



Data analytics approach for melt-pool geometries in metal additive manufacturing

Seulbi Lee , Jian Peng , Dongwon Shin & Yoon Suk Choi

To cite this article: Seulbi Lee , Jian Peng , Dongwon Shin & Yoon Suk Choi (2019) Data analytics approach for melt-pool geometries in metal additive manufacturing, Science and Technology of Advanced Materials, 20:1, 972-978, DOI: [10.1080/14686996.2019.1671140](https://doi.org/10.1080/14686996.2019.1671140)

To link to this article: <https://doi.org/10.1080/14686996.2019.1671140>



© 2019 The Author(s). Published by National Institute for Materials Science in partnership with Taylor & Francis Group.



[View supplementary material](#)



Published online: 16 Oct 2019.



[Submit your article to this journal](#)



Article views: 3043



[View related articles](#)



[View Crossmark data](#)



Citing articles: 7 [View citing articles](#)

Data analytics approach for melt-pool geometries in metal additive manufacturing

Seulbi Lee^a, Jian Peng^b, Dongwon Shin^b and Yoon Suk Choi^a

^aSchool of Materials Science and Engineering, Pusan National University, Busan, Korea;

^bMaterials Science and Technology Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA

ABSTRACT

Modern data analytics was employed to understand and predict physics-based melt-pool formation by fabricating Ni alloy single tracks using powder bed fusion. An extensive database of melt-pool geometries was created, including processing parameters and material characteristics as input features. Correlation analysis provided insight for relationships between process parameters and melt-pools, and enabled the development of meaningful machine learning models via the use of highly correlated features. We successfully demonstrated that data analytics facilitates understanding of the inherent physics and reliable prediction of melt-pool geometries. This approach can serve as a basis for the melt-pool control and process optimization.

ARTICLE HISTORY

Received 31 July 2019

Revised 19 September 2019

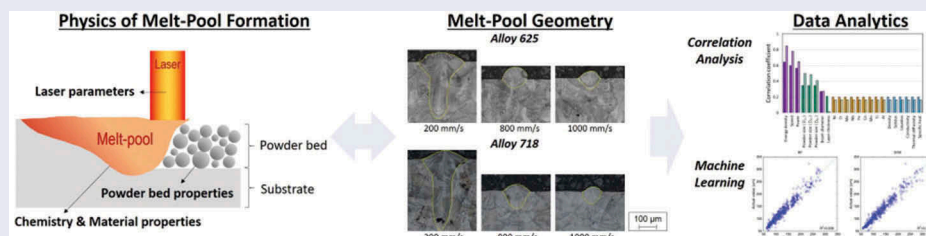
Accepted 19 September 2019

KEYWORDS

Powder bed fusion (PBF) process; melt-pool; single track; machine learning; correlation analysis

CLASSIFICATION

106 Metallic materials; 404 Materials informatics / Genomics



1. Introduction

Additive manufacturing (AM) offers many key benefits, which could change the industrial paradigm in various fields, as a tool-free, cost-efficient and digital approach to manufacturing [1]. Among the various metal AM processes, powder bed fusion (PBF) offers advantages for the fabrication of fully dense, near-net-shape metallic parts directly using high-energy heat sources, such as a laser and an electron beam. The PBF process comprises a sequence of layer depositions by selectively melting a powder layer (together with the previously deposited layer (or a substrate)), which leads to the formation of a series of melt-pools on a micrometer spatial scale, formed by harsh solidification conditions on a millisecond temporal scale. Here, increasing the reliability and stability of additively manufactured products as a function of processing conditions is a grand challenge.


The melt-pool geometry is important as a primary criterion for optimizing processing conditions because it is an indirect evidence of interactions between processing parameters and intrinsic materials properties. Also, there is a need to understand geometry-related phenomena, such as defect formation, microstructure evolution,

thermo-mechanical properties, and so on. Extensive research, as well as computational analysis, has been performed to clarify the effects of materials and process parameters on melt-pool characteristics, and the underlying physics [2–7]. However, there are more than 130 processing parameters associated with the PBF process [8,9], which interact intricately with various phase changes over a wide temperature range during the process. It is difficult to understand the overall relationships among the processing variables through well-controlled experiments or simulations using only a few variables. Also, it is challenging to quantitatively determine which features have utmost priorities, and how relevant they are to the characteristics of AM-processed parts.

Herein, we introduce an emerging data analytics approach to predict the multi-physics-based phenomena in AM, along with a scientific insight into the underlying mechanisms [10–15]. To predict melt-pool geometries, machine learning training was performed using a database of PBF-processed single tracks of Alloy 625 and Alloy 718 powders. Through the correlation analysis, the machine learning models were improved using highly correlated features, which enabled us to interpret the underlying physical

CONTACT Yoon Suk Choi  choiys@pusan.ac.kr  School of Materials Science and Engineering, Pusan National University, Busan, Korea

This article has been republished with minor changes. These changes do not impact the academic content of the article.

 The supplementary data for this article can be accessed [here](#).

© 2019 The Author(s). Published by National Institute for Materials Science in partnership with Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

mechanisms. Finally, the results demonstrated that the reasonably trained data analytics approach prioritizes key materials/process parameters of the melt-pool formation, with the physical relevance, and facilitates the accurate prediction of melt-pool geometries for the AM process optimization.

2. Experimental and data analytics details

2.1. Data acquisition through systematic microstructural analysis

Pre-alloyed spherical powders of Alloy 625 (21.4Cr-9.0Mo-3.3Nb-3.6Fe-0.2Co-0.3Mn-0.3Ti-0.2Al-Ni, by Chang Sung Co.) and of Alloy 718 (19.2Cr-3.1Mo-5.0Nb-18.4Fe-0.2Co-0.05Mn-1.0Ti-0.5Al-Ni, by Carpenter Inc.), were used to fabricate single tracks on a substrate of the same material. The PBF machines used in this study were Mlab and M2 cusing (Concept laser) with maximum laser powers of 100 W and 400 W, respectively. The laser power, scan speed,

beam diameter and layer thickness were used as process parameters for the PBF operation. Table 1 summarizes the laser parameters and powder bed properties under a wide range of process windows. All processes were performed at room temperature in an inert argon atmosphere.

For reliable data acquisition, sections in the middle of the 3-mm long single tracks were prepared for the metallography. Figure 1 presents the representative optical metallographs in a cross-section view, showing the melting mode from conduction to key-hole. As shown in Figure 2, the target output parameters (will be referred to as targets, hereafter) were set up such that they represent characteristics of melt-pool geometries: those measured in the substrate (width (w), depth (d), and area within the substrate (A_{sub}) in Figure 2) and in the powder bed (height (h) and area based on the height (A_h) in Figure 2). By measuring these target values through image analysis, a geometry database of 472 melt-pools was constructed (Table 1).

Table 1. List of PBF process parameters for fabricating single tracks of alloy 625 and alloy 718 powders.

Processing parameter	Material (machine)		
	Alloy 625 (M2)	Alloy 718 (Mlab)	Alloy 718 (M2)
Power (W)	120, 180, 240	70, 80, 90, 100	120, 180, 240
Scan speed (mm/s)	200, 400, 600, 800, 1000	500, 700, 800, 900, 1100	200, 400, 600, 800, 1000
Beam diameter (μm)	50, 100, 150	50	50, 100, 150
Layer thickness (μm)	25, 50	25, 35, 45	25, 50
Powder size distribution – D_{10} , D_{50} , D_{90} (μm)	18, 31, 49	23, 34, 45	15, 30, 47
Total no. of the data set	175	117	180

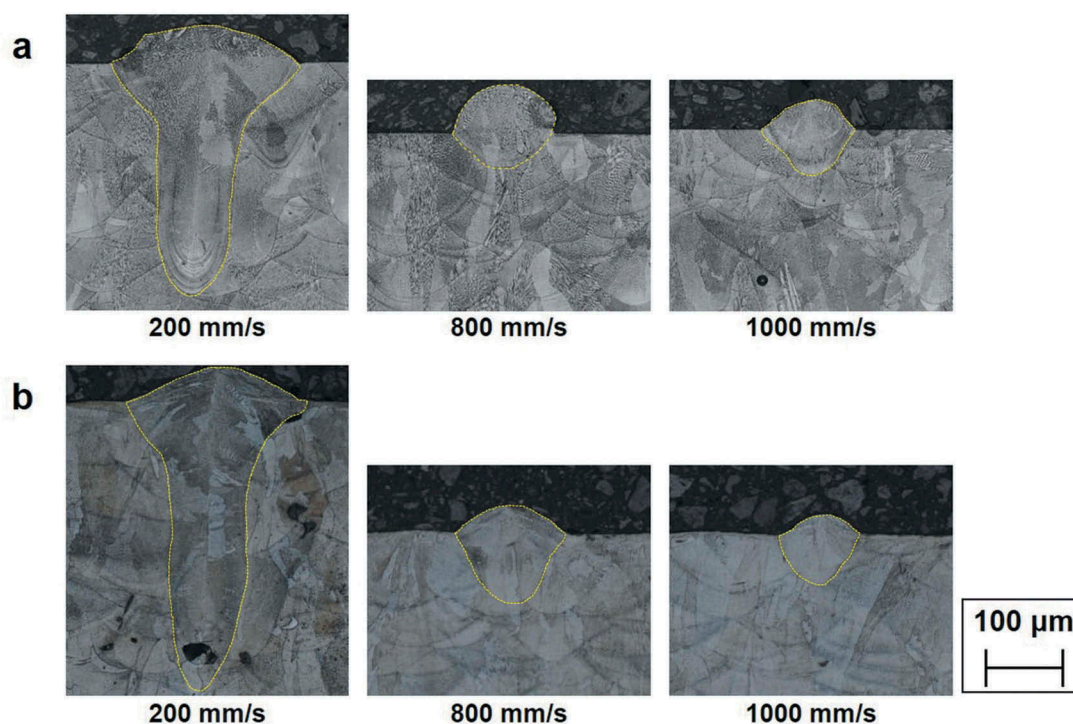


Figure 1. Metallographs of melt-pool cross sections for (a) Alloy 625 and (b) Alloy 718 as a function of the scan speed. Here, the laser power, beam diameter and layer thickness are 180 W, 50 μm and 25 μm , respectively.

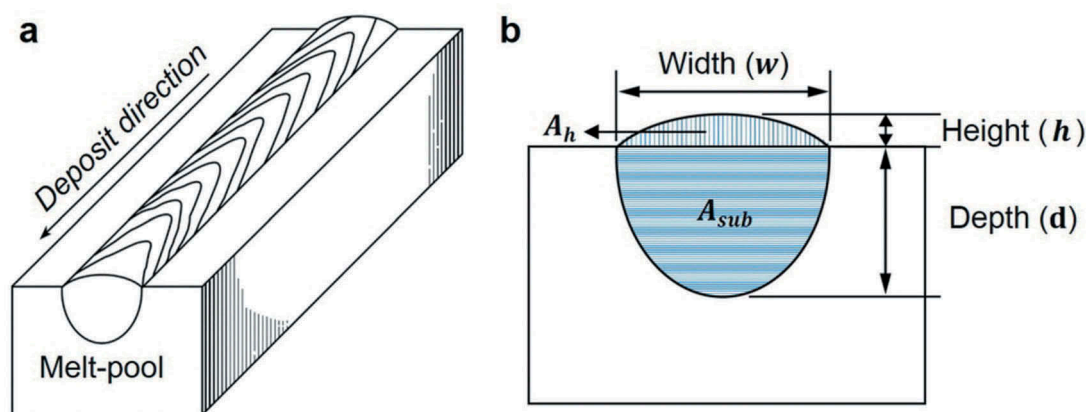


Figure 2. Schematic illustration of (a) single track deposition by the PBF process and (b) the melt-pool geometry parameters (targets) from the cross-section view.

2.2. Data analytics

(1) Correlation analysis

A total of 5 targets and 23 input features categorized by process parameters and material-related properties are listed in Table 2.

The correlation analysis between input features and targets was performed using both the advanced maximal information coefficient (MIC) approach and the conventional Pearson's correlation coefficient (PCC) approach. The PCC approach is suitable for seeking the linear relationship between input features and target, while MIC approach can identify non-linear relationships of high-dimensional large dataset [10,15–17]. By using two different approaches, we expect to have insights for different statistical aspects between input features and target properties and then to inspire domain expert for understanding the physical phenomena [15]. In the PCC approach, coefficient value has between -1 and $+1$, where a positive value indicates a direct relationship and a negative correlation coefficient indicates a reciprocal relationship between input features and targets. The MIC belongs to the maximal information-based nonparametric exploration (MINE) [17], having between 0 to 1 coefficient value. To compare the results of both approaches, MIC values and absolute PCC values were used. Based on the quantitative comparison of these values, we can identify which physical phenomenon is predominant for each melt-pool formation and develop

meaningful machine learning models using highly ranked features.

(2) Machine learning

In this study, six machine learning algorithms were employed for generality (i.e. not limited to a specific data analytic algorithm): Bayesian ridge regression (BR) [18], kernel ridge regression (KR) [19], linear regression (LR) [20], nearest neighbors regression (NN) [21], random forest regression (RF) [22], and support vector machine (SVM) [23]. All of these are in the open-source data analytic toolkit ASCENDS [24]. Based on a 5-fold cross-validation, the predictability of machine learning models was evaluated by calculating the coefficient of determination (R^2) between the actual and predicted values as a function of the number of top-ranked features (5, 10, 15, 20 and all) associated with each correlation approach.

3. Results and discussion

Figure 3 shows the results of the correlation analysis for the five targets using both MIC and PCC approaches. Here, asterisk marks on PCC correlation bars indicate negative correlation coefficients, which mean inverse influence. The rank order of the features shows a distinct difference between the melt-pool geometries associated with the substrate (w , d and A_{sub}) and with the

Table 2. List of input features and targets used in the present study.

Classification	Constituents	
Input features	Chemistry of powders	Ni, Cr, Mo, Nb, Fe, Co, Mn, Ti, Al
	Materials thermal property	Solidus, liquidus, density, conductivity, thermal diffusivity, specific heat
Targets for the melt-pool geometry	Information of the powder bed	Powder size distribution (D_{10} , D_{50} , D_{90}), layer thickness
	Laser parameters	Power, scan speed, energy density*, beam diameter
Targets for the melt-pool geometry	Measured in the substrate	Width, depth, area within the substrate
	Measured in the powder bed	Height, area based on the height

* Energy density is given by $E_p = \frac{P}{v \cdot l}$ (herein, E_p is the surface energy density, P is the power, v is the scan speed, and l is the layer thickness of powder bed.)

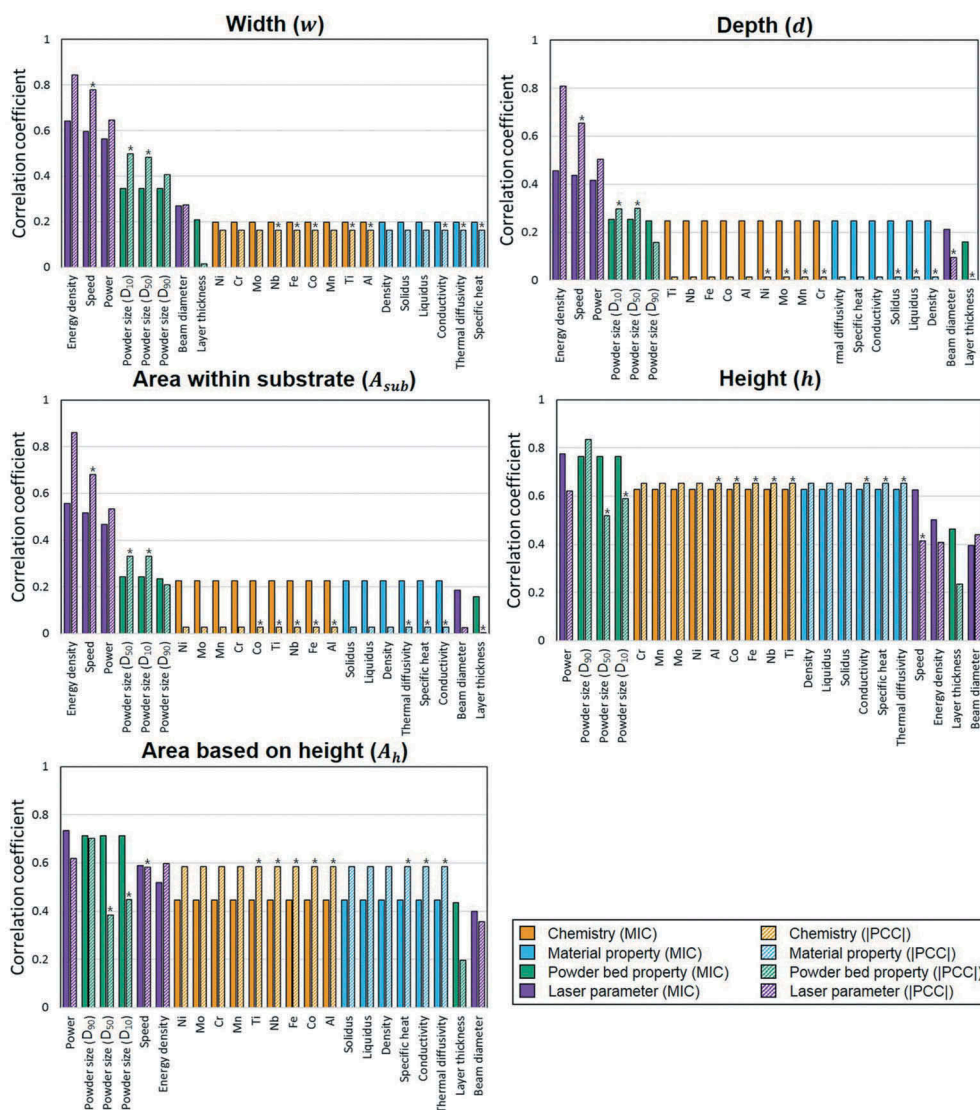


Figure 3. Results of correlation analysis using the MIC approach, compared with the absolute PCC values (bars marked by patterns), for the five targets (w , d , A_{sub} , h and A_h). In the plots, asterisk marks on the PCC bars indicate negative values.

powder bed (h and A_h). Also, the rank orders are different depending on the correlation approach and targets, as listed in supplementary Table 1.

In Figure 3, the energy density, laser power, scan speed, and powder size distribution were highly ranked for most targets. Here, first three parameters are associated with how much energy is emitted by the heat source. The powder size distribution (in the powder bed) is related to the energy absorptivity. The laser beam interacts with the powder bed first, and multiple scattering between the powders causes a higher laser absorptivity than on the bare substrate. In particular, the laser absorptivity tends to be chiefly dependent on the powder size distribution in the top layer, which affected the primary reflection of the laser [25].

To identify the dependence for selected features based on correlation analysis, we conducted the accuracy analysis of 5 machine learning training cycles, as shown in Figure 4 (numeric values with standard deviations are listed in supplementary Table2).

In Figure 4, the accuracies (R^2 -values) show higher predictability for the melt-pool geometries associated with the substrate (w , d and A_{sub}), when the PCC approach is employed, and higher predictability for the melt-pool geometries associated with the powder bed (h and A_h) under the MIC approach (although w and A_{sub} are highly predictable using both approaches). In general, the PCC approach is suitable for capturing linear relationships between the input features and targets; whereas the MIC approach is suitable for capturing a wide range of relationships, including linear, exponential, or periodic, or even all functional relationships [17]. A possible interpretation of the results is that melt-pool geometries measured in the substrate have linear relationships with the most highly ranked features, whereas those measured in the powder bed have complex relationships with the most highly ranked features. Also, the equitability of the MIC approach seems to facilitate interpretation of the relationships with noises [17] caused by surface defects,

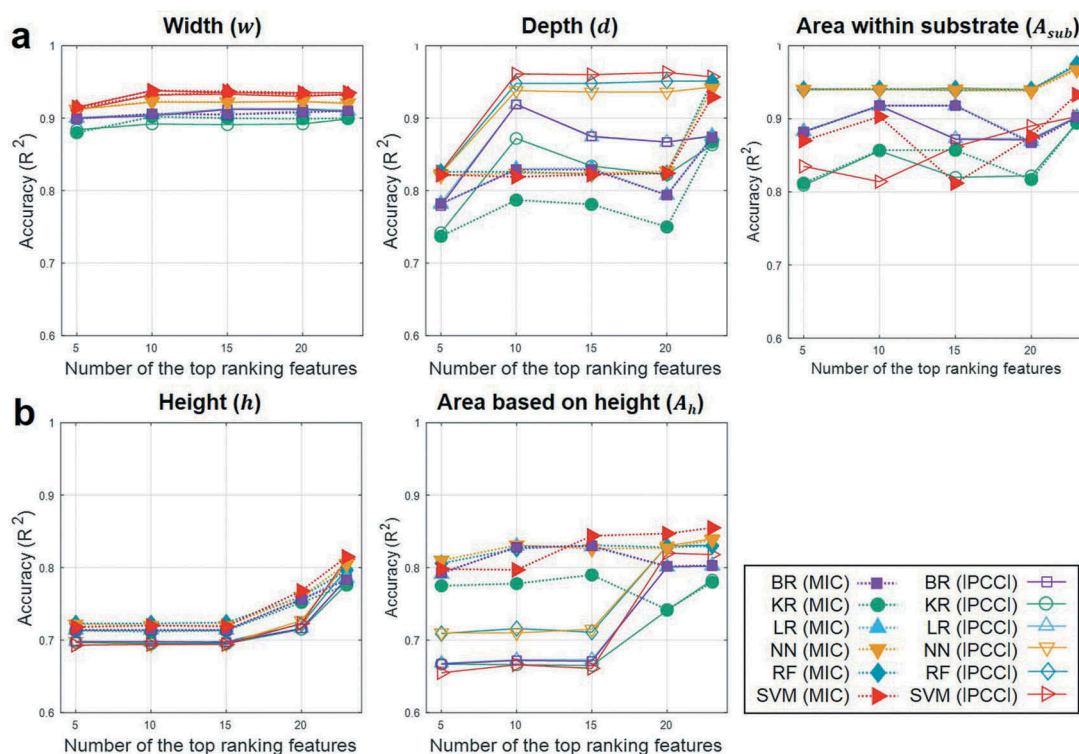


Figure 4. Accuracy analysis of the five targets (w , d , A_{sub} , h and A_h) as a function of the number of the top-ranked features (5, 10, 15, 20 and all), determined by the two correlation approaches (MIC and PCC), for the six different machine learning algorithms (BR, KR, LR, NN, RF and SVM). Here, R^2 -values are averages after 5 machine learning training cycles.

vaporization losses, height roughness, and so on. In the following, the main discussion will focus on correlation approaches with better machine learning accuracies (PCC for w , d and A_{sub} , and MIC for h and A_h).

In Figure 3, the energy density, laser power, scan speed, and powder size distribution ranked high for w , d and A_{sub} based on the PCC approach. Chemistries and material properties followed in the rankings; they show different PCC correlation values between w (about 0.17), and d and A_{sub} (≤ 0.03). Here, the similar correlation tendency for d and A_{sub} is probably due to a proportionality between two targets. Considering the melt-pool formation under the laser irradiation, d is determined by direct laser penetration (the laser cavity) resulting from the largest recoil pressure. However, w is decided by the heat transfer, mainly through the convection of liquid flow surrounding the laser cavity [2,3,7].

This phenomenon was observed by Zhao et al [2], who conducted in situ observations of the melt-pool formation through the synchrotron X-ray imaging technique. The experimental evidence shows that the depth of the laser cavity is similar to the d value of the melt-pool, while the w value of the melt-pool continuously increases because the cavity has a constant width (similar to the beam diameter) during melt-pool evolution [2]. Herein, it can be seen that the PCC correlation value of the beam diameter for w is higher than that for d , also having a positive value (Figure 3). The computational study by Khairallah et al. [3] showed

that the liquid metal under an incident laser moves away from the bottom to the rear part of the melt-pool through a dynamical vortex fluid for heat dissipation. In brief, the direct laser penetration is a major phenomenon that determines d , and the heat transfer through the fluid convection is a major physical phenomenon that determines w . The correlation analysis performed in the present study accurately captured such a difference.

It is interesting in Figure 3 that the powder size distribution, chemistries, and material properties have much higher correlation values for h and A_h than do the targets associated with the substrate (w , d and A_{sub}) for both PCC and MIC approaches. Since the melt volume above the substrate is determined by the amount of consolidated powders, the powder size distribution is the highly correlated feature for h and A_h in terms of the powder packing density [26]. After the laser passes, the final h and A_h values will be decided by the fluid characteristics during cooling. Since fluid characteristics are intrinsically based on material properties and chemistries, these features have high correlation values with melt-pool geometries associated with powder bed.

Based on the accuracies of the trained models in Figure 4, the minimum numbers of top-ranked features, excluding noise-inducing features and including physically key features [15], were determined to be 10 and 20 for melt-pool geometries measured in the substrate (w , d and A_{sub}) and the powder bed (h and A_h),

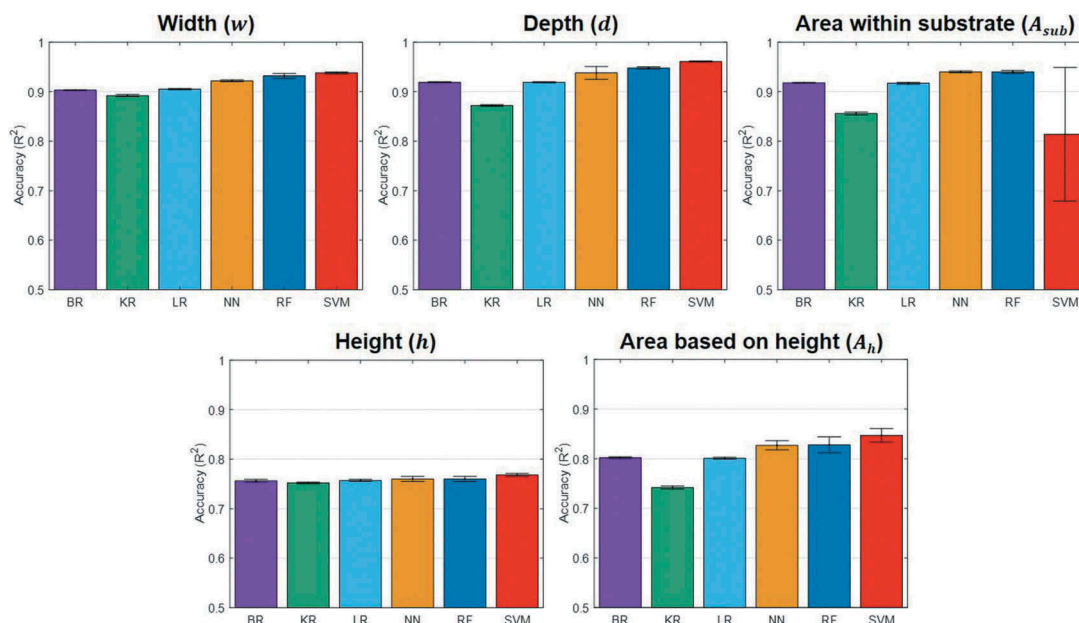


Figure 5. Accuracy analysis of the five targets (w , d , A_{sub} , h and A_h) for the six optimized machine learning algorithms (BR, KR, LR, NN, RF and SVM). Here, average R^2 -values and standard deviations are plotted after 5 machine learning training cycles.

respectively. Figure 5 shows the resulting accuracies (R^2 -values) of the six optimized machine learning models.

In Figure 5, for w , d and A_{sub} , the machine learning models show fairly high predictability, having a highest R^2 -value of more than 0.9, which is striking, considering a limited number of melt-pool data available for training. However, for h and A_h , the results show rather low predictability, having R^2 -values between 0.75 and 0.85. The reason for the relatively low predictability of h and A_h seems to be associated with the complexity of dynamic powder motions. During the laser irradiation, vapor-driven flow causes powder spatter ejection backward, upward, or even forward with respect to the scan direction [3–5]. In particular, the development of dynamic vapor flow as a result of uneven temperature distribution in the melt-pool makes it even more difficult to predict powder motion [2–5]. The powder entrainment toward the melt-pool caused by the interaction between the metal vapor and the surrounding gas pressure causes more complex powder motion [2,4,5]. Also, the difficulty of controlling the layer thickness uniformity over the powder bed seems to give an additional effect on the uncertainty of the powder bed properties [27]. Therefore, for the improvement of predictability for h and A_h , it is suggested to incorporate input features representing the gas- and vapor-driven powder behaviors, and related to the material properties associated with optics and rheological characteristics. Also, we will consider the integrated input features (such as D_{90}/D_{10} for the powder distribution, new energy density formula, and so on) and the new representative targets for defining the melt-pool geometries having physically meaningful insights.

4. Conclusions

Modern data analytics approaches were employed to check the feasibility of predicting melt-pool geometries of PBF single-tracks and the physical relevance to the real phenomena, and the following conclusions were drawn.

- (1) Machine learning models better predicted melt-pool geometries measured in the substrate with the PCC approach, and those measured in the powder bed with the MIC approach. It implies different interactions of input features and output targets by the laser irradiation to the powder bed and substrate.
- (2) The correlation analysis suggested that the depth, width and height of the melt-pool are physically relevant with the laser penetration, fluid convection, and fluid properties of melted powders, respectively.
- (3) Machine learning models optimized with high ranking features accurately predicted melt-pool geometries measured in the substrate, but showed relatively low predictability for those measured in the powder, which seemed to be associated with complex and dynamic powder motions.
- (4) Input features prioritized by the correlation analysis and machine learning training well captured the relevant physical phenomena to form melt-pools. A strong possibility of using modern data analytics to accelerate the PBF process optimization for the melt-pool control was demonstrated.

Acknowledgments

The authors are grateful to Dr. J.-K. Hong of the Korea Institute of Materials Science (KIMS) and Dr. Y. Kim of KAMI Co. Ltd. for the sample fabrication throughout the project.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported by the Industrial Strategic Technology Development Program [10077677] and the Technology Innovation Program [20000201] funded by the Ministry of Trade, Industry and Energy (MOTIE, Korea).

ORCID

Dongwon Shin  <http://orcid.org/0000-0002-5797-3423>

References

- [1] Gorsse S, Hutchinson C, Gouné M, et al. Additive manufacturing of metals: a brief review of the characteristic microstructures and properties of steels, Ti-6Al-4V and high-entropy alloys. *Sci Technol Adv Mater.* 2017;18(1):584–610.
- [2] Zhao C, Fezzaa K, Cunningham RW, et al. Real-time monitoring of laser powder bed fusion process using high-speed X-ray imaging and diffraction. *Sci Rep.* 2017;7:3602.
- [3] Khairallah SA, Anderson AT, Rubenchik A, et al. Laser powder-bed fusion additive manufacturing: physics of complex melt flow and formation mechanisms of pores, spatter, and denudation zones. *Acta Mater.* 2016;108:36–45.
- [4] Bidare P, Bitharas I, Ward RM, et al. Fluid and particle dynamics in laser powder bed fusion. *Acta Mater.* 2018;142:107–120.
- [5] Matthews MJ, Guss G, Khairallah SA, et al. Denudation of metal powder layers in laser powder-bed fusion processes. *Acta Mater.* 2016;114:33–42.
- [6] Leung CLA, Marussi S, Atwood RC, et al. In situ X-ray imaging of defect and molten pool dynamics in laser additive manufacturing. *Nat Commun.* 2018;9:1355.
- [7] Ki H, Mohanty PS, Mazumder J. Modeling of laser keyhole welding: part II. Simulation of keyhole evolution, velocity, temperature profile, and experimental verification. *Metall Mater Trans A Phys Metall Mater Sci.* 2002;33A:1831–1842.
- [8] Yadroitsev I. Selective laser melting: direct manufacturing of 3D-objects by selective laser melting of metal powders. Saarbrücken: LAP Lambert Academic Publishing; 2009.
- [9] Rehme O, Emmelmann C. Reproducibility for properties of selective laser melting products. *Proceedings of the Third International WLT-Conference on Lasers in Manufacturing*; 2005 Jun 13–16; Munich, Germany.
- [10] Shin D, Lee S, Shyam A, et al. Petascale supercomputing to accelerate the design of high-temperature alloys. *Sci Technol Adv Mater.* 2017;18:828–838.
- [11] Kamath C. Data mining and statistical inference in selective laser melting. *Int J Adv Manuf Technol.* 2016;86:1659–1677.
- [12] Agrawal A, Deshpande PD, Cecen A, et al. Exploration of data science techniques to predict fatigue strength of steel from composition and processing parameters. *Integr Mater Manuf Innov.* 2014;3:1–19.
- [13] Mishra S, DebRoy T. Tailoring gas tungsten arc weld geometry using a genetic algorithm and a neural network trained with convective heat flow calculations. *Mater Sci Eng A.* 2007;454–455, 477–486.
- [14] Wang J, Nobakht AY, Blanks JD, et al. Machine learning for thermal transport analysis of aluminum alloys with precipitate morphology. *Adv Theory Simul.* 2019;1800196:1–10.
- [15] Shin D, Yamamoto Y, Brady MP, et al. Modern data analytics approach to predict creep of high-temperature alloys. *Acta Mater.* 2019;168:321–330.
- [16] Shin D, Shyam A, Lee S, et al. Solute segregation at the Al/θ'-Al₂Cu interface in Al-Cu alloys. *Acta Mater.* 2017;141:327–340.
- [17] Reshef DN, Reshef YA, Finucane HK, et al. Detecting novel associations in large data sets. *Science.* 2011;334:1518–1524.
- [18] Yu K, Tresp V, Schwaighofer A. Learning Gaussian processes from multiple tasks. *Proceedings of the 22nd International Conference on Machine Learning (ICML-05)*; 2005; Bonn, Germany. p. 1012–1019.
- [19] Zhang Y, Duchi J, Wainwright M. Divide and conquer kernel ridge regression: a distributed algorithm with minimax optimal rates. *J. Mach. Learn. Res.* 2015;16:3299–3340.
- [20] Neter J, Kutner MH, Nachtsheim CJ, et al. *Applied linear statistical models*. Vol. 4, Chicago: Irwin; 1996.
- [21] Cherkassky V, Ma Y. Comparison of model selection for regression. *Neural Comput.* 2003;15:1691–1714.
- [22] Segal MR. Machine learning benchmarks and random forest regression. *Biostatistics.* 2004;1–14.
- [23] Hsu C-W, Lin C-J. A comparison of methods for multi-class support vector machine. *IEEE Trans Neural Netw.* 2002;13:415–425.
- [24] Lee S, Shin D, Peng J. ASCENDS: a data SciENce toolkit for non-data scientists; 2019. Available from: <https://github.com/ornlpmcp/ASCENDS>.
- [25] King WE, Anderson AT, Ferencz RM, et al. Laser powder-bed fusion additive manufacturing of metals: physics, computational, and materials challenges. *Appl Phys Rev.* 2015;041304:1–26.
- [26] Lee YS, Nandwana P, Zhang W. Dynamic simulation of powder packing structure for powder bed additive manufacturing. *Int J Adv Manuf Technol.* 2018;96(1–4):1507–1520.
- [27] Moges T, Ameta G, Witherell P. A review of model inaccuracy and parameter uncertainty in laser powder bed fusion models and simulations. *J Manuf Sci Eng.* 2019;141:1–14.