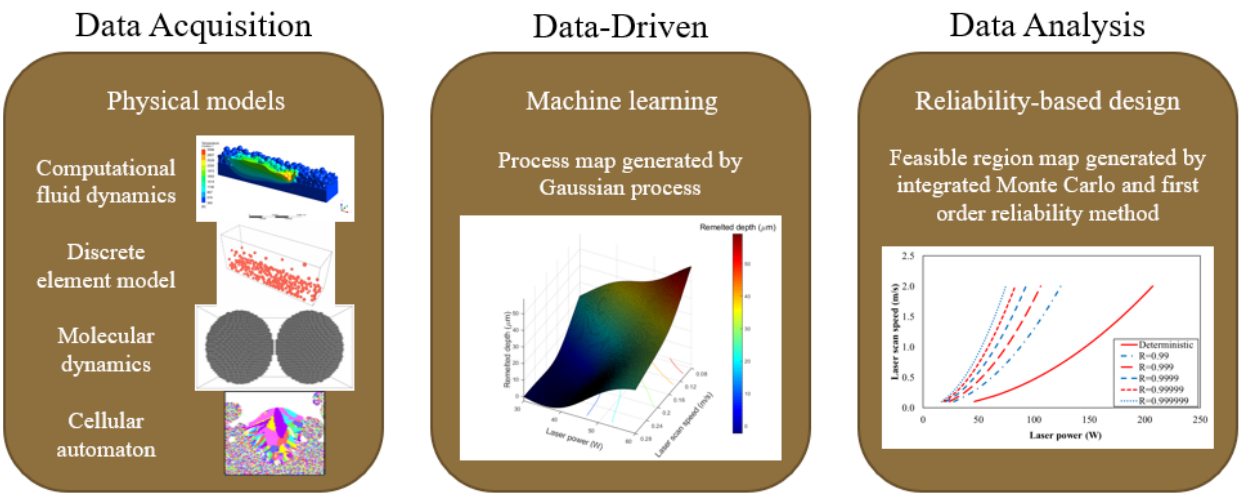


Figure 1.7. Probabilistic design framework for reliability-based design in the L-PBF process

1.4 Aims and Objectives of the Research

The aim of this thesis is to develop a probabilistic design framework for reliability-based design in the L-PBF process. The framework can serve as a one-stop tool for probabilistic design and optimization for the future L-PBF process.

The specific objectives include:

1. Development of a physical-based computational model that can simulate the spattering phenomenon, provide guidance to avoid it and generate reliable training dataset for machine learning model. A SPH model is developed to address this objective.
2. Development of a machine learning model that is capable of make predictions at any given settings within the design space after trained by reliable dataset, in order to enhance process optimization. A Gaussian process regression model is developed to achieve this objective.
3. Validation of the physical-based computational model and the machine learning model by comparison with experimental measurements.
4. Development of an efficient probabilistic model that can predict the feasible design region within the design space. The design requirement will be satisfied within the feasible design region with a probability higher or equal to the specified reliability. An integrated MCS and FORM model is developed for the reliability-based design.

1.5 Structure of the Thesis

This thesis is organized as follows:

Chapter 1 illustrates the background of laser powder bed fusion and presents the research problem, literature review, and the objectives of this thesis.

Chapter 2 proposes a new feasibility of simulating the spattering phenomenon in the L-PBF process, in order to provide more physical insight in the process and guidance of avoiding spatters. A smoothed particle hydrodynamics (SPH) model is developed in this thesis to simulate the spattering phenomenon. The simulation includes the interaction between the metal vapor and the metal powders, the melting and solidification of the metal powders and the sticking phenomenon observed in the experiment. The model is validated by a similar scenario in the experiment.

Chapter 3 provides a comprehensive guide for ML applications in AM field. The available data in the P-S-P relation chain with their corresponding machine learning applications are presented. The existing applications in literature are classified into different types of ML tasks, including regression, classification, and clustering. Model selections based on ML task, data type, size and dimension are suggested. Model assessment methods for classification task are presented. Methods monitoring and solving the overfitting issue in supervised learning are discussed and assessed. Future research directions related to machine learning applications are suggested.

Chapter 4 elaborates a Gaussian process (GP) regression model to assist the design and optimization for the L-PBF process. After trained by data acquired from a series of CFD simulations, the model can serve as a surrogate model that generate a regression surface and make predictions at any combinations of the input parameters in a very efficient manner. The predictions are compared against experimental data for validation of the GP model. The applications of the machine learning models in process optimization are demonstrated by calibration and verification of the convenient normalized enthalpy criterion for 316L and 17-4 PH stainless steel.

Chapter 5 proposes a novel algorithm by integrating the Monte Carlo simulation and first order reliability method for probabilistic process design. The model efficiently determines a feasible region of the design space for given design requirements at specified reliability levels. If a design point falls into the feasible region, the design requirement will be satisfied with a probability higher or equal to the specified reliability.

Chapter 6 presents the final conclusions and recommended future work directions.

2. SIMULATION OF THE SPATTERS STICKING PHENOMENON IN LASER POWDER BED FUSION PROCESS USING THE SMOOTHED PARTICLE HYDRODYNAMICS METHOD

Abstract: In this chapter, a smoothed particle hydrodynamics (SPH) method is developed to simulate the spattering phenomenon in the laser powder bed fusion (L-PBF) process. First, an experiment using the high-speed synchrotron X-ray full-field imaging is conducted to acquire in-situ images during the L-PBF process. Then, a scenario is selected from the X-ray image as a case study of the SPH model. In the case study, a particle is ejected and melted by the metal vapor, impacts with another particle, solidifies, and sticks to the other particle to form a rigid body. As a result, the trajectories of the two particles match well with the experimental observation. The evolution of velocity and temperature of the particle is extracted from the simulation for analysis. The SPH model can be an excellent alternative of computational models of simulating the spattering phenomenon of L-PBF and provide guidance of avoiding it.

2.1 Introduction

Laser powder bed fusion (L-PBF) is a popular additive manufacturing (AM) technique for metallic materials, due to the capability of producing parts with improved density, resolution, and surface finish that require less post-processing compared to other AM processes such as binder jetting [2]. In a typical L-PBF process, a recoating blade first pushes a layer of fresh powder from the powder tank to the top of the previously built surface or the substrate, which forms the powder bed. Then, a laser beam passes through a system of lenses and is reflected by a mirror that controls the laser beam spot moving along the designed path to selectively sinter the metal powders. This process is repeated in a layer-by-layer manner.

Spattering is a physical phenomenon in the L-PBF process observed from experiments [26, 27] and is found to be one of the major causes of the structural defects in the printed products [28, 29].. To alleviate it, forerunners have made unremitting efforts in experiments to investigate the spattering phenomenon. Zhao *et al*. [18] applied high-speed synchrotron X-ray to monitor the L-PBF process of Ti-6Al-4V *in situ*, and found that some spatters tend to merge together and form larger particles. Andani *et al*. [30] used SEM images to quantify the spatters sizes and their distribution during PBF process with multi-laser technology. Through the comparison between

multi-laser and single laser scanning, they found that a higher number of working laser beams induces greater recoil pressure above the melting pools and thus more spatters are ejected from the molten pool. In addition, the spatter particles size was found to be much larger than the raw powders. Large spatters tend to create an inclusion or remain as a non-melted region into the printed part, and thus degrade its quality [30]. Guo *et al*. [31] used in-situ high-speed high-energy x-ray imaging to detailedly study the spattering mechanism and quantification as a function of time, environment pressure, and location in the L-PBF process. Under stationary laser impulse, as time goes by, the vaporization occurs after melting and generates an intensive vapor jet which ejects the metal powders, and then argon gas flow forms and forces the particles surrounding the molten pool to move toward the molten pool. If the pressure of the surrounding argon gas is decreased, the intensity of the argon gas flow will also decrease, and thus powders are ejected with a large divergence angle as the vapor can expand freely. Regarding the effect of location, the experiment was conducted under moving laser beam at normal pressure. They found that particles behind the laser beam tend to move towards laser beam entrained by argon gas flow, particles around the laser beam tend to be ejected by the evaporation recoil pressure, and particles ahead of the laser beam tend to incline towards the laser beam. Yin *et al*. [32] applied a laser confocal microscope to characterize the morphological features of Inconel 718 sample. The correlation between the laser power and the ejection angle was revealed by their study. Overall, the spattering mechanisms and its effect are well exposed by experiments.

In spite of the experimental efforts, simulation and modelling of the spattering phenomenon is still a challenge due to its multi-phase nature and complexity in physics. Pioneers have made efforts to simulate the L-PBF process using numerous computational models. These models range from nanoscopic model such as molecular dynamics (MD), and mesoscopic models such as computational fluid dynamics (CFD).

MD is a numerical method for studying the atomic motion and has been used to investigate the sintering behavior for nanoparticles [33-36]. For L-PBF, to achieve simulation using MD, metal particles are assumed to be scaled down to nanoscale. Yi Zhang *et al*. [17] applied different constant uniform temperature field to represent different laser power in their MD simulation of the sintering process of 41 well stacked nickel particles, and concluded that higher laser power will lead to higher resultant tensile strength. Yue Zhang *et al*. [1] applied non-translational kinetic energy to a moving laser spot using the “fix heat” method in LAMMPS to represent the laser

energy in their MD simulation of $Cu_{50}Zr_{50}$ metallic glass in a simulated powder bed of 5*5 particles with a fix power density of 15.4 J/mm^2. The multi-layer sintering process was successfully simulated by their model. Despite of these efforts by forerunners, spattering phenomenon was not simulated by their MD models. Even under the scaling assumption, the scope of spattering phenomenon is still beyond the simulation capability of MD, due to its physical complexity.

CFD is a class of numerical methods aiming at solving problems involve fluid flows. As the L-PBF process involves molten pool dynamics, CFD models are frequently applied to simulate the sintering process in L-PBF. Khairallah *et al*. [40] developed a numerical simulation software tool, ALE3D, based on LBM and the arbitrary Lagrangian-Eulerian (ALE) techniques, and applied it to simulate the spattering phenomenon in L-PBF process [19, 41, 42]. Though liquid-phase spatters driven by the recoil pressure can be observed from their simulation results, solid particles are stationary and cannot be ejected in CFD models. In addition, the recoil pressure in all the simulations mentioned above is simulated by adding a boundary condition that reflects its effect, instead of simulating the vapor expansion. In this regard, Bidare *et al*. [43] simulated the argon gas flow induced by the metal vapor using a 2D finite element analysis (FEA) model, but powder bed and molten pool dynamics were not included in their model, as it's not a CFD based model. Overall, though CFD models are suitable to simulate the molten pool dynamics, the fact that the solid metal powders can only be stationary in CFD model limits its application in simulating the spattering behavior in L-PBF process.

This thesis proposes a new feasibility of simulating the spattering phenomenon in the L-PBF process, in order to provide more physical insight in the process and guidance of avoiding spatters. A smoothed particle hydrodynamics (SPH) model is developed in this chapter to simulate the spattering phenomenon. The simulation includes the interaction between the metal vapor and the spatters, the melting and solidification of the metal powders and the sticking phenomenon observed in the experiment. The model is validated by a similar scenario in the experiment.

The structure of the rest of this chapter is organized as follows. Section 2.2 elaborates the SPH model and the experiment done by the Argonne National Laboratory. Section 2.3 focuses on a case study with similar configuration in the experiment. The simulation result is compared with the experimental result for validation. Section 2.4 summarizes the chapter and provides the future work suggestions.

2.2 Methodology

This chapter develops a SPH model to simulate the spatter sticking phenomenon observed from the experiment. For simplification, two particles from the experiment are selected to be simulated, and their trajectories from the simulation are then compared with the experimental results for model validation. In this section, the theory of the SPH model is briefly introduced in Section 2.2.1. Then, the experiment setting is described in Section 2.2.2.

2.2.1 Smoothed Particle Hydrodynamics Model

SPH is a method to solve problems in Lagrangian continuum mechanics. It was developed by Gingold and Monaghan [45] and Lucy [46] in 1977, initially for astrophysical problems. Then, it was extended to solid mechanics by Libersky and Petschek [47, 48].

The governing equation of SPH fluid is the Euler equations of mass conservation and momentum balance:

$$\frac{d\rho}{dt} = -\rho\nabla\cdot\mathbf{v}; \tag{2.1}$$

$$\frac{d\mathbf{v}}{dt} = -\frac{1}{\rho}\nabla\cdot P; \tag{2.2}$$

$$\frac{de}{dt} = -\frac{1}{\rho}P:\nabla\mathbf{v} - \frac{1}{\rho}\nabla\cdot Q, \tag{2.3}$$

where ρ is the density, t is the time, $\mathbf{v}$ is the velocity vector, P is the pressure tensor, e is the energy per unit mass, and the heat-flux vector $Q = \kappa\nabla T$, with thermal conductivity κ and temperature gradient ∇T. SPH interpolates the set of field variables by means of kernel interpolation [49]. For any variable $f(r)$, a local average at each coordinate $\mathbf{r_i}$ is calculated according to [49]

$$f(\mathbf{r_i}) = \sum_j m_j \frac{f_j}{\rho_j} W(\mathbf{r_i} - \mathbf{r_j}). \tag{2.4}$$

Here, m_j is the mass of particle j. f_j is the value of the field $f(r)$ at position $\mathbf{r_j}$. ρ_j is the value of the mass density at $\mathbf{r_j}$. $W(\mathbf{r_i} - \mathbf{r_j})$, denoted by W_{ij}, is the kernel function of compact support and decays to zero within a range h comparable to a few typical inter-particle spacings. Since m_j, f_j, ρ_j are particle properties and are not affected by the gradient operator ∇, the only term that is affected by ∇ is W_{ij}.

The governing equation of SPH solid is the total Lagrangian formulation given by [50]:

$$\rho J = \rho_0; \tag{2.5}$$

$$\ddot{\mathbf{u}} = \frac{1}{\rho_0}\nabla_0 \boldsymbol{P}^T; \tag{2.6}$$

$$\dot{e} = \frac{1}{\rho_0}\dot{\boldsymbol{F}}\colon \boldsymbol{P}. \tag{2.7}$$

Here, $\boldsymbol{P}$ is the first Piola-Kirchhoff stress tensor. The subscript 0 and the absence of the subscript indicate that a quantity is evaluated in the reference configuration and the current configuration, respectively. J is the determinant of the deformation gradient $\boldsymbol{F}$, given by [50]

$$F = \frac{d\mathbf{x}}{d\mathbf{X}} = \frac{d\mathbf{u}}{d\mathbf{X}} + I, \tag{2.8}$$

where $\mathbf{x}$ and $\mathbf{X}$ are the current coordinates and the reference coordinates, respectively, $\mathbf{u} = \mathbf{x} - \mathbf{X}$ is the displacement, and I is the diagonal unit matrix. Same as the SPH fluid, SPH solid also uses Eq. (4) to evaluate variables in terms of the reference coordinates. For more detailed introduction to the SPH model, the reader is referred to the excellent book by Liu [51] and the guides by Ganzenmüller [49, 50].

This work utilizes the USER-SPH and USER-SMD packages [49, 50] in the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) [52]. A missing capability in a typical SPH simulation is the phase change. In this work, a batch file in shell language is used to change the particle type to represent the melting and solidification. The melting is assumed to finish once the liquidus is reached, and the solidification is assumed to finish once the solidus is reached. In other words, the rate of melting and solidification is assumed to be infinity. The latent heat is added to the particle for solidification and deducted from the particle for melting. The particle motion is simulated by the USER-SMD package and the heat transfer is simulated by the USER-SPH package.

2.2.2 Experimental Setting

This work applies the high-speed synchrotron X-ray to monitor the L-PBF process. The high-speed synchrotron X-ray full-field imaging experiments are performed on beamline 32-ID-B at the Advanced Photon Source (APS), Argonne National Laboratory [13]. The recording rate is 50000 Hz. The laser system consists of an ytterbium fiber laser source (IPG YLR-500-AC, IPG Photonics, Oxford, Massachusetts, USA) integrated with a laser scanner (Intelli*SCAN*$_{\text{de}}$ 30, SCANLAB GmbH, Puchheim, Germany). The reader is referred to Ref. [13] for detailed information of the experimental apparatus. The Inconel 718 powders fabricated by the Praxair Inc. is used as the material.

A sample high-speed X-ray image acquired from the experiment is shown in Figure 2.1. A lot of powders ejected from the molten pool can be observed (Figure 2.1(b)). Among these ejected powders, some are melted and merged together to form large spatters (Figure 2.1(a)), and some stick with other powders (Figure 2.1(d)). The SPH model developed in this work focuses on simulating the sticking phenomenon observed from the experiment, in order to demonstrate its capability of simulating the phase change, heat transfer and particle motion in the L-PBF process. Particle A and B are selected from this image for simulation, as they will stick with each other after several frames.

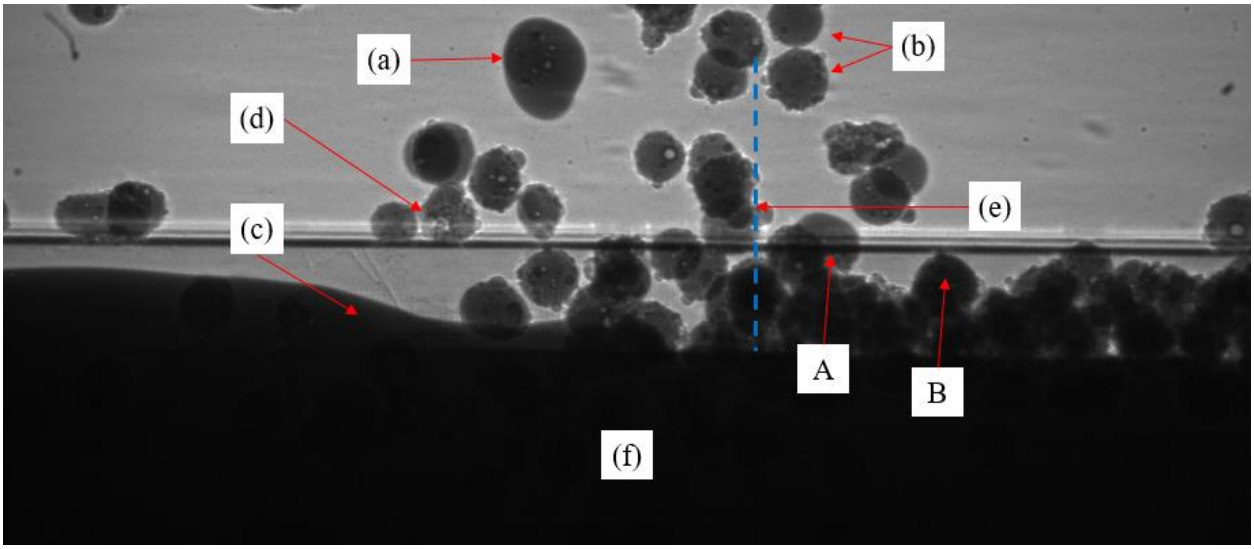


Figure 2.1. A sample high-speed X-ray image. (a) large spatter or molten metal ejected from the melt pool, (b) ejected powder, (c) molten pool, (d) stuck powders, (e) current laser scanning position, and (f) substrate. A and B represent to two particles selected from the experiment for simulation.

2.3 Case study

2.3.1 Initial Configuration

This work focuses on a 2-D case for simplification. The initial configuration is shown in Figure 2.2. The Open Visualization Tool (OVITO) [53] is used for visualization. Here, the yellow, red, and blue particles represent the substrate, metal vapor and metal powders. The parameters related to the initial magnitude and direction of the velocity of particle A are defined such that the configuration can resemble the typical scenario shown in Figure 2.1. These parameters including: powder diameter $d_p = 80$ μm, angle between vapor moving direction and horizontal line $\theta = 45°$,

vapor initial speed $v_i = 500$ m/s, initial center to center distance between particle A and B $d_i = 240$ μm, initial temperature of vapor $T_g = 3000$ K, initial temperature of powders $T_A = 1300$ K, $T_B = 300$ K, vapor density $\rho_g = 0.02$ g/cm^3, and gravity $F_g = 9.81$ N/kg. Since the parameters related to the velocity of particle A has been used as input in the simulation, only the interaction between particle A and B and their trajectories will be used for model validation. The material parameters of Inconel 718 used in this work are listed as follows [54]: density $\rho = 8.19$ g/cm^3, Young's modulus $E = 204.9$ GPa, Poisson's ratio $\nu = 0.3$, thermal diffusivity $\alpha = 5.0$ mm^2/s, thermal conductivity $k = 11.4$ W/m · K, liquidus $T_l = 1609$ K, solidus $T_s = 1533$ K. The air is not modeled in this simulation, as the diffusion of gas is difficult to simulate in a typical SPH model. Therefore, the effect of air resistance and air cooling is ignored in this work.

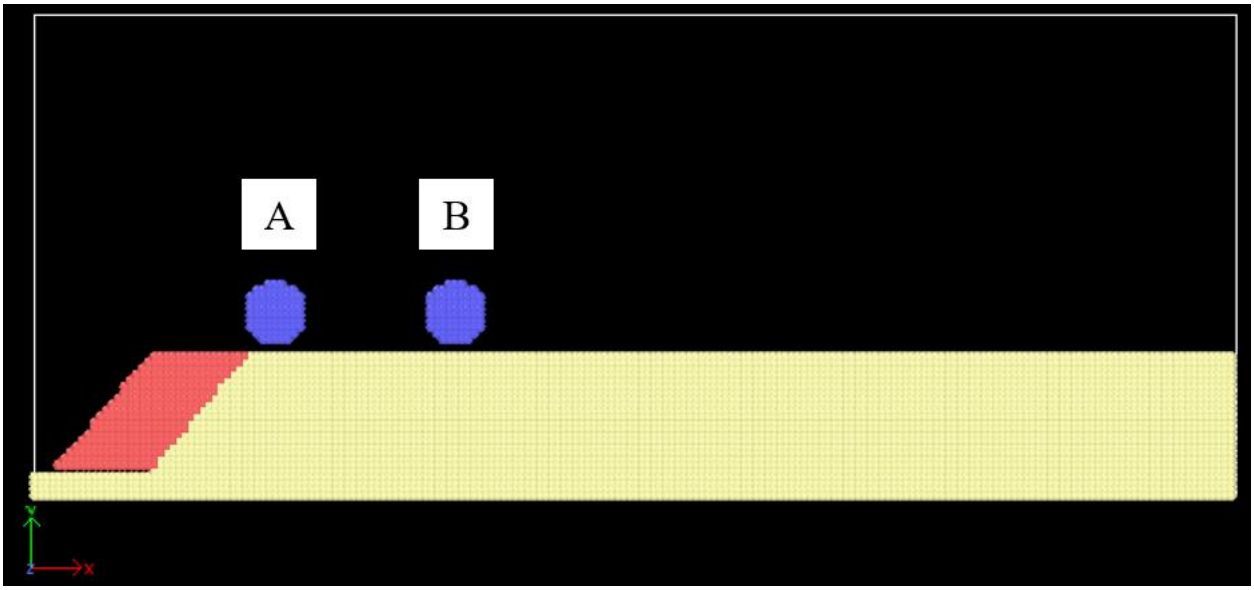


Figure 2.2. Initial configuration of the SPH model

2.3.2 Particle Trajectories

The trajectories of particle A and B of the scenario in Section 2.3.2 are displayed in Figure 2.3. The time at Figure 2.3(a) is set to be $t_0 = 0$ ms. At 0 ms, particle A is interacting with the metal vapor, acquiring both kinetic energy and internal energy from the vapor. At 0.08 ms, particle A is melted, and keeps moving towards particle B. At 0.14 ms, particle A gets in contact with particle B and starts heat diffusion. It can be observed that the shape of particle A has changed a little since it is in liquid phase. At 0.24 ms, particle A solidifies and sticks to particle B. From then

on, they become a rigid body. At 0.32 ms and 0.44 ms, they keep moving towards right with some angular velocity.

Note that the X-ray image is 2D. Therefore, some overlapped particles can be observed. However, they may have different distance from the camera, which means that they may not contact even though they seem overlapping. The overlapping phenomenon is severe at the powder bed, but the observation of particle A and B is not disturbed by the overlapping after particle A is ejected.

With the well-defined initial configuration which can retrieve the initial velocity of particle A, the SPH model can resemble the rest of the trajectories of both particles in the experiment. Thus, the matching result verities the validation of the SPH model, as well as its capability of simulating the sticking phenomenon in the L-PBF process.

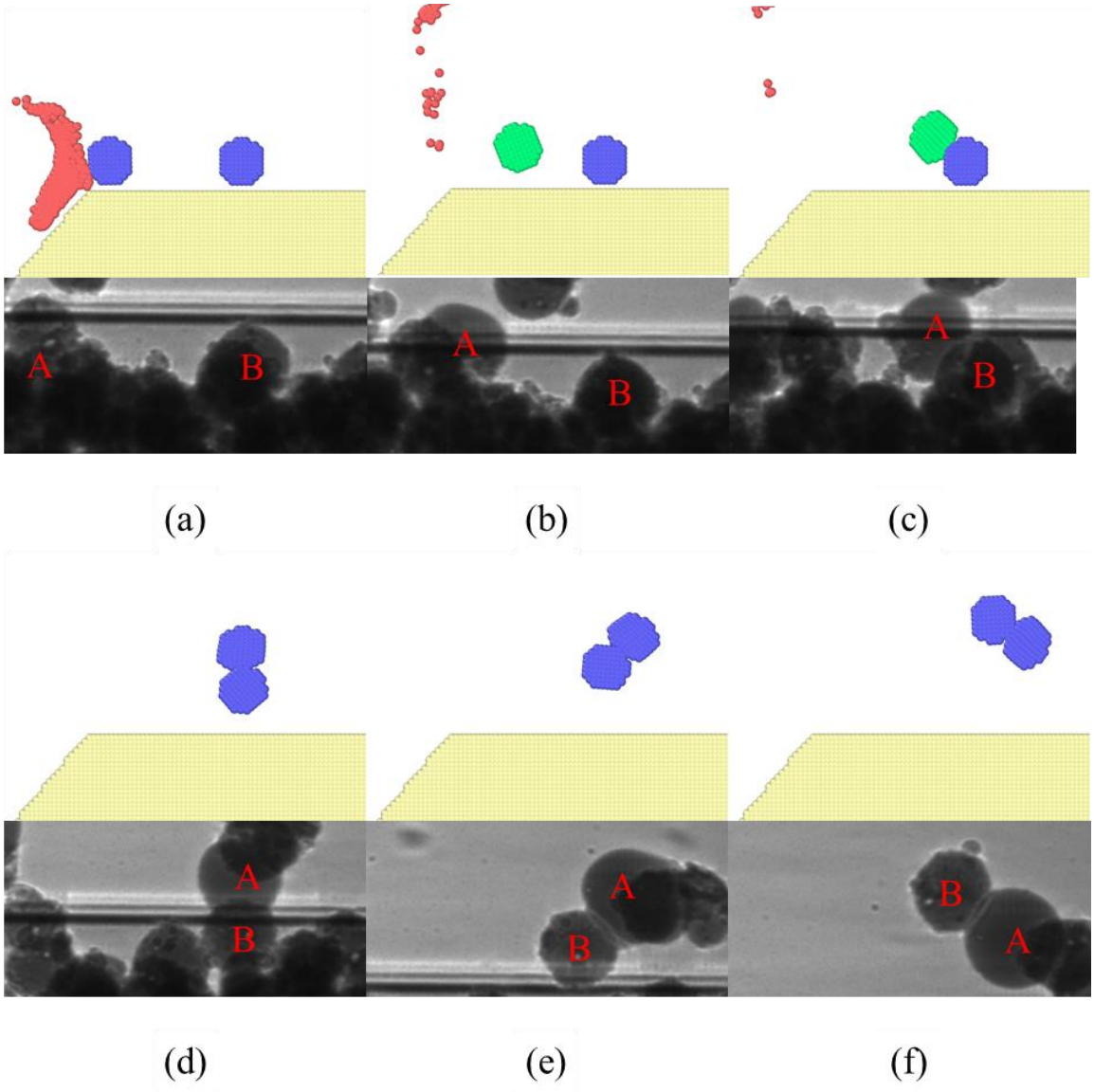


Figure 2.3. Trajectories of particle A and B in simulation (top) and experiment (bottom). Green particle means melted. (a) 0 ms, (b) 0.08 ms, (c) 0.14 ms, (d) 0.24 ms, (e) 0.32 ms, (f) 0.44 ms.

2.3.3 Velocity Evolution

The velocity of particle A during the simulation is plotted in Figure 2.4, with the same time domain in Section 2.3.2. At 0.02 ms, the interaction between the vapor and the particle ends, and therefore the particle keeps a constant velocity for several microseconds. At 0.14 ms, particle A impact with particle B, and a significant change on the velocity of particle A can be observed. From then on, particle A and particle B become a rigid body. Their motion is a combination of

translation and rotation. As a result, the velocity of particle A in the rest of the simulation changes like a sine function.

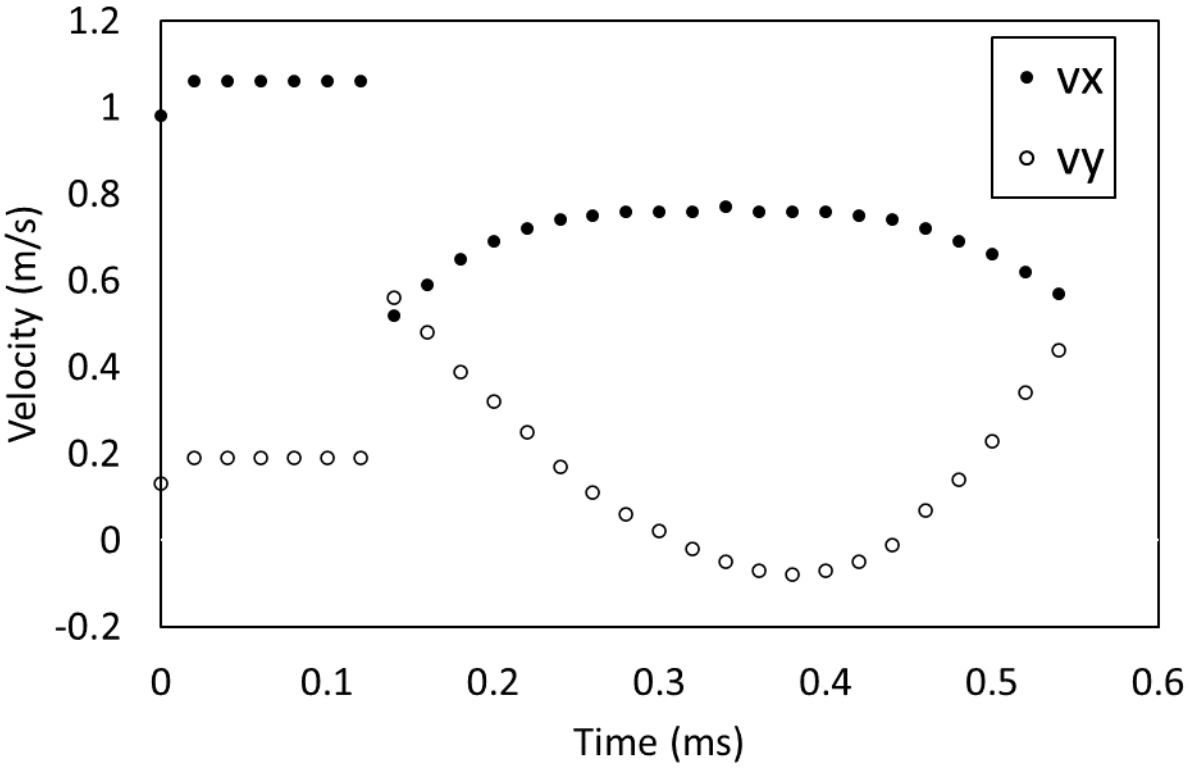


Figure 2.4. Evolution of velocity of particle A.

The effect of gravity cannot be obviously observed in the simulation, as the magnitude of time is in microsecond. This phenomenon has also been explained by Guo *et al.* [31] using equivalent pressure. However, according to their quantification of angular powder velocity profile driven by metal vapor, the air resistance plays an important role in the deacceleration of the spatters. This phenomenon is not captured in this simulation as the effect of air is ignored.

2.3.4 Temperature Evolution

Temperature evolution of the particles is difficult to be acquired from the experiment. Therefore, extracting temperature evolution is a desirable capability of computational models. Here, the average temperature of particle A is extracted from the simulation, shown in Figure 2.5.

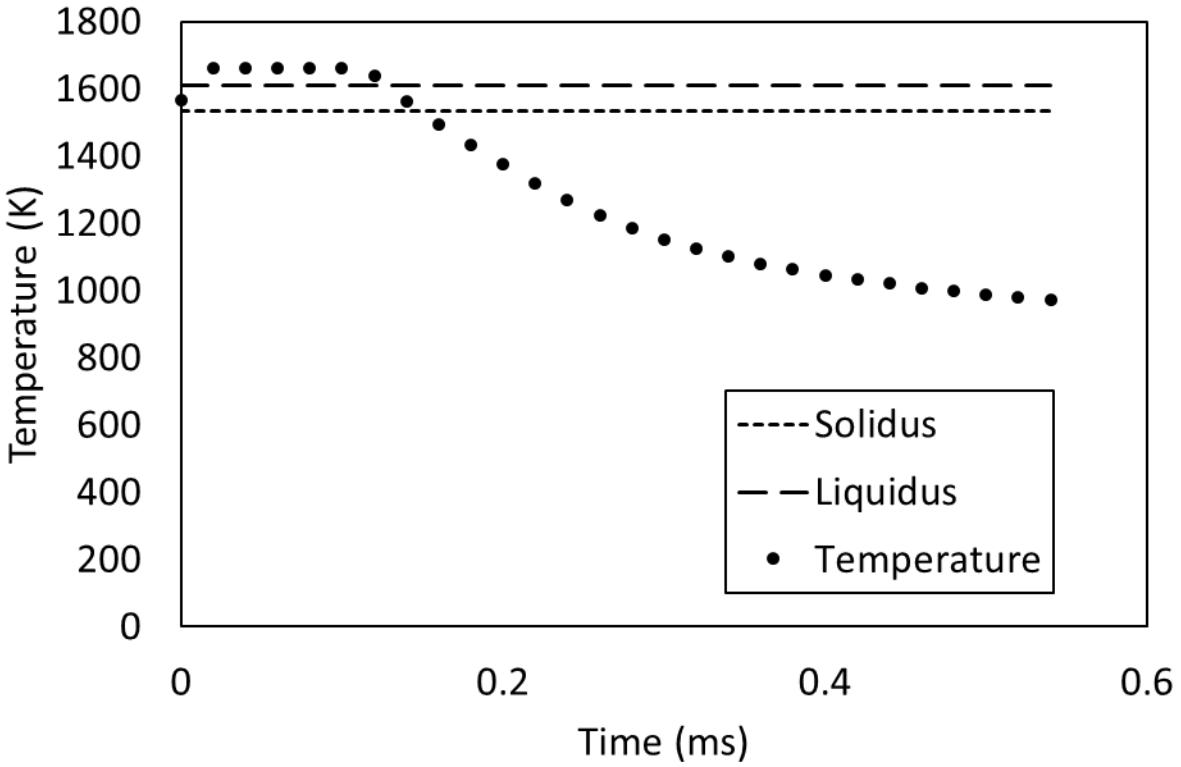


Figure 2.5. Evolution of the average temperature of particle A.

At 0 ms, the average temperature of particle A exceeds the solidus, but it does not reach the liquidus, and therefore the particle has not melted yet. After the interaction between metal vapor and particle A, the particle melts, and keeps a constant temperature as the air cooling is not captured in this simulation. Upon impacts with particle B, particle A cools down through heat diffusion, and therefore it solidifies right after the impact. After the solidification, particle A and B become a rigid body. In the rest of the simulation, the heat diffusion continues, with a decreasing cooling rate as the temperature difference between particle A and B decreases.

So far, the capability of simulating heat transfer, particle motion and phase change of the SPH model is demonstrated. The sticking phenomenon observed from the experiment is simulated. The trajectories of particles in the simulation match well with experiment. Though there are limitations in the model shown in this work, the SPH method can be an excellent alternate of the computational models used to simulate the spattering phenomenon.

2.4 Summary

In this chapter, the SPH method is proposed to simulate the spattering phenomenon in the L-PBF process. A model is developed to perform a case study with the same scenario as the experiment. The major finding is summarized as follows:

1. The SPH model can resemble the trajectories of the particles and simulate the sticking phenomenon observed in the experiment. The simulation result matches well with experimental result.
2. The L-PBF process involves physical phenomena including heat transfer, phase change and particle motion, which is difficult to be simulated in one computational model like CFD. The chapter demonstrates the capability of SPH in simulating these physical phenomena.
3. The phase change is done by the particle type change capability in LAMMPS when certain conditions (solidus and liquidus) are matched. In spite of the assumption of infinite rate of phase change, the SPH models with adjustable rate of phase change has been reported [55-57].

There are still some limitations in the model demonstrated in this chapter. These limitations include the missing effects of air resistance and air cooling, the difficulty in defining boundary conditions like fluid inlet and outlet, the assumed infinite rate of phase change, and the lack of the third dimension. Future work on improving the SPH model to overcome these limitations will be recommended.

3. A GUIDE FOR MACHINE LEARNING APPLICATIONS IN ADDITIVE MANUFACTURING

A version of this chapter has been published in *Journal of The Minerals, Metals & Materials Society* (doi: 10.1007/s11837-020-04155-y)

Abstract: In this chapter, the latest applications of machine learning (ML) in additive manufacturing (AM) field are reviewed. These applications, such as parameter optimization and anomaly detection, are classified into different types of ML tasks, including regression, classification, and clustering. The available data in the P-S-P relation chain with their corresponding machine learning applications are discussed. Model selections based on ML task, data type, size and dimension are suggested. Model assessment methods for classification task are presented. Methods monitoring and solving the overfitting issue in supervised learning are discussed and assessed. Future research directions related to machine learning applications are suggested.

Nomenclature	
Abbreviation	**Meaning**
3D	Three dimensional
AE	Acoustic emission
AI	Artificial intelligence
AM	Additive manufacturing
BJ	Binder jetting
BoW	Bag of words
BP	Backpropagation
CNN	Convolutional neural network
CT	Computed tomography
CV	Cross-validation
DA	Discriminant Analysis
DBN	Deep belief network
DED	Direct energy deposition
DT	Decision tree
FFF	Fused filament fabrication
FN	False negative
FP	False positive

Nomenclature continued	
GP	Gaussian process
KNN	k-nearest neighbors
LOOCV	Leave-one-out cross-validation
L-PBF	Laser powder bed fusion
ME	Material extrusion
MJ	Material jetting
ML	Machine learning
NN	Neural network
PBF	Powder bed fusion
PCA	Principal component analysis
PSP	Process-structure-property
RF	Random forest
RMSE	Root mean square error
RT	Regression tree
SL	Sheet lamination
SOM	Self-organizing map
SVM	Support vector machine
TN	True negative
TP	True positive
UQ	Uncertainty Quantification

3.1 Introduction

Machine learning (ML), a subset of artificial intelligence (AI), has increasingly become popular in additive manufacturing (AM) research. Additive manufacturing is defined as a group of layer-upon-layer fabrication processes controlled by a computer-aided design (CAD) model [58, 59]. Machine learning is defined as computer programming to optimize a performance criterion using example data or past experience [60]. For machine learning in additive manufacturing, besides the typical application of making predictions through data fitting, the research community is exploring new and innovative approaches to integrate ML and AI methods into AM. ML algorithms, applications, and platforms are helping AM practitioners improve product quality, optimize manufacturing process, and reduce costs.

A major challenge in current AM field is the inconsistency of the quality of the printed products, which are highly dependent on numerous processing parameters, such as printing speed and layer thickness. These process-structure-property (PSP) relationships have been discussed in

many review articles [1, 6, 7]. One method to address this challenge is conducting experiments or high-fidelity simulations [61, 62] to obtain reliable data and help optimize the processing parameters, but both of them are either time-consuming or expensive, and sometimes both. Another method to ensure part quality and process reliability is the application of *in situ* monitoring systems [63], but an efficient way for defect detection using the *in situ* data such as images is needed. In both methods, there is a critical need of an effective and efficient tool for data analysis and data mining. This need is being addressed by a subset of AI known as ML.

With a reliable training dataset, the ML models learn knowledge from the training set and make inference based on the knowledge. On one hand, the trained machine learning models can make predictions and determine the optimal processing parameters in an efficient way. On the other hand, it can also deal with *in situ* data for defect detection in real time. Some other ML applications, such as geometric deviation control, cost estimation, and quality assessment, are also reported in recent literature. In general, the ML applications can be regarded as the art of data manipulation. This capability makes ML a key aspect of Industry 4.0 [44].

Machine learning tasks can be divided into three main categories: supervised learning, unsupervised learning, and reinforcement learning [60]. Figure 3.1 displays the taxonomy of ML with the corresponding applications in AM field. In supervised learning, each input datum is labeled with an output Y, and the training set consists of many input-output pairs. Each input is a vector contains all involved features, $X_1, X_2, \dots, X_n$, that may affect its output. Each output can be a target classification such as quality assessment (good or bad) and the corresponding ML category is classification, or a target parameter such as porosity and tensile strength and the corresponding ML category is regression. In unsupervised learning, each input datum does not come with an output, and the model will study the relationship among input data. A typical application of unsupervised learning is clustering, in which all data are clustered into groups based on their similarity. Reinforcement learning, on the other hand, is learning how to map situations to actions so as to maximize a numerical reward signal [64], the applications of which include self-driving car and chess. Figure 3.1 illustrates some example applications in AM field with their corresponding ML categories. In AM field, most of the ML applications fall into the supervised learning category.

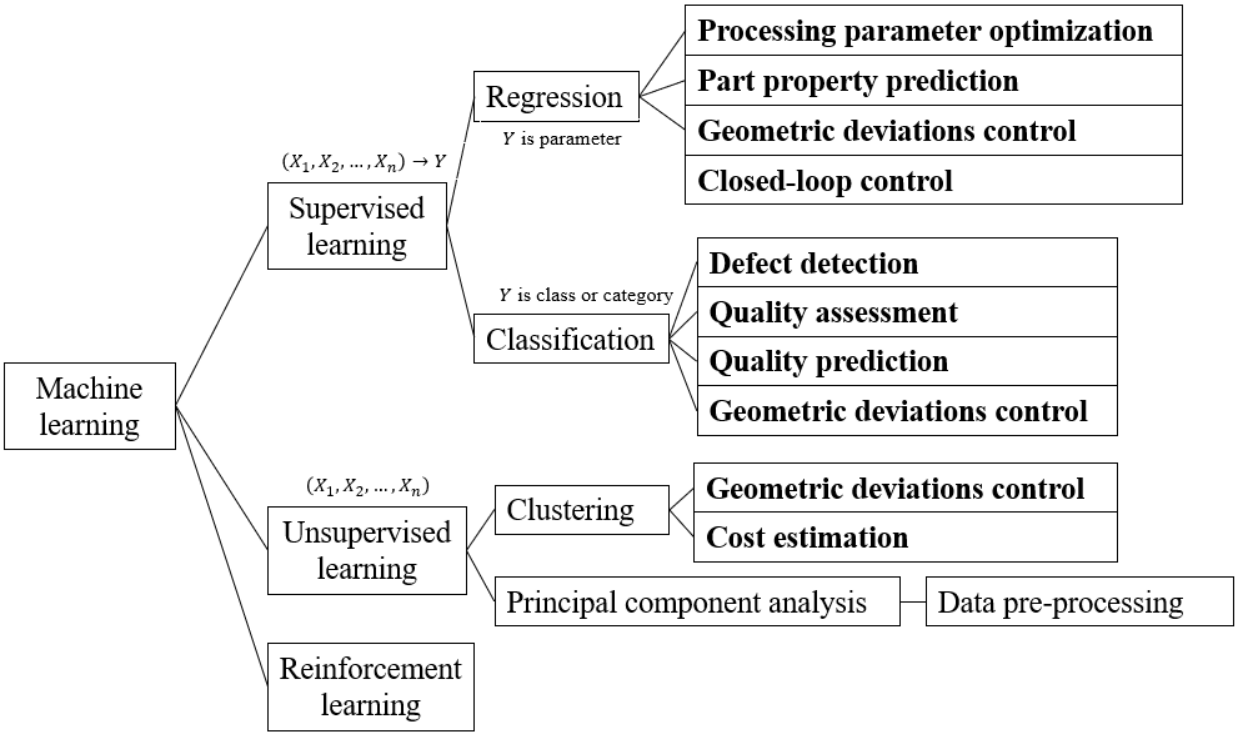


Figure 3.1. Taxonomy of ML applications in AM field. Text outside box is the data type. Text in bold is the ML applications in AM. $(X_1, X_2, \dots, X_n)$ is the input vector containing all input features and Y is the output.

The objective of this review article is to present the latest applications of machine learning in AM field, and thus provide a starting point for AM practitioners and researchers who are interested in applying ML. Very recently, there are a few related review articles [65] and reports [66] available. While Ref. [65] focuses on the applications and challenges of only neural network (NN), and Ref. [66] focuses more on the data acquisition in AM field, this chapter focuses on providing guidance on how to generate ideas in applying ML in AM field, distinguish the type of ML tasks, and make selection of different ML models. The whole chapter is organized as follows: The latest applications of ML in AM field are first reviewed from the perspective of AM in Section 3.2, and then classified into different categories of machine learning tasks, including supervised learning (Section 3.3) and unsupervised learning (Section 3.4). The performance of various machine learning algorithms applied in recent literature are compared and evaluated in Section 3.3 and Section 3.4. Finally, in Section 3.5, the chapter is summarized, and several future research directions are suggested.

3.2 ML Applications in AM

ML is a data manipulation tool. Figure 3.2 demonstrates various types of data available to be analyzed and utilized in the PSP relation chain. There are many relationships between these data, including but not limited to: (1) the processing parameters, such as extruder temperature in ME, laser power in laser powder bed fusion (L-PBF), printing speed, and layer thickness, significantly affect the structure of the printed parts, and thus dominate their quality and performance; (2) the designed shape play a crucial role in the printing cost and the geometric deviation of the printed products; (3) the *in situ* images and acoustic emission (AE) acquired by the monitoring systems are available to detect the occurrence of defect and its type in real time. Therefore, if a dataset, which consists of at least two types of related data in the PSP relation chain, is used to train ML models, the ML models will be able to make inference based on these data. This is the general procedure to apply ML models.

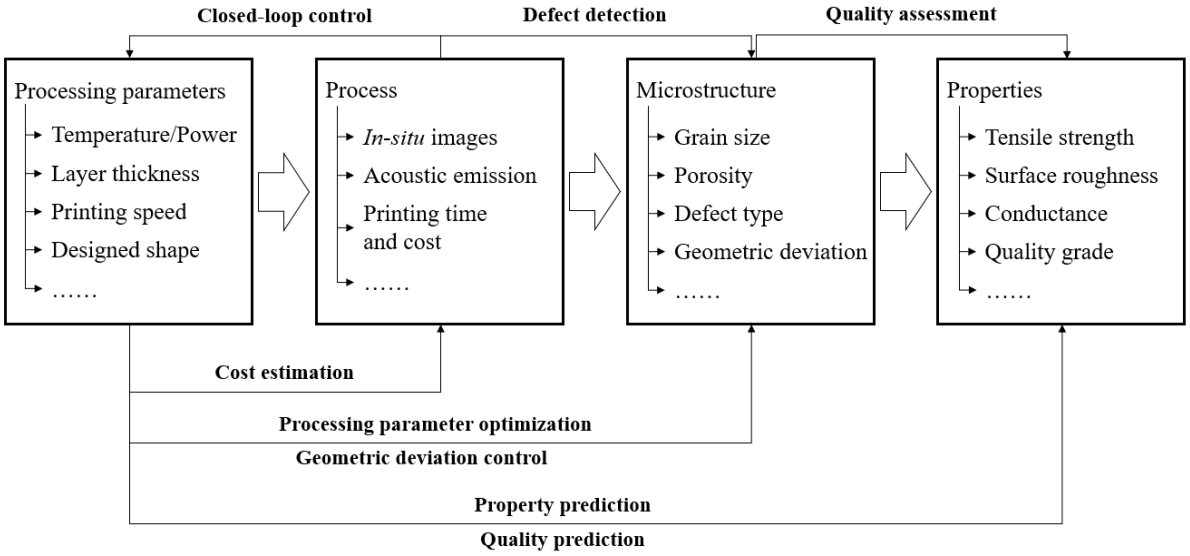


Figure 3.2. Process-structure-property (PSP) relation chain in additive manufacturing. The "process" term in the widely used PSP relationships is partitioned into two terms, "processing parameter" and "process data", to distinguish data available before the process and during the process. Texts in the boxes represent the available data that can be used in machine learning. Bold texts represent some existing machine learning applications in additive manufacturing field. The origin and the end of each arrow represent the input and output data, respectively.

3.2.1 Processing Parameters Optimization and Property Prediction

For designers, the quality of a part using a certain combination of processing parameters will remain uncertain until it is finally printed. Therefore, a series of efforts, such as printing some samples and testing their performance, have to be made to ensure the part quality, which makes the design process expensive, time-consuming and dynamic. In this regard, a direct relationship between the processing parameters and part quality is strongly desirable. To this end, experiments and simulations are useful methods to help construct such a relationship, but it's impractical to obtain optimal processing parameters using the two methods when a large amount of input features is involved. ML models, on the other hand, can be applied as surrogate models to assist process optimization.

Given a series of reliable training data of the property of interest (output) at some combinations of processing parameters (input), a process map can be generated by these discrete data points using ML regression models. Figure 3.3(a) demonstrates a process map of melt pool depth (output) in terms of laser power and scan speed (input) of 316L strainless steel in L-PBF process [67]. The applications of the process map is twofold: (1) it can make predictions to the output at any combinations of input features as a surrogate model and therefore reduce the demand of experimental and computational study, and (2) it can provide the relevance of each input feature to the output so as to obtain optimal input combination. Figure 3.3(b) plots the uncertainty and the discrete data points used to generate the process map. The uncertainty from the ML model is part of epistemic uncertainties in uncertainty quantification (UQ) [68]. This process map enables designers to achieve property prediction and process optimization efficiently. Since the process map is a typical production of ML regression models, the recent applications from literature in this topic are reviewed in Section 3.3.1.

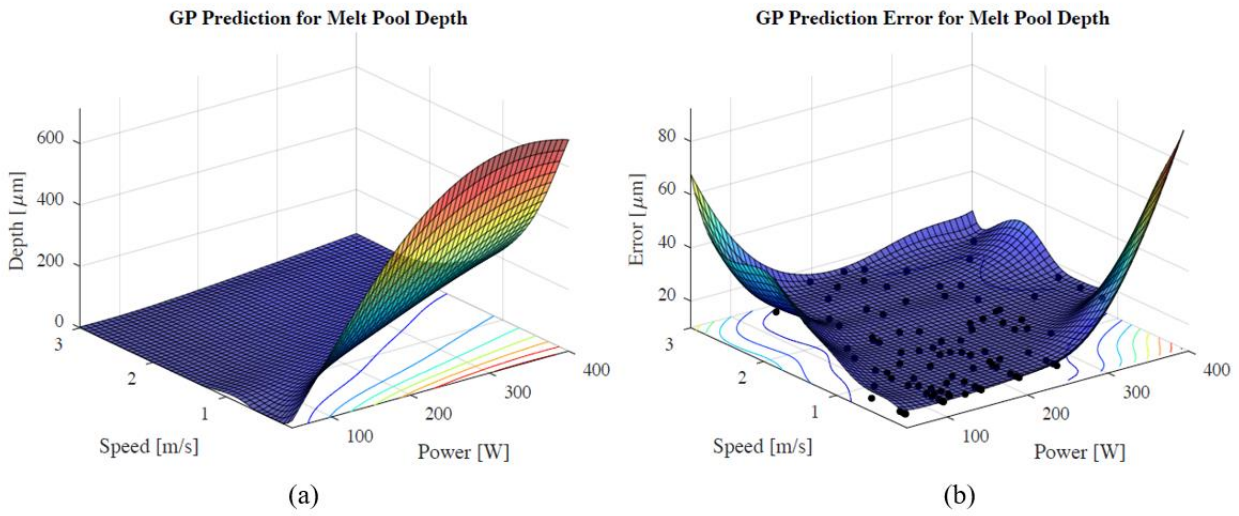


Figure 3.3. The process map of melt pool depth in terms of laser power and scan speed of 316L strainless steel in L-PBF process (a) and the corresponding uncertianty (b) [67]. The data points used to generate the process map are plotted in (b).

3.2.2 Defect Detection, Quality Prediction and Closed-loop Control

The development of the *in situ* monitoring systems enables the acquisition of real time data that can be used for defect detection and closed-loop control for AM [63]. These real time data, including spectra, images, AE and computed tomography (CT), can be utilized by ML models in several ways: (1) label these data with defect (possibly with defect types) or not by experimental results or human knowledge, and then use the labelled data to train supervised learning models for defect detection and quality prediction in real time, which is a typical application of ML classification models and will be discussed in Section 3.3.2; (2) conduct cluster analysis using unsupervised learning models to cluster the abnormal data so as to achieve defect detection without the labelling process, which will be discussed in Section 3.4.1; (3) train the ML regression models using the data in (1) along with some real-time controllable processing parameters, and therefore the ML models detect the occurrence of defects and tune these controllable processing parameters in real time. An example of the third way is the voltage level control in MJ process by Wang *et al.* [69]. Their process control framework consists of three main parts, as demonstrated in Figure 3.4. First, a charge-coupled device (CCD) camera is used to capture the dynamic images for the droplet. Second, four properties (satellite, ligament, volume, and speed) of the droplet are extracted from the images to train a neural network (NN) ML model along with the current voltage. Third, the

trained ML model is then used to determine the optimized voltage level and send it to the voltage adjustment system to control the droplet jetting behavior.

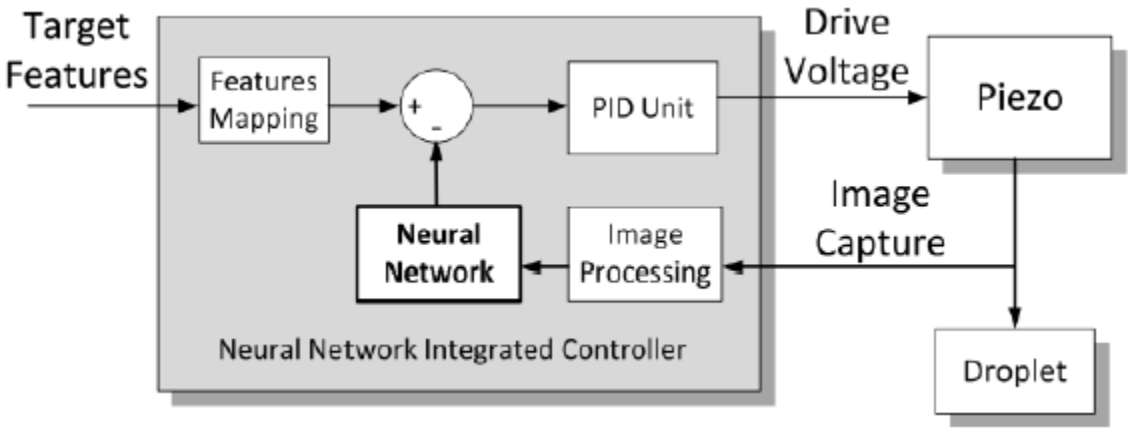


Figure 3.4. The closed-loop voltage control framework in MJ process [69].

3.2.3 Geometric Deviation Control

Low geometric accuracy and poor surface integrity are common defects of AM parts [4]. These geometric defects impede the applications of AM in several industries, such as aerospace and medical [70]. In this regard, ML models are capable of identifying the occurrence of geometric defect, quantifying the geometric deviation, and providing guidance of geometric error compensation. For instance, Francis *et al.* [71] developed a geometric error compensation framework for L-PBF process using convolutional neural network (CNN) ML model, shown in Figure 3.5. Using thermal history and some processing parameters as input and distortion as output, the trained ML model is capable of predicting distortion which is then imported reversely to the CAD model to achieve error compensation. By this means, the geometric accuracy of parts fabricated by the compensated CAD model will be significantly improved.

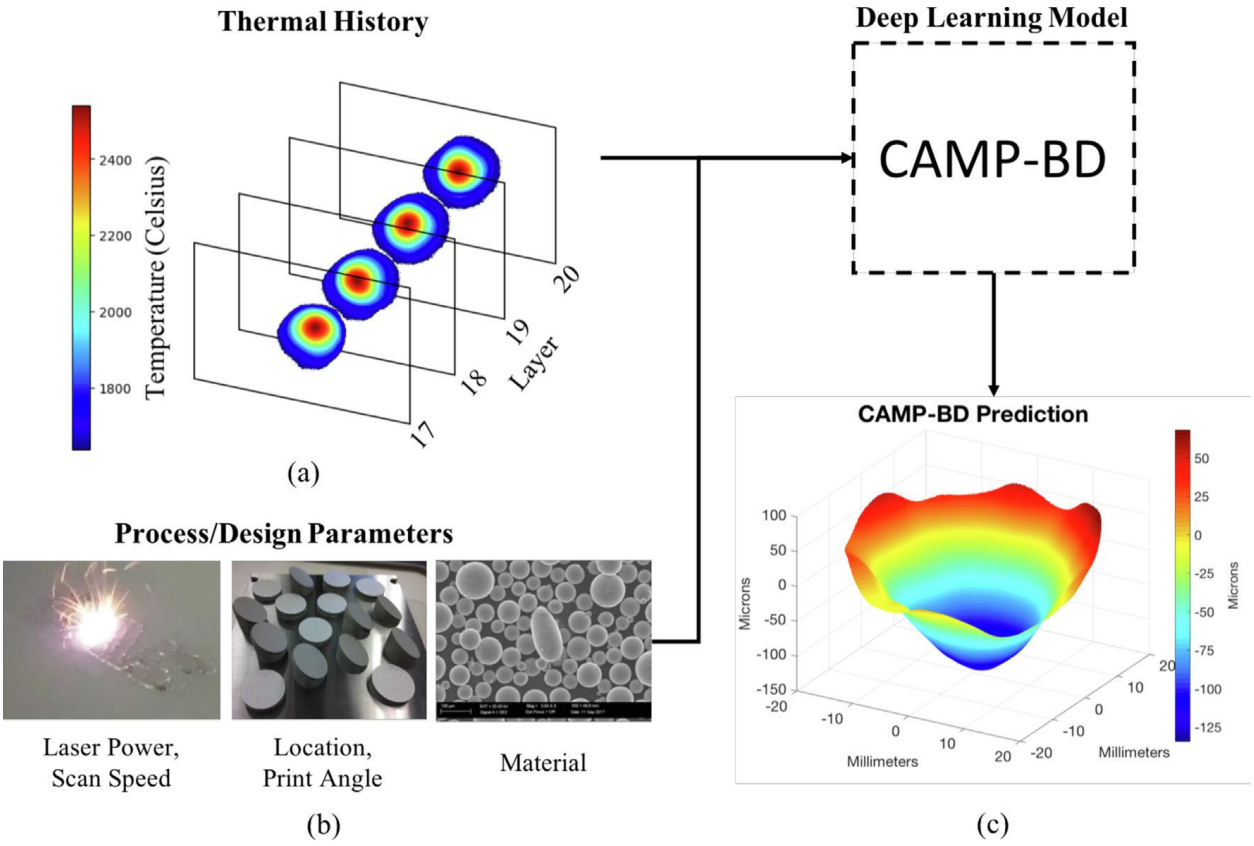


Figure 3.5. Procedure of geometric error compensation of Ti-6Al-4V in L-PBF process [71]. The input data are the thermal history and some processing parameters. The output data are the distortion. Error compensation is achieved by reverse the distortion in the CAD model. CAMP-BD represents Convolutional and artificial neural network for Additive Manufacturing Prediction using Big Data.

3.2.4 Cost Estimation

The printing cost and time are significant components of information shared between the manufacturers, clients and other stakeholders within the supply chain. Although they can be roughly estimated by the volume of the designed shape, there is still a need of a more accurate and efficient tool for cost estimation. Recently, an application of cost estimation by Chan *et al.* [72] is reported. Figure 3.6 demonstrates the cost estimation framework they proposed: (1) a client submits a manufacturing job with a 3D model; (2) features are generated from the 3D model and form the input vector, which is then imported to the trained ML models for cost prediction based on similar jobs using clustering analysis; (3) if client prefers or the training dataset size for ML models is small, the 3D model will be forwarded to simulation models to predict the cost, which

will also become training data for ML models; (4) the final predicted cost is estimated by combining the ML and simulation predictions; (5) the final prediction is forwarded to the client.

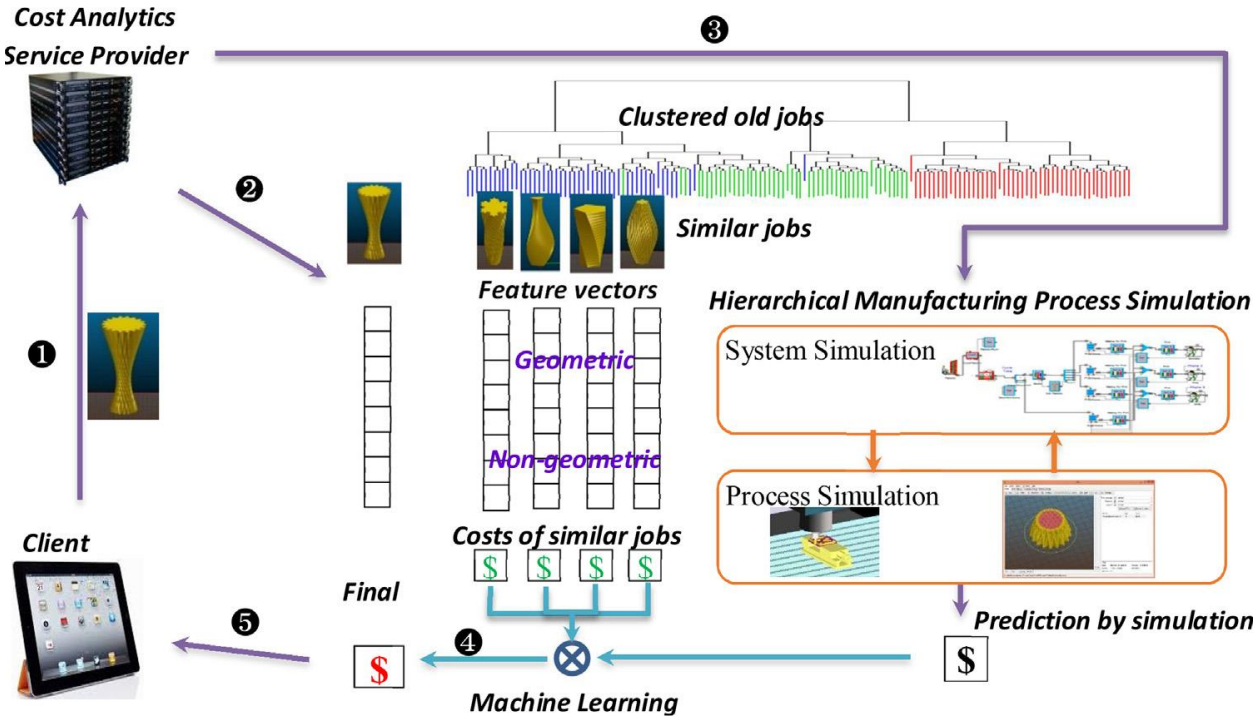


Figure 3.6. The cost estimation framework based on geometry and process similarities [72].

3.3 Supervised Learning

So far, the latest applications of ML in AM have been reviewed in Section 3.2 from the perspective of AM. From now on, these applications are classified into different categories of ML tasks in Figure 3.1. This is important to make selection of ML models for some reasons: (1) even with same applications, the ML models applied may be different with different data type, such as defect detection which can be achieved by both supervised learning and unsupervised learning models; (2) ML models tend to possess similar performance in the same categories of ML tasks using similar data type and dataset size. In this regard, Section 3.3 and Section 3.4 aim at providing guidance of making selection of data type and corresponding ML models.

A must-consider factor in applying ML in AM field is the data acquisition. A ML model requires sufficient data to make accurate predictions. The required number of training data also increases exponentially with the increasing number of input features. However, in many applications, the acquisition and labelling of data requires high experimental, computational, and/or laboring costs. Therefore, before a ML model is applied, the dimension of a ML task should

be determined carefully considering the number of available data and the cost to obtain them. Section 3.3 and Section 3.4 also list many examples of ML applications in literature that can help determine the task dimension.

In supervised learning, all input data are labeled with an output. The output can be either parameters and the corresponding ML task is regression, or classes and the corresponding ML task is classification. Since most of the ML applications in AM aim at predicting a target parameter or class, supervised learning is the major type of ML applications in AM field.

3.3.1 Regression

In regression tasks, the output of each input is parameters, such as porosity of the printed products, efficiency, melt pool depth, mechanical property, etc. The AM algorithm learns the relevance between the input and output parameters from the training dataset, and then makes inference from a new input to its output using the relevance it learns.

Regression Applications in AM

The major functionality of ML regression models in AM field is the generation of process map, which has been discussed in Section 3.2.1. Therefore, processing parameters optimization and property prediction will be the two major applications of ML regression models. In addition, since the targets in geometric deviation control and cost estimation are all parameters, they may also be the applications of ML regression models.

Regression Models Assessment in AM Applications

Table 3.1 shows the recent regression applications along with the ML models in AM field. According to Table 3.1, the two major ML models for regression tasks applied in AM field in recent literature are neural network (NN) and Gaussian process (GP).

Table 3.1. ML regression applications in AM

Applications	Inputs	Outputs	Models
Geometric deviations control [73]	Shape parameters	Shape deviation parameters	Gaussian process (GP)
Processing parameters optimization [67]	Laser power and scan speed	Melt pool depth	GP
Processing parameters optimization [74]	Laser power and scan speed	Porosity	GP
Processing parameters optimization [9]	Laser power and scan speed	Melt pool depth	Regression tree (RT), GP
Processing parameters optimization [75]	Laser power and scan speed	Melt pool depth and width depth ratio	GP
Trace geometry prediction [76]	Laser power and scan speed, and powder feeding rate	Deposited trace cross-section geometrical parameter	NN
Processing parameters optimization [77]	Layer thickness, layer power, hatch spacing, laser speed, interval time, surroundings temperature, and scanning mode	Shrinkage ratio	NN
Property prediction [78]	Material property, extruder temperature, printing speed, layer thickness	Tensile strength	Support vector regression, random forest (RF), recurrent NN
Property prediction [79]	108 input features including extruder temperature, printing speed, and layer thickness	Surface roughness	RF, AdaBoost, RT, support vector regression (SVR), Ridge regression, NN, and ensemble of them
Thermal history prediction [80]	Toolpath feature, the time of deposition, closest distance to the boundary of the build, layer height, laser intensity, and laser state	Thermal history	Recurrent NN

Table 3.1 continued

Real-time composition monitoring [81]	Spectral line-intensity-ratio and spectral integrated intensity	Element composition	SVR, NN,
Geometric deviations control [71]	Thermal history, laser power and scan speed, printing location and angle, and material	Distortion	CNN
Geometric deviations control [82]	Deformed note locations	Original note locations	NN
Closed-loop control [69]	Droplet features (satellite, ligament, volume, and speed)	Voltage level	NN

Artificial neural networks, inspired by biological neural systems, are computing systems consist of massively parallel interconnected networks of simple (usually adaptive) elements and their hierarchical organizations [83]. All the "neural network" or "NN" in this chapter refer to artificial neural networks, instead of biological neural systems. A typical neural network contains an input layer, one or more hidden layers and one or more output layers. Each layer is made of numerous neurons. The information of each neuron is propagated to the next layer based on the its weight. A NN will be categorized to recurrent NN when the propagation of its neurons forms cycles, and feedforward NN otherwise. During training, the weight of each neuron is optimized by the learning rule as soon as a new observation is imported into the NN. The most popular learning rule for NN is the backpropagation (BP) algorithm [84], which adjusts the weights based on the gradient descent. However, due to the strong learning ability of BP algorithm, NN usually suffers from overfitting issue (More discussions in Section 3.3.3), which can be alleviated by either early stopping method or regularization [85, 86]. For more knowledge about NN, refer to Ref. [87].

Caiazzo *et al.* [76] applied BP-NN for trace geometry prediction with RMSE of around 5% using 30 tranining data. Rong-Ji *et al.* [77] tested the performance of BP-NN with 5 to 10 hidden neurons and their results exposed the trend that more hidden neurons tend to make better predictions. Zhang *et al.* [78] used recurrent NN in ME process to predict the tensile strength of the printed products and the RMSE was around 2%. Figure 3.7 [78] illustrates the NN they constructed: during training (bottom), the output of each input combination are progagating backward in the NN to adjust the relavance of each input feature, and after training (top), new

combinations of input features are propagating forward to predict tensile strength. Their result exposed that recurrent NN has outperformed the random forest and support vector regression algorithms in this application. Overall, NN displayed excellent performance in regression tasks, but also required tuning a series of hyperparameters such as number of hidden neurons and layers [65].

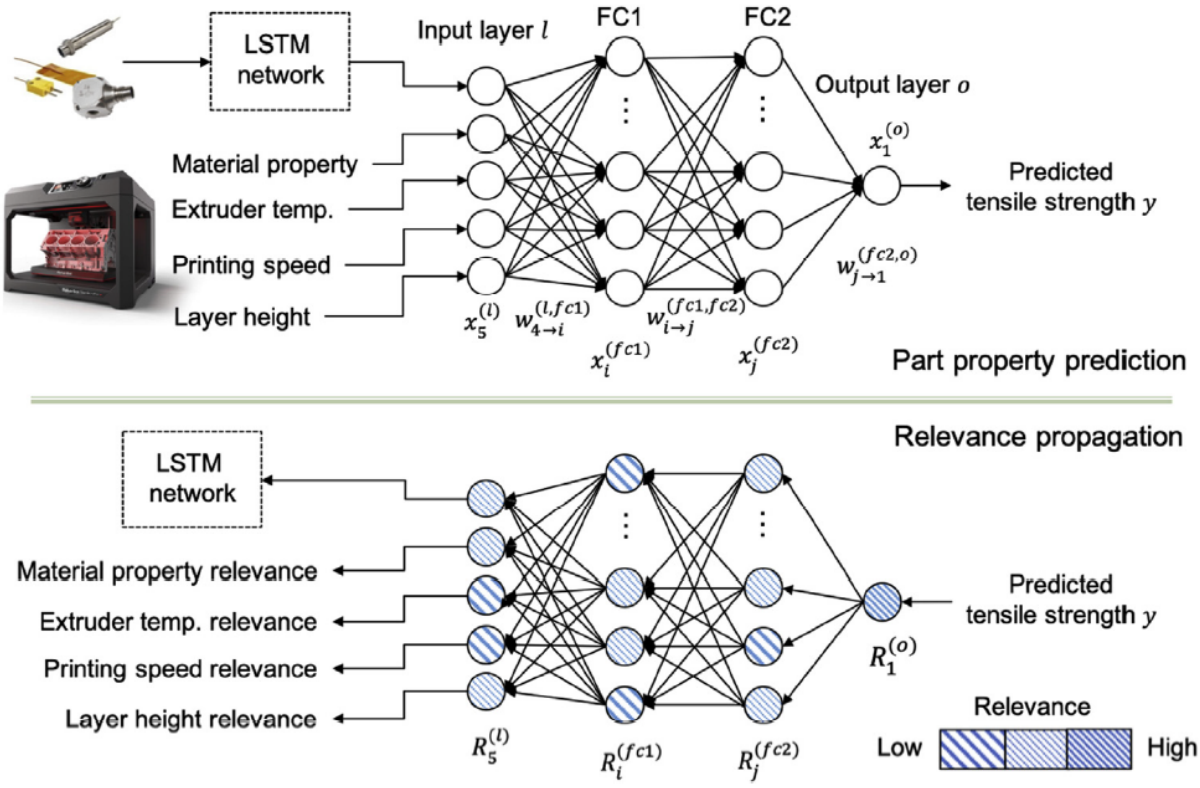


Figure 3.7. Layer-wise relevance propagation through the trained neural network for polylactic acid (PLA) in fused filament fabrication (FFF) [78]: porpagation forward for prediction of tensile strength (top) and propagation backward for training the relevance of each input feature (bottom).

A Gaussian process is defined as a collection of random variables, any finite number of which has a joint Gaussian distribution [88]. Similar to the mean value and variance in Gaussian distribution, a GP is completely specified by a mean function $m(\boldsymbol{X})$ and a covariance function $C(\boldsymbol{X}, \boldsymbol{X}^*)$, where $\boldsymbol{X}$ is the input vector containing all input features. The covariance function is defined by a single or a combination of kernel functions and is critical to the performance of GP as it captures the spatial dependence between two different locations, $\boldsymbol{X}$ and $\boldsymbol{X}^*$. The selection of kernel functions should be based on practical applications and has been discussed in Ref. [88].

Tapia *et al.* [67, 74] applied GP to make predictions of porosity and melt pool depth in L-PBF process. The process map of melt pool depth in terms of laser power and scan speed with the corresponding uncentainty are plotted in Figure 3.3. GP shows excellent regression performance in noisy environment and with limited number of training data. The GP can be used calibrate a convinient criterion [67] to aviod porous formation due to keyhole mode [21] in L-PBF process. An exmaple of this calibration is demonstrated in Ref. [89], where the process maps of 316L and 17-4 PH stainless steels are generated by GP model using experimental dataset and used to compare anaginst and calibrate the normalized enthalpy criterion [67]. Zhu *et al.* [73] applied GP to make predictions on shape deviation and the RMSE is around 3% using 75% of the whole dataset as the training set. Overall, GP is efficient and effective in regression tasks with a few input features and a small dataset. GP may lose its efficiency when the number of involved input features is large or the size of the training dataset is too large, due to the high computational costs in performing matrix inversion.

In general, both GP and NN are capable of handling regression tasks in AM field. NN is more complicated than GP and requires more knowledge to tune the hyperparameters. The selection of ML models should be based on the complexity of the training dataset (i.e. the number of training data points and input features). For low complexity tasks, GP is recommended. For high complexity tasks, NN is recommended. The application of an ensemble of multiple algorithms [79] (including NN) are also reported, which predict more accurately than NN and can be regarded as an alternative.

3.3.2 Classification

In classification tasks, the output of each input is a class or a category, such as different defect types or quality assessment grade. Similar to regression tasks, the ML models learn how to make classification from the training set, and then use the knowledge to classify new input.

Classification Applications in AM

In AM field, there are various classes with different criteria that can be used to distinguish part quality, such as defect and non-defect, quality is good or bad, quality grade assessment on a scale of 1 to 10 to quantify the quality, etc. If a ML model is trained by some classification

examples at different input settings, it will be able to make classification to new input henceforth. Therefore, ML classification models can be used in AM field in three main aspects: (1) it can use *in situ* data, such as images and AE, to make predictions of defects so as to help defect detection in real time, (2) it can predict the part quality at different processing parameters, and (3) it can assist quality assessment using the geometric information of printed parts. As the geometric deviation can be described by several types, such as translation and rotation, ML classification models can also achieve geometric deviation control.

Classification Performance Assessment Method

An assessment method is necessary to quantify the performance of a classification model. Classification tasks can be further divided into two subgroups: (1) binary problems, in which only two categories are involved, and (2) multiclass problems, in which at least three categories are involved. The performance of ML algorithms in classification tasks is usually assessed by precision, recall, or F_1 score in binary problems, and accuracy in multiclass problems.

Table 3.2 displays the confusion matrix of binary classification problems. Precision is defined as $\frac{TP}{TP+FP}$ and represents the ability of a model to identify only the relevant instances, whereas recall is defined as $\frac{TP}{TP+FN}$ and represents the ability of a model to find all the relevant instances. As there is usually a trade-off between precision and recall, F_1 score is defined as $2 \times \frac{Precision \times Recall}{Precision+Recall}$ and represents the overall performance of a model. The range of F_1 score is from 0 to 1, and the larger the F_1 score, the better the performance. Accuracy is defined as the total number of correct predictions over all predictions, or $\frac{TP+TN}{TP+TN+FP+FN}$ in binary problems, but it may not be appropriate in binary problems when the number of positive and negative samples is imbalanced.

Table 3.2. Confusion matrix of binary classification problems

Ground truth / Prediction	Positive	Negative
Positive	True positive (TP)	False positive (FP)
Negative	False negative (FN)	True negative (TN)

Classification Models Assessment in AM applications

Table 3.3 shows recent classification applications along with the ML models in AM field. Typical ML algorithms for classification tasks are decision trees (DT), support vector machines (SVM), and convolutional neural networks (CNN).

Table 3.3. ML classification applications in AM.

Applications	Inputs	Outputs	Models
Design feature recommendation [90]	Target components	Recommended AM feature	Support vector machines (SVM), dendrogram
Defect detection [91]	Regions of interest of spatters, plume and melt pool	Class 1,2 or 3	SVM, convolutional neural networks (CNN)
Defect detection [92, 93]	*In situ* images	Defect type	Bag of words (BoW), CNN
Defect detection [94]	Spectral intensity graph	Defect or not	SVM
Defect detection [95]	Melt pool characteristics	Porous or not	Decision trees (DT), k-nearest neighbors (KNN), SVM, Discriminant Analysis (DA)
Quality assessment [96]	Dimensional variation	Infill classes	KNN, naive Bayes, NN, SVM, DT
Quality prediction [97]	Energy density, particle distribution and surface morphology	Quality: Good or bad	SVM

Table 3.3 continued

Defect detection [98]	AE	Defect type	Deep belief network (DBN), SVM, NN
Geometric deviation control [99]	Voxel grid	Deformation type	CNN
Defect detection [100]	25 thermal features	Porosity label and normalized porosity size	KNN, NN, Self-organizing error-driven neural networks (SOEDNN)
Defect detection [101]	CT image layers	Defect or not	SVM
Defect detection [102]	In situ images	Defect or not	CNN
Quality prediction [103]	AE	Poor, medium or high quality	Spectral convolutional neural networks (SCNN)

Decision trees [104] are a type of common ML algorithm for classification tasks. Compared with NN, decision trees are more interpretable. Khanzadeh *et al.* [95] and Tootooni *et al.* [96] applied multiple ML models including DT for defect detection and quality assessment, respectively. In both articles, DT shows medium performance among many classifiers. Overall, DT is a relatively simple method and is capable of dealing with classification tasks in AM field. Though it may not perform the best, it is recommended as a contrast when applying other models to better show the performance of other models.

Support vector machine is designed to deal with binary classification problems [105], but it can also be generalized to multiclass problems [106]. In binary problems, as each input-output pair in training set consists of a high dimensional input vector containing all input features and a target category as output, SVM uses a hyperplane in the high dimensional space to partition the two groups. According to Table 3.3, SVM is a very popular classifier in AM applications. In the comparison of multiple classifiers [95, 96], SVM shows comparable performance with other algorithms.

While SVM is good at handling inputs consist of only parameters or classes, it can also be applied in image-based problems. Figure 3.8 [95] demonstrates a procedure using images as input

for defect detection of Ti-6Al-4V in L-PBF process. For each thermal image labeled with either porous or not porous, some geometric features are extracted from the image and used to train the ML models. Zhang *et al.* [91] applied SVM for defect detection using *in situ* images as input. In their article, though CNN performs better (92.8% accuracy), SVM shows 90.1% accuracy in this three-group classification task. Ye et al. [98] applied SVM for defect detection using AE as input, which also requires a feature extraction procedure like images. In this binary classification problem, SVM (98.01% accuracy) outperformed the deep belief network (95.87%). Gobert et al. [101] applied SVM for defect detection using CT image layers as input, and the F_1 score (refer to Section 0) of their optimized SVM model is 0.62. Overall, SVM is a great alternative in classification problems.

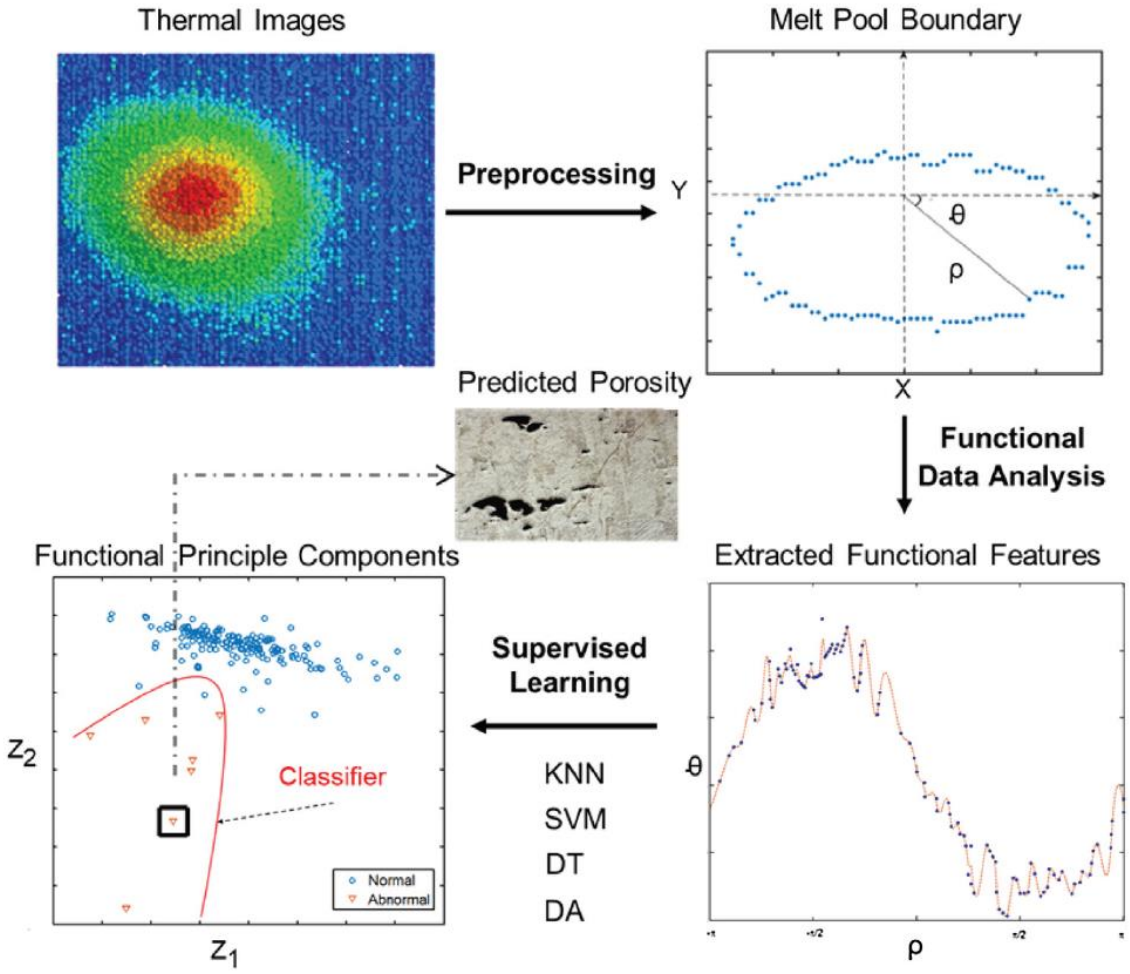


Figure 3.8. The procedure from thermal images (input) to porosity predictions (output) of Ti-6Al-4V in L-PBF process. Some geometric features are extracted from the thermal images to train the ML models, which can then classify whether the printed product is porous (abnormal) or not (normal).

Neural network tends to be the most popular algorithm in classification problems. While normal NN is usually applied in problems with input consists of only parameters and classes, a special type of NN, known as convolutional neural network (CNN), is designed to handle problems with images and AE [107]. Scime *et al.* [92, 93] applied mutli-scale CNN for defect detection using *in situ* images and the overall, anomaly detection, and anomaly differentiation accuracies are 97%, 85%, and 93%, respectively. The multi-scale CNN they implemented is demonstrated in Figure 3.9. The information of images is propagated in the NN using convolution. Shen *et al.* [99] applied CNN for geometric error compensation using voxel grid as geometric input feature and got an overall F_1 score (refer to Section 0) of 0.95. Overall, NN is a complex but strong model among the existing algorithms for classification tasks in AM field. NN is applicable in most classification tasks.

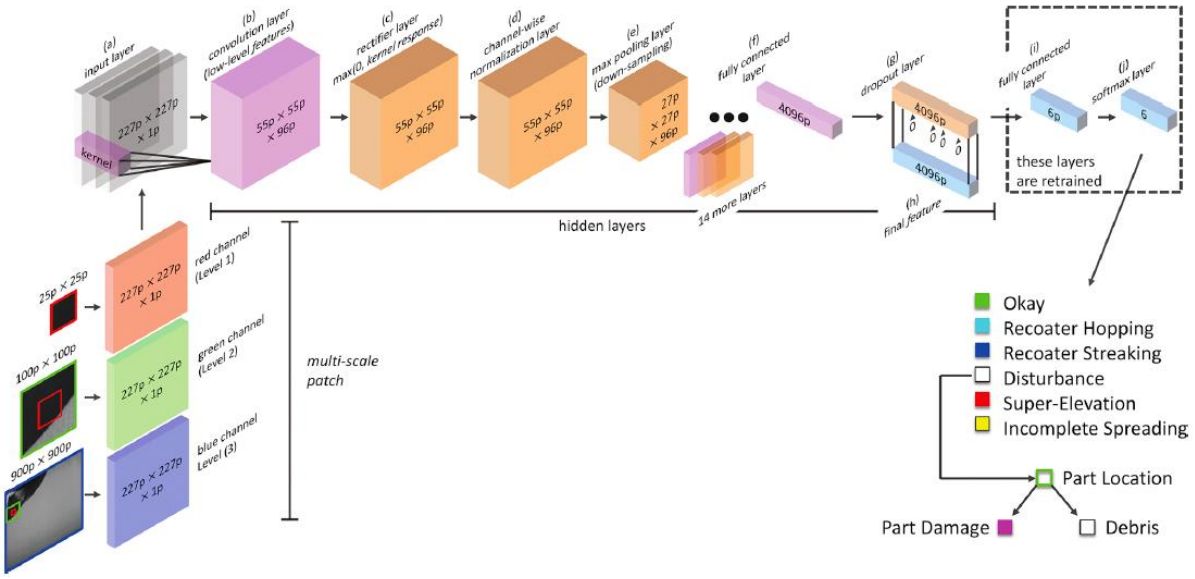


Figure 3.9. The flowchart of multi-scale CNN in defect detection in L-PBF process using multiple materials [93]. Images labeled with okay or 5 types of defect are used to train multi-scale CNN. The information of images is propagated in the multi-scale CNN using convolution and finally used to classify the type of defect.

In general, the selection of different classification models should be based on the type of input features. While most of classifiers including DT, KNN, NN, SVM can deal with common parametric input problems in AM, SVM and CNN are recommended for images or AE based problems.

3.3.3 Overfitting Issue and Solutions

ML models learn knowledge from training data, and then use the knowledge to make predictions. Therefore, if the training data is used for testing the performance of ML models, the models tend to make perfect predictions in these training data, which seems wonderful but may trap in the overfitting issue. Figure 3.10 demonstrates an example of this situation in melt pool depth predictions using GP in L-PBF process. It reflects the fact that validation should never be done with the training dataset. Another example of this situation is the 100% F_1 score of NN in Ref. [96].

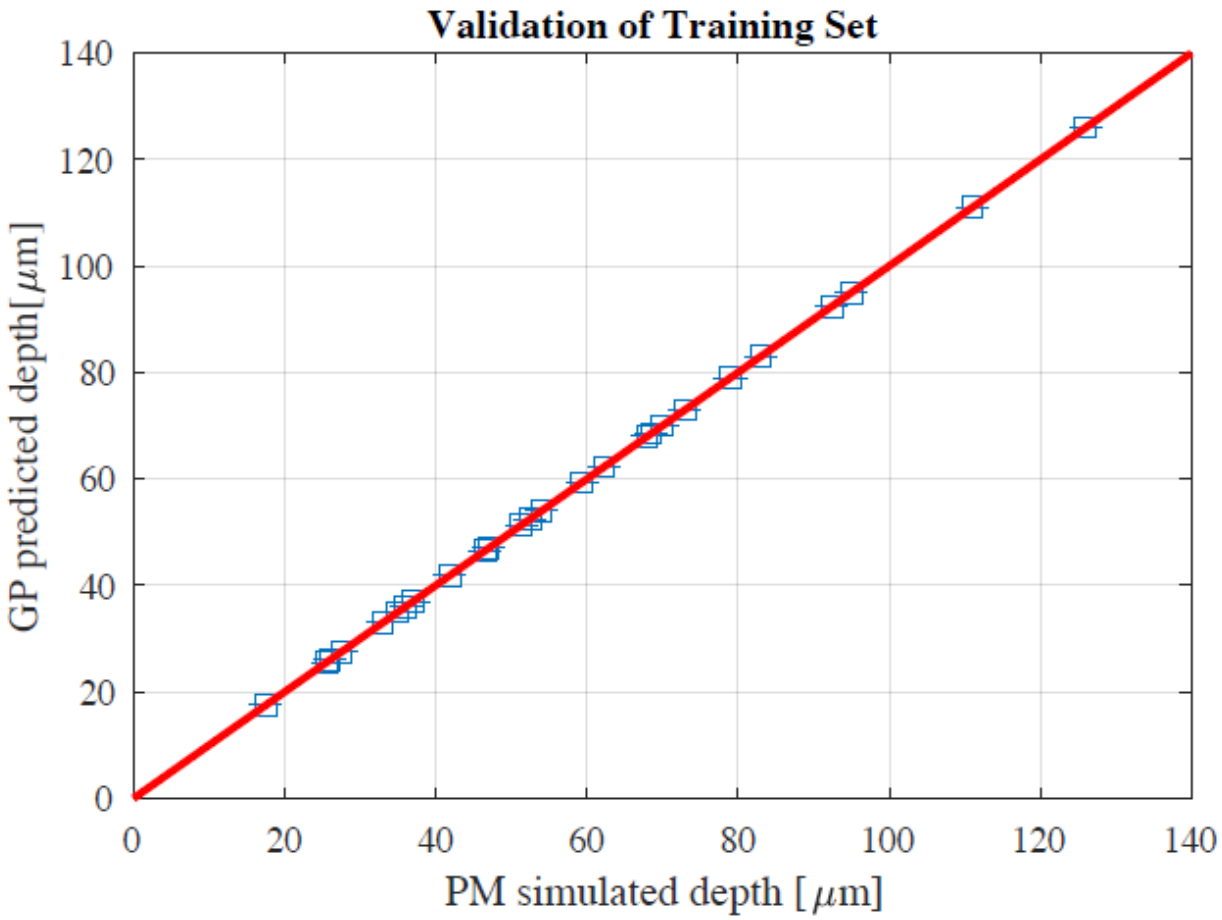


Figure 3.10. Validation plot of GP in prediction of melt pool depth in terms of laser power and speed of 316L strainless steel in L-PBF process [67]. GP predictions are comparing against its training dataset. The closer from each point to the ideal $y = x$ line, the more accurate the prediction is. In this plot, all points are exactly lying on the ideal line without any uncertainty, indicating that the GP perfectly represents the data in the training set, and also the fact that validation should never be done with the training dataset.

Overfitting is a phenomenon that the model adjusts itself to fit the training dataset too exactly. In other words, with decreasing training error, the prediction error for future observations tends to increase. This is a common issue in supervised learning and should be avoided by some means. Three popular methods to help detect and avoid the overfitting issue in AM field are the hold-out method, the *k*-fold cross-validation method, and the regularization method.

The hold-out method, also known as data splitting, is a simple method to monitor overfitting. It partitions the whole dataset into two subsets, training set and testing set. The training set is used to train the model and the testing set is used to test the performance of the model. By this means, data in the testing set will not be used to train the model and is useful to test the performance of the model and whether overfitting occurs. An appropriate size of the training set is usually around 70% of the whole dataset. However, this method has a main drawback: it will further reduce the size of the training dataset when initially the number of data points is limited, which is the common situation in additive manufacturing as the cost, consumed time, and human labor to obtain each data point is usually high. This method is commonly applied in most of the applications mentioned above.

The *k*-fold cross-validation (CV) method is an iterative procedure which can monitor the overfitting issue and enhance the utilization of data. It partitions the whole dataset into *k* subsets of roughly same size. In each iteration, one subset is left out as the testing set and all other subsets are used to train the model. The iteration is repeated until all subsets have been left out once. A special case of this method is $k = n$, n being the number of data points, which is also called *n*-fold cross-validation method or leave-one-out cross-validation (LOOCV). Compared with the simple hold-out method, the CV method alleviates the common issue of the limited size of dataset in AM field.

The regularization method is a process which discourages the ML model to become too complex by adding information during training [108, 109]. In general, the goal of a ML model is to minimize the loss function:

$$E = \frac{1}{m}\sum_{k=1}^{m} E_k, \tag{3.1}$$

where E is the accumulative error, m is the number of training data points, and E_k is the error at each training data point. However, if noise exists in the training data, the ML model will also learn the noise using Eq. (3.1) and tend to overfit. To avoid this situation, the regularization method adds a term to the loss function, to penalize the complexity of the model. A commonly used

regularization method is called L_2 regularization [110], which encourages the sum of the squares of the parameters to be small. For example, the loss function using L_2 regularization for neural network is:

$$E = \frac{1}{m}\sum_{k=1}^{m} E_k + (1-\lambda)\sum_i w_i^2, \tag{3.2}$$

where $\lambda \epsilon (0, 1)$ is the tuning parameter that determines how much penalty is added to the model complexity, and w_i is the weight of each neuron. The tuning parameter should be carefully selected and is usually estimated by cross-validation [108, 109].

Many algorithms, due to their learning mechanisms or strong learning ability, tend to overfit the training data. For instance, a characteristic of GP is that it will pass through all training points in regression tasks (Figure 3.10) in a noise free environment. For another instance, Hornik et al. [111] has shown that multilayer feedforward NN can approximate any function to any desired degree of accuracy, provided sufficiently many hidden units are available. Such strong learning ability of NN makes it likely overfit. To counter overfitting, many applications using above methods are reported in literature, such as the 10-fold [67] and *n*-fold cross-validation method [9, 75, 112] for GP, and L2 regularization [86] for NN. Some algorithms also have their own methods against overfitting, such as the dropout method [113] for NN. Overall, the applications of one or more methods to monitor and avoid overfitting issue is necessary to make the ML model robust.

3.4 Unsupervised Learning

In unsupervised learning, all data are not labeled with an output. The most common task in unsupervised learning is the clustering analysis, in which the data are separated into groups based on their similarity. Another main type of unsupervised learning is principal component analysis (PCA), which converts a dataset with many possibly correlated variables into a smaller set of values of linearly uncorrelated variables called principal components by orthogonal transformation.

3.4.1 Clustering Analysis in AM

In clustering analysis, all data are separated into groups based on their similarity. In general, a clustering analysis usually requires a large dataset size. However, the dataset size in AM field is usually limited, which impedes the application of clustering analysis. Therefore, only a few applications of clustering analysis in AM are reported recently.

In the cost estimation framework proposed by Chan et al., the ML models applied are the least absolute selection and shrinkage operator (LASSO) and elastic net (EN) models. Another typical ML model for clustering analysis is the self-organizing map (SOM), which is a type of NN for unsupervised learning. Recently, an application for geometric accuracy analysis using SOM by Khanzadeh *et al.* [114] is reported. Using SOM, millions of data of geometric deviation are separated into clusters, and the overall geometric accuracy of the part fabricated using each combination of processing parameters can then be assessed: the more clusters that appear, the more types of deviations in terms of direction and magnitude it has. It should be noted that the same dataset in Ref. [114] is also used in Ref. [96] for quality assessment using supervised learning. This reveals that one dataset can have multiple applications in machine learning.

Khanzadeh *et al.* [115] also applied their SOM model for defect detection of Ti-6Al-4V in DED process. Under the assumptions that (1) an abnormal melt pool has low correlation with others, and (2) the percentage of abnormal melt pool is much smaller compared with normal melt pools, the data representing the melt pool temperature distribution characteristics are clustered and therefore the cluster with low correlation to all others is considered anomaly and porosity tends to occur at the corresponding locations. Another recent application of clustering analysis for defect detection using SOM is conducted by Wu *et al.* [116]. Features are extracted from AE signals acquired by AE sensors and then imported into SOM for clustering, as demonstrated in Figure 3.11. As the AE signals in the abnormal (failure) cases tend to be different to the ones in normal cases, the cluster of abnormal cases can be extracted from the normal cluster, and thus the defect detection is achieved. These two applications for defect detection in AM are a main functionality of clustering analysis known as anomaly detection. Comparing with the supervised learning-based defect detection, this method possesses a significant advantage that it doesn’t require human interaction to label data. Therefore, clustering analysis can be a strong alternative for defect detection.

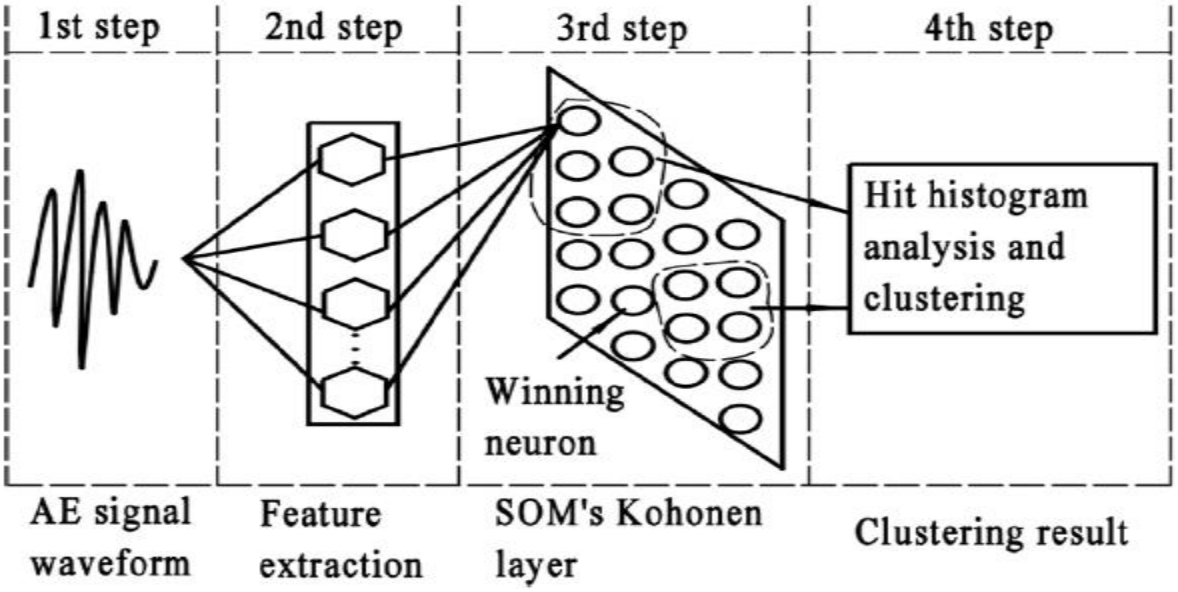


Figure 3.11. Procedure of a clustering analysis of AE signals of acrylonitrile butadiene styrene (ABS) in FFF process [116].

3.4.2 Principal Component Analysis in AM

Sometimes the number of features in a dataset is very large, especially when the input data type is image. In this case, to avoid the problem being too complicated, PCA is usually applied as a data pre-processing tool in AM to reduce the number of features so as to simplify the data. Khanzadeh *et al.* [95] applied PCA to simplify the features from melt pool characteristics and finally obtained nine principal components that account for almost 99.52% of variation in the data. Yang et al. [117] also demonstrates how to extract geometric features for energy consumption estimation in mask image projection stereolithography using PCA. In an application using *in situ* images as input data [91], the PCA increases the accuracy of SVM from 89.6% to 90.1% using 33 input features extracted from the image. However, in their 17-feature case, PCA is reported to weaken the performance of SVM, which indicate that PCA may also have negative effects on the performance of the coupled model, since too many features are eliminated and too much information is lost. Overall, when dealing with image-based problems, PCA is a great alternative to simplify the data.

3.5 Summary

In this review article, the latest applications of ML in AM field are reviewed in terms of the type of learning tasks: supervised learning and unsupervised learning. For each specific types of tasks, including regression, classification, clustering and PCA, the corresponding applications and some popular algorithms are discussed, and the performance of some popular algorithms are assessed. The following is the recommended future research directions:

1. While ML has been developing for several decades, the applications of ML in AM field have only been discovered for several years. These applications span processing parameters optimization, property prediction, defect detection, geometric deviation control, quality prediction and assessment, etc. Firstly, ML models can learn the relevance between the processing parameters and property using existing data, so as to provide guidance of optimizing these processing parameters. Secondly, ML models can predict the geometric deviation based on the designed geometry after training and provide guidance of geometric error compensation. Thirdly, ML models are good at dealing with *in situ* images and acoustic emission during printing and detecting defect formation in real time. However, the available data that can be extracted from the processing parameter-process-microstructure-property map have not been fully utilized. In this regard, exploiting more data acquisition methods, exploring more ML applications and developing better algorithms will be the main research directions in this infancy research field.
2. A missing but useful functionality in supervised learning in recent literature is active learning. In AM field, labelling the output of each input data point is usually expensive in terms of the consumed time, cost and human labor, because it requires conducting an experiment or a simulation at each input setting to make this observation. Active learning is a method that can alleviate this issue. In recent literature, the common procedure in ML models is acquiring enough input-output pairs first and then using them to train ML models without further query of labelling new data. On the contrary, the procedure in active learning is that the ML models can make query interactively for labelling new data during training so as to maximize its performance. By this means, the ML models may use fewer data points to achieve

better performance. Therefore, active learning is strongly recommended in the case that a dataset to be used to train the ML model has not been acquired.

3. Another potential research field is the uncertainty quantification (UQ), which is critical for a robust design. The uncertainty in AM field has been reviewed in Ref. [68]. In regression tasks, ML models like GP provide not only the mean value at a certain input as the prediction of its output, but also standard deviation which represents the uncertainty at that point. Also, in classification tasks, ML models will also provide confidence when they make a classification. These uncertainties are part of the epistemic uncertainty and have not been utilized in recent literature. In addition, a typical UQ procedure [118] may require hundreds of data points, which is impractical to obtain from experiments or simulations. In this regard, a ML-based surrogate model is very helpful in obtaining the required data and increasing the efficiency of the UQ procedure. Overall, UQ in ML applications in AM field is a good research direction that has not been investigated in depth.

4. PROCESS DESIGN OF LASER POWDER BED FUSION OF STAINLESS STEEL USING A GAUSSIAN PROCESS BASED MACHINE LEARNING MODEL

A version of this chapter has been published in *Journal of The Minerals, Metals & Materials Society* (doi: 10.1007/s11837-019-03792-2)

Abstract: In this chapter, a Gaussian process (GP) based machine learning model is developed to predict the remelted depth of single tracks, as a function of combined laser power and laser scan speed in laser powder bed fusion (L-PBF) process. The GP model is trained by both simulation and experimental data from the literature. The mean absolute prediction error magnified by the GP model is only 0.6 μm for a powder bed with layer thickness of 30 μm, suggesting the adequacy of the GP model. Then the process maps of two metals, 316L and 17-4 PH stainless steel, are developed using the trained model. The normalized enthalpy criterion of identifying keyhole mode is evaluated for both stainless steel. For 316L, the result suggests that the $\frac{\Delta H}{h_s} \geq 30$ criterion should be related to the powder layer thickness. For 17-4 PH, the criterion should be revised to $\frac{\Delta H}{h_s} \geq 25$.

4.1 Introduction

Additive manufacturing (AM) is defined as a group of layer-upon-layer fabrication processes and previously applied for rapid prototyping. In the last decade, AM technology has rapidly evolved into a decent alternative for direct fabrication of tools and functional parts. Laser powder bed fusion (L-PBF) is an AM technique that uses a laser as the power source to selectively melt or sinter a metallic powder bed. In a typical L-PBF process, for each new layer, a recoating blade pushes a layer of fresh powder from the powder tank to the top of the previously built surface or the substrate. Then, a laser beam passes through a system of lenses and reflected by a mirror that controls the laser beam spot moving along the designed path. This process is repeated until the product is printed. Among AM techniques, L-PBF process tends to be the most common for metallic materials, due to their capability of producing parts with improved density, resolution, and surface finish that require less post-processing compared to other processes such as binder jetting [2]. One of the major challenges in L-PBF technique is defect formation in printed products.

The quality of the fused powder is affected by numerous processing parameters, including laser power, laser scan speed, laser beam size, powder size, layer thickness, and material absorptivity, especially those parameters related to energy input.

In laser processing field, there are two modes of heat transfer mechanisms related to defect formation: (1) conduction mode, in which the melting is controlled by heat conduction, and (2) keyhole mode, in which the input power density of the laser beam is sufficient to vaporize the metal and drill a much deeper cavity than in conduction mode [21]. Since the collapse of the cavities in keyhole mode may leave voids in the printed parts [22], the conduction mode is preferred in laser additive manufacturing. Therefore, for L-PBF process, if the processing parameters are not well tuned, the energy input can be either insufficient to cause lack of fusion issue, or overlarge to cause pores formation in keyhole mode. Consequently, the optimization of processing parameters in L-PBF is critical to for a robust design.

Forerunners have made lots of efforts in process optimization by experimental [8-11, 31] and computational study [14-17, 119-122]. For designers, on one hand, experiments and testing are helpful to determine the preferred conduction-controlled region of processing parameters, but they are also time-consuming and expensive. On the other hand, the modeling and simulation tools enable designers to make predictions of the printed products of L-PBF so as to optimize the processing parameters and decrease the need for conducting experiments. In addition, those mesoscopic simulation models are extremely helpful in understanding the physical meaning of the process. In spite of these advantages, models with high accuracy are usually computationally intensive. It is sometimes impractical to optimize the processing parameters using either experimental or computational study, especially when the number of involved processing parameters is large. Consequently, there is a need of a more efficient way for process optimization.

An attractive alternative to this dilemma is the application of the data-driven tool, *e.g.*, machine learning (ML) models. Machine learning can be defined as computer programming to optimize a performance criterion using example data or past experience [60]. In general, ML tasks can be divided into three groups, supervised learning, unsupervised learning, and reinforcement learning [60]. In AM field, the type of applied ML technique is usually the supervised learning, in which each input datum is labeled with an output, and predictions are made from an unobserved input to an output based on example input-output pairs. The input is a vector contains all involved processing parameters. The output can be either classes and categories such as defect types and

the corresponding ML task is classification, or parameters such as porosity and the corresponding ML task is regression. With a reliable training dataset, the ML models learn knowledge from the training set and make inference based on the knowledge. After training, ML models can make prediction at any new input points without making more observations, and thus reduce the need of experimental and computational study. In this work, the remelted depth will be used as a pointer of the quality of the printed parts, as it indicates how well the successive layers bond, whether the energy from the laser is sufficient, and whether the keyhole mode occurs. As the remelted depth, the output in this work, is a parameter, this work focuses on regression ML models. In a word, the objective of this chapter is to develop an efficient ML based method to assist process optimization.

Pioneers have applied supervised machine learning models to various AM regression tasks. Z. Zhu *et al.* [73] developed a Gaussian process (GP) regression model to make predictions on shape deviation based on the input geometry with an average R-square of around 0.97. G. Tapia *et al.* [74] developed a GP regression model to makes predictions using the experimental data of 17-4 PH stainless steel, and demonstrated how to determine the optimal processing parameters at desired porosity using GP model trained by an experimental dataset with high noise level. G. Tapia *et al.* [67] also applied the GP model to make predictions of remelted depth based on laser power and laser scan speed using experimental data of 316L stainless steel. Their model predicts consistent processing windows with the normalized enthalpy > 30 criterion to identify keyhole mode in L-PBF [10], which confirms the validation of their model. In these pioneers' efforts, GP shows its excellent regression performance in AM field. Neural network (NN), on the other hand, also plays a role in regression tasks in AM field. F. Caiazzo *et al.* [76] developed a NN regression model for trace geometry prediction and the root mean square errors (RMSE) is around 5%. J. Zhang *et al.* [78] developed a recurrent NN model in material extrusion process to predict the tensile strength of the printed products and the RMSE was around 2%. M. Mozaffar *et al.* [80] developed a recurrent NN model to predict the thermal history in directed energy deposition processes and achieved a mean square error of 2.97e-5. Overall, NN is a powerful model in regression tasks in AM field, but it also requires tuning a series of hyperparameters such as number of hidden neurons and layers [87], which makes NN complicated to apply. Therefore, GP will be applied as the regression model in this work.

In this chapter, a Gaussian process regression model is developed to assist the process design of L-PBF process. The whole chapter is divided into two parts. The first part, Section 4.2

and Section 4.3, focuses on developing and validating an efficient GP regression model which can generate a regression surface based on a training dataset and make predictions at any input settings. To obtain the training dataset, a computational fluid dynamics (CFD) model is applied to perform a set of single-track simulations of the melting process of laser powder bed fused 316L stainless steel. The remelted depth is then measured from the simulations at various combinations of laser scan speed (v) and laser power (P), which forms the training dataset. Subsequently, a GP regression model is developed to generate a regression surface to make predictions at the unobserved settings based on observed data from CFD. The n-fold cross validation is employed. The cross-predictions from the GP model are used to compare against experimental measurements for model validation. The second part, Section 4.4, demonstrates the applications of machine learning models in process optimization. Two datasets of experimental observations of 316L and 17-4 PH stainless steel are used to train the GP model. Then, the conduction mode regions of both materials on the plane of laser scan speed and laser power are predicted and compared against the regions computed by normalized enthalpy. The criteria for identifying keyhole mode for different materials can be obtained by this comparison. Designers can make use of these predicted criteria to optimize the processing parameters, thus reducing the need for experimental and computational observations.

4.2 Methodology

4.2.1 Gaussian Process Regression Model

A Gaussian process is defined as a collection of random variables, such that any finite number of which has a joint Gaussian distribution [123]. In this work, the GP is applied as a regression method, which is different from an interpolation method. A regression task is aiming at finding a regularized way to fit a certain dataset so as to make better predictions for future data, whereas an interpolation task is aiming at perfectly fitting the certain dataset regardless of whether it predicts well or bad for future data [60].

Similar to the mean value and variance in Gaussian distribution, a GP is completely specified by a mean function $m(\boldsymbol{X}_i)$ and a covariance function $C(\boldsymbol{X}_i, \boldsymbol{X}_j)$, where i and j range from 1 to n, n being the number of observation points. $\boldsymbol{X}$ is the input vector containing all processing parameters, and $\boldsymbol{X} = (P, v)$.

In this work, the GP model is expressed as:

$$f(\boldsymbol{X}) = GP\left(m(\boldsymbol{X}_i), C\left(\boldsymbol{X}_i, \boldsymbol{X}_j\right)\right). \tag{4.1}$$

The statistical model can then be defined as:

$$Y(\boldsymbol{X}) = f(\boldsymbol{X}) + e(\boldsymbol{X}), \tag{4.2}$$

where Y is the output parameter of interest (*i.e.*, the remelted depth d in this work), and e models the observation error which captures the inherent noise associated with the measurements. This observation error is usually treated as an independent term to the input vectors and represented by a Gaussian distribution with zero mean, *i.e.*, $e \sim \mathcal{N}(0, \sigma^2)$. In a CFD model, no random seed is involved, which means that the measured remelted depth remains unchanged no matter how many simulations are conducted in the same setting. Thus, this error term can be disregarded. This omission will be verified by the fact that the noise level term approaches 0 after optimization in the result section.

In GP, the covariance function is defined by a single or a combination of kernel functions, which capture the spatial dependence between two different locations, $\boldsymbol{X}_i$ and $\boldsymbol{X}_j$. The selection of kernel functions should be based on practical applications, and has been discussed in Ref. [123]. In this work, the two input dimensions, P and v, of the input vector $\boldsymbol{X}$ possess different units. Therefore, the Matérn 5/2 kernel is employed [123], which allows different lengthscales for different input dimensions. The white-noise regularization is also added to the covariance function, in order to capture the noise level. As a result, 4 kernel hyperparameters are involved in this GP model, including the variances of the two kernel functions and the lengthscales of the two input dimensions. The initial values of both variances are set as 1. The lengthscales of the laser scan speed and the laser power are reasonably set to be the mesh size, (i.e. 0.04, 10), respectively. These four hyperparameters are updated once a new observation is made and converge to the values that better predict future data. The GP regression model is performed using the GPy software package [124].

The performance of the GP model is quantitatively assessed by mean absolute prediction error (MAPE), which is defined as:

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|Y_{obs}(P_i, v_i) - Y_{pre}(P_i, v_i)\right|, \tag{4.3}$$

where n is the number of observation points from the CFD model.

4.2.2 Simulation of Melting Process Using Computational Fluid Dynamics Model

A coupled solid-fluid-thermal model [122] is applied to simulate the melting process of laser powder bed fused 316L stainless steel. In this model, the initial configuration of the powder bed is generated using the discrete element method (DEM) [16]. A laser heat source is applied to the top surface of the powder bed, moving along *x*-direction in order to make a single track. The power of the laser heat source is represented by the Gaussian distribution [122]:

$$S_{laser} = \frac{\alpha P}{\pi {r_0}^2} exp\left\{-\frac{2\left[(x-vt-x_i)^2+(y-y_i)^2\right]}{{r_0}^2}\right\}, \tag{4.4}$$

where P is the laser power, α is the absorption coefficient of the material, r_0 is the half of the laser beam diameter D4σ (D4σ is the beam diameter expressed by four times the standard deviation of the Gaussian distribution, and will be expressed as σ thereafter), v is the laser scan speed, and x_i, y_i is the initial position of the laser focal center. The laser power is projected to the top surface of the metal powders, such that all the heat source is absorbed by the top surface at the first interaction [19]. This laser model provides more physical phenomena, such as shadowing, than the volumetric heat source method. The reflection is ignored, a reasonable assumption as discussed in Ref. [122].

In this model, the conservation of mass, momentum, energy and volume is coupled with some complicated physical phenomena, including melting, buoyancy-driven flow, surface tension, Marangoni convection, and metal evaporation induced recoil pressure [122]. These conservation equations are modeled in User Defined Functions (UDF) in ANSYS Fluent. The laminar solver is applied with the assumption that the Reynolds number within the melt pool is small enough [125-127]. The mesh size is set as 3 μm [122]. This CFD model is validated in Ref. [122] by comparing the simulation results with experimental results at the same settings.

In this work, a set of simulations using the CFD model is conducted to generate an observation dataset, $\boldsymbol{Y}_{obs}(P, v)$, where $\boldsymbol{Y}_{obs}$ is the predicted remelted depth as a function of laser power P and laser scan speed v. The layer thickness and the laser beam diameter σ keeps constant as 50 μm and 70 μm, respectively, so as to be consistent with the experimental settings [128]. The remelted depth for a single track is defined as d in Figure 4.1. These observation points are then imported to the GP regression model as training data and validation data. Typically, the simulation time of a 1 mm single track simulation will be more than 20 hours in an Intel Core i7-8750H CPU with 32.0 GB installed memory [122]. The computational costs also reveal the necessity of the application of machine learning models to the process optimization.

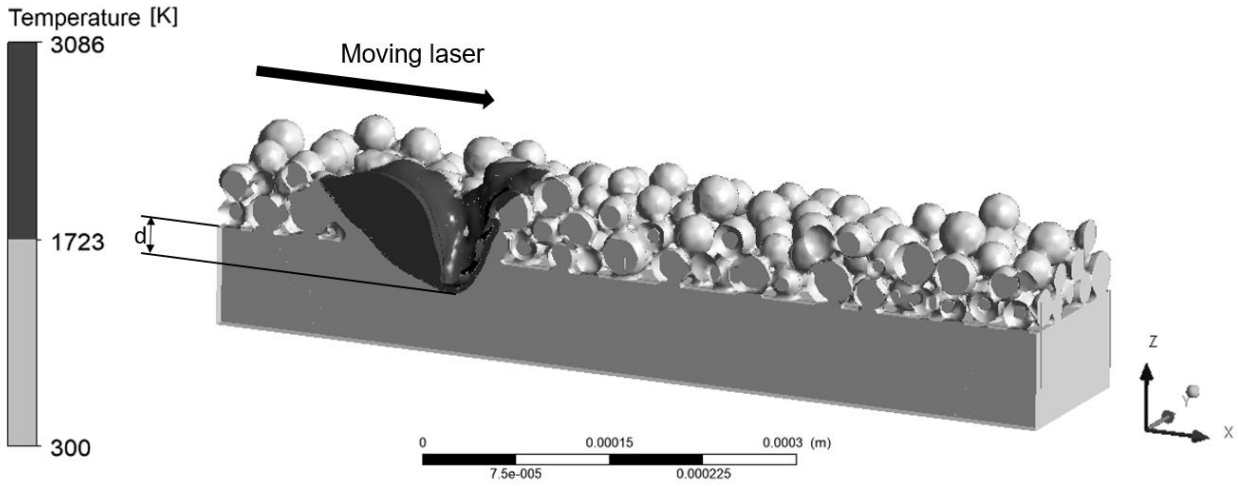


Figure 4.1. Remelted region (black) is shown in the temperature contour of the CFD model cut through view at the plane $y = 0$. 1723K is the liquidus of 316L stainless steel. The remelted depth *d* is illustrated as the average distance between the deepest melted region and the interface of the solid substrate and metal powders during the laser scanning.

4.2.3 Validation Using n-fold Cross-validation Method

Although the MAPE is able to quantitatively assess the performance of the machine learning model, a cross-validation method is still necessary because the MAPE will be always nearly 0 and make no sense if the GP model is trained by all data in a noiseless dataset. This is a characteristic of the GP models. In general, to assess the performance of a supervised learning model, the entire dataset is usually partitioned into at least two subsets, training set and testing set. The training set is used to optimize the hyperparameters of the model and make predictions at the same input settings of the testing data. The data in testing set will then be compared against the model predictions. If the model predictions are consistent with the testing set, then it confirms the validation of the machine learning model. In our case, the number of observation points is limited, as the CFD model is computationally expensive. Besides, a general trend of machine learning models is that the predictive performance has a positive correlation with the number of training data. Therefore, to make the utmost use of the limited data, the *n*-fold cross-validation method is employed in this work.

The *n*-fold cross-validation, also called leave-one-out cross-validation (LOOCV), is an iterative procedure which can monitor the overfitting issue and maximize the utilization of observation points. This is a special case of the widely known *k*-fold cross-validation method. In a *k*-fold cross-validation process, the entire dataset is randomly separated into *k* subsets of roughly

same size. In each iteration, one subset is left out as the testing set and all other subsets form the training set. This iteration is repeated until all subsets have been used as the testing set once. The *n*-fold cross-validation is the case when *k* is set as the number of observation points, or $k = n$. By this means, all data are fully utilized, and the intrinsic overfitting issue of GP is also resolved.

In this work, each GP prediction at a certain input setting, $\boldsymbol{X}_i = (P_i, v_i)$, is made by the model trained by all data except $\boldsymbol{X}_i$. Figure 4.2 shows the flow chart of our GP model with the *n*-fold cross-validation method, where *n* is the number of CFD observation points. These GP cross-predictions are compared against both the CFD observations and experimental measurements. Then, the difference between the MAPEs of these two comparisons will be the major performance indicator of our model, as it captures the error magnified by the Gaussian process model.

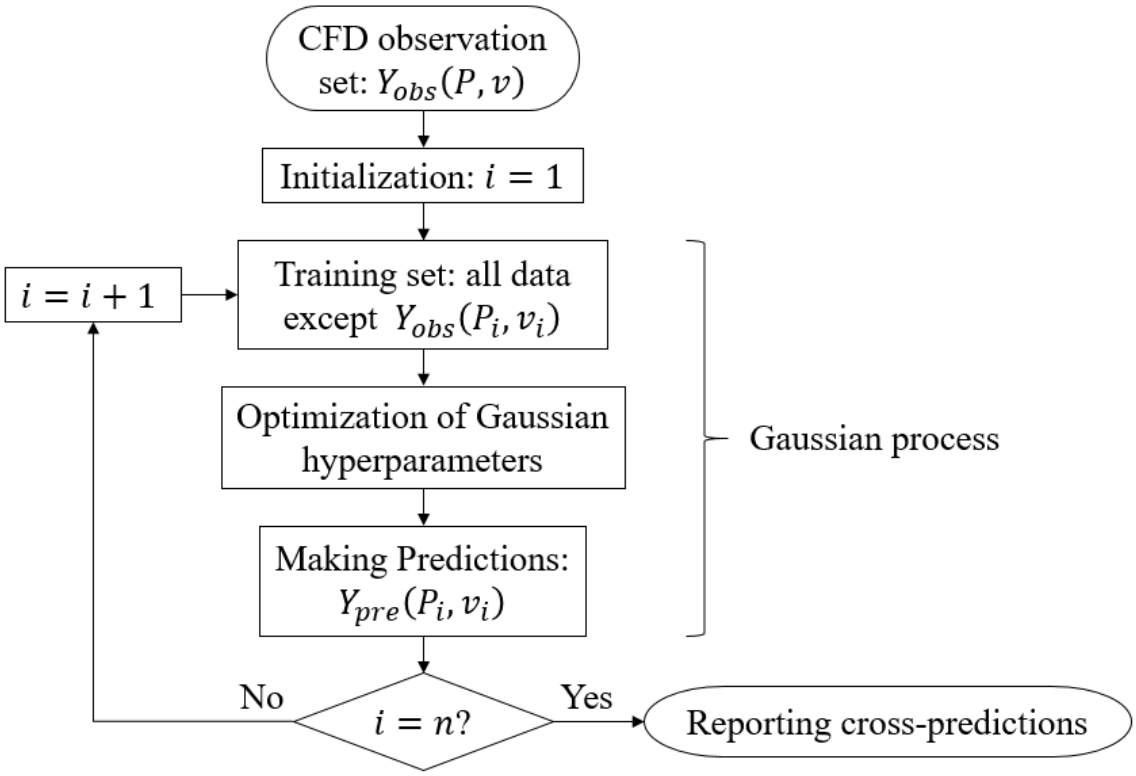


Figure 4.2. Flow chart of the GP model with *n*-fold cross-validation method.

4.3 Gaussian Process Model Results and Validation

4.3.1 Computational Fluid Dynamics Observations

In this work, 24 observation points in total are made using the CFD model, with laser power ranging from 30 W to 60 W, and laser scan speed from 0.08 m/s to 0.28 m/s. The resultant remelted

depth in each case is measured, as summarized in Table 4.1. Those data points with 0 remelted depth indicate that the melted region at this input setting doesn't get in touch with the previous layer. In other words, the energy input in these settings is insufficient to sinter different powder layers. These observation points are then imported into the GP regression model to generate a regression surface.

Table 4.1. CFD predictions and experimental results [32] at various combinations of laser scan speed and laser power

Laser scan speed (m/s)	Laser power (W)	Remelted depth (μm)	
		Experiment [32]	CFD (this work)
0.08	30	/	21.5
0.08	40	/	36.8
0.08	50	36.9	43.8
0.08	60	/	59.5
0.12	30	/	10.5
0.12	40	/	25.3
0.12	50	31.8	35
0.12	60	/	46.2
0.16	30	/	0.8
0.16	40	/	12.6
0.16	50	27	26.5
0.16	60	/	36.1
0.2	30	/	0
0.2	40	/	1
0.2	50	21.8	20.1
0.2	60	/	31.2
0.24	30	/	0
0.24	40	/	0
0.24	50	15.2	16.8
0.24	60	/	27.6
0.28	30	/	0
0.28	40	/	0
0.28	50	20.1	15
0.28	60	/	24

4.3.2 Gaussian Process Regression Surface

Using the GP regression model in Section 4.2.1 and the 24 observations points in Section 4.3.1, a regression surface is generated, as shown in Figure 4.3. This plot follows the general trend

that the remelted depth is increasing with increased energy input. The updated Matérn kernel variance, white kernel variance, laser scan speed lengthscale and laser power lengthscale are 1267.9, 3.5456×10^{-211}, 0.23962 and 28.517, respectively. It can be verified that the white kernel variance, which captures the noise level of the input dataset, is approaching 0 after optimization. This is the aforementioned expected result that the noise level of the CFD predictions is nearly 0 due to the inexistence of randomness in the computational models and can be reasonably disregarded.

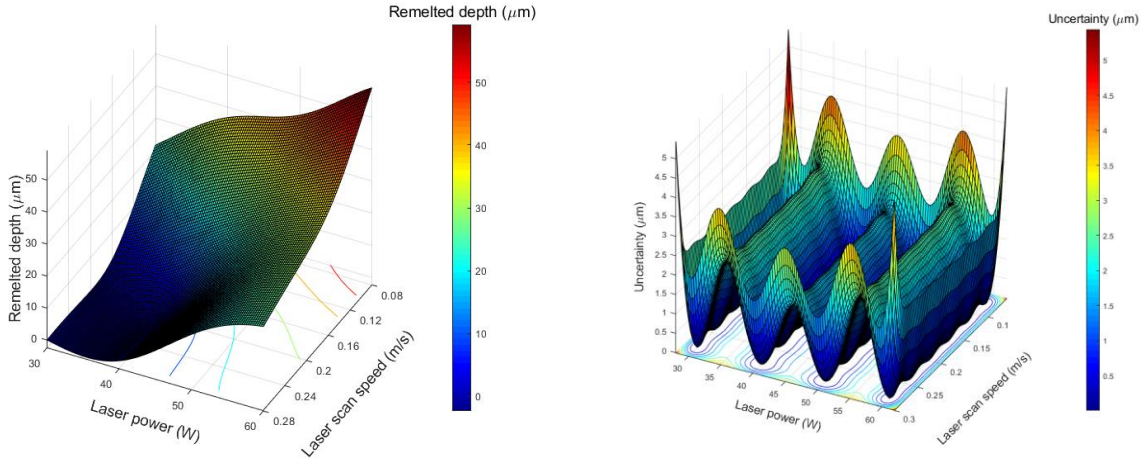


Figure 4.3. Regression surface of remelted depth as a function of laser power and scan speed made by the GP regression model (left) and the corresponding uncertainty based on the 95% confidence interval (right). The grid in left figure represents the mesh of the data points from CFD.

The regression surface in Figure 4.3 can then be used to predict the remelted depth with the corresponding uncertainty at any combinations of laser power and scan speed. Designers can make use of this process map to predict the remelted depth at any settings without conducting more experimental or computational study. The uncertainty within the range of training dataset are relatively low, whereas the uncertainty outside the training range will be larger due to the extrapolation error. Since all 24 observation points are used to train the GP model and generate the regression surface, it can be observed that the uncertainty at the input settings of the 24 training points approach 0. In other words, if the 24 training points are used again to test the performance of GP model, the model will make perfect predictions without any uncertainty. This is a

characteristic of GP models in a noiseless environment. This exposes the significance of the cross-validation method.

4.3.3 Cross-predictions and Validations

In this Section, the cross-validation method in Section 4.2.3 is employed. The *n*-fold cross-validation plot is presented in Figure 4.4. It represents a comparison between the cross-predictions made by the GP model and the CFD predictions. All data points are close to the ideal y = x line, with MAPE = 1.4 μm. The low MAPE indicates the adequacy of the GP predictions. The total computational time for the whole iterative process is less than one minute, indicating the high efficiency of the GP model.

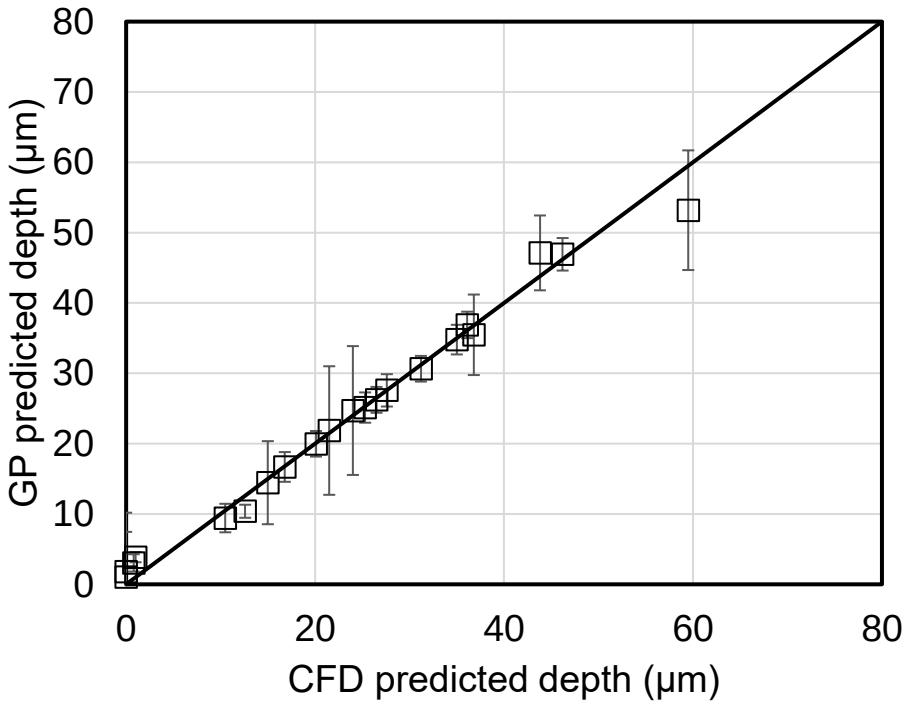


Figure 4.4. Results of the *n*-fold cross-validation using the GP model and CFD model data

Up to now, the adequacy of the GP model in regression tasks in AM field are validated. To finally validate our model, the cross-predictions made by the GP model need to be compared against experimental results. This step is critical, because the GP regression model coupled with high-fidelity simulation models will become an effective tool for designers to reduce the need of expensive and time-consuming experiments. The validation plot is shown in Figure 4.5, where the

GP predicted depths are compared against experimentally observed depths. The quality of the predicted values is not as high as the CFD model as expected, but the GP model still provides a reasonable prediction.

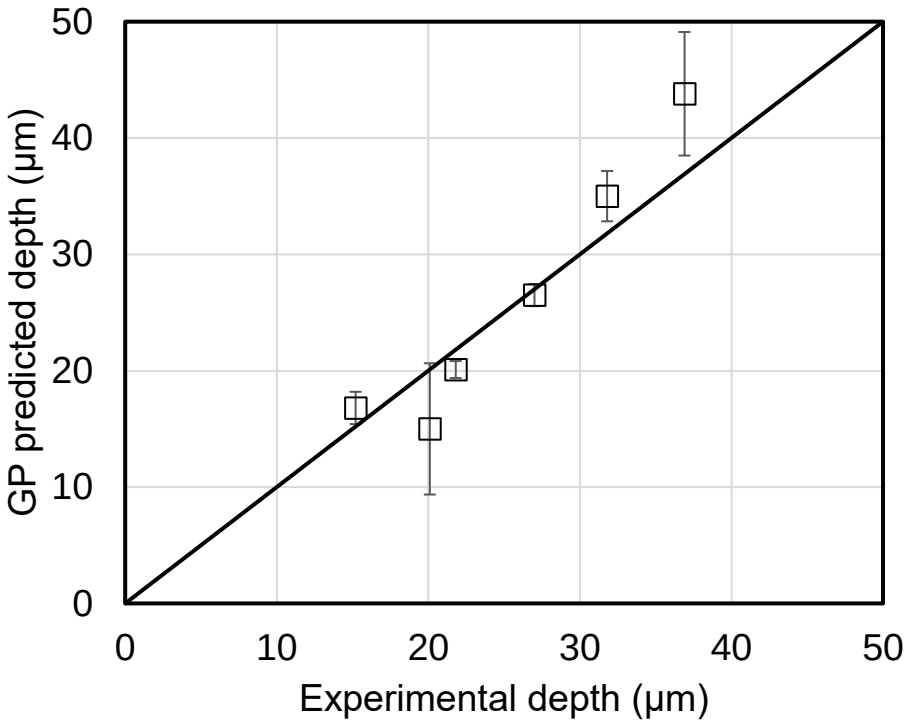


Figure 4.5. Validation plot of GP predictions and experimental results

Figure 4.6 shows the remelted depth comparison among the GP predictions, experimental results [128], and CFD predictions. For all the cases, with an increased speed and other parameters fixed, the energy input is decreased and should have led to a decreased remelted depth. In the experimental dataset, the laser scan speed is ranging from 0.08 m/s to 0.28 m/s with a fixed laser power of 50 W. It should be noted that the noise level of the experimental dataset is not 0, which can be observed between the data point of 0.24 m/s and 0.28 m/s.

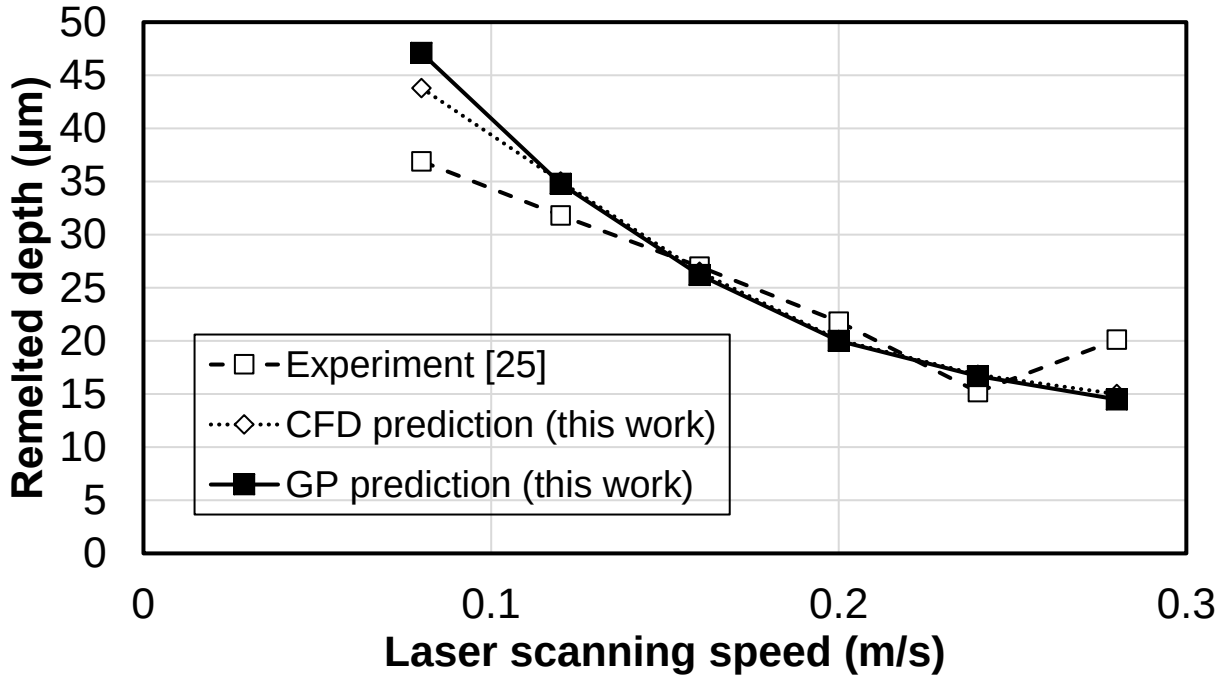


Figure 4.6. Comparison of the remelted depth among GP predictions, experimental results [128], and CFD predictions

In terms of the performance of the GP model, the MAPE between the GP predictions and experimental data is 3.8 μm, whereas the MAPE between the CFD predictions and experimental data is 3.2 μm. Since the GP model is trained by CFD predictions, the fact that the MAPE is only magnified by 0.6 μm, indicating the excellent performance of the GP model. It is worth mentioning that those data points with large error bars are all lying between the boundaries of the input range. The large standard variations are due to the extrapolation errors of the GP regression model. It can be avoided by conducting additional observations until these data points of interest are not lying at the boundary. Overall, the GP predictions are in good agreement with the experimental results. This confirms that the GP model developed in this work is capable of generalizing the predictions made by the CFD simulations to any other processing parameter settings within the predefined design space.

The predicted error magnified by GP is considerably small, indicating that as long as the training data is reliable, the GP model will be adequate for making accurate predictions, even with only a limited number of computationally expensive simulations.

4.4 Process Map and Process Optimization Using Gaussian Process Model

In this section, the process maps of two metals, 316L and 17-4 PH stainless steel, are developed using our Gaussian process regression model. According to previous experimental and computational studies [3, 10, 23, 24], the basic criterion of identifying the keyhole mode is

$$\frac{2d}{w} \geq 1, \tag{4.5}$$

where d is the remelted depth , w is the melt pool width, and the quantity, $\frac{d}{w}$, is the normalized depth. With a reliable dataset of the normalized depth with respect to laser power and laser scan speed, the GP model is able to predict the conduction mode region where $\frac{2d}{w} < 1$, and keyhole mode where $\frac{2d}{w} \geq 1$.

Besides this criterion, Hann *et al.* [25] proposed another empirical method to identify the keyhole mode. By experimental results, they first found that the normalized depth is proportional to the product of power density multiplied by the square root of the laser interaction time, such that:

$$\frac{d}{w} \propto P_d\sqrt{\tau}, \tag{4.6}$$

where $P_d = \frac{P}{\pi\sigma^2}$ is the power density, and $\tau = \frac{2\sigma}{v}$ is the laser interaction time. Then, the slope of the relationship was determined by dimensional analysis. Finally, the normalized enthalpy term is proposed to identify the keyhole mode is based, such that:

$$\frac{\Delta H}{h_s} = \frac{AP}{\rho h_s\sqrt{\pi D v\left(\frac{\sigma}{4}\right)^3}}, \tag{4.7}$$

where ΔH is the specific enthalpy, h_s is the enthalpy at melting, A is the material absorptivity of the laser power, P is the laser power, ρ is density, D is thermal diffusivity, v is the laser scan speed, σ is the laser beam size D4σ. After that, King *et al.* [10] quantified a threshold valud of 30 for 316L stainless steel based on experimental study. Therefore, it predicts the occurance of keyhole mode when $\frac{\Delta H}{h_s} \geq 30$, and condcution mode when $\frac{\Delta H}{h_s} < 30$. As all the involved terms in this criterion are processing parameters, it enables designers to identify the preferred conduction mode in L-PBF conveniently. However, this criterion is an empirical method, it is crucial to test it to other materials before applying it.

In this work, two sets of experimental data of 316L and 17-4 PH stainless steels will be used to train our GP model. To apply the $\frac{2d}{w} \geq 1$ criterion, the output of the training dataset has changed from remelted depth to normalized depth, $\frac{d}{w}$. After training, GP model will then make predictions on the region of keyhole mode using the $\frac{2d}{w} \geq 1$ criterion. The GP predicted region will then be compared against the region predicted by the normalized enthalpy criterion, thus testing the adequacy of the criterion.

4.4.1 Case 1: 316L Stainless Steel

The normalized enthalpy criterion of 316L stainless steel is tested in this section. Table 4.2 shows the experimental dataset from Kamath *et al.* [129]. Table 4.3 shows the material properties used for the melt pool size and enthalpy calculations.

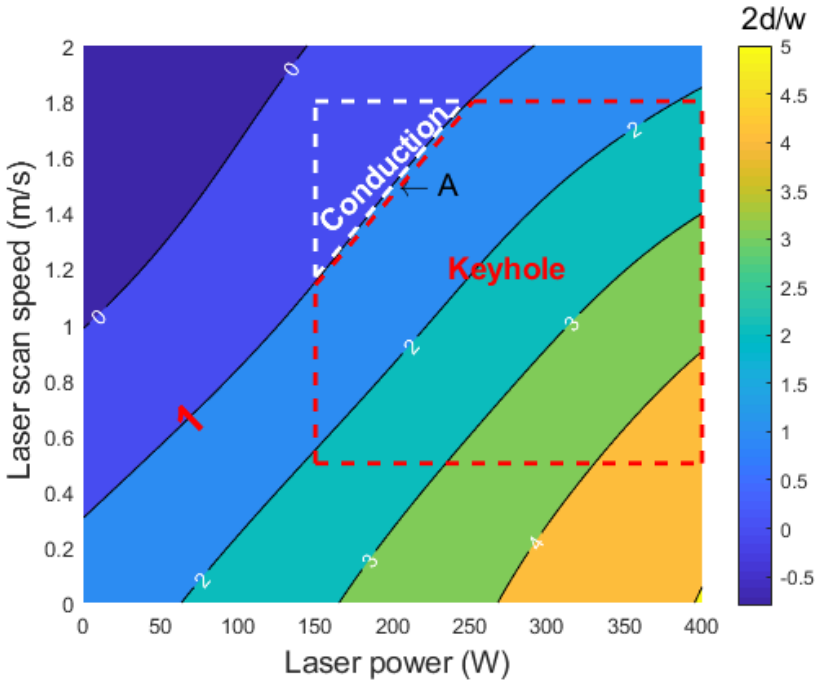
Table 4.2. Experimental data of 316L stainless steel from Ref. [36]

Laser power (W)	Laser scan speed (m/s)	Remelted depth (μm)	Melt pool width (μm)
400	1.8	105	112
400	1.5	119	103
400	1.2	182	83
300	1.8	65	94
300	1.5	94	83
300	1.2	114	111
300	0.8	175	118
200	1.5	57	84
200	1.2	68	104
200	0.8	116	123
200	0.5	195	121
150	1.2	30	79
150	0.8	67	109
150	0.5	120	115

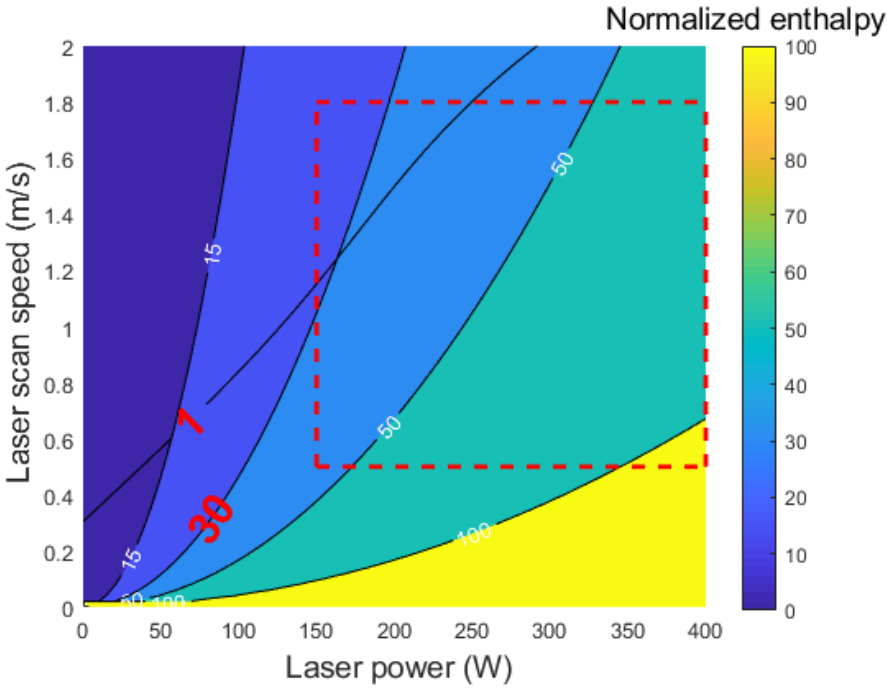
Table 4.3. Property values of 316L and 17-4 PH stainless steels used for calculations

Property	316L SS [reference]	17-4 SS [reference]
A, absorptivity	0.4 [6]	0.4 [6]
ρ, density (kg/m^3)	7980 [6]	7740 [37]
h_s, enthalpy at melting (kg/m^3)	1.2×10^6 [6]	8.84×10^5 [37]
D, thermal diffusivity (m^2/s)	5.38×10^{-6} [6]	5.13×10^{-6} [37]
σ, laser beam size (μm)	52 [36]	100 [34]

Using the trained GP model with the normalized depth computed from the dataset in Table 4.2, a contour plot of the quantity, $\frac{2d}{w}$, can be obtained, as shown in Figure 4.7 (a). The preferred conduction mode region can be readily identified from Figure 4.7 (a). For example, the point labeled with "A" marked in the figure illustrates how to use it for process optimization: if other settings are kept the same as the experiments and the expected laser scan speed is 1.5 m/s, then the laser power should not exceed 204 W. Note that though the contour plot displays the entire range of laser scan speed and laser power, the predictions outside the marked boundary or extrapolated by the GP model may accompany with relatively large uncertainty and should be used with caution.



(a)



(b)

Figure 4.7. Process map of 316L stainless steel showing the keyhole mode and conduction mode regions. (a) Using $\frac{2d}{w} \leq 1$ criterion. (b) Using normalized enthalpy criterion. The red dashed square represents the range of the laser parameters used in the experiment. In (a), the dark blue region is the preferred conduction mode region. In (b), the contour line in red text is $\frac{2d}{w} = 1$ from (a), which is to compare against the normalized enthalpy criterion.

The conduction mode region in Figure 4.7 (a) is then compared against Figure 4.7 (b), which is the contour plot of the normalized enthalpy calculated by Eq. (4.7). The boundary separating the two modes, *i.e.*, the contour line of $\frac{2d}{w} = 1$ in Figure 4.7 (a), is also plotted in Figure 4.7 (b) to better compare the two criteria. Within the red dashed square region, the normalized enthalpy criterion provides acceptable predictions with a little deviation. Therefore, the empirical normalized enthalpy criterion can distinguish the two heat transfer modes, but additional experimental data points are recommended when normalized enthalpy is around 30.

It should be noted that Tapia *et al.* [67] also conducted a similar test to the normalized enthalpy criterion of 316L stainless steel using a GP model. In their result, the predictions made by both criteria closely matched. However, there are several differences between our work. First, the dataset they used to train their GP model contained those which used to quantify the $\frac{\Delta H}{h_s} = 30$ criterion [10]. Therefore, the matching of both predictions verifies the validation of their GP model, but cannot confirm the accuracy of the normalized enthalpy criterion. Second, the covariance function of GP they used was the power exponential kernel, whereas in our model the Matérn 5/2 kernel is employed. While in most case the selection of kernel functions doesn't significantly affect the GP predictions, it can be a potential source of small deviation. Third, the powder layer thickness of our dataset is 30 μm, whereas the layer thickness of their dataset was 50 μm. Consequently, our result exposes that the normalized enthalpy criterion is dependent on the layer thickness, which is currently not included in its formula. Hence, the normalized enthalpy criterion for 316L stainless steel is in need of calibration with the consideration of layer thickness.

4.4.2 Case 2: 17-4 PH Stainless Steel

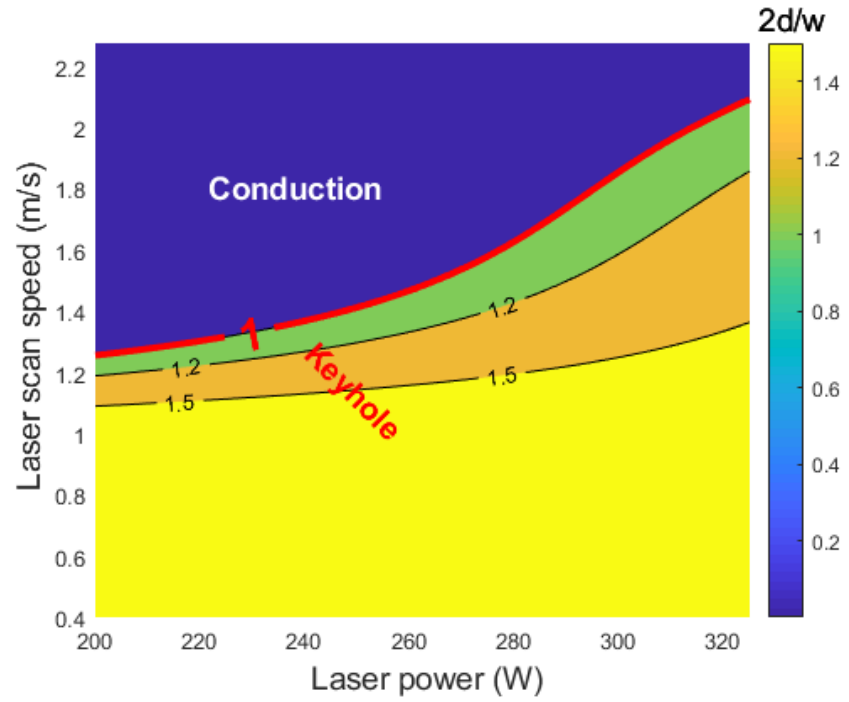
This section tests the normalized enthalpy criterion of another material, 17-4 PH stainless steel. The training dataset is displayed in

Table 4.4. Similar to the case of 316L stainless steel, the contour plot of the quantity $\frac{2d}{w}$ is generated using the GP model trained by the dataset in

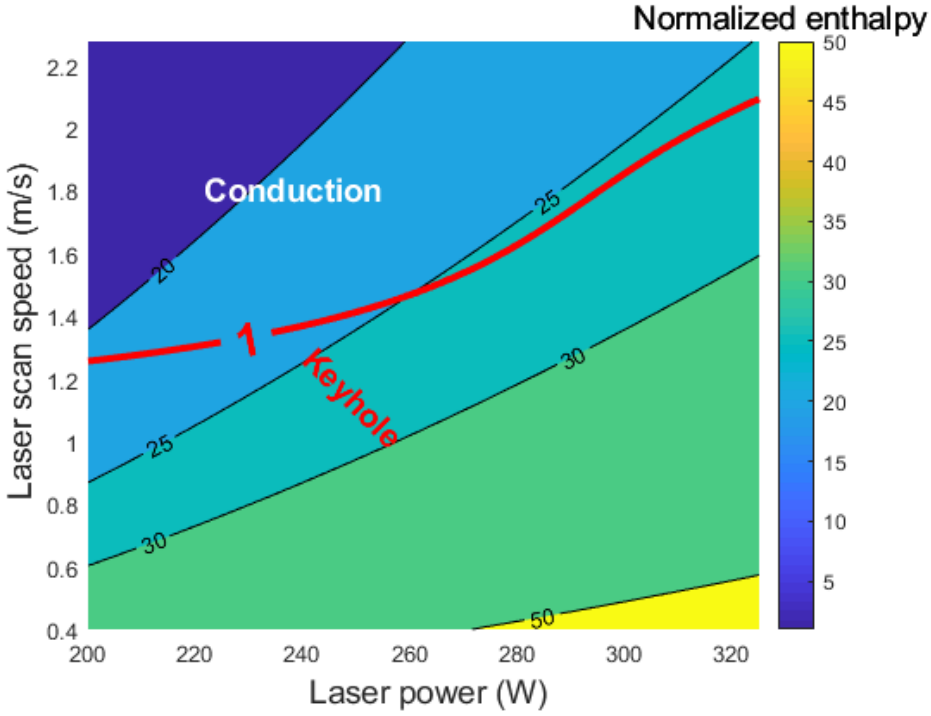
Table 4.4, shown in Figure 4.8 (a). It should be noted that in this case, the number of levels of laser power in training dataset is only two, which may result in a larger uncertainty of the predictions along the path changing the laser power. In addition, the range of the laser scan speed and laser power displayed in Figure 4.8 (a) is set to be the same as the experimental dataset.

Table 4.4. Experimental observations of 17-4 PH stainless steel [34]

Laser power (W)	Laser scan speed (m/s)	Remelted depth (μm)	Melt pool width (μm)
200	1.6	52	75
200	1.4	49	67
200	1.2	35	58
200	1	75	82
200	0.8	108	100
200	0.6	123	87
200	0.4	248	122
325	2.275	31	83
325	1.95	47	79
325	1.625	58	72
325	1.3	82	106
325	0.975	111	124
325	0.65	149	146



(a)



(b)

Figure 4.8. Process map of 17-4 PH stainless steel showing the keyhole mode and conduction mode regions. (a) Using $\frac{2d}{w} \geq 1$ criterion. (b) Using normalized enthalpy criterion. In (a), the dark blue region is the preferred conduction mode region. In (b), the contour line in red is $\frac{2d}{w} = 1$ from (a), which is to compare against the normalized enthalpy criterion.

Figure 4.8 (b) shows the process map using the normalized enthalpy. It can be observed that the contour line of $\frac{2d}{w} = 1$ is close to $\frac{\Delta H}{h_s} = 25$, indicating that the normalized enthalpy criterion should be revised to $\frac{\Delta H}{h_s} \geq 25$ for 17-4 PH stainless steel. It also suggests that the normalized enthalpy criterion is sensitive to the selection of the materials. Therefore, to make use of the convenient normalized enthalpy criterion, calibration is still needed when switching materials. In summary, machine learning models can be applied to process design. Machine learning models trained by reliable data can predict the preferred conduction mode region and will be very helpful to calibrate this criterion and thus take advantage of it.

4.5 Summary

In this study, a Gaussian process regression model is developed to assist the process optimization of the L-PBF additive manufacturing process. The major findings are summarized as follow:

1. The predictions made by the trained GP model are in good agreement with the experimental observations. With a limited number (24) of data points, the mean absolute prediction error is only magnified by 0.6 μm using the GP model for a powder layer thickness of 30 μm, and can be further decreased by obtaining more training data.
2. The process maps of two metals, 316L and 17-4 PH stainless steel, are developed using our trained Gaussian process regression model. The preferred conduction mode regions of laser power and laser scan speed for 316L and 17-4 PH stainless steel are predicted by the GP model trained by experimental datasets using the $\frac{2d}{w} < 1$ criterion.
3. The normalized enthalpy criterion of identifying keyhole mode is evaluated for both 316L and 17-4 PH stainless steel. For 316L, the result suggests that the $\frac{\Delta H}{h_s} \geq 30$ criterion should be related to the powder layer thickness. For 17-4 PH, the criterion should be revised to $\frac{\Delta H}{h_s} \geq 25$.

5. PROBABILISTIC FEASIBILITY DESIGN OF LASER POWDER BED FUSION PROCESS USING INTEGRATED FIRST ORDER RELIABILITY AND MONTE CARLO METHODS

Abstract: Quality inconsistency due to uncertainty hinders the extensive applications of laser powder bed fusion (L-PBF) additive manufacturing process. To address this issue, this study proposes a new and efficient probabilistic method for the reliability analysis and design of the L-PBF process. The method determines a feasible region of the design space for given design requirements at specified reliability levels. If a design point falls into the feasible region, the design requirement will be satisfied with a probability higher or equal to the specified reliability. Since the problem involves the inverse reliability analysis that requires calling the direct reliability analysis repeatedly, directly using Monte Carlo simulation (MCS) is computationally intractable, especially for a high reliability requirement. In this chapter, a new algorithm is developed to combine MCS and the First Order Reliability Method (FORM). The algorithm finds the initial feasible region quickly by FORM and then updates it with higher accuracy by MCS. The method is applied to several case studies, where the normalized enthalpy criterion is used as a design requirement. The feasible regions of the normalized enthalpy criterion are obtained as contours with respect to the laser power and laser scan speed at different reliability levels, accounting for uncertainty in seven processing and material parameters. The results show that the proposed method dramatically alleviates the computational cost while maintaining high accuracy. This chapter provides a guidance for the process design with required reliability.

5.1 Introduction

Additive manufacturing (AM) is a group of layer-upon-layer fabrication processes based on computer-aided design (CAD) models [130]. AM has become a popular fabrication method of metal components due to its capability of producing parts with complex geometries. Among the processes of metal components, laser powder bed fusion (L-PBF) is a popular AM technique that uses a high-energy laser to selectively melt or sinter a metallic powder bed. For each layer in a typical L-PBF process, a recoating blade first pushes a layer of fresh powder from the powder tank to the top of the previously built surface or the substrate. Then, a laser beam passes through a

system of lenses and is reflected by a mirror that controls the laser beam spot moving along the designed path. This process is repeated in a layer-by-layer manner [1].

A major challenge in the current L-PBF field is the process variability and quality inconsistency. This is primarily due to the physical complexity involved in the laser melting process and the uncertainty involved in the processing parameters and material properties. To resolve this issue, efforts have been made in many aspects, including parametric experimental study [10, 13, 31, 128], *in situ* experimental monitoring [98, 100, 102], process modeling [19, 120-122, 131], surrogate modeling [67, 71, 89], and uncertainty quantification (UQ) [68, 132-134]. A recent review of the latest applications of machine learning (ML) in the AM field is also available in the authors' work [135].

To date, the mechanisms of laser melting process have been extensively studied from experiments and simulations, which provide the guidance for selecting processing parameters in the process-structure-property (P-S-P) relation chain. For example, two major causes of pores formation, keyhole mode [10] and spattering [28, 29], are found to be related to the laser energy density controlled by some processing parameters, such as laser power and scan speed. With this guidance, surrogate models, usually based on machine learning, can then generate a process map with a clear relationship between processing parameters and target properties or performance indicators such as melt pool dimension. Built upon this knowledge, the uncertainty, the main cause of process inconsistency, unpredictability and unrepeatability, has been investigated using Monte Carlo simulation (MCS) [118] and polynomial chaos expansions (PCE) [132]. Despite of this effort, the link between uncertainty quantification and optimal process design is still missing. Efficient tools for reliability analysis and process design in the L-PBF field are needed.

With the above discussions in mind, we propose a new probabilistic method that links uncertainty quantification with the L-PBF process design and reliability. Instead of directly quantifying uncertainty and reliability, we specify a feasible design region for given design requirements at a specified reliability level. The feasible design region allows researchers to select process design variables in the feasible design region so that the required reliability can be ensured. It also provides constraints on design variables when optimization is used.

The structure of the chapter is arranged as follows. Probabilistic design and reliability analysis methods are reviewed in Section 5.2. Then the proposed new method using the combined First Order Reliability Method and Monte Carlo simulation is presented in Section 5.3, followed

by a case study in Section 5.4, where a criterion based on normalized enthalpy [10, 25] is used to identify the conduction mode in the L-PBF process. The feasible design regions are determined by the given reliability levels. Section 5.5 provides conclusions and suggested future work.

5.2 Review of Probabilistic Design and Reliability Analysis

This work focuses on the probabilistic process design of the L-PBF process and is based on FORM and MCS. These methods are briefly reviewed here. The normalized enthalpy criterion is also discussed, since it is the design requirement in the case study for the L-PBF design.

5.2.1 Probabilistic Design

In an L-PBF process, the design and manufacture of a component are subject to one or multiple requirements, such as strength, geometric accuracy, and other desired properties. These quantities (output) corresponding to the requirements are determined by design variables (input), such as the laser power and scan speed. Since uncertainty exists in the input variables, the requirements may not always be satisfied. If the probability of satisfying a design requirement is denoted by reliability, the reliability may not be always 100% or 1.0. A typical probabilistic process design identifies the design variables so that the reliability is equal to or higher than the target reliability [136-138].

Probabilistic design uses limit-state functions. A limit-state function specifies the functional relationship between an output variable Y and input variables $\mathbf{X}$ and is given by [139]

$$Y = g(\mathbf{X}), \tag{5.1}$$

where Y is the output quantity of interest, and $\mathbf{X} = (X_1, X_2, \dots, X_n)$ is a vector of input variable with n being the size of $\mathbf{X}$. When uncertainty is associated with the input variables, they are modeled as random variables with their probability density function (PDF) $f_{X_i}(x_i),\ i = 1, \dots, n$. In this study, we assume the input random variables are independent. Without losing generality, we assume that a design requirement is satisfied if $Y = g(\mathbf{X}) > 0$; in other words, a failure occurs if $Y = g(\mathbf{X}) \leq 0$. The reliability R is then given by

$$R = \Pr\{g(\mathbf{X}) > 0\}, \tag{5.2}$$

where $\Pr\{\cdot\}$ denotes a probability.

For the L-PBF process, Y can be one of the properties of the printed parts, as well as some quality indicator such as the melt pool dimension. $\mathbf{X}$ usually includes processing parameters, such as laser power, laser scan speed, laser beam size, and layer thickness, as well as powder properties, including absorption coefficient, powder size distribution, and thermal conductivity.

A limit-state function is usually derived from physical principles, and it may be a black-box model. It may also be obtained from experiments. If the limit-state function, such as a finite element analysis model, is computationally expensive, its surrogate can be constructed. The computational cost of a surrogate model is low, typically in seconds or minutes. For a high dimensional and highly nonlinear limit-state function, the surrogate model can be built by machine learning methods [67, 71, 89].

5.2.2 Uncertainty Quantification and Reliability Analysis

During the probabilistic design, the reliability in Eq. (5.2) should be calculated, and this is the task of reliability analysis, whose objective is to find R given PDFs $f_{X_i}(x_i), i = 1, \dots, n$ and $g(\mathbf{X})$. Data acquisition is essential to constructing a limit-state function $g(\mathbf{X})$ and estimating PDFs $f_{X_i}(x_i), i = 1, \dots, n$. One way for data acquisition is to use experimental measurements, and this may be time consuming and costly. Another way is by means of computational models, which could alleviate the need of the expensive experiments; but a model still needs experimental validation. It should be noted that experimental measurements will introduce aleatory (random) uncertainty whereas computational models will have their model uncertainty [133]. The uncertainty sources and data acquisition are briefly discussed below, and will also be elaborated further for the case studies in Section 5.4.1.

The sources of uncertainty involved in the L-PBF process have been summarized in Ref. [133]. They can be classified into two categories: (1) aleatory uncertainty, which refers to the irreducible natural variability, and (2) epistemic uncertainty, which refers to the reducible uncertainty due to lack of knowledge of parameters, as well as approximations and assumptions introduced during the modeling process. The epistemic uncertainty can be further classified into two subgroups: (1) data uncertainty due to the imprecise measurements, and (2) model uncertainty due to the assumptions, simplifications, and numerical discretizations involved in the model [133].

An example is the uncertainty in the laser power and scan speed, which strongly affect the input energy density and the melt pool dynamics, due to the variation of the laser 3D printer. Another example is the imprecise measurements of the material properties, such as the absorption coefficient to the laser energy and thermal diffusivity. There are many uncertainty and reliability analysis methods, and two of them are reviewed in Sections 5.2.3 and 5.2.4.

5.2.3 Monte Carlo Simulation

Monte Carlo simulation (MCS) [68, 118, 132, 133, 140] is a numerical algorithm that relies on repeated random sampling. N samples of $\mathbf{X}$ are first generated from the distributions of $\mathbf{X}$, and then the limit-state function is evaluated at the samples of $\mathbf{X}$, resulting in samples of Y. The number of failures n_f is obtained by counting the number of samples of Y in the failure region $Y < 0$. The probability of failure is estimated by

$$p_f = \frac{n_f}{N}, \tag{5.3}$$

and the reliability is given by $R = 1 - p_f$. It is easy to use MCS, but with a computational cost when the probability of failure is low. For instance, if $p_f = 10^{-6}$, for sufficient accuracy, 10^9 samples are needed, and this means that the limit-state function has to be called 10^9 times.

5.2.4 First Order Reliability Method (FORM)

FORM is one of the most commonly used reliability methods [139]. It predicts the reliability $= \Pr\{g(\mathbf{X}) > 0\}$, given the limit-state function and distributions of $\mathbf{X}$. In this work, we use the inverse FORM [141]. It solves an inverse problem: find the limit state $g(\mathbf{X}) = c$, given the limit-state function, distributions of $\mathbf{X}$, and reliability R, so that

$$\Pr\{g(\mathbf{X}) > c\} = R. \tag{5.4}$$

The key idea is to linearize the limit-state function at the most probable point (MPP) so that the error of linearization can be minimized. The first step is to transform the random variables X_i $(i = 1, \dots, n)$ into standard normal variables U_i by

$$F_{X_i}(X_i) = \Phi(U_i), \tag{5.5}$$

where $F_{X_i}(X_i)$ is the cumulative density function (CDF) of X_i, and $\Phi(U_i)$ is the CDF of U_i.

Denoting the transformation from the X-space to U-space by $\mathbf{X} = \mathrm{T}(\mathbf{U})$, we have the limit-state function in the U-space space as follows:

$$Y = g(\mathbf{X}) = g\big(\mathrm{T}(\mathbf{U})\big). \tag{5.6}$$

The MPP $\mathbf{u}^*$ is found by solving the following optimization problem:

$$\begin{cases} \min\limits_{\mathbf{u}} g\big(\mathrm{T}(\mathbf{U})\big) \\ \text{subject to } \|\mathbf{u}\| = \beta \end{cases}, \tag{5.7}$$

where $\|\cdot\|$ denotes the magnitude of a vector, and β is the reliability index, given by

$$\Phi(\beta) = R. \tag{5.8}$$

The limit state is then found by

$$c = g\big(\mathrm{T}(\mathbf{u}^*)\big). \tag{5.9}$$

Since is $\mathbf{u}^*$ is the shortest distance point to $g\big(\mathrm{T}(\mathbf{U})\big) = c$, $\mathbf{u}^*$ is perpendicular to $g\big(\mathrm{T}(\mathbf{U})\big) = c$. In other words, $\mathbf{u}^*$ is in the opposite direction of the gradient of $g\big(\mathrm{T}(\mathbf{U})\big)$ at $\mathbf{u}^*$, where $\nabla g\big(\mathrm{T}(\mathbf{U})\big) = \left(\frac{\partial g}{\partial U_1}, \dots, \frac{\partial g}{\partial U_n}\right)\Big|_{\mathbf{u}^*}$. This gives

$$\frac{\mathbf{u}^*}{\beta} = -\frac{\nabla g(\mathrm{T}(\mathbf{U}))}{\|\nabla g(\mathrm{T}(\mathbf{U}))\|} = -\alpha(\mathbf{u}^*), \tag{5.10}$$

where

$$\alpha(\mathbf{u}^*) = \frac{\nabla g(\mathrm{T}(\mathbf{U}))}{\|\nabla g(\mathrm{T}(\mathbf{U}))\|}, \tag{5.11}$$

which is a unit vector along the gradient direction.

The goal of FORM now becomes searching for the MPP point $\mathbf{u}^*$. According to Eq. (5.10), the MPP point $\mathbf{u}^*$ satisfies

$$\mathbf{u}^* = -\beta\alpha(\mathbf{u}^*). \tag{5.12}$$

The MPP search process is iterative, and the search algorithm includes the Hasofer-Lind and Rackwitz-Fiessler (HL-RF) algorithm [139, 142]. For the work in this study, since β is known, for the *i*-th iteration, we have

$$\mathbf{u}_i = -\beta\alpha(\mathbf{u}_{i-1}). \tag{5.13}$$

5.2.5 Normalized Enthalpy Criterion

To use probabilistic analysis and design, we need to create limit-state functions. We now discuss the construction of limit-state functions for the L-PBF process. We take the normalized

enthalpy criterion as an example, which was proposed by Hann *et al.* [25] and quantified by King *et al.* [10] for keyhole mode identification for 316L stainless steel. The criterion was later applied in a Gaussian process (GP)-based machine learning model in Ref. [89] to predict the remelted depth of single tracks, as a function of combined laser power and laser scan speed in a laser powder bed fusion process.

The criterion of the occurrence of keyhole mode is

$$\frac{\Delta H}{h_s} = \frac{AP}{\rho h_s \sqrt{\pi D v \left(\frac{r}{2}\right)^3}} \geq 30, \tag{5.14}$$

where ΔH is the specific enthalpy, h_s is the enthalpy at melting, A is the material absorption coefficient to the laser power, P is the laser power, ρ is density, D is thermal diffusivity, v is the laser scan speed, r is the laser beam radius. The preferred conduction mode region is $\frac{\Delta H}{h_s} < 30$.

The limit-state function according to the criterion is constructed by

$$Y = g(\mathbf{X}) = g(A, P, \rho, h_s, D, v, r) = 30 - \frac{AP}{\rho h_s \sqrt{\pi D v \left(\frac{r}{2}\right)^3}}, \tag{5.15}$$

where $\mathbf{X} = (A, P, \rho, h_s, D, v, r)$. if $Y > 0$, we would have the preferred conduction mode region, and if $Y \leq 0$, we would get a failure if we consider the keyhole mode as a failure. Note that in reality a state in keyhole mode may not be necessarily a failed state, but it is the state we want to avoid. If the laser scan speed v and laser power P are the two processing parameters we can control during L-PBF process, we can plot the safe (conduction) region and failure (keyhole) region specified by the limit-state function as shown in Figure 5.1.

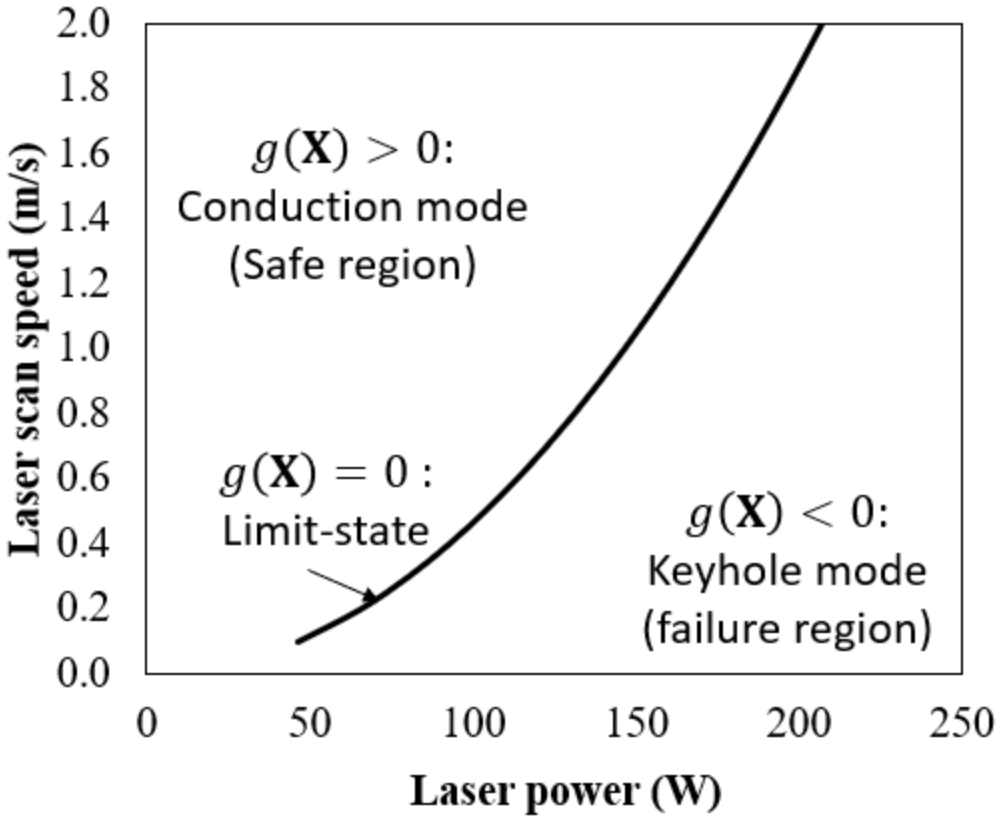


Figure 5.1. Limit-state function constructed by the normalized enthalpy criterion.

To avoid the keyhole mode, we can select the design variables (v, P) in the safe region. The curve that divides the safe and failure regions, however, is not fixed due to the randomness in **X**. To accommodate the uncertainty, we develop a probabilistic method, so that the safe region can ensure reliability equal to or higher than the target reliability.

5.3 A Probabilistic Method for Identifying a Feasible Design Region

The objective of this work is to identify a feasible design region for a given requirement at a desired reliability level. If a design point falls into the feasible design region, the probability of satisfying the requirement will be satisfied with a probability that is no less than the desired reliability. Identifying the feasible design region is extremely computationally expensive if we use MCS to do so. This is the reason in this study that we develop an efficient method to generate a feasible design region for process design with sufficient accuracy.

5.3.1 Overview

A limit-state function $Y = g(\mathbf{X})$ is available for a design requirement $Y > 0$, and the input variables $\mathbf{X}$ are random. To satisfy the requirement $Y = g(\mathbf{X}) > 0$ at reliability $R = \Pr(g(\mathbf{X}) > 0)$, we need to change the design variables, which are part of $\mathbf{X}$. Let the design variables be $\mathbf{d} = (d_1, d_2, \dots, d_m)$, which are deterministic variables, for instance, the means of some random variables in $\mathbf{X}$, with m being the number of design variables. Taking the keyhole mode requirement as an example, we could include the means of the laser scan speed and laser power as design variables. By considering the design variables, the limit-state function is rewritten as $Y = g(\mathbf{X}; \mathbf{d})$. Since the X-to-U transformation also depends on $\mathbf{d}$ if it contains distribution parameters $\mathbf{X}$, the transformation is expressed by $\mathbf{X} = \mathrm{T}(\mathbf{U}; \mathbf{d})$. The limit-state function in the U-space is then given by $Y = g(\mathrm{T}(\mathbf{U}; \mathbf{d}); \mathbf{d})$. In this study, we assume all the random variables in $\mathbf{X}$ are independent. For the general case where dependent random variables exist, they can be transformed into independent variables, and consequently, the proposed method can still be used. The task is summarized as follows:

Given: the distributions of $\mathbf{X}$, limit-state function $Y = g(\cdot)$, and required reliability R

Find: domain of $\mathbf{d}$ such that $\Pr\{g(\mathbf{X}; \mathbf{d}) > 0\} = R$

The domain of $\mathbf{d}$ specified above is the reliability-based feasible design region. Denote the feasible design region by $H(\mathbf{d}) = 0$. Once feasible region is available, design variables can be selected from it. We are particularly interested in plotting the region so that it is convenient to use the graph. Plotting $H(\mathbf{d}) = 0$ or $\Pr\{g(\mathbf{X}; \mathbf{d}) > 0\} = R$, however, is computationally demanding.

Take a two-dimensional problem as an example. Let $\mathbf{d} = (d_1, d_2)$, and therefore $H(d_1, d_2) = 0$, or $\Pr\{g(\mathbf{X}; d_1, d_2) > 0\} = R$. Suppose we use 100 points to plot the curve, and we then discretize the range of d_1 with 100 points. For each point of d_1, we solve for d_2 from $\Pr\{g(\mathbf{X}; d_1, d_2) > 0\} = R$. This is an inverse probabilistic analysis, meaning that given the probability R, find d_2, which may be a distribution parameter of one random variable in $\mathbf{X}$. One inverse probabilistic analysis will need a number of direct probabilistic analyses, which find the probability $\Pr\{g(\mathbf{X}; d_1, d_2) > 0\}$ for a given value of d_2. If the probability does not equal to R, d_2 will be changed. This process continues until $\Pr\{g(\mathbf{X}; d_1, d_2) > 0\} = R$. Suppose we use MCS and the required reliability is $R = 0.99999$ (the probability of failure $p_f = 10^{-5}$). Also assuming that we use a sample size of 10^7 for one direct probabilistic analysis, an inverse probabilistic

analysis needs 10 direct probabilistic analyses, and solving for d_2 needs 10 function calls. Then the total number of calling the limit-state function $g(\mathbf{X}; d_1, d_2)$ will be $100 \times 10 \times 10 \times 10^7 = 10^{11}$ times. Such a high computational cost will be unaffordable for many applications.

To overcome the above mentioned obstacles, in this study, we develop a new efficient method based on the inverse FORM. The feasible design region can be quickly generated by the proposed method with much less function calls. If a higher accuracy is needed for critical applications, the accuracy will then be checked with MCS. If the accuracy is not satisfactory, the MCS result will be used to refine the design region. The direct use of the inverse FORM will involve a double-loop procedure, as illustrated in Table 5.1 in Section 5.3.2. The outer loop is to search for the MPP $\mathbf{u}^*$ for a given set of $\mathbf{d}$, and the inner loop is solved for $\mathbf{d}$ from $H(\mathbf{d}) = g(\mathrm{T}(\mathbf{u}^*; \mathbf{d}); \mathbf{d}) = 0$. To further improve the efficiency, we propose to combine the two loops so that solving $H(\mathbf{d}) = g(\mathrm{T}(\mathbf{u}^*; \mathbf{d}); \mathbf{d}) = 0$ is embedded in the MPP search algorithm.

5.3.2 FORM-based Method for Generating a Feasible Design Region

As discussed above, design variables are solved from $H(\mathbf{d}) = g(\mathrm{T}(\mathbf{u}^*; \mathbf{d}); \mathbf{d}) = 0$ where $\mathbf{u}^* = \beta\alpha(\mathbf{u}^*)$ is indicated in Eq. (5.12). Eq. (5.11) gives

$$\alpha(\mathbf{u}^*) = \frac{\nabla g(\mathrm{T}(\mathbf{U};\mathbf{d});\mathbf{d})}{\|\nabla g(\mathrm{T}(\mathbf{U};\mathbf{d});\mathbf{d})\|}. \tag{5.16}$$

Function $H(\mathbf{d})$ becomes

$$H(\mathbf{d}) = g\left(\mathrm{T}\left(\beta \frac{\nabla g(\mathrm{T}(\mathbf{U};\mathbf{d});\mathbf{d})}{\|\nabla g(\mathrm{T}(\mathbf{U};\mathbf{d});\mathbf{d})\|}; \mathbf{d}\right); \mathbf{d}\right) = 0. \tag{5.17}$$

Solving Eq. (5.17) requires an iterative procedure. The first step is to find the reliability index by [139]

$$\beta = \Phi^{-1}(R), \tag{5.18}$$

where the required reliability R is greater than 0.5.

Suppose at iteration i, the design variables are obtained from the previous or the (i-1)-th iteration, and the MPP is $\mathbf{u}$ in the U-space. We convert them into the original random variables in the X-space by $\mathbf{x} = T^{-1}(\mathbf{u})$, where the transformation $T^{-1}(\cdot)$ is given in Eq. (5.6). We calculate the derivative of the limit-state function at $\mathbf{x}$ and obtain $\frac{\partial g}{\partial X_j}$, where $j = 1,2,\dots,n$. The derivative of the limit-state function at $\mathbf{u}$ is then

$$\frac{\partial g}{\partial u_j} = \frac{\partial g}{\partial X_j}\frac{dx_j}{\partial u_j}, \tag{5.19}$$

where $\frac{dx_j}{\partial u_j}$ can be derived from Eq. (5.6).

$$\frac{dx_j}{\partial u_j} = \frac{\phi(u_j)}{f_{X_j}(x_j)}, \tag{5.20}$$

where $\phi(\cdot)$ and $f_{X_j}(\cdot)$ are the PDF of U_j and X_j, respectively. Then the gradient is $\nabla g\left(\mathrm{T}\left(\mathbf{u}^i;\mathbf{d}\right);\mathbf{d}\right) = \left(\frac{\partial g}{\partial U_1},\ldots,\frac{\partial g}{\partial U_n}\right)\Big|_{\mathbf{u}}$, and the associated unit vector is $\alpha(\mathbf{u}) = \frac{\nabla g(\mathrm{T}(\mathbf{u}^i;\mathbf{d});\mathbf{d})}{\|\nabla g(\mathrm{T}(\mathbf{u}^i;\mathbf{d});\mathbf{d})\|}$.

Then the MPP is updated to

$$\mathbf{u} = \beta\alpha(\mathbf{u}). \tag{5.21}$$

Once the $\mathbf{u}$ is updated, we solve $H(\mathbf{d}) = g(\mathrm{T}(\mathbf{u}^*;\mathbf{d});\mathbf{d}) = 0$ and obtain a new set of design variables $\mathbf{d}$. Thus, one iteration is complete. Then convergence is checked. If the distance between the new MPP on the left-hand side of Eq. (5.19) is sufficiently close to the previous MPP on the right-hand side of the equation, the MPP is found; otherwise, we calculate the gradient of the limit-state function at the updated MPP and then find a new MPP. We repeat this process until convergence criterion is met. Once the iterative process is complete, we obtain a design point $\mathbf{d}$ for the feasible design region. By repeating this iterative process multiple times, we obtain a number of design points, which will then result in a boundary of the feasible design region.

The convergence is judged by

$$\varepsilon = \|\mathbf{u}_{\mathrm{new}} - \mathbf{u}_{\mathrm{old}}\|, \tag{5.22}$$

where $\mathbf{u}_{\mathrm{new}}$ and $\mathbf{u}_{\mathrm{old}}$ are the MPPs at the current iteration and previous iteration, respectively. The convergence criterion is: given a tolerance δ, we will terminate the search if $\varepsilon < \delta$. δ can be set to 0.01 or 0.001.

The procedure of the FORM-based feasible design for solving one design point is summarized in Table 5.1.

Table 5.1. FORM-Based Feasible Design

Procedure: FORM-based feasible design
Input
Required reliability R
CDFs $F_{X_i}(\cdot)$, $i = 1,2,\ldots,n$
Initial design point $\mathbf{d}_0$
Initial MPP $\mathbf{u}_0$
Convergence tolerance δ
Output: **d**
$\mathbf{u} = \mathbf{u}_0$
$\mathbf{d} = \mathbf{d}_0$
$\beta = \Phi^{-1}(R)$
$convergence = false$
while $convergence = false$
$\mathbf{u}_{\text{old}} = \mathbf{u}$
$\mathbf{x} = T^{-1}(\mathbf{u}; \mathbf{d})$
$\frac{\partial g}{\partial u_i} = \frac{\partial g}{\partial X_j}\frac{dx_i}{\partial u_i}$
$\nabla g = \left(\frac{\partial g}{\partial u_i}\right)_{i=1,n}$
$\alpha = \frac{\nabla g}{\|\nabla g\|}$
$\mathbf{u} = \beta\alpha$
$\mathbf{d} =$ solution to $g(\mathrm{T}(\mathbf{u}; \mathbf{d}); \mathbf{d}) = 0$
$\varepsilon = \|\mathbf{u} - \mathbf{u}_{\text{old}}\|$
if $\varepsilon < \delta$ **then**
$convergence = true$
end if
end while

If we use the above algorithm to produce a sufficient number of design points, we can then generate a feasible design region. The algorithm requires evaluating the gradient of the limit-state function and solving the equation given by $g(\mathrm{T}(\mathbf{u}; \mathbf{d}); \mathbf{d}) = 0$. It does not involve any random sampling and is therefore computationally efficient. Since the method employs the most commonly used FORM, its accuracy should be acceptable for most applications. The accuracy, however, may deteriorate if the limit-state function is highly nonlinear around the MPP in the U-space.

There are two possible ways to improve the accuracy. The first way is to use MCS based on the result from the proposed method, while the second way is to use the Second Order Reliability Method (SORM), which also uses the second derivatives of the limit-state function. In this study, we choose the former method for efficiency.

5.3.3 Accuracy Improvement by MCS

The strategy is to use MCS to accurately calculate the reliability at the design point **d**, which is obtained from the above FORM based feasibility design. Denote the reliability from MCS by R_{MCS}, and let the difference from the required reliability be $\varepsilon_R = R_{MCS} - R_{Req}$. If $|\varepsilon_R|$ is greater than a tolerance δ_R, we will adjust the required reliability by

$$R_n = \frac{R_{n-1}R_{Req}}{R_{MCS}}, \tag{5.23}$$

where R_n denotes the input reliability of FORM at the *n*th iteration. Then the algorithm in Table 1 is then called. To reduce the computational cost, we use the design point **d** and MPP **u** from the last feasibility design as the initial design point and MPP, respectively. This process is repeated until the convergence criterion is met, and convergence can be reached with a few iterations. As a result, no inverse MCS is needed, and the number of direct MCS is minimized. The tolerance δ_R is determined by the accuracy requirement for a specific application. In this work, we employ a relative tolerance, such that

$$\delta_R = cp_r, \tag{5.24}$$

where c is a relative coefficient. For example, we set $c = 0.5$. If $R_{Req} = 0.99999$, then $p_r = 10^{-5}$, and $\delta_R = 5 \times 10^{-6}$. Therefore, upon convergence, the actual reliability will be 0.99999 ± 0.000005. The algorithm of the d FORM and MCS is provided in Table 2.

Table 5.2. Integrated FORM and MCS

Procedure: FORM/MCS-based feasible design
Input
Required reliability R_{Req}
CDFs $F_{X_i}(\cdot), i = 1,2, \dots, n$
Initial design point $\mathbf{d}_0$
Initial MPP $\mathbf{u}_0$
Convergence tolerance δ for FORM
Convergence tolerance δ_R for reliability
Output: **d**
$convergence = false$
while $convergence = false$
$\mathbf{u} = \mathbf{u}_0$
$\mathbf{d} = \mathbf{d}_0$
$\beta = \Phi^{-1}(R)$
Call FORM-based feasibility design (see Table 1)
Obtain d, **u**
Call MCS
Obtain R_{MCS}
$\varepsilon_R = R_{MCS} - R_{Req}$
if $|\varepsilon_R| < \delta_R$ **then**
$convergence = true$
end if
$R = \dfrac{RR_{Req}}{R_{MCS}}$
$\mathbf{u}_0 = \mathbf{u}$
$\mathbf{d}_0 = \mathbf{d}$
end while

5.4 Case Studies

In this section, the proposed method is demonstrated by its application for the reliability-based L-PBF process design.

5.4.1 Case 1: Normal Distribution and Standard Deviations are Fractions of Their Means

Limit-state function and model input

In this case, standard deviations change with respect to the means, and are fractions of their means. The limit-state function for the normalized enthalpy criterion is given in Eq. (5.15). There are seven random input variables, and three of which are design variables, including the laser

power P, laser scan speed v and laser beam radius r. During the process design, the means of the three random variables are to be changed. In this study, we assume the mean of the laser beam radius r is fixed due to its low flexibility. Then the design problem becomes two-dimensional, and the design variables are the means μ_P and μ_v, of P and v, respectively, or their nominal values; namely, $\mathbf{d} = (\mu_P, \mu_v)$. Note that the reliability analysis is seven-dimensional because seven random variables are involved.

We identify the feasible design region for the means $\mathbf{d} = (\mu_P, \mu_v)$. Since we need to plot a two-dimensional curve for the feasible design region, we discretize one variable and solve for the other at the discretized points of the first one. We divide the nominal values of v, or μ_v, into 0.1, 0.2, ..., 2.0 m/s and then solve for the corresponding nominal values of P, or μ_P. Since there are in total 20 points, we call the algorithm in Table 5.2 20 times, searching for a set of μ_P for a given set of μ_v; namely, 0.1, 0.2, ..., 2.0 m/s. For the stop criteria, we use $c = 0.5$ in Eq. (5.24) and $\delta = 0.01$ in Eq. (5.22).

We need to know the distributions of all the seven input random variables. For this case study, we use the distributions from previous research and experiments reported in literature. Table 5.3 summarizes the distributions with the associated references. All the seven random variables follow normal distributions, and their standard deviations are fractions of their means. For specific applications, the distributions will be different, and the effect of distribution types is analyzed in Sections 5.4.2 and 5.4.3.

Table 5.3. Distributions of input random variables

	Unit	Mean μ	Standard deviation σ	Distribution	References
P	W	μ_P	$0.025\mu_P$	Normal	[132, 134, 143]
v	m/s	μ_v	$0.015\mu_v$	Normal	[132, 134, 143]
A	-	0.4	$0.2\mu_A$	Normal	[10, 40, 144]
ρ	kg/m^3	7980	$0.01\mu_\rho$	Normal	[89, 118, 143]
D	m^2/s	5.38×10^{6}	$0.1\mu_D$	Normal	[89, 134]
r	m	2.70×10^{-5}	$0.04\mu_r$	Normal	[132]
h_s	J/kg	1.20×10^{6}	$0.1\mu_{h_s}$	Normal	[89, 145]

Reliability-based feasibility design

We first generate a feasible design region for an allowable probability of failure of $p_f = 10^{-6}$, or a target reliability of $R = 0.999999$. The feasible design region obtained from FORM, with the dotted line, is shown in Figure 5.2, where the region above the curve is the feasible design region for the L-PBF process. If a design point (μ_P, μ_v) is chosen above the curve, the probability of in the keyhole mode will be less than 10^{-6}. If the design point is on the curve, the probability of in the keyhole mode will be 10^{-6}.

For comparison, we also plot the deterministic feasible design region generated using the means of all the seven random input variables. If the design point is on the curve of the deterministic feasible design, the probability of the occurrence of keyhole mode will be very high, around 0.5. The deterministic feasible design region occupies most of the entire design space while the reliability feasible design region is reduced significantly, suggesting that the effect of uncertainty is significant.

The proposed method with FORM is extremely efficient. The computational time for the entire curve is only 0.07 seconds with an Intel Core i7-8750H processor and 32 GB RAM. If higher accuracy is preferred, the FORM/MCS method can be formed. It runs MCS after the feasible design region is found by FORM. Figure 5.2 shows that the curve from the FORM/MCS method with a sample size of 10^9 is very close to the one from FORM, and this indicates that FORM is not only efficient, but also accurate. The computational cost of the FORM/MCS method is 43,209 seconds, much higher than that of FORM. If direct MCS is used, the computational cost is 371,085 seconds, and this indicates that the proposed FORM and FORM/MCS methods can cut the computational cost significantly.

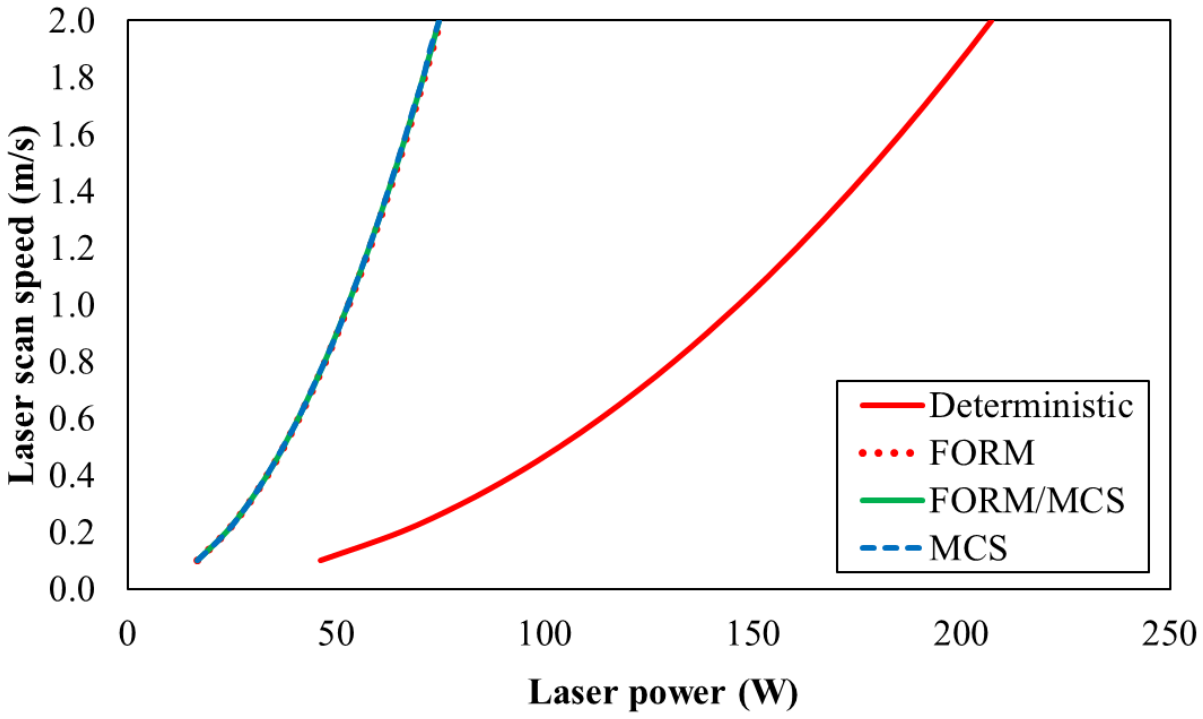


Figure 5.2. Reliability contours at $p_f = 10^{-6}$ from FORM, FORM/MCS, MCS and deterministic methods in Case 1

Different applications may require different reliability targets. We also perform the proposed method for other levels of required reliability, and the associated probabilities of failure are 10^{-2}, 10^{-3}, 10^{-4}, and 10^{-5}, respectively. The reliability feasible design regions from FORM/MCS are plotted in Figure 5.3.

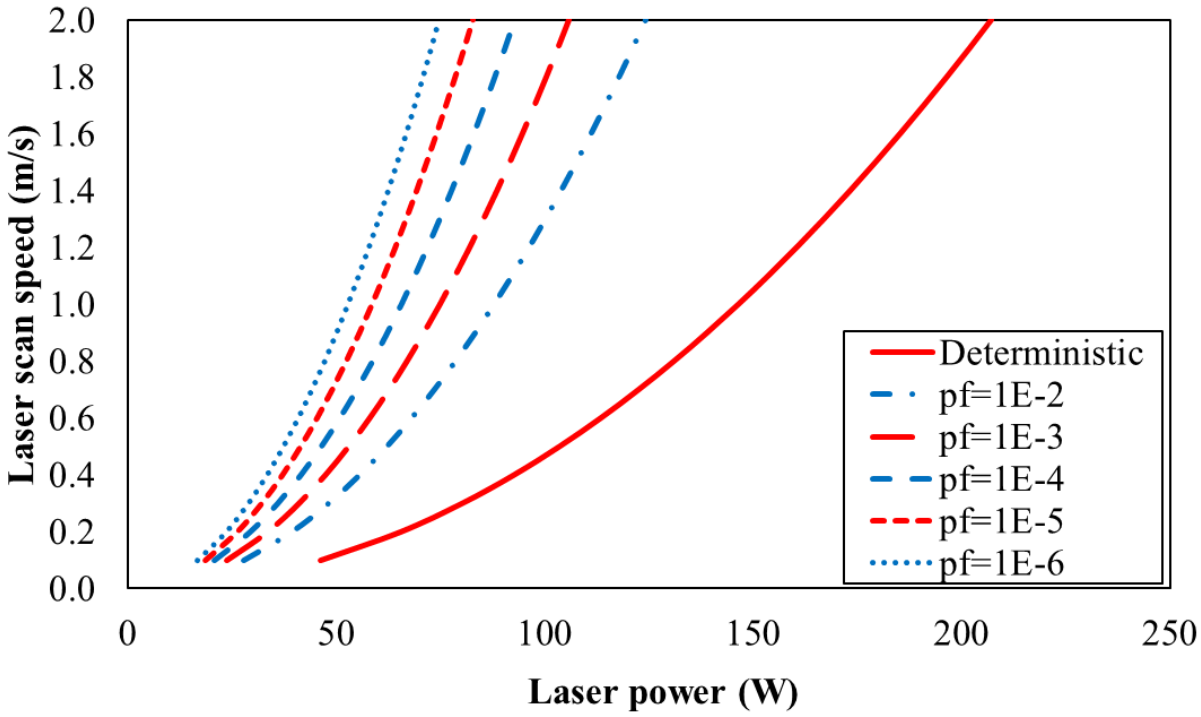


Figure 5.3. Reliability feasible design regions for different probabilities of failure (pf) in Case 1

Figure 5.3 also demonstrates that the boundaries of the reliability feasible design region are nonlinear and are increasing functions. It also shows that the increasing reliability reduces the feasible design region, or the boundaries shift to left of the figure. In other words, higher reliability results in a narrower design space.

The computational costs for different reliability targets are summarized in Table 5.4. It is clear that the cost of FORM/MCS or the direct MCS increases when the required reliability increases, since the number of Monte Carlo samples needed for higher reliability is larger for sufficient accuracy. The cost of FORM remains almost constant, not affected by the required reliability. Overall, FORM shows its excellent efficiency and good accuracy. If FORM/MCS method is unaffordable, the reliability feasible design region can be used; to meet the required reliability with high accuracy, a design point can be selected no close to the boundary of the reliability feasible design region.

Table 5.4. Comparison of computational costs among FORM, FORM/MCS, and MCS methods

Method	$p_f = 10^{-2}$	$p_f = 10^{-3}$	$p_f = 10^{-4}$	$p_f = 10^{-5}$	$p_f = 10^{-6}$
FORM	0.07s	0.07s	0.07s	0.07s	0.07s
FORM/MCS	5s	47s	440s	4,351s	43,209s
MCS	41s	387s	3,786s	37,249s	371,085s

5.4.2 Case 2: Normal Distribution and Constant Standard Deviations

In Case 1, the standard deviations change with respect to the means. In some L-PBF processes, however, the standard deviations are constant for certain parameters, such as the uncertainty due to imprecise measurements of laser power and scan speed. To evaluate the proposed method, we study one case with fixed standard deviations, $\sigma_P = 1$ W and $\sigma_v = 0.1$ m/s. The other distributions remain unchanged. The permitted probability of failure is still 10^{-6}.

The reliability feasible design regions of Case 2 from FORM and FORM/MCS are plotted in Figure 5.4. Similar to Case 1, Figure 5.4 shows that the curve from FORM is still very close to the one from FORM/MCS, which indicates the high accuracy of FORM in this scenario. The computational cost of FORM is still 0.07s, indicating its high efficiency.

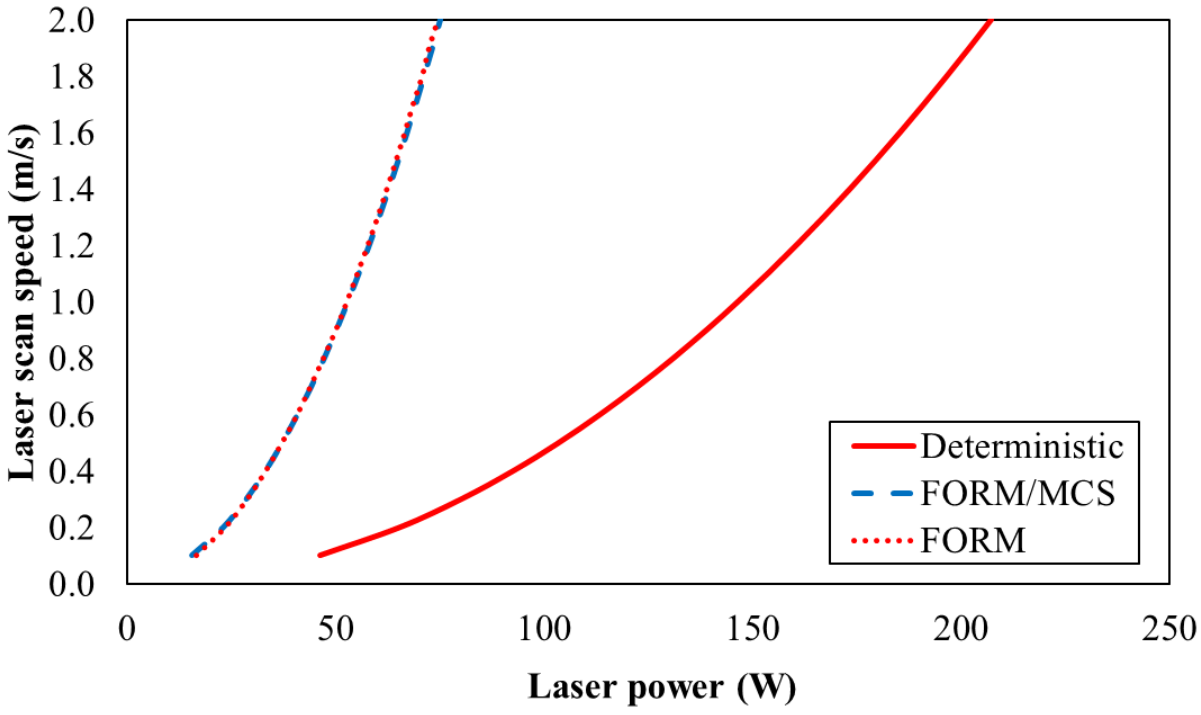


Figure 5.4. Reliability contours at $p_f = 10^{-6}$ from FORM, FORM/MCS, and deterministic methods in Case 2

5.4.3 Case 3: Non-normal Distribution

Only normal distributions are involved in the above two cases. In practice, non-normal distributions may also exist. The proposed method is also applicable for the non-normal distributions. To demonstrate this, we study a new case where the laser scan speed follows a uniform distribution, as reported in Ref. [132]. The lower and upper bounds of the uniform distribution are $0.97\mu_v$ and $1.03\mu_v$, respectively, where μ_v is given in Table 5.3. The permitted probability of failure is still 10^{-6}. The other distributions remain the same as Case 1.

Figure 5.5 displays the reliability feasible design regions of Case 3 from FORM and FORM/MCS at $p_f = 10^{-6}$. As mentioned in Section 5.2.4, the transformation from non-normal distribution to normal distribution in FORM is non-linear, and therefore an error is expected to occur in Case 3. However, as shown in Figure 5.5, the curve from FORM still matches very well with the one from FORM/MCS. This verifies the outstanding accuracy and flexibility of FORM.

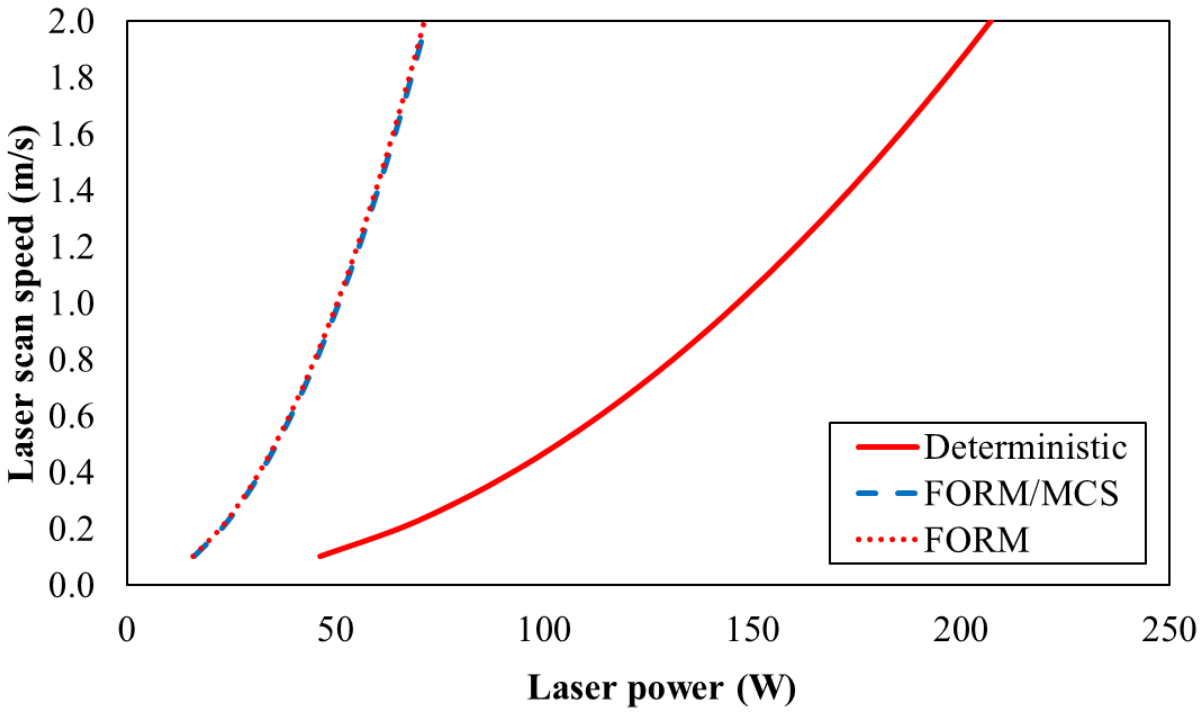


Figure 5.5. Reliability contours at $p_f = 10^{-6}$ from FORM, FORM/MCS, and deterministic methods in Case 3

Case 2 and Case 3 demonstrate the flexibility of the proposed method. The reliability feasibility regions in the three cases are different, showing the impact of distribution types and

distribution parameters. The three cases also indicate the high accuracy and efficiency of the proposed FORM-based method.

5.5 Summary

Uncertainty exists in all aspects of additive manufacturing, including its process design. If the uncertainty is significant, it is imperative to account for uncertainty to ensure that the reliability and quality requirements are satisfied. This work develops a reliability-based feasibility design to identify a feasible design region for selecting design variables for L-PBF additive manufacturing process design. The major conclusions from this study are summarized below.

1. The input of the proposed method includes the following:
    a. The limit-state function for a given requirement. It can be an explicit model or a black-box model.
    b. Distributions of input random variables.
    c. Design variables, which are the means of process design variables (part of the input random variables).
    d. The target reliability.
2. The output of the proposed method is a feasible design region.
3. If a design point is selected from the feasible design region, the probability of satisfying the requirement is no less than the target reliability.
4. The proposed method is based on the first order reliability method (FORM) and Monte Carlo simulation (MCS). For higher efficiency, only FORM can be performed when the computational cost of calling the limit-state function is high. Both the FORM and FORM/MCS are much efficient than the direct MCS.
5. The three case studies in L-PBF process demonstrate that the impact of uncertainty on the process design is significant, since the feasible design region is reduced significantly, especially when high reliability is required. The results show that the required reliability of 99.9999% can narrow the design space more than half.
6. The proposed method is not limited to additive manufacturing. It can be used to generate reliability feasible design regions for other applications if the input indicated in 1) is available.

For the future work, the efficiency can be further enhanced for computationally expensive limit-state functions, especially for those from multiphysics and multiscale simulations. The future work may also include improving the efficiency of the MPP search and using surrogate models for the original limit-state functions. Another future work can be including a lower bound, e.g., lack of fusion, in the analysis. This task requires to develop a relevant limit-state function.

6. CONCLUSIONS AND FUTURE WORK

6.1 Conclusions

In this thesis, a probabilistic design modeling framework has been developed for the reliability-based design in the L-PBF process. The modeling framework can serve as a one-stop tool for designers to tune the design variables in an uncertain environment. The major conclusions are summarized as follows:

1. A new feasibility of simulating the spattering phenomenon in the L-PBF process is proposed. A SPH model is developed which can resemble the particle trajectories with a similar scenario in the experiment and simulate the sticking phenomenon.
2. A guide for ML applications in AM field is provided. The existing applications in literature are summarized. Suggestions on feature selections, model selections, model assessment methods, and monitoring and solving the overfitting issue in the AM field are elaborated. Future research directions related to machine learning applications are suggested.
3. A Gaussian process-based machine learning model is developed. After trained by a limited number of data points, the ML model is capable of making predictions efficiently with comparable errors with the intrinsic experimental noise. The application of the GP model for process optimization is demonstrated. The preferred conduction mode region of laser power and laser speed can be predicted by the GP model, which can be used to calibrate a convenient criterion of identifying keyhole mode in L-PBF.
4. A novel probabilistic model integrating the MCS and FORM is proposed. After input the target reliability as well as the limit-state function which defines the design requirement, the distribution of all random variables and the design variables, the framework outputs the feasible design region, within which the design requirement will be satisfied with a probability higher or equal to the specified reliability. The proposed method is much efficient than the conventional MCS model, while possesses high accuracy same as MCS.

6.2 Contributions

The major contributions of the thesis are summarized as follows:

1. A probabilistic design modeling framework has been developed for the reliability-based design in the L-PBF process. The modeling framework can serve as a one-stop tool for designers to tune the design variables in a noisy environment in the L-PBF process.
2. At this time, a physical-based model that can simulate the spattering phenomenon is still missing. The SPH model developed in this thesis uniquely provide a feasibility of simulating the spattering phenomenon where phase change, heat transfer and particle motion are involved.
3. Currently, though there are a lot of research articles applying ML models in various L-PBF applications, there is no articles that can provide a comprehensive guide for the machine learning applications in the L-PBF field. The guide provided in this thesis elaborates the important components of a successful ML application, including data feature selection, task taxonomy, model selection, model assessment, overfitting monitoring and avoidance. Engineers with minor knowledge towards ML models can refer to this work as a quick start of ML applications in the L-PBF process.
4. Although there are machine learning applications in the L-PBF field, none of them focuses on the construction of a limit-state function for probabilistic design. The ML model developed in this thesis predicts the width depth ratio of the melt pool, which can serve as a design requirement to avoid the keyhole mode. The limit-state function is essential for a probabilistic process design in an uncertain environment.
5. Although some uncertainty quantification models have been developed by previous work in the literature, the connection between the quantified uncertainty and the design requirement is still missing. The outcome of a typical uncertainty quantification process is merely the distribution of the output quantities, which have not been linked to the reliability. In addition, the probabilistic design problem is inverse, whose desired output is the feasible design region. The model developed in this thesis uniquely provide the feasible region for reliability-based design.
6. Though conventional MCS is available to address the probabilistic design problem, its high computational cost impedes a quick design process. FORM, on the other

hand, can obtain the feasible design region efficiently, but it loses accuracy when non-normal distribution is involved due to the nonlinear transformation. In this thesis, the novel method integrating the MCS and FORM is much more efficient than MCS while maintain high accuracy.

6.3 Future Work

Modeling of L-PBF is an emerging research topic that is full of challenges and opportunities. Beyond the work in this thesis, the following aspects remain interesting to explore in future work.:

1. Many limitations exist in the SPH model demonstrated in this work, including the missing effects of air resistance and air cooling, the difficulty in defining boundary conditions like fluid inlet and outlet, the assumed infinite rate of phase change, and the lack of the third dimension. For better simulation of the spattering phenomenon, future work on improving the SPH model to overcome these limitations will be recommended.
2. A missing but useful functionality in supervised learning in recent literature is active learning. In AM field, labelling the output of each input data point is usually expensive in terms of the consumed time, cost and human labor, because it requires conducting an experiment or a simulation at each input setting to make this observation. Active learning is a method that can alleviate this issue. In recent literature, the common procedure in ML models is acquiring enough input-output pairs first and then using them to train ML models without further query of labelling new data. On the contrary, the procedure in active learning is that the ML models can make query interactively for labelling new data during training so as to maximize its performance. By this means, the ML models may use fewer data points to achieve better performance. Therefore, active learning is strongly recommended in the case that a dataset to be used to train the ML model has not been acquired.
3. A barrier in the ML applications in the L-PBF process is the limitation of data points. Though a constructed surrogate ML model is very efficient, the number of required training data points is exponentially increasing with the increased input features. Thus, current research usually focuses on a few input features for training. In this

regard, physical models with high-fidelity and high efficiency are desirable for a faster design process.

4. The adjustment algorithm of input reliability in the integrated MCS and FORM model and the searching algorithm of MPP in FORM have not been optimized. When non-normal distribution involves, the integrated model loses some efficiency due to the accuracy reduction in FORM. Development of more efficient algorithms is a good research direction to further enhance the efficiency of the integrated model.
5. Though this thesis demonstrates the application of the proposed method on a reliability-based design based on keyhole mode avoidance, this is not the only design requirement in the L-PBF process. For example, lack of fusion is an opposite case of the keyhole mode when the laser power density is insufficient to sinter the metal powders. It will also constrain some region in the design space. A practical probabilistic design considering all the design requirements with carefully identified uncertainty of each random variables is significant for a reliable process.

REFERENCES

[1] Y. Zhang, L. Wu, X. Guo, S. Kane, Y. Deng, Y.-G. Jung, J.-H. Lee, and J. Zhang, "Additive manufacturing of metallic materials: A review," *Journal of Materials Engineering and Performance,* vol. 27, no. 1, pp. 1-13, 2018.

[2] G. Tapia and A. Elwany, "A review on process monitoring and control in metal-based additive manufacturing," *Journal of Manufacturing Science and Engineering,* vol. 136, no. 6, p. 060801, 2014.

[3] Y. Zhang, "Multi-Scale Multi-Physics Modeling Of Laser Powder Bed Fusion Process Of Metallic Materials With Experiment Validation," Ph.D., Mechanical Engineering, Purdue University, West Layayette, Indiana, US, 2018.

[4] M. Grasso and B. M. Colosimo, "Process defects and in situ monitoring methods in metal powder bed fusion: a review," *Measurement Science and Technology,* vol. 28, no. 4, p. 044005, 2017.

[5] S. Tammas-Williams, H. Zhao, F. Léonard, F. Derguti, I. Todd, and P. Prangnell, "XCT analysis of the influence of melt strategies on defect population in Ti–6Al–4V components manufactured by Selective Electron Beam Melting," *Materials Characterization,* vol. 102, pp. 47-61, 2015.

[6] W. J. Sames, F. List, S. Pannala, R. R. Dehoff, and S. S. Babu, "The metallurgy and processing science of metal additive manufacturing," *International Materials Reviews,* vol. 61, no. 5, pp. 315-360, 2016.

[7] W. E. Frazier, "Metal additive manufacturing: a review," *Journal of Materials Engineering and Performance,* vol. 23, no. 6, pp. 1917-1928, 2014.

[8] S. Bontha, N. W. Klingbeil, P. A. Kobryn, and H. L. Fraser, "Thermal process maps for predicting solidification microstructure in laser fabrication of thin-wall structures," *Journal of Materials Processing Technology,* vol. 178, no. 1-3, pp. 135-142, 2006.

[9] C. Kamath, "Data mining and statistical inference in selective laser melting," *The International Journal of Advanced Manufacturing Technology,* vol. 86, no. 5-8, pp. 1659-1677, 2016.

[10] W. E. King, H. D. Barth, V. M. Castillo, G. F. Gallegos, J. W. Gibbs, D. E. Hahn, C. Kamath, and A. M. Rubenchik, "Observation of keyhole-mode laser melting in laser powder-bed fusion additive manufacturing," *Journal of Materials Processing Technology,* vol. 214, no. 12, pp. 2915-2925, 2014.

[11] M. Yan and P. Yu, "An Overview of densification, microstructure and mechanical property of additively manufactured Ti-6Al-4V—Comparison among selective laser melting, electron beam melting, laser metal deposition and selective laser sintering, and with conventional powder," in *Sintering techniques of materials*: InTech, 2015.

[12] R. Cunningham, C. Zhao, N. Parab, C. Kantzos, J. Pauza, K. Fezzaa, T. Sun, and A. D. Rollett, "Keyhole threshold and morphology in laser melting revealed by ultrahigh-speed x-ray imaging," *Science,* vol. 363, no. 6429, pp. 849-852, 2019.

[13] N. D. Parab, C. Zhao, R. Cunningham, L. I. Escano, K. Fezzaa, W. Everhart, A. D. Rollett, L. Chen, and T. Sun, "Ultrafast X-ray imaging of laser–metal additive manufacturing processes," *Journal of synchrotron radiation,* vol. 25, no. 5, 2018.

[14] A. Gusarov, I. Yadroitsev, P. Bertrand, and I. Smurov, "Model of radiation and heat transfer in laser-powder interaction zone at selective laser melting," *Journal of heat transfer,* vol. 131, no. 7, p. 072101, 2009.

[15] C. Panwisawas, C. Qiu, M. J. Anderson, Y. Sovani, R. P. Turner, M. M. Attallah, J. W. Brooks, and H. C. Basoalto, "Mesoscale modelling of selective laser melting: Thermal fluid dynamics and microstructural evolution," *Computational Materials Science,* vol. 126, pp. 479-490, 2017.

[16] J. Zhang, Y. Zhang, W. H. Lee, L. Wu, H.-H. Choi, and Y.-G. Jung, "A multi-scale multi-physics modeling framework of laser powder bed fusion additive manufacturing process," *Metal Powder Report,* vol. 73, no. 3, pp. 151-157, 2018.

[17] Y. Zhang and J. Zhang, "Sintering phenomena and mechanical strength of nickel based materials in direct metal laser sintering process—a molecular dynamics study," *Journal of Materials Research,* vol. 31, no. 15, pp. 2233-2243, 2016.

[18] C. Zhao, K. Fezzaa, R. W. Cunningham, H. Wen, F. De Carlo, L. Chen, A. D. Rollett, and T. Sun, "Real-time monitoring of laser powder bed fusion process using high-speed X-ray imaging and diffraction," *Scientific reports,* vol. 7, no. 1, pp. 1-11, 2017.

[19] S. A. Khairallah, A. T. Anderson, A. Rubenchik, and W. E. King, "Laser powder-bed fusion additive manufacturing: Physics of complex melt flow and formation mechanisms of pores, spatter, and denudation zones," *Acta Materialia,* vol. 108, pp. 36-45, 2016.

[20] Z. Chen, H. Wang, X. Wang, P. Chen, Y. Liu, H. Zhao, Y. Zhao, and Y. Duan, "Low-temperature remote plasma enhanced atomic layer deposition of ZrO 2/zircone nanolaminate film for efficient encapsulation of flexible organic light-emitting diodes," *Scientific reports,* vol. 7, no. 1, pp. 1-9, 2017.

[21] R. Rai, J. Elmer, T. Palmer, and T. DebRoy, "Heat transfer and fluid flow during keyhole mode laser welding of tantalum, Ti–6Al–4V, 304L stainless steel and vanadium," *Journal of physics D: Applied physics,* vol. 40, no. 18, p. 5753, 2007.

[22] J. D. Madison and L. K. Aagesen, "Quantitative characterization of porosity in laser welds of stainless steel," *Scripta Materialia,* vol. 67, no. 9, pp. 783-786, 2012.

[23] T. Eagar and N. Tsai, "Temperature fields produced by traveling distributed heat sources," *Welding journal,* vol. 62, no. 12, pp. 346-355, 1983.

[24] V.-P. Matilainen, H. Piili, A. Salminen, and O. Nyrhilä, "Preliminary investigation of keyhole phenomena during single layer fabrication in laser additive manufacturing of stainless steel," *Physics Procedia,* vol. 78, pp. 377-387, 2015.

[25] D. Hann, J. Iammi, and J. Folkes, "A simple methodology for predicting laser-weld properties from material and laser parameters," *Journal of Physics D: Applied Physics,* vol. 44, no. 44, p. 445401, 2011.

[26] L. E. Criales, Y. M. Arısoy, B. Lane, S. Moylan, A. Donmez, and T. Özel, "Laser powder bed fusion of nickel alloy 625: experimental investigations of effects of process parameters on melt pool size and shape with spatter analysis," *International Journal of Machine Tools Manufacture,* vol. 121, pp. 22-36, 2017.

[27] M. Simonelli, C. Tuck, N. T. Aboulkhair, I. Maskery, I. Ashcroft, R. D. Wildman, and R. Hague, "A study on the laser spatter and the oxidation reactions during selective laser melting of 316L stainless steel, Al-Si10-Mg, and Ti-6Al-4V," *Metallurgical and Materials Transactions A,* vol. 46, no. 9, pp. 3842-3851, 2015.

[28] V. Gunenthiram, P. Peyre, M. Schneider, M. Dal, F. Coste, I. Koutiri, and R. Fabbro, "Experimental analysis of spatter generation and melt-pool behavior during the powder bed laser beam melting process," *Journal of Materials Processing Technology,* vol. 251, pp. 376-386, 2018.

[29] Y. Liu, Y. Yang, S. Mai, D. Wang, and C. Song, "Investigation into spatter behavior during selective laser melting of AISI 316L stainless steel powder," *Materials & Design,* vol. 87, pp. 797-806, 2015.

[30] M. T. Andani, R. Dehghani, M. R. Karamooz-Ravari, R. Mirzaeifar, and J. Ni, "Spatter formation in selective laser melting process using multi-laser technology," *Materials & Design,* vol. 131, pp. 460-469, 2017.

[31] Q. Guo, C. Zhao, L. I. Escano, Z. Young, L. Xiong, K. Fezzaa, W. Everhart, B. Brown, T. Sun, and L. Chen, "Transient dynamics of powder spattering in laser powder bed fusion additive manufacturing process revealed by in-situ high-speed high-energy x-ray imaging," *Acta Materialia,* vol. 151, pp. 169-180, 2018.

[32] J. Yin, D. Wang, L. Yang, H. Wei, P. Dong, L. Ke, G. Wang, H. Zhu, and X. Zeng, "Correlation between forming quality and spatter dynamics in laser powder bed fusion," *Additive Manufacturing,* vol. 31, p. 100958, 2020.

[33] B. Cheng and A. H. Ngan, "The sintering and densification behaviour of many copper nanoparticles: A molecular dynamics study," *Computational materials science,* vol. 74, pp. 1-11, 2013.

[34] Q. Mao, Y. Ren, K. H. Luo, and S. Li, "Sintering-induced phase transformation of nanoparticles: A molecular dynamics study," *The Journal of Physical Chemistry C,* vol. 119, no. 51, pp. 28631-28639, 2015.

[35] J. Xu, R. Sakanoi, Y. Higuchi, N. Ozawa, K. Sato, T. Hashida, and M. Kubo, "Molecular dynamics simulation of Ni nanoparticles sintering process in Ni/YSZ multi-nanoparticle system," *The Journal of Physical Chemistry C,* vol. 117, no. 19, pp. 9663-9672, 2013.

[36] F. Rahmani, J. Jeon, S. Jiang, and S. Nouranian, "Melting and solidification behavior of Cu/Al and Ti/Al bimetallic core/shell nanoparticles during additive manufacturing by molecular dynamics simulation," *Journal of Nanoparticle Research,* vol. 20, no. 5, p. 133, 2018.

[37] Y. Lee and W. Zhang, "Mesoscopic simulation of heat transfer and fluid flow in laser powder bed additive manufacturing," in *Int. Solid Free Form Fabr. Symp. Austin, TX*, 2015, pp. 1154-1165.

[38] C. Körner, E. Attar, and P. Heinl, "Mesoscopic simulation of selective beam melting processes," *Journal of Materials Processing Technology,* vol. 211, no. 6, pp. 978-987, 2011.

[39] C. Körner, A. Bauereiß, and E. Attar, "Fundamental consolidation mechanisms during selective beam melting of powders," *Modelling and Simulation in Materials Science and Engineering,* vol. 21, no. 8, p. 085011, 2013.

[40] S. A. Khairallah and A. Anderson, "Mesoscopic simulation model of selective laser melting of stainless steel powder," *Journal of Materials Processing Technology,* vol. 214, no. 11, pp. 2627-2636, 2014.

[41] S. Ly, A. M. Rubenchik, S. A. Khairallah, G. Guss, and M. J. Matthews, "Metal vapor micro-jet controls material redistribution in laser powder bed fusion additive manufacturing," *Scientific reports,* vol. 7, no. 1, p. 4085, 2017.

[42] M. J. Matthews, G. Guss, S. A. Khairallah, A. M. Rubenchik, P. J. Depond, and W. E. King, "Denudation of metal powder layers in laser powder bed fusion processes," *Acta Materialia,* vol. 114, pp. 33-42, 2016.

[43] P. Bidare, I. Bitharas, R. Ward, M. Attallah, and A. J. Moore, "Fluid and particle dynamics in laser powder bed fusion," *Acta Materialia,* vol. 142, pp. 107-120, 2018.

[44] H. Lasi, P. Fettke, H.-G. Kemper, T. Feld, and M. Hoffmann, "Industry 4.0," *Business & information systems engineering,* vol. 6, no. 4, pp. 239-242, 2014.

[45] R. A. Gingold and J. J. Monaghan, "Smoothed particle hydrodynamics: theory and application to non-spherical stars," *Monthly notices of the royal astronomical society,* vol. 181, no. 3, pp. 375-389, 1977.

[46] L. B. Lucy, "A numerical approach to the testing of the fission hypothesis," *The astronomical journal,* vol. 82, pp. 1013-1024, 1977.

[47] L. D. Libersky and A. G. Petschek, "Smooth particle hydrodynamics with strength of materials," in *Advances in the free-Lagrange method including contributions on adaptive gridding and the smooth particle hydrodynamics method*: Springer, 1991, pp. 248-257.

[48] F. A. Allahdadi, T. C. Carney, J. R. Hipp, L. D. Libersky, and A. G. Petschek, "High strain lagrangian hydrodynamics: a three dimensional SPH code for dynamic material response," Phillips Lab Kirtland AFB NM, 1993.

[49] G. C. Ganzenmüller and M. O. Steinhauser. "The implementation of Smooth Particle Hydrodynamics in LAMMPS." https://lammps.sandia.gov/doc/PDF/SPH_LAMMPS_userguide.pdf (accessed May 4th, 2020).

[50] G. C. Ganzenmüller. "Smooth Mach Dynamics." https://lammps.sandia.gov/doc/PDF/SMD_LAMMPS_userguide.pdf (accessed May 4th, 2020).

[51] G.-R. Liu and M. B. Liu, *Smoothed particle hydrodynamics: a meshfree particle method*. World scientific, 2003.

[52] S. Plimpton, "Fast parallel algorithms for short-range molecular dynamics," Sandia National Labs., Albuquerque, NM (United States), 1993.

[53] A. Stukowski, "Visualization and analysis of atomistic simulation data with OVITO–the Open Visualization Tool," *Modelling and Simulation in Materials Science and Engineering,* vol. 18, no. 1, p. 015012, 2009.

[54] G. Pottlacher, H. Hosaeus, E. Kaschnitz, and A. Seifter, "Thermophysical properties of solid and liquidInconel 718 Alloy," *Scandinavian Journal of Metallurgy,* vol. 31, no. 3, pp. 161-168, 2002.

[55] J. Leinonen and A. von Lerber, "Snowflake melting simulation using smoothed particle hydrodynamics," *Journal of Geophysical Research: Atmospheres,* vol. 123, no. 3, pp. 1811-1825, 2018.

[56] D. Tarwidi, "Smoothed Particle Hydrodynamics Method for Two-dimensional Stefan Problem," *The 5th International Symposium on Computational Sciences,* 2016.

[57] A. Farrokhpanah, M. Bussmann, and J. Mostaghimi, "New smoothed particle hydrodynamics (SPH) formulation for modeling heat conduction with solidification and melting," *Numerical Heat Transfer, Part B: Fundamentals,* vol. 71, no. 4, pp. 299-312, 2017.

[58] *Standard terminology for additive manufacturing technologies*. ASTM Committee F42 on Additive Manufacturing Technologies, 2012.

[59] I. Gibson, D. W. Rosen, and B. Stucker, *Additive manufacturing technologies*. Springer, 2014.

[60] E. Alpaydin, *Introduction to machine learning*. MIT press, 2009.

[61] M. M. Francois, A. Sun, W. E. King, N. J. Henson, D. Tourret, C. A. Bronkhorst, N. N. Carlson, C. K. Newman, T. S. Haut, and J. Bakosi, "Modeling of additive manufacturing processes for metals: Challenges and opportunities," *Current Opinion in Solid State Materials Science,* vol. 21, no. LA-UR-16-24513; SAND-2017-6832J, 2017.

[62] M. Markl and C. Körner, "Multiscale modeling of powder bed–based additive manufacturing," *Annual Review of Materials Research,* vol. 46, pp. 93-123, 2016.

[63] S. K. Everton, M. Hirsch, P. Stravroulakis, R. K. Leach, and A. T. Clare, "Review of in-situ process monitoring and in-situ metrology for metal additive manufacturing," *Materials & Design,* vol. 95, pp. 431-445, 2016.

[64] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[65] X. Qi, G. Chen, Y. Li, X. Cheng, and C. Li, "Applying Neural-Network-Based Machine Learning to Additive Manufacturing: Current Applications, Challenges, and Future Perspectives," *Engineering,* 2019.

[66] J. Patel, "Data-Driven Modeling for Additive Manufacturing of Metals: Proceedings of a Workshop," The National Academies Press, Washington, DC, 2019.

[67] G. Tapia, S. Khairallah, M. Matthews, W. E. King, and A. Elwany, "Gaussian process-based surrogate modeling framework for process planning in laser powder-bed fusion additive manufacturing of 316L stainless steel," *The International Journal of Advanced Manufacturing Technology,* vol. 94, no. 9-12, pp. 3591-3603, 2018.

[68] Z. Hu and S. Mahadevan, "Uncertainty quantification and management in additive manufacturing: current status, needs, and opportunities," *The International Journal of Advanced Manufacturing Technology,* vol. 93, no. 5-8, pp. 2855-2874, 2017.

[69] T. Wang, T.-H. Kwok, C. Zhou, and S. Vader, "In-situ droplet inspection and closed-loop control system using machine learning for liquid metal jet printing," *Journal of manufacturing systems,* vol. 47, pp. 83-92, 2018.

[70] M. Mahesh, Y. Wong, J. Fuh, and H. Loh, "Benchmarking for comparative evaluation of RP systems and processes," *Rapid Prototyping Journal,* vol. 10, no. 2, pp. 123-135, 2004.

[71] J. Francis and L. Bian, "Deep Learning for Distortion Prediction in Laser-Based Additive Manufacturing using Big Data," *Manufacturing Letters,* vol. 20, pp. 10-14, 2019.

[72] S. L. Chan, Y. Lu, and Y. Wang, "Data-driven cost estimation for additive manufacturing in cybermanufacturing," *Journal of manufacturing systems,* vol. 46, pp. 115-126, 2018.

[73] Z. Zhu, N. Anwer, Q. Huang, and L. Mathieu, "Machine learning in tolerancing for additive manufacturing," *CIRP Annals,* vol. 67, no. 1, pp. 157-160, 2018.

[74] G. Tapia, A. Elwany, and H. Sang, "Prediction of porosity in metal-based additive manufacturing using spatial Gaussian process models," *Additive Manufacturing,* vol. 12, pp. 282-290, 2016.

[75] L. Meng and J. Zhang, "Process Design of Laser Powder Bed Fusion of Stainless Steel Using a Gaussian Process-Based Machine Learning Model," *JOM,* pp. 1-9, 2019.

[76] F. Caiazzo and A. Caggiano, "Laser direct metal deposition of 2024 Al alloy: Trace geometry prediction via machine learning," *Materials,* vol. 11, no. 3, p. 444, 2018.

[77] W. Rong-Ji, L. Xin-hua, W. Qing-ding, and W. Lingling, "Optimizing process parameters for selective laser sintering based on neural network and genetic algorithm," *The International Journal of Advanced Manufacturing Technology,* vol. 42, no. 11-12, pp. 1035-1042, 2009.

[78] J. Zhang, P. Wang, and R. X. Gao, "Deep learning-based tensile strength prediction in fused deposition modeling," *Computers in Industry,* vol. 107, pp. 11-21, 2019.

[79] Z. Li, Z. Zhang, J. Shi, and D. Wu, "Prediction of surface roughness in extrusion-based additive manufacturing with machine learning," *Robotics and Computer-Integrated Manufacturing,* vol. 57, pp. 488-495, 2019.

[80] M. Mozaffar, A. Paul, R. Al-Bahrani, S. Wolff, A. Choudhary, A. Agrawal, K. Ehmann, and J. Cao, "Data-driven prediction of the high-dimensional thermal history in directed energy deposition processes via recurrent neural networks," *Manufacturing letters,* vol. 18, pp. 35-39, 2018.

[81] L. Song, W. Huang, X. Han, and J. Mazumder, "Real-time composition monitoring using support vector regression of laser-induced plasma for laser additive manufacturing," *IEEE Transactions on Industrial Electronics,* vol. 64, no. 1, pp. 633-642, 2016.

[82] S. Chowdhury and A. Sarn, "Artificial neural network based geometric compensation for thermal deformation in additive manufacturing processes," University of Cincinnati, 2016.

[83] T. Kohonen, "An introduction to neural computing," *Neural networks,* vol. 1, no. 1, pp. 3-16, 1988.

[84] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Cognitive modeling,* vol. 5, no. 3, p. 1, 1988.

[85] A. R. Barron, "Complexity regularization with application to artificial neural networks," in *Nonparametric functional estimation and related topics*: Springer, 1991, pp. 561-576.

[86] F. Girosi, M. Jones, and T. Poggio, "Regularization theory and neural networks architectures," *Neural computation,* vol. 7, no. 2, pp. 219-269, 1995.

[87] S. Haykin, *Neural networks: a comprehensive foundation*. Prentice Hall PTR, 1994.

[88] C. K. Williams and C. E. Rasmussen, *Gaussian processes for machine learning* (no. 3). MIT Press Cambridge, MA, 2006.

[89] L. Meng and J. Zhang, "Process Design of Laser Powder Bed Fusion of Stainless Steel Using a Gaussian Process-Based Machine Learning Model," *JOM,* vol. 72, no. 1, pp. 420-428, 2020.

[90] X. Yao, S. K. Moon, and G. Bi, "A hybrid machine learning approach for additive manufacturing design feature recommendation," *Rapid Prototyping Journal,* vol. 23, no. 6, pp. 983-997, 2017.

[91] Y. Zhang, G. S. Hong, D. Ye, K. Zhu, and J. Y. Fuh, "Extraction and evaluation of melt pool, plume and spatter information for powder-bed fusion AM process monitoring," *Materials Design,* vol. 156, pp. 458-469, 2018.

[92] L. Scime and J. Beuth, "Anomaly detection and classification in a laser powder bed additive manufacturing process using a trained computer vision algorithm," *Additive Manufacturing,* vol. 19, pp. 114-126, 2018.

[93] L. Scime and J. Beuth, "A multi-scale convolutional neural network for autonomous anomaly detection and classification in a laser powder bed fusion additive manufacturing process," *Additive Manufacturing,* vol. 24, pp. 273-286, 2018.

[94] J. Mazumder, "Design for metallic additive manufacturing machine with capability for “Certify as You Build”," *Procedia CIRP,* vol. 36, pp. 187-192, 2015.

[95] M. Khanzadeh, S. Chowdhury, M. Marufuzzaman, M. A. Tschopp, and L. Bian, "Porosity prediction: Supervised-learning of thermal history for direct laser deposition," *Journal of manufacturing systems,* vol. 47, pp. 69-82, 2018.

[96] M. S. Tootooni, A. Dsouza, R. Donovan, P. K. Rao, Z. J. Kong, and P. Borgesen, "Classifying the dimensional variation in additive manufactured parts from laser-scanned three-dimensional point cloud data using machine learning approaches," *Journal of Manufacturing Science Engineering*, vol. 139, no. 9, p. 091005, 2017.

[97] K. Aoyagi, H. Wang, H. Sudo, and A. Chiba, "Simple method to construct process maps for additive manufacturing using a support vector machine," *Additive Manufacturing*, vol. 27, pp. 353-362, 2019.

[98] D. Ye, G. S. Hong, Y. Zhang, K. Zhu, and J. Y. H. Fuh, "Defect detection in selective laser melting technology by acoustic signals with deep belief networks," *The International Journal of Advanced Manufacturing Technology*, vol. 96, pp. 2791-2801, 2018.

[99] Z. Shen, X. Shang, M. Zhao, X. Dong, G. Xiong, and F.-Y. Wang, "A Learning-Based Framework for Error Compensation in 3-D Printing," *IEEE transactions on cybernetics*, vol. 49, no. 11, pp. 4042-4050, 2019.

[100] R. Jafari-Marandi, M. Khanzadeh, W. Tian, B. Smith, and L. Bian, "From in-situ monitoring toward high-throughput process control: cost-driven decision-making framework for laser-based additive manufacturing," *Journal of Manufacturing Systems*, vol. 51, pp. 29-41, 2019.

[101] C. Gobert, E. W. Reutzel, J. Petrich, A. R. Nassar, and S. Phoha, "Application of supervised machine learning for defect detection during metallic powder bed fusion additive manufacturing using high resolution imaging," *Additive Manufacturing*, vol. 21, pp. 517-528, 2018.

[102] A. Caggiano, J. Zhang, V. Alfieri, F. Caiazzo, R. Gao, and R. Teti, "Machine learning-based image processing for on-line defect recognition in additive manufacturing," *CIRP Annals*, vol. 68, no. 1, pp. 451-454, 2019.

[103] S. A. Shevchik, C. Kenel, C. Leinenbach, and K. Wasmer, "Acoustic emission for in situ quality monitoring in additive manufacturing using spectral convolutional neural networks," *Additive Manufacturing*, vol. 21, pp. 598-604, 2018.

[104] J. R. Quinlan, "Induction of decision trees," *Machine learning*, vol. 1, no. 1, pp. 81-106, 1986.

[105] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273-297, 1995.

[106] C.-W. Hsu and C.-J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE transactions on Neural Networks*, vol. 13, no. 2, pp. 415-425, 2002.

[107] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097-1105.

[108] P. Bühlmann and S. Van De Geer, *Statistics for high-dimensional data: methods, theory and applications*. Springer Science & Business Media, 2011.

[109] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An introduction to statistical learning*. Springer, 2013.

[110] A. Y. Ng, "Feature selection, L 1 vs. L 2 regularization, and rotational invariance," in *Proceedings of the twenty-first international conference on Machine learning*, 2004: ACM, p. 78.

[111] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural networks*, vol. 2, no. 5, pp. 359-366, 1989.

[112] E. Popova, T. M. Rodgers, X. Gong, A. Cecen, J. D. Madison, and S. R. Kalidindi, "Process-structure linkages using a data science approach: application to simulated additive manufacturing data," *Integrating Materials and Manufacturing Innovation*, vol. 6, no. 1, pp. 54-68, 2017.

[113] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929-1958, 2014.

[114] M. Khanzadeh, P. Rao, R. Jafari-Marandi, B. K. Smith, M. A. Tschopp, and L. Bian, "Quantifying geometric accuracy with unsupervised machine learning: using self-organizing map on fused filament fabrication additive manufacturing parts," *Journal of Manufacturing Science and Engineering*, vol. 140, no. 3, p. 031011, 2018.

[115] M. Khanzadeh, S. Chowdhury, M. A. Tschopp, H. R. Doude, M. Marufuzzaman, and L. Bian, "In-situ monitoring of melt pool images for porosity prediction in directed energy deposition processes," *IISE Transactions*, vol. 51, no. 5, pp. 437-455, 2019.

[116] H. Wu, Z. Yu, and Y. Wang, "Experimental study of the process failure diagnosis in additive manufacturing based on acoustic emission," *Measurement*, vol. 136, pp. 445-453, 2019.

[117] Y. Yang, M. He, and L. Li, "A new machine learning based geometry feature extraction approach for energy consumption estimation in mask image projection stereolithography," *Procedia CIRP*, vol. 80, pp. 741-745, 2019.

[118] Z. Wang, P. Liu, Y. Xiao, X. Cui, Z. Hu, and L. Chen, "A Data-driven Approach for Process Optimization of Metallic Additive Manufacturing under Uncertainty," *Journal of Manufacturing Science and Engineering*, vol. 141, no. 8, p. 081004, 2019.

[119] Y. Zhang and J. Zhang, "Finite element simulation and experimental validation of distortion and cracking failure phenomena in direct metal laser sintering fabricated component," *Additive Manufacturing*, vol. 16, pp. 49-57, 2017.

[120] Y. Zhang, X. Xiao, and J. Zhang, "Kinetic Monte Carlo simulation of sintering behavior of additively manufactured stainless steel powder particles using reconstructed microstructures from synchrotron X-ray microtomography," *Results in Physics*, vol. 13, p. 102336, 2019.

[121] J. Zhang, L. Wu, Y. Zhang, and L. Meng, "Phase field simulation of dendritic microstructure in additively manufactured titanium alloy," *Metal powder report*, vol. 74, no. 1, pp. 20-24, 2019.

[122] Y. Zhang and J. Zhang, "Modeling of solidification microstructure evolution in laser powder bed fusion fabricated 316L stainless steel using combined computational fluid dynamics and cellular automata," *Additive Manufacturing*, 2019.

[123] C. K. Williams and C. E. Rasmussen, "Gaussian processes for machine learning," *the MIT Press,* vol. 2, no. 3, p. 4, 2006.

[124] GPy, "GPy: A gaussian process framework in python," ed, 2012.

[125] I. Choquet, A. J. Shirvan, and H. Nilsson, "On the choice of electromagnetic model for short high-intensity arcs, applied to welding," *Journal of Physics D: Applied Physics,* vol. 45, no. 20, p. 205203, 2012.

[126] M. Courtois, M. Carin, P. Le Masson, S. Gaied, and M. Balabane, "A new approach to compute multi-reflections of laser beam in a keyhole for heat transfer and fluid flow modelling in laser welding," *Journal of Physics D: Applied Physics,* vol. 46, no. 50, p. 505305, 2013.

[127] W. Tan, N. S. Bailey, and Y. C. Shin, "Investigation of keyhole plume and molten pool based on a three-dimensional dynamic model with sharp interface formulation," *Journal of Physics D: Applied Physics,* vol. 46, no. 5, p. 055501, 2013.

[128] I. Yadroitsev, P. Krakhmalev, I. Yadroitsava, S. Johansson, and I. Smurov, "Energy input effect on morphology and microstructure of selective laser melting single track from metallic powder," *Journal of Materials Processing Technology,* vol. 213, no. 4, pp. 606-613, 2013.

[129] C. Kamath, B. El-dasher, G. F. Gallegos, W. E. King, and A. Sisto, "Density of additively-manufactured, 316L SS parts using laser powder-bed fusion at powers up to 400 W," *The International Journal of Advanced Manufacturing Technology,* vol. 74, no. 1-4, pp. 65-78, 2014.

[130] J. Zhang and Y.-G. Jung, *Additive manufacturing: materials, processes, quantifications and applications*. Butterworth-Heinemann, 2018.

[131] A. Verma, S. Tyagi, and K. Yang, "Modeling and optimization of direct metal laser sintering process," *The International Journal of Advanced Manufacturing Technology,* vol. 77, no. 5-8, pp. 847-860, 2015.

[132] G. Tapia, W. King, L. Johnson, R. Arroyave, I. Karaman, and A. Elwany, "Uncertainty propagation analysis of computational models in laser powder bed fusion additive manufacturing using polynomial chaos expansions," *Journal of Manufacturing Science and Engineering,* vol. 140, no. 12, p. 121006, 2018.

[133] Z. Hu and S. Mahadevan, "Uncertainty quantification in prediction of material properties during additive manufacturing," *Scripta materialia,* vol. 135, pp. 135-140, 2017.

[134] F. Lopez, P. Witherell, and B. Lane, "Identifying uncertainty in laser powder bed fusion additive manufacturing models," *Journal of Mechanical Design,* vol. 138, no. 11, p. 114502, 2016.

[135] L. Meng, B. McWilliams, W. Jarosinski, H.-Y. Park, Y.-G. Jung, J. Lee, and J. Zhang, "Machine Learning in Additive Manufacturing: A Review (DOI: 10.1007/s11837-020-04155-y)," *JOM,* 2020.

[136] X. Du, "Reliability‐based design optimization with dependent interval variables," *International Journal for Numerical Methods in Engineering,* vol. 91, no. 2, pp. 218-228, 2012.

[137] X. Du, A. Sudjianto, and B. Huang, "Reliability-based design with the mixture of random and interval variables," 2005.

[138] X. Zhuang, R. Pan, and X. Du, "Enhancing product robustness in reliability-based design optimization," *Reliability Engineering & System Safety*, vol. 138, pp. 145-153, 2015.

[139] X. Du, W. Chen, and Y. Wang, "Most probable point-based methods," in *Extreme Statistics in Nanoscale Memory Design*: Springer, 2010, pp. 179-202.

[140] I. Gholaminezhad, H. Assimi, A. Jamali, and D. A. Vajari, "Uncertainty quantification and robust modeling of selective laser melting process using stochastic multi-objective approach," *The International Journal of Advanced Manufacturing Technology*, vol. 86, no. 5-8, pp. 1425-1441, 2016.

[141] X. Du and W. Chen, "Sequential optimization and reliability assessment method for efficient probabilistic design," *J. Mech. Des.*, vol. 126, no. 2, pp. 225-233, 2004.

[142] A. M. Hasofer and N. C. Lind, "Exact and invariant second-moment code format," *Journal of the Engineering Mechanics division*, vol. 100, no. 1, pp. 111-121, 1974.

[143] L. Ma, J. Fong, B. Lane, S. Moylan, J. Filliben, A. Heckert, and L. Levine, "Using design of experiments in finite element modeling to identify critical variables for laser powder bed fusion," in *International solid freeform fabrication symposium*, 2015: Laboratory for Freeform Fabrication and the University of Texas Austin, TX, USA, pp. 219-228.

[144] M. Badrossamay and T. Childs, "Further studies in selective laser melting of stainless and tool steel powders," *International Journal of Machine Tools and Manufacture*, vol. 47, no. 5, pp. 779-784, 2007.

[145] C. S. Kim, "Thermophysical properties of stainless steels," Argonne National Lab., Ill.(USA), 1975.

VITA

Lingbin Meng was born in Zhuhai, Guangdong, Peoples Republic of China in 1994. He graduated from Sun Yet-Sen University (China) and Indiana University - Purdue University Indianapolis (IUPUI) with B.S. degree in Mechanical Engineering in 2016. Then he obtained an M.S. degree in Mechanical Engineering in 2017. He is currently a Ph.D. candidate in the School of Mechanical Engineering at Purdue University. His research under the direction of Professor Jing Zhang and Professor Benxin Wu focuses on the machine learning and probabilistic process design for the laser powder bed fusion process.

PUBLICATIONS

Journal papers

1. **Lingbin Meng**, Xiaoping Du, Brandon McWilliams, Jing Zhang, Probabilistic Feasibility Design of Laser Powder Bed Fusion Process Using Integrated First Order Reliability and Monte Carlo Methods, submitted
2. **Lingbin Meng**, Brandon McWilliams, William Jarosinski, Hye-Yeong Park, Yeon-Gil Jung, Jehyun Lee, Jing Zhang, Machine Learning in Additive Manufacturing: A Review, *Journal of The Minerals, Metals & Materials Society*, 2020. https://doi.org/10.1007/s11837-020-04155-y
3. **Lingbin Meng**, Xuehui Yang, Eduardo Salcedo, Dong-Cheon Baek, Jong E. Ryu, Zhe Lu, Jing Zhang, A Combined Modeling and Experimental Study of Tensile Properties of Additively Manufactured Polymeric Composite Materials, *Journal of Materials Engineering and Performance*, 2020. https://doi.org/10.1007/s11665-020-04746-5
4. **Lingbin Meng**, Jing Zhang, Process Design of Laser Powder Bed Fusion of Stainless Steel Using a Gaussian Process-Based Machine Learning Model, *Journal of The Minerals, Metals & Materials Society*, Vol. 72, pp. 420-428 (2020).
5. **Lingbin Meng**, Yi Zhang, Xuehui Yang, Jing Zhang, Atomistic Modeling of Resistivity Evolution of Copper Nanoparticle in Intense Pulsed Light Sintering Process, *Physica B: Condensed Matter*, Vol. 554, pp. 31-34 (2019).
6. Jing Zhang, Linmin Wu, Yi Zhang, and **Lingbin Meng**, Phase Field Simulation of Dendritic Microstructure in Additively Manufactured Titanium Alloy, *Metal powder report*, Vol. 74(1), pp. 20-24 (2019).

Book chapters

1. Yi Zhang, Weng Hoh Lee, Linmin Wu, **Lingbin Meng**, Yeon-Gil Jung, Jing Zhang, Chapter 7: Multi-Scale Multi-Physics Modeling of Laser Powder Bed Fusion Process, *Additive Manufacturing: Materials, Modeling, Processes, Quantifications and Applications*, Elsevier, 2018, ISBN: 9780128121559.