Additive Manufacturing Letters 6 (2023) 100137

Contents lists available at ScienceDirect

Additive Manufacturing Letters

journal homepage: www.elsevier.com/locate/addlet

Short Communication

Synchrotron imaging derived relationship between process parameters and build quality for directed energy deposition additively manufactured IN718

S.V. Notley[a,*], Y. Chen[b,c,d,e], N.A. Thacker[c], P.D. Lee[b], G. Panoutsos[a]

[a] University of Sheffield, Western Bank, Sheffield, S10 2TN, United Kingdom
[b] School of Mechanical Engineering, University College London, Gower Street, London, WC1E 6BT, United Kingdom
[c] School of Materials, University of Manchester, Oxford Road, Manchester, M13 9PL, United Kingdom
[d] School of Engieering, RMIT University, La Trobe St, Melbourne, VIC 3000, Australia
[e] The European Synchrotron Radiation Facility, 27 rue de Martyre, Grenoble, 38000, France

ARTICLE INFO

Keywords:
Laser additive manufacturing
Directed energy deposition
Neural networks
Meltpool geometry
In-situ x-ray imaging

ABSTRACT

Laser additive manufacturing is transforming several industrial sectors, especially the directed energy deposition process. A key challenge in the widespread uptake of this emerging technology is the formation of undesirable microstructural features such as pores, cracks, and large epitaxial grains. The trial and error approach to establish the relationship between process parameters and material properties is problematic due to the transient nature of the process and the number of parameters involved. In this work, the relationship between process parameters, melt pool geometry and quality of build measures, using directed energy deposition additive manufacturing for IN718, is quantified using neural networks as generalised regressors in a statistically robust manner. The data was acquired using *in-situ* synchrotron x-ray imaging providing unique and accurate measurements for our analysis. An analysis of the variations across repeated measurements show heteroscedastic error characteristics that are accounted for using a principled nonlinear data transformation method. The results of the analysis show that surface roughness correlates with melt pool geometry while the track height directly correlates with process parameters indicating a potential to directly control efficiency and layer thickness while independently minimising surface roughness.

1. Introduction

Laser Additive Manufacturing (LAM) enables the direct fabrication of complex geometries layer-by-layer from powder or wire feedstock. Its flexibility in manufacturing is transforming several industrial sectors including aerospace, biomedical, and automotive. Directed Energy Deposition Additive Manufacturing (DED-AM) is one of the most industrial appealing LAM methods due to its capability for producing large near-net-shape free-form components. However, one of the key challenges which restrict the widespread industrialisation of DED-AM is the formation of undesirable microstructural features such as pores, cracks or large epitaxial grains. Trial and error experiments have been used to establish the relationship between process parameters and resulting material properties, however, an accurate description is still unavailable due to the highly transient nature of the process and the many parameters involved.

In 2019, Chen et al. [1] implemented the world's first L-DED process replicator on the Diamond Light Source, UK, and reported first-hand information including the melt pool behavior, defect formation, and micro-structural revolution of the process. *In-situ* and *operando* fast acquisition Synchrotron X-ray imaging, as a promising approach, has been applied to LAM, assisting the interpretation of the processes [2–4] with a series of previously unseen phenomena have been revealed in both real and reciprocal space. Chen et al. [5] used X-ray radiography to visualise how keyhole formation connects to porosity during laser powder bed fusion (L-PBF). These Synchrotron experiments have focused on the observation of physical phenomena including surface roughness, and defect density during LAM directly correlate with the processing parameters. However, due to the constraint of synchrotron time and the complexity of the experiment, limited process maps and repeats can be realised. Many other groups have performed in-situ imaging of LAM; however, most have been on laser powder bed fusion. One group, Wolff et al. [6], performed a study identifying the mechanisms by which laser-matter interaction influences powder flow and porosity formation. However, the study did not explore the parameter-properties relationship.

* Corresponding author.
E-mail addresses: s.v.notley@sheffield.ac.uk (S.V. Notley), yunhui.chen@rmit.edu.au (Y. Chen), neil.a.thacker@manchester.ac.uk (N.A. Thacker), peter.lee@ucl.ac.uk (P.D. Lee), g.panoutsos@sheffield.ac.uk (G. Panoutsos).

https://doi.org/10.1016/j.addlet.2023.100137
Received 23 January 2023; Received in revised form 4 March 2023; Accepted 6 March 2023

The aim of this work is to quantify, for IN718 super alloy, the relationship between process parameters, melt pool geometry and two build properties: Surface Roughness (SR) and Resultant Track Height (RTH). To this end we employ neural networks, specifically multi-layer perceptrons, as generalised regressors and subsequently assess the potential for finding sets of process parameters that optimise the material properties, in this case, minimising surface roughness while maximising efficient use of feedstock material.

Neural networks, in particular the multilayer perceptron (MLP), have been employed in the analysis of the LAM process [7–9]. However, there are particular fundamental issues, related to small data samples and error characteristics, that need to be accounted for with AM data when optimising models. Qi et al. [10] systematically reviewed the recent work investigating process-property-performance linkage, using neural networks, for a number of different additive manufacturing methods. One of the challenges identified, as noted above, is the limited data available to constrain models and give useful predictions in unexplored areas of the input domain (generalisation). In the machine learning literature methods, such as data augmentation, have been used to increase dataset sizes, however, these methods provide no more independent information to refine model parameters with respect to knowledge discovery. Rather, these methods are a brute force method to enforce robustness and are more akin to regularisation via noise injection [11].

For robust inferences to be concluded with confidence, using small data, it is imperative that information is used efficiently to give the best accuracy of model parameters. In this study, we employ a methodological framework [12] that accounts for the properties of stochastic variations, derived from repeated measurements, to enable the full use of available data during optimisation of model parameters and further enables results to be interpreted from a statistically motivated perspective.

The experimental design, with repeated measurements, allows stochastic variations to be characterised using Bland-Altman plots [13]. The stochastic variations in the measurements were found to be heteroscedastic and correlate with the measurement value itself. It is well known, for linear systems, that the ordinary least squares (OLS) solution results in the best linear unbiased estimate (BLUE) of the model parameters when the uncertainties are uncorrelated, the expectation value of zero and equal variance (homoscedastic). Therefore, we use power law functions to transform measurements to a domain where the variance is stabilised and the stochastic variations are homoscedastic. The transformation functions are optimised with a novel approach (see Section 2.3) that is independent of the model and maximises the likelihood of the stochastic variations being normally distributed.

Stabilisation and normalisation (in the sense of transforming to a normal distribution) allow the distribution of the stochastic variations to be robustly summarised with a single statistic; the variance (assuming expectation to be zero). In terms of a neural network, the dependent output variable may be standardised by the standard deviation of the *variations* resulting in an output space that has a z-score style of statistical interpretation; the units may now be expressed in standard deviations of 'noise'. Standardisation of the dependent variable in this manner also has implications in interpreting the reliability of the regression found by the neural network in that the squared error cost function is now expected to follow a reduced chi-square distribution. This property may be used to objectively assess the goodness of fit of the neural network and may further be used, in this sense, as a stopping criterion for the training process without the need for cross-validation and thus maximising the amount of training data available.

The results of the analysis show that, while surface roughness may be optimised through process parameters settings, this is an indirect relationship. We hypothesise that the determining factor in surface roughness is the geometry and dynamics of the melt pool while efficiency in the use of feedstock is determined by the relationship between energy density and powder feed rate.

2. Methods and data

2.1. Data

Synchrotron in situ X-ray imaging of the DED-AM process was performed at I12: Joint Engineering, Environmental, and Processing (JEEP) beamline at the Diamond Light Source (Beam time number LTP MT20096). A DED-AM process replicator, which faithfully replicates a commercial DED-AM system, was designed to fit onto a synchrotron beamline to capture the key physics during the laser-matter interaction under X-ray. Figure 1 shows a schematic of the experimental set-up and details of the process may be found in [14]. A process map was explored with a range of process parameters including laser power, traverse speed, and powder feed rate listed in Table 1. Measurements of the melt pool morphology (length and height), top surface roughness, and resultant track height were mapped from the radiographs.

To estimate top surface roughness and resultant track height the radiographs were binarised. The resultant track height was calculated as the mean value of the top surface height variations. The top surface roughness was then estimated as the arithmetic mean deviation around the resultant track height as given by

$$R_a = \frac{1}{N}\sum_{n=1}^{N} |h_i| \tag{1}$$

where h_i is the deviation between a point on the profile and the resultant track height (see [14] for more details of the image processing).

For each experiment, 3 repeated measurements were made from the radiographs for length, the height of the melt pool, surface roughness and the resultant track height of the deposition. Each measurement was made at fixed points, along each track, spatially separated to ensure no correlation between measurements due to dynamics.

2.2. Regression analysis

Neural network models were employed, as generalised regressors, to investigate the relationship between metal additive manufacturing (AM) process parameters, the resultant build properties (roughness and material deposition) and melt pool geometrical measurements (length and height).

As with many cases in AM, the amount of data available is at a premium [7,10]. The approach adopted in this work, to optimise model parameters, exploits repeated measurements to ensure good fits to data, based on statistical criteria, while maximising the use of the data available (as described in [12]). The method ensures data meet the assumptions of using a least-squares cost function, via stabilising the variance of the noise estimated from the difference of repeats. Scaling of the data, by stable noise estimates, allows optimisation of model parameters to be stopped based on the criteria of the distribution of the squared error having a chi-squared per-degree-of-freedom (χ^2_{red}) equal to 1. The model parameters were optimised using Levenburg-Marquardt error back propagation [15,16].

All neural networks trained were multi-layer perceptrons with tanh activation functions in the hidden layer and a linear output neuron. For models trained to find mappings from process parameters to melt pool geometry, surface roughness, or RTH the input space is three dimensions consisting of laser power, head speed, and feed rates. During training, each combination of the input factors is presented in combination with every repeat of the output target data.

For the mappings between the melt pool geometry to the surface roughness or RTH, the input space is two-dimensional consisting of the length and height measurements. Every measurement and repeat of the input factors were presented in combination with every repeat of the output target data.

The model order selected for each network (the number of hidden units) was selected as the lowest model order capable of finding an

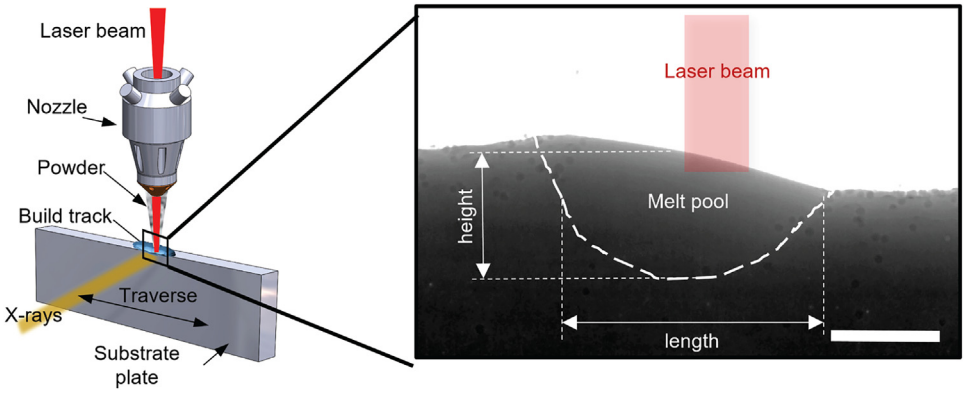

Fig. 1. Schematic of the melt pool measurement process from the x-ray images.

Table 1
Process parameters settings. The dataset consists of samples for all combinations of these values.

Laser Power (W)	50	100	150	200
Substrate Traverse Speed (mms^{-1})	1, 2.5 & 5	1, 2.5 & 5	1, 2.5 & 5	1,2.5 & 5
Powder Feed rate ($gmin^{-1}$)	1,2, & 3	1,2, & 3	1, 2, & 3	1, 2, & 3

optimal parameter solution that gave a square-error distribution with $\chi^2_{red} = 1$.

For all fits, the least-squares cost function, residual distributions, and regression plots between training target data and model outputs are shown in the supplementary material.

2.3. Data transformation and variance stabilisation

Bland-Altman plots (see supplementary material) show that the melt pool height, roughness, and surface addition measurements have noise distributions whose variance is correlated with the mean of the measurements[1]; similar observations were found in [12]. For datasets with unequal variance, minimization of an ordinary least-squares cost function is no longer guaranteed to give the best linear unbiased estimate (BLUE) of the optimal model parameters.

To overcome this issue, and ensure data is being used to best efficiency, we consider variance stabilization [17–20] via a non-linear transformation of the form:

$$f(x) = x^{\theta} \tag{2}$$

to transform the stochastic variations to a domain such that the data meets the ordinary least squares assumptions. The parameter θ is optimised by minimising, with respect to θ, the negative log-likelihood function given by:

$$-\log L(v_i, r_i|\theta) = \sum_{i=1}^{I} \Big[\frac{(f(v_i;\theta) - f(r_i;\theta))^2}{2\sigma_f^2} + \log(\sqrt{2\pi}\sigma_f) - \log\Big(\Big|\frac{\partial f(v_i;\theta)}{\partial v_i} + \frac{\partial f(r_i;\theta)}{\partial r_i}\Big|\Big)\Big] \tag{3}$$

where v_i is the ith measurement of a variable and r_i is a repeat of the ith measurement for the same variable. The function is minimised by an exhaustive search over the parameter range $-3 \le \theta \le 3$. Minimisation of this function maximises the likelihood of the measurement noise having a Gaussian distribution and stabilises the variance. A full derivation of Eq. (A.1) may be found in the appendix (Section Appendix A).

The negative log-likelihood functions for each variable (melt pool length and height, roughness, and material deposition) are shown in the supplementary material, with the optimal values for θ shown in Table 2.

After stabilisation, each variable is normalised by the standard deviation of the transformed repeatability noise distribution[2]. In this domain, each variable has the property of being in units of standard deviations of noise. Distances computed over spaces constructed with variables in this domain have a Euclidean geometry and a statistical interpretation similar to a z-score.

3. Results

3.1. Data pre-processing

The melt pool length, melt pool height, surface roughness and RTH measurements were transformed using the power law of Eq. (2) with θ optimised by minimising the negative log-likelihood given by Eq. (3). The negative log-likelihood functions, and the associated pre- and post-transform Bland-Atlamn plots, are shown in the supplementary material with the value of θ at the minimum indicated.

Table 2 shows the Shapiro-Wilks test results and kurtosis for the distribution of the difference of repeats, pre- and post-transformation, for each variable. The value of θ found corresponding to the minimum of the negative log-likelihood is also shown. Prior to transformation all the p-values, for all variables, indicate that the noise distributions are significantly different to a unit Gaussian; the kurtosis scores also reflect this with values greater than 3 indicating super-Gaussian distributions. Post-transformation the melt pool height, surface roughness, and RTH all show distributions that are no longer significantly different from that of a unit Gaussian with kurtosis scores that are closer to 3.

For the melt pool length the transform parameter, θ, is close to 1 and has minimal effect. The Shapiro-Wilks scores remain significant and the distribution of the noise remains super-Gaussian. Inspection of the Bland-Altman plot (supplementary material figure 3), for melt pool length, shows that pre-transform there is little indication of heteroscedasticity. In this case, the power law function is not able to transform the variable to a form where normality is increased. It may be noted that the normality of the residual distribution is not a requirement to attain the BLUE estimator, however, homoscedasticity is. For non-linear estimation, as in the estimation of neural network parameters, many of the properties derived from linear estimation theory hold, at least asymptotically [21], and thus post-transformation the variables meet these criteria.

[1] Heteroscedasticity.

[2] $f_{scaled}(x) = \frac{f(x)}{\sigma_{f(x)}}$

Table 2
Shapiro-Wilk scores, p-values, and kurtosis for raw and stabilised datasets. Stabilised datasets are indicated with a T in parenthesis. The estimated stabilisation parameter for the function $f(x) = x^{\theta}$ is also shown.

	S-W	p-value	kurtosis	θ	S-W (T)	p (T)	kurtosis (T)
length	0.95	0.02	4.7	0.92	0.95	0.02	4.6
height	0.81	<0.001	6.9	0.25	0.98	0.64	2.6
Roughness	0.97	0.05	4.2	0.15	0.99	0.74	2.7
RTH	0.92	<0.001	5.3	0.54	0.99	0.87	2.9

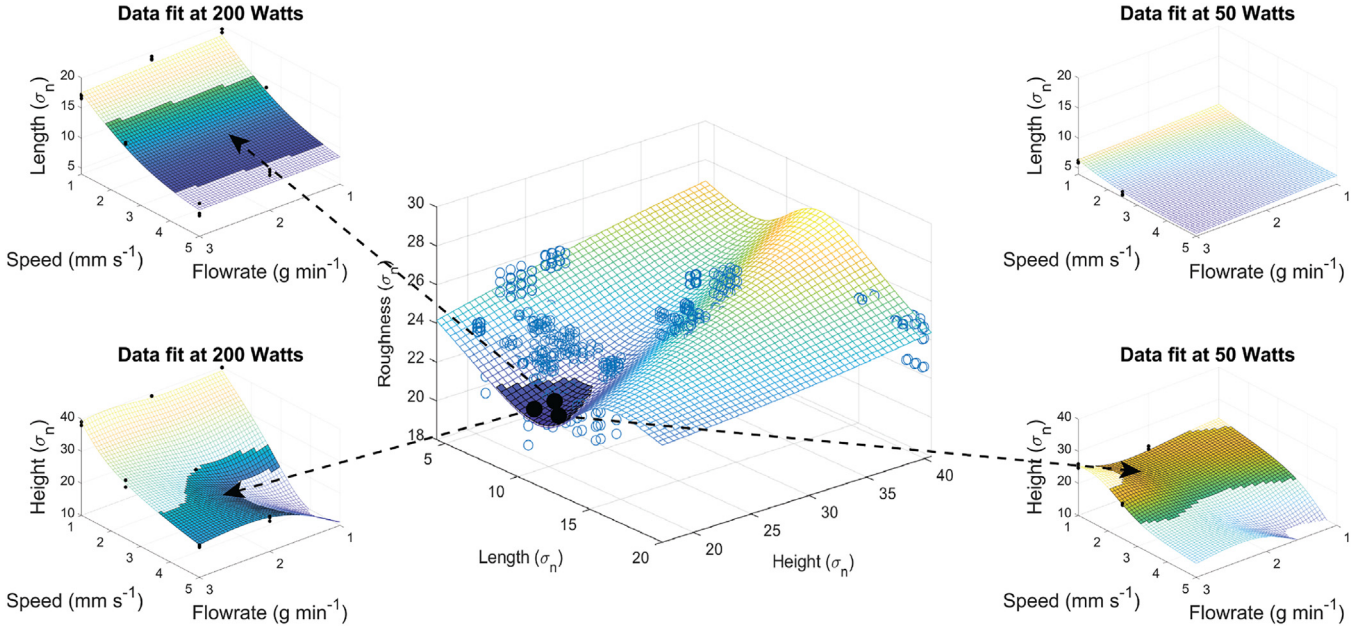

Fig. 2. Regression map of melt pool geometry to surface roughness with shaded region showing the area within a standard deviation of the minimum (central figure; The blue circles show data locations). Regression maps of the process parameters to for height and length (subplots) for two laser power settings (200W and 50W). Shaded areas indicate parameter combinations that correspond to the surface roughness minimum (black dots show data locations). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.2. Regression fits

3.2.1. Process parameters to melt pool geometry mappings

Low-order regression maps from process parameters to melt pool geometrical measures were found, that meet the goodness-of-fit criteria, from the full range of the process parameters used experimentally (see supplementary material for regression maps at 4 power settings).

The results of the analysis show that melt pool length is dependent on laser power and traverse speed alone; powder feed rate has no significant effect. The melt pool length shows a clear proportional relationship to the energy density, increasing with increased power and decreased head speed.

The results for melt pool height are more complicated. For traverse speeds of (1 and 2 $mm\,s^{-1}$) there is evidence to show that melt pool height has a relationship dependent on both energy density and powder feed rate; for traverse speed of 2 $mm\,s^{-1}$ significant change with powder feed-rate only occurs for high laser powers ($> 150W$). For head speeds of 5 $mm\,s^{-1}$ no significant change in the melt pool height with laser power or feed rate is supported by the data.[3]

3.2.2. Meltpool to roughness mapping

The central plot, of Fig. 2, shows the lowest order (3 hidden units) regression map found, that meets the goodness-of-fit criteria, between melt pool geometry and surface roughness. The blue circles show the locations of the measured data[4] and the shaded area of the plot corresponds to a region of a standard deviation (of repeatability noise) around the minimum. The plot shows that melt pools with a longer length and lower height correlate with lower levels of surface roughness. There is some evidence to show that the ratio of height to length is determining factor but more evidence of this is required. It may also be noted, from the pre-transformation Bland-Altman plots, that lower height melt pools have significantly lower repeatability variations than taller melt pools, suggesting that melt pool dynamics, and not just static geometry, have a role in determining surface roughness.

The subplots of Fig. 2 show regression maps generated directly between process parameters (speed and feed rate) and melt pool length and height. The shaded areas indicate the mapping of the minimum, of melt pool geometry to surface roughness, to the process parameters for two laser power settings (200W and 50W). For 50W there is no speed setting that enables a melt pool length and height combination to be realised that corresponds to the minimum area of surface roughness. For all other power settings, the minimum surface roughness was attainable to within error.

3.2.3. Process parameters to roughness mapping

Figure 3 shows regression maps directly from process parameters to surface roughness for 4 laser power settings (full regression meets goodness-of-fit criteria with 5 hidden nodes). The figures show that, for areas well constrained by data (black dots), the roughness does not sig-

[3] The apparent drop in melt pool height shown in the regressions, for low powder feed-rate, is not constrained by any data and is unreliable.

[4] Regions of the plot with no data are unconstrained and no meaningful inference may be made from these regions.

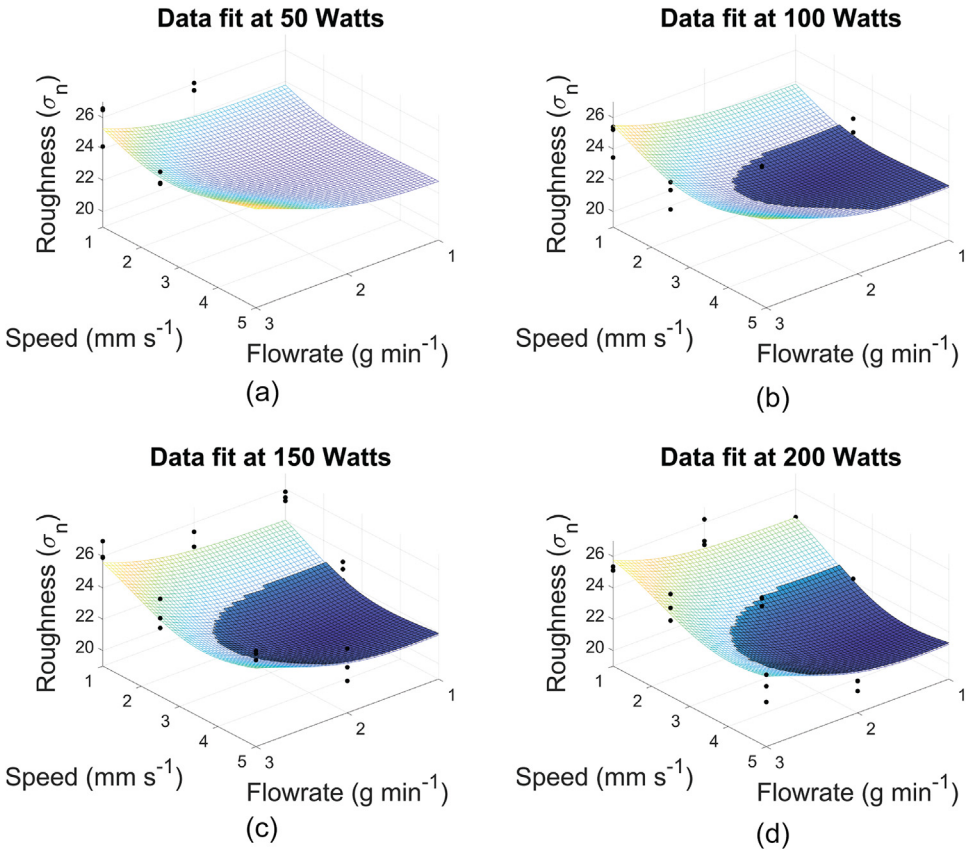

Fig. 3. Mapping of process parameters to surface roughness measured at the 4 laser power settings. Shading indicates ±1 standard deviation around the minimum surface roughness found in mapping melt pool geometry to surface roughness (The black dots show the data locations). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

nificantly change with laser power although it should be noted that the regression is over a 3d input space and areas that appear unconstrained in two dimensions may be constrained by data in the z direction.

The shaded areas of the plots indicate ±1 standard deviation around the minimum surface roughness found in mapping melt pool geometry to surface roughness (Fig. 2). The process parameters indicated by these areas, although not directly matching, correlate with the results associated with mappings to and from the melt pool geometry measures.

3.2.4. Process parameters to resultant track height

Figure 4 shows regression maps from process parameters to surface addition for 4 laser power settings (full regression meets goodness-of-fit criteria with 8 hidden nodes). A low-order mapping between melt pool geometry and RTH could *not* be found that met the goodness-of-fit criteria.

The figures show a significant increase in RTH with an increase in both powder feed rate and energy density. It may also be noted that RTH significantly increases with laser power for all combinations of head speed and powder feed rate. Crucially, this suggests, for the range of parameters studied, that RTH may be maximised independently of the melt pool geometry thus enabling minimum surface roughness to be achieved simultaneously.

4. Discussion and conclusions

This work has applied a rigorous machine learning framework to analyse a repeated measures data set collected from Synchrotron in situ X-ray imaging of the DED-AM process. The data set is small and care has been taken to ensure that the regression models fit the data and are reliable via variance stabilisation and the use of reduced chi-square (chi-square per degree of freedom) as a goodness-of-fit criterion.

Bland-Altman plots show that, except for melt pool length, the measured data have heteroscedastic repeatability errors and are thus are not directly suitable for optimisation of model parameters via a least squares cost function. The data was stabilised via a non-linear function optimised by maximising a novel log-likelihood function that maximises the normality of the distribution of the difference of repeats. This is in contrast to methods such as *Box and Cox* which is model dependent and considers the distribution of residuals.

Stabilised measurements were scaled by the standard deviation of the stabilised repeatability noise (difference of repeats). In this domain, the measurements have an interpretation that is in units of the standard deviation of repeatability noise. Thus, regression maps may be used to relate statistically significant changes in the input space to statistically significant changes in the output space e.g. melt pool geometry to surface roughness.

Neural networks were used, as generalised regressors, to map between: (a) process parameters and melt pool geometry, (b) process parameters and surface roughness, (c) melt pool geometry and surface roughness, and (d) process parameters and RTH. In each case the model, with the lowest model order, that was able to meet the reduced chi-square goodness-of-fit criteria was used. For mapping between melt pool geometry and RTH a low order model could *not* be found that met the goodness-of-fit criteria.

Figure 3 shows that, across all powers, the minimum surface roughness is achieved at the highest velocities and lowest feed rates. We hypothesise that high velocity and low powder feed rate gives a smaller melt pool with smaller surface waves and hence lower surface undulations. Contrary to a general understanding, the results show that process parameter do not have a direct relationship to surface roughness. Rather than an indirect relationship exists through control of the melt pool geometry, and possibly dynamics, through appropriate selection of traverse speed and powder feed rate. The framework allowed for an area of minimum surface roughness to be identified within the limits of the data and mapped back to the process parameters. Crucially, it was found that laser power had no significant effect on surface rough-

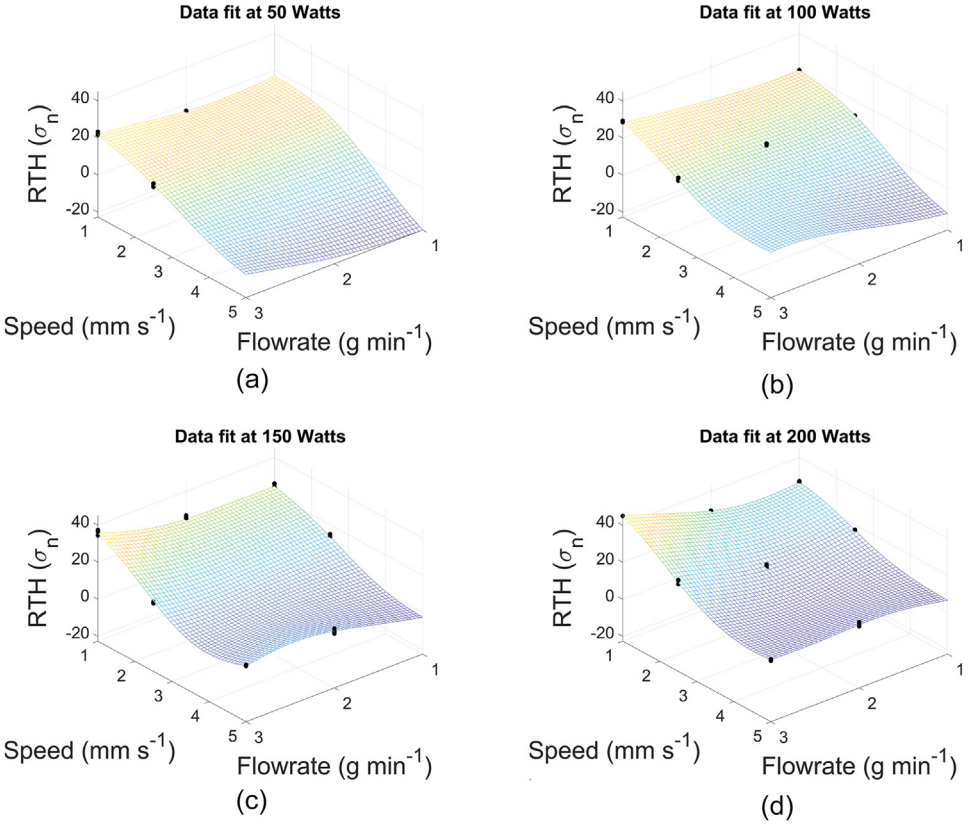

Fig. 4. Mapping of process parameters to surface addition at the 4 power settings (The black dots show the data locations). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

ness over the range of process parameters used and that longer length and lower height melt pools significantly correlated with the minimum in the surface roughness. In terms of melt pool dynamics, lower-height melt pools have less variability in the difference of repeated measurements which may have a direct correlation with surface roughness but this requires further evidence and analysis.

It was found that RTH had a significant correlation with all three process parameters used. In contrast, laser power had no significant correlation with surface roughness and thus, in terms of optimisation of the DED process, it is feasible to maximise RTH over the range of head speed and powder flow rates that minimise the surface roughness. As an example for a traverse speed of $3.5\,mm\,s^{-1}$ and feed-rate of $5g\,min^{-1}$ (predicted minimum roughness by the models) the RTH is predicted to increase from $40\mu m$ to $336\mu m$ with an increase in laser power from 150W to 200W.

5. Declaration of competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was supported under MAPP: EPSRC Future Manufacturing Hub in Manufacture using Advanced Powder Processes (EP/P006566/1), a Royal Academy of Engineering Chair in Emerging Technology (CiET1819/10), Rolls-Royce plc. via the Horizon 2020 Clean Sky 2 WP5.8.1 programme. Laboratory space and facilities were provided by the Research Complex at Harwell. The authors thank Diamond Light Source for providing beamtime (MT20096) and the staff at I12 beamline for technical assistance.

Y.Chen would like to acknowledge the funding support from the RMIT University Vice Chancellor's Senior Research Fellowship and the travel funding provided by the International Synchrotron Access Program (ISAP) managed by the Australian Synchrotron, part of ANSTO, and funded by the Australian Government.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

S.V. Notley: Conceptualization, Formal analysis, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. **Y. Chen:** Data curation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. **N.A. Thacker:** Methodology, Writing – review & editing. **P.D. Lee:** Conceptualization, Funding acquisition, Methodology, Project administration, Supervision, Writing – review & editing. **G. Panoutsos:** Conceptualization, Formal analysis, Methodology, Funding acquisition, Supervision, Writing – review & editing.

Data availability

Data will be made available on request.

Appendix A. Log-Likelihood derivation

For continuous variables (random variable following a continuous probability distribution) the likelihood is defined in terms of a probability density. Thus, the log-likelihood of a sampled continuous variable having a Gaussian probability density is commonly defined as

$$-\log L = \sum_{i=1}^{I} \left[\frac{x_i^2}{2\sigma^2} + \log(\sqrt{2\pi}\sigma) \right] \tag{A.1}$$

where x_i is the ith sample from a continuous variable.

However, in some circumstances an estimation process may need to adjust the data variables and the use of the density is not appropriate. For scientific validity we must work with a cost function based on the probability of a given event, this is independent of arbitrary experimental choices. i.e.

$$P(a \leq x_i \leq b) = P(f(a;\theta) \leq f(x_i;\theta) \leq f(b;\theta)) \tag{A.2}$$

for any continuous and differentiable monotonic function $f(x)$.

The relationship between probabilities and probability densities is given by

$$P(x) = \int_{x_-}^{x_+} p(x)dx \approx p(x)\Delta x \tag{A.3}$$

where $P(x)$ is the probability of x being in the interval (x_-, x_+), Δx is the interval given by $x_+ - x_-$ and $p(x)$ is the probability *density* over x. In the limit of Δx tending to zero, the approximation becomes exact. In terms of the reproducibility noise, we have

$$P(a \leq (v_i - r_i) \leq b) = p(v_i - r_i)\Delta(v_i - r_i) \tag{A.4}$$

$$= p(f(v_i;\theta) - f(r_i;\theta))\Delta(f(v_i;\theta) - f(r_i;\theta)) \tag{A.5}$$

where v_i is a measurement, for a build, i, with a given a set of process parameters and r_i is a repeated measurement for the same build.

Thus

$$P(a \leq (v_i - r_i) \leq b) = p(f(v_i;\theta) - f(r_i;\theta))\frac{\delta(f(v_i;\theta) - f(r_i;\theta))}{\delta(v_i - r_i)}\Delta(v_i - r_i) \tag{A.6}$$

This probability definition may be used to define a cost function, in terms of θ, that maximises the overall probability of the reproducibility noise being drawn from a Gaussian distribution as

$$\begin{aligned} -\log P(v_i, r_i|\theta) = \sum_{i=1}^{I} \Big[&\frac{(f(v_i;\theta) - f(r_i;\theta))^2}{2\sigma_f^2} + \log(\sqrt{2\pi}\sigma_f) \\ &- \log\left(\left|\frac{\delta(f(v_i;\theta) - f(r_i;\theta))}{\delta(v_i - r_i)}\right|\right) - \log(\Delta(v_i - r_i))\Big] \end{aligned} \tag{A.7}$$

where $\sigma_f = \sigma_{f(v_i;\theta)-f(r_i;\theta)}$. The perturbations of v_i and r_i are independent and thus $\delta(v_i - r_i) = \delta v_i - \delta r_i$; eliminating the final constant term gives a Gaussian log-likelihood function (in terms of probability density) for the transformed data as

$$\begin{aligned} -\log L(v_i, r_i|\theta) = \sum_{i=1}^{I} \Big[&\frac{(f(v_i;\theta) - f(r_i;\theta))^2}{2\sigma_f^2} + \log(\sqrt{2\pi}\sigma_f) \\ &- \log\left(\left|\frac{\partial f(v_i;\theta)}{\partial v_i} + \frac{\partial f(r_i;\theta)}{\partial r_i}\right|\right)\Big] \end{aligned} \tag{A.8}$$

It may be noted that this has additional differential terms to the original log-likelihood definition (Eq. (A.1)), which are needed in order to compensate for the changes in density arising from data transformation.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.addlet.2023.100137.

References

[1] Y. Chen, S. J. Clark, L. Sinclair, C. L. A. Leung, S. Marussi, T. Connolley, O. V. Magdysyuk, R. C. Atwood, G. J. Baxter, M. A. Jones, et al., In situ and operando X-ray imaging of directed energy deposition additive manufacturing, 2020, arXiv preprint arXiv:2006.09087.

[2] Y. Chen, S. Clark, A.C.L. Leung, L. Sinclair, S. Marussi, R. Atwood, T. Connoley, M. Jones, G. Baxter, P.D. Lee, Melt pool morphology in directed energy deposition additive manufacturing process, in: IOP Conf. Ser.: Mater. Sci. Eng, Vol. 861, 2020, p. 012012, doi:10.1088/1757-899X/861/1/012012.

[3] S. Clark, C. Leung, Y. Chen, L. Sinclair, S. Marussi, P. Lee, Capturing Marangoni flow via synchrotron imaging of selective laser melting, in: IOP Conf. Ser.: Mater. Sci. Eng, Vol. 861, 2020, p. 012010, doi:10.1088/1757-899X/861/1/012010.

[4] L. Sinclair, C.L.A. Leung, S. Marussi, S.J. Clark, Y. Chen, M.P. Olbinado, A. Rack, J. Gardy, G.J. Baxter, P.D. Lee, In situ radiographic and ex situ tomographic analysis of pore interactions during multilayer builds in laser powder bed fusion, Addit. Manuf. 36 (2020) 101512, doi:10.1016/j.addma.2020.101512.

[5] Y. Chen, S.J. Clark, C.L.A. Leung, L. Sinclair, S. Marussi, M.P. Olbinado, E. Boller, A. Rack, I. Todd, P.D. Lee, In-situ synchrotron imaging of keyhole mode multi-layer laser powder bed fusion additive manufacturing, Appl. Mater. Today 20 (2020) 100650, doi:10.1016/j.apmt.2020.100650.

[6] S.J. Wolff, H. Wu, N. Parab, C. Zhao, K.F. Ehmann, T. Sun, J. Cao, In-situ high-speed X-ray imaging of piezo-driven directed energy deposition additive manufacturing, Sci. Rep. 9 (1) (2019) 962.

[7] C. Wang, X. Tan, S.B. Tor, C.S. Lim, Machine learning in additive manufacturing: state-of-the-art and perspectives, Addit. Manuf. 36 (2020) 101538.

[8] F. Caiazzo, A. Caggiano, Laser direct metal deposition of 2024 Al alloy: trace geometry prediction via machine learning, Materials 11 (3) (2018).

[9] A. Singh, D. Cooper, N. Blundell, G. Gibbons, D. Pratihar, Modelling of direct metal laser sintering of EOS DM20 bronze using neural networks and genetic algorithms, in: Proceedings of the 37th International MATADOR Conference, 2012, p. 395.

[10] X. Qi, G. Chen, Y. Li, X. Chen, C. Li, Applying neural-network-based machine learning to additive manufacturing: current applications, challenges, and future perspectives, Engineering 5 (2019) 721–729.

[11] C.M. Bishop, Training with noise is equivalent to Tikhonov regularization, Neural Comput. 7 (1) (1995) 108–166.

[12] S.V. Notley, Y. Chen, P.D. Lee, G. Panoutsos, Variance stabilised optimisation of neural networks: a case study in additive manufacturing, in: Proceedings of the 2021 International Joint Conference on Neural Networks (IJCNN), 2012, pp. 1–7. China

[13] D.G. Altman, J.M. Bland, Measurement in medicine: the analysis of method comparison studies, J. R. Stat. Soc. Ser. D 32 (3) (1983) 307–317.

[14] Y. Chen, S.J. Clark, Y. Huang, L. Sinclair, C.L.A. Leung, S. Marussi, T. Connolley, O.V. Magdysyuk, R.C. Atwood, G.J. Baxter, M.A. Jones, I. Todd, P.D. Lee, In situ X-ray quantification of melt pool behaviour during directed energy deposition additive manufacturing of stainless steel, Mater. Lett. 286 (2021) 129205.

[15] K. Levenburg, A method for the solution of certain non-linear problems in least squares, Q. Appl. Math. 2 (2) (1944) 164–168.

[16] D. Marquardt, An algorithm for least-squares estimation of nonlinear parameters, SIAM J. Appl. Math. 11 (2) (1963) 431–441.

[17] F.J. Anscombe, The transformation of poisson, binomial and negative-binomial data, Biometrika 35 (3-4) (1948) 246–254.

[18] F.J. Anscombe, J.W. Tukey, The examination and analysis of residuals, Technometrics 5 (2) (1963).

[19] M.S. Bartlett, The use of transformations, Biometrics 3 (1) (1947).

[20] G.E.P. Box, D.R. Cox, An analysis of transformations, J. R. Stat. Soc. 26 (2) (1964) 211–252.

[21] H. Bunke, Parameter estimation in nonlinear regression, in: P.R. Krishnaiah (Ed.), Handbook of Statistics, 1, North-Holland Publishing Company, 1980, pp. 593–615.