



This is a repository copy of *Methods for rapid pore classification in metal additive manufacturing*.

White Rose Research Online URL for this paper:  
<https://eprints.whiterose.ac.uk/151123/>

Version: Published Version

---

**Article:**

Snell, R., Tamas-Williams, S., Chechik, L. et al. (5 more authors) (2020) Methods for rapid pore classification in metal additive manufacturing. JOM Journal of the Minerals, Metals and Materials Society, 72 (1). pp. 101-109. ISSN 1047-4838

<https://doi.org/10.1007/s11837-019-03761-9>

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:  
<https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>



# Methods for Rapid Pore Classification in Metal Additive Manufacturing

ROBERT SNELL <sup>1,5</sup> SAM TAMMAS-WILLIAMS,<sup>2</sup> LOVA CHECHIK,<sup>1</sup>  
ALISTAIR LYLE,<sup>1</sup> EVERTH HERNÁNDEZ-NAVA,<sup>1</sup> CHARLOTTE BOIG,<sup>1,3</sup>  
GEORGE PANOUTSOS,<sup>4</sup> and IAIN TODD<sup>1</sup>

1.—Department of Materials Science and Engineering, University of Sheffield, Sheffield, UK. 2.—Department of Maritime and Mechanical Engineering, Liverpool John Moores University, Liverpool, UK. 3.—Cummins Turbo Technologies, St Andrew's Road, Huddersfield, UK. 4.—Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield, UK. 5.—e-mail: r.m.snell@sheffield.ac.uk

The additive manufacturing of metals requires optimisation to find the melting conditions that give the desired material properties. A key aspect of the optimisation is minimising the porosity that forms during the melting process. A corresponding analysis of pores of different types (e.g. lack of fusion or keyholes) is therefore desirable. Knowing that pores form under different thermal conditions allows greater insight into the optimisation process. In this work, two pore classification methods were trialled: unsupervised machine learning and defined limits. These methods were applied to 3D pore data from X-ray computed tomography and 2D pore data from micrographs. Data were collected from multiple alloys (Ti-6Al-4V, Inconel 718, Ti-5553 and Haynes 282). Machine learning was found to be the most useful for 3D pore data and defined limits for the 2D pore data; the latter worked by optimising the limits using energy densities.

## INTRODUCTION

Additive manufacturing (AM) is the catch-all term for a number of different technologies that melt material in an additive way rather than removing material from a larger piece. The AM of metal forms a significant fraction of AM research, with most metal processes based around the powder bed fusion (PBF) method. PBF entails the melting of successive layers of metal powder.<sup>1</sup> There are many benefits to AM regarding less material wastage and the potential for lightweight designs and rapid prototyping.<sup>1</sup> As AM technologies mature, research has focused on developing the technologies from a research and small-scale production tool into an industrial production method.<sup>2</sup>

One of the biggest technological challenges of metal-based AM is porosity. The defects are often the cause of failure in AM parts<sup>3</sup> and are one of the biggest challenges to overcome in getting AM parts into aerospace,<sup>4</sup> medical<sup>5</sup> and many other applications.<sup>6</sup> AM parts can receive a hot isostatic press treatment to reduce porosity<sup>7</sup> but the majority of pore reduction can be achieved by optimising the process parameters.

Compared with conventional metal production processes, such as casting or rolling, the process parameters in AM are all relatively recent developments. For PBF these process parameters concern the power, velocity and pattern taken by the melting beam, all of which control the melting of metal powder. For any new powder composition or morphology, new parameters must be developed. Optimising the process parameters is usually guided by the material properties. The density, roughness and phases are all properties that can be considered as well as practical concerns, such as powder disturbances and swelling.

Even with all these considerations, porosity is still one of the most critical properties for optimising the process parameters.<sup>8</sup> Mechanical properties of AM-produced parts can depend heavily on the porosity with fatigue,<sup>9</sup> strength<sup>10</sup> and Young's modulus<sup>11</sup> all decreasing with increased porosity. It is also worth noting that pores can vary significantly in size and shape and are formed because of different thermodynamic conditions. As large pores are more likely to cause structural failures, they are a higher priority for removal than small pores.<sup>12</sup>

Optimising parameters usually involves some form of Design of Experiment (DoE) or other experiment scheme. The process involves trialling different parameters and then measuring the results to find which parameters worked well (low porosity) and which did not (high porosity). The best parameters will be the ones that fully melt the metal powder without causing porosity via overheating or any other means. Quickly and accurately classifying pores would enable optimisation to be carried out on the most critical types of pores, not just on the total pore volume.

## BACKGROUND

### Types of Pores

Three main types of pores are thought to form during AM processes: gas, keyholes (KH) and lack of fusion (LF). Each has a distinct shape, size and formation mechanism.<sup>13-15</sup>

LF pores are caused by a lack of input energy to the powder bed during the melting process. The lower input energy fails to fully melt the metal powder and so leaves voids in the resultant structure. LF pores tend to be large, similar in scale to the melt pool size and irregular<sup>13</sup> as shown in Fig. 1b.

KH pores are caused by an excess of input energy during the melting process. Excess beam power causes excessive penetration of the metal powder, which after solidification leaves a pore near the bottom of the melt pool. The result is a relatively large pore that is usually circular horizontally and elongated in the vertical direction as shown in Fig. 1a. KH pores can also be wider at the top than the bottom, looking somewhat like a keyhole.<sup>14,15</sup>

Gas pores are the smallest and most spherical of all pores. These pores are connected to trapped gas either already present in the metal powder or trapped during the melting process. These pores are usually the most prevalent pore type in AM.<sup>13</sup>

### Unsupervised Machine Learning Methods

Unsupervised machine learning, or clustering, is an algorithmic method of grouping data together without prior knowledge about the underlying data structure. The methods take data, with any number of variables, and collect them together according to a set of numerical metrics. Most simple methods rely on distance-based metrics, for example the Euclidean distance between data points. More sophisticated methods would also include metrics of similarity of groups, overlap, information density measures, etc. The clustering in this research utilises exclusively the size, shape and relative dimensions of pores as the variables to be clustered, with the hypothesis that there is an underlying relationship that links pore morphology. In the literature, there are example of analyses related to AM that use the process parameters and pore locations but these are outside the scope of this work.

There are many clustering methods that each has different strengths and weaknesses in terms of performance, as well as differences in the level of complexity and sophistication in the way the clusters are estimated. Some hierarchical algorithms ‘grow’ clusters from the available data, while others estimate potential cluster centres, then assign data to each cluster. The differences in performance usually relate to the computational speed of the

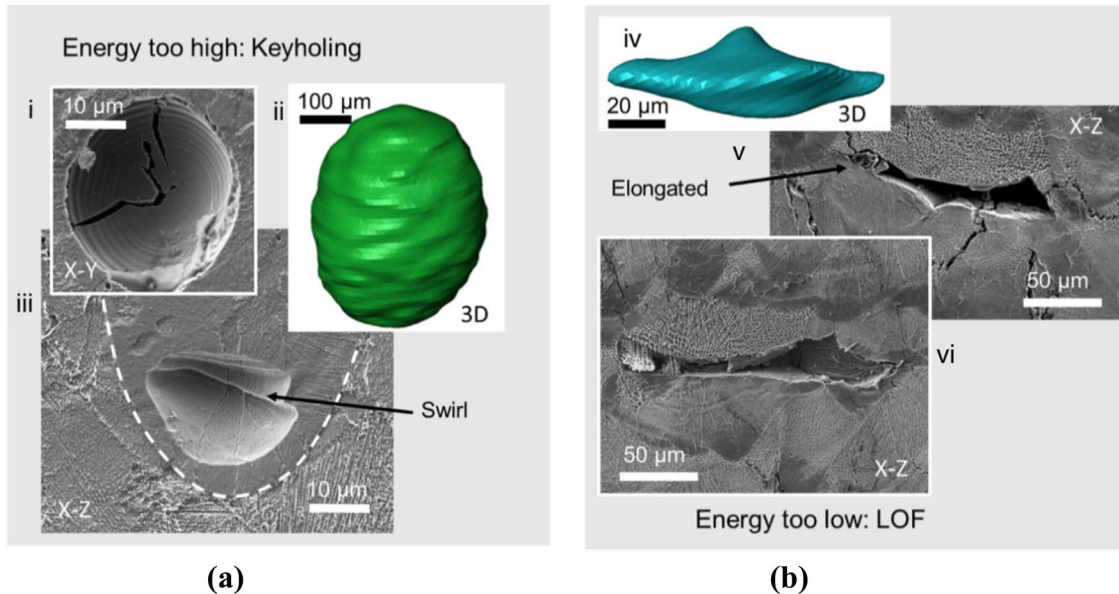


Fig. 1. Taxonomy of KH and LF pores. These images are from samples of AM-produced IN713C. (a) Image i is an SEM of a KH pore in the XY plane. Image ii shows an XCT reconstruction. Image iii shows an SEM image from the XZ plane. (b) Image iv shows an XCT reconstruction of a LF pore. Image v and image vi show SEM image from XZ plane. Source: Charlotte Boig PhD thesis, Sheffield.<sup>16</sup>

algorithms and how they cope with specific patterns of data. One example, Density-Based Spatial Clustering of Applications with Noise (DBSCAN), has the advantage that the number of clusters does not need to be specified. However, this method can suffer if there are large differences in the data density.<sup>17</sup> This is pertinent to pore data where there may be relatively few large pores but classifying them is important.

Other examples include mixture models that can be modified into clustering algorithms. These models work best when the data fit a type of pattern (e.g. Gaussian). Given that the shape of the pore data is unknown and likely to vary significantly across data sets, a more general method was selected.

This research uses a K-means algorithm. This is one of the simplest clustering methods, based on using distance metrics to estimate potential cluster centres. While this algorithm may not be the most sophisticated computationally, it is very robust and with numerous examples in the literature with very good performance. The K-means algorithm starts by assigning randomly placed cluster centres,  $\mu$ , amongst the data set such that

$$\mu_1, \mu_2, \mu_3 \dots \mu_K \quad (1)$$

where  $K$  is the number of clusters. The data points are then labelled according to which cluster centre each data point is nearest to across  $n$ -dimensions (where  $n$  is the number of variables). The closest cluster centre is found by taking each point,  $x_i$ , and measuring the distance to each cluster centre

$$\|x_i - \mu_k\| \quad (2)$$

for all cluster centres. The  $\mu$  with the minimum value will label the point  $x_i$  to that cluster. Once this has been completed the new cluster centres are then calculated using the average position of the newly labelled data. This whole process is iterated until every data label and cluster centre is static.<sup>18</sup>

The random initialisation of cluster centres may produce a result that is not optimal. To prevent this the algorithm of random initialisations and clustering is repeated many times. A sensible number of repeats will balance computational time and ensure that at least one initialisation produces the best clustering result. Comparisons between cluster results are done using a cost function,  $J$ , that measures the distance between data points and cluster centres. This can be expressed as

$$J = \frac{1}{m} \sum_{i=1}^m \|x_i - \mu_c\| \quad (3)$$

where  $m$  is the total number of data points and  $\mu_c$  is the cluster centre that each  $x_i$  is assigned to. The cluster results with the minimum value of  $J$  are selected as the optimal results.

The number of clusters that are present may be known or unknown. If the number of groups is

unknown the most appropriate number of clusters can be found. This is done by plotting the cost function,  $J$ , against number of clusters and looking for an ‘elbow’ or ‘kink’ where  $J$  drops suddenly.<sup>18</sup>

## METHOD

### Pore Measurement Techniques

The two techniques that were used for data collection were XCT for 3D pore data and optical microscopy for 2D pore data. For the XCT, samples were analysed using the 320/225 kV Nikon XTEK bay at the Henry Moseley X-ray Imaging facility within the University of Manchester. A cone-shaped beam of X-rays was generated using an electron beam with accelerating voltage of 160 kV and current of 150  $\mu$ A directed onto a silver reflection target. A 1.5-mm copper filter was used to filter out lower energy photons and reduce the effect of beam hardening. The samples were positioned to give a geometric magnification of approximately 22.5 before 3142 radiographs were collected with a 1-s exposure time using a square 4-megapixel detector with 0.2-mm pixel pitch. Three-dimensional volumes with 8.9  $\mu$ m voxel size were reconstructed from the radiographs using proprietary Nikon software employing a filtered back projection algorithm. The 32-bit 3D volume was imported into Avizo image processing software and converted to 8 bits to reduce computational requirements. Data sets were also segmented and quantified in Avizo using built-in automatic thresholding and quantification tools. Numerical data, including size, morphology and position, regarding individual pores were exported for further analysis.

For optical microscopy the samples were removed from the base plates using wire electrical discharge machining 1 mm from the base. The samples were ground approximately 1 mm and then polished to a 9-micron finish. Measurements were carried out using an Olympus BX51 Clemex microscope with Clemex Vision PE software, which allows roundness and horizontal dimensions to be measured. An example micrograph is shown in Fig. 2.

In terms of classifying different types of pores the 3D pore data acquired using XCT provide more information per pore than optical microscopy. However, in terms of practicality, optical microscopy offers a quicker and more cost-effective way of collecting data for multiple samples. This is particularly relevant for AM parameter optimisation, which often requires dozens of samples to find the optimal process parameters.

### Production Methods

Samples were produced using laser beam powderbed SLM machines. The Ti-6Al-4V sample was built using a Renishaw 125. The sample was a cuboid, 10.61 mm in length and width and 15 mm in height. The process parameters used were a laser beam of

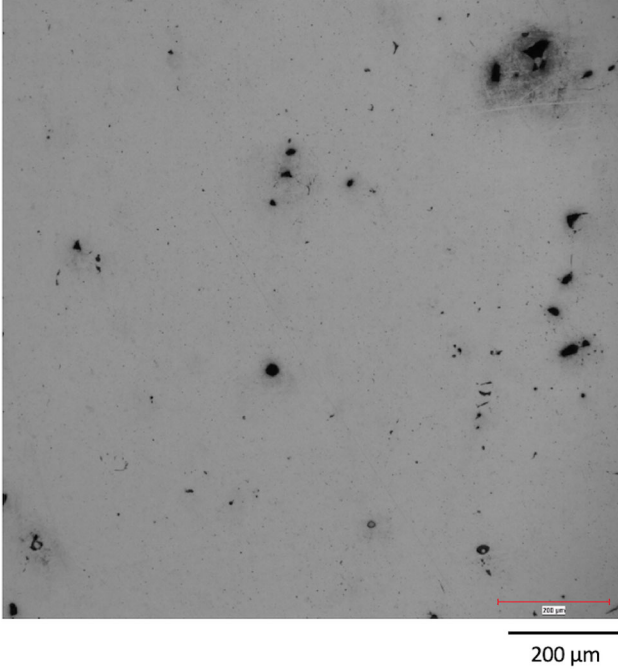


Fig. 2. An example micrograph from an Inconel 718 sample with an energy density of  $49.4 \text{ J/mm}^3$ .

200 W, point distance of  $70 \mu\text{m}$  and exposure time of  $150 \mu\text{s}$ . The hatch spacing was  $70 \mu\text{m}$  and the layer thickness was  $50 \mu\text{m}$ , which are optimised for melting Ti-6Al-4V on this machine.

The Inconel 718 samples were produced using an Aconity Mini with a range of process parameters. The samples were cubes with a length  $10 \text{ mm}$ . Laser beam power varied from  $60 \text{ W}$  to  $187 \text{ W}$ , laser velocity was varied between  $0.3 \text{ m/s}$  and  $2.3 \text{ m/s}$  and hatch distance was varied between  $25 \mu\text{m}$  and  $140 \mu\text{m}$ . The thickness of the powder layer was  $30 \mu\text{m}$ , which produced an energy density range of  $11 \text{ J/mm}^3$  to  $200 \text{ J/mm}^3$ .

The Haynes 282 (a nickel superalloy) and Ti-5553 samples were also cubes with a length of  $10 \text{ mm}$  but were produced on a Renishaw 125. Process parameters were varied between energy densities of  $2 \text{ J/mm}^3$  and  $45 \text{ J/mm}^3$  for the Haynes 282 and  $0.4 \text{ J/mm}^3$  and  $267 \text{ J/mm}^3$  for the Ti-5553. Full details of process parameters can be found in the supplementary data.

All samples were produced with gas atomised powder that has a size range of  $15 \mu\text{m}$  to  $45 \mu\text{m}$ . The Ti-6Al-4V, Ti-5553 and Inconel 718 were all purchased from LPW Carpenter Additive, while the Haynes 282 is from Praxair.

There are cases where the porosity and quality of the as-received powder have a direct influence on the porosity that is found in subsequent builds.<sup>19</sup> However, the small particle size limits the pores to  $45 \mu\text{m}$ , so it is unlikely that these pores will be incorrectly classified as LF or KH pores.

### Classification Method: Defined Limits

A defined limits approach sets limits on variables to identify types of pores. Knowing that pores tend to be small or large, irregular or round, etc., allows an unlabelled set of pore data to be classified. These limits for pores can be found in published values.<sup>14,15,20,21</sup> References to the length of a pore is the longest linear distance within a pore. Pore height is the length in the build direction and pore width refers to the longest distance in the horizontal plane.

LF pores are large and irregular and can be found using values for pore length and sphericity. The criteria selected were pores that were  $> 100 \mu\text{m}$  in length<sup>20</sup> and  $< 0.6$  in sphericity.

KH pores are relatively large and tend to be elongated in the vertical direction. These pores can be identified using a minimum size criterion (above  $100 \mu\text{m}$ <sup>21</sup>) and a comparison of the height against the width of the pore (ensuring the height is at least double the width).<sup>14,15</sup>

Similar criteria limits were applied to 2D pore data using the roundness and pore length. However, limitations with the 2D pore data meant that the data were better analysed using defined limits with optimisation.

### Classification Method: Defined Limits with Optimisation

This method uses defined limits to classify pores but finds the limits using optimisation. Instead of using existing data the limits are generated by analysing multiple data sets and finding the limits that best fit the data. This is particularly useful for 2D pore data where manually defining accurate limits is difficult but acquiring multiple data sets is relatively easy. For 2D pore data a defined limits method uses roundness and pore length, as shown in Fig. 3.

There is the opportunity to use the multiple data sets to find the pore length and roundness limits that best classify KH and LF pores. Given the energy density will dictate the formation of LF and KH pores there should be a relationship between energy density and pore frequency. Energy density,  $E$ , can be measured using

$$E = \frac{P}{v \times h \times t} \quad (4)$$

where  $P$  is the power of the laser beam,  $v$  the velocity of the beam,  $h$  the hatch spacing and  $t$  the thickness of the layer.<sup>22</sup> This equation is suitable for Laser AM machines that have a continuous laser. For machines that have a pulse laser a slightly modified equation is required,

$$E = \frac{P}{\frac{\text{Point distance}}{\text{Exposure time}} \times h \times t} \quad (5)$$

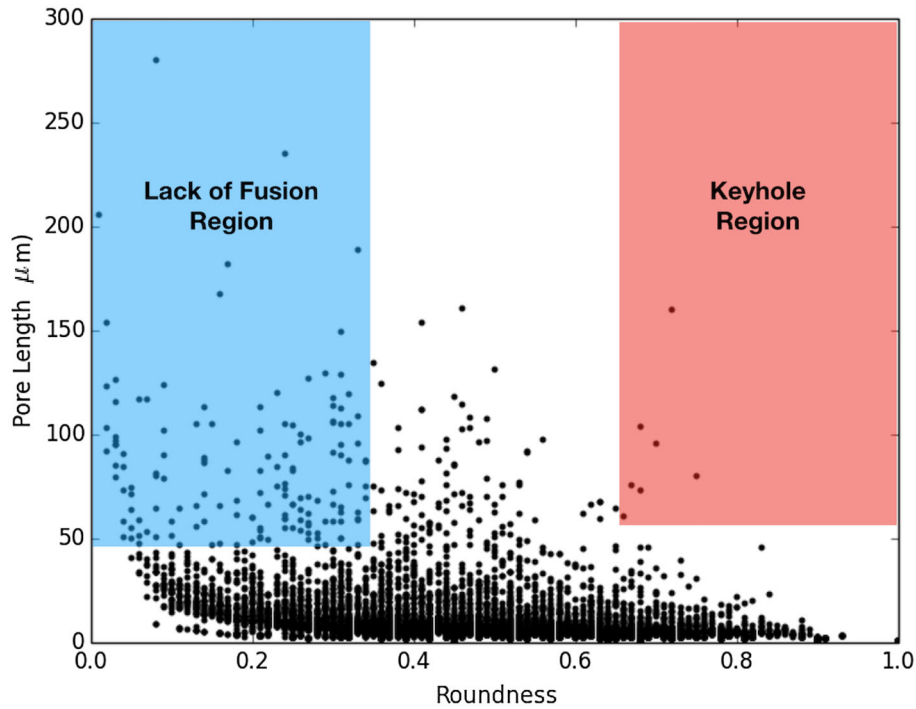


Fig. 3. Roundness and pore length for a sample set of 2D pore data. The red and blue regions represent example regions where the pores within will be classified as LF or KH. The values shown for the regions are illustrative only and are to be determined based on the data analysis (Color figure online).

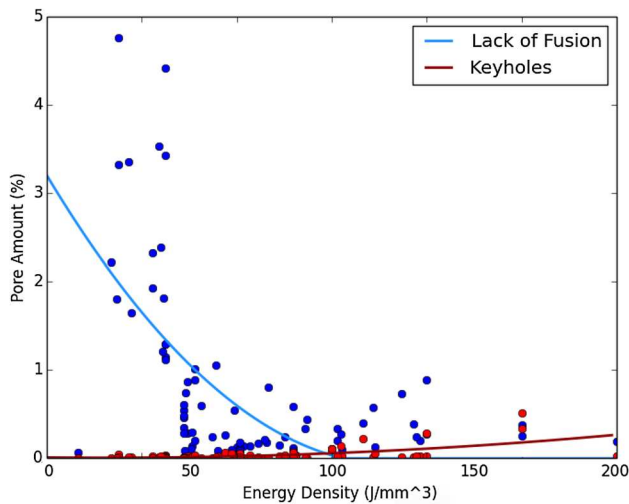


Fig. 4. Percentage of pores classified plotted against the energy density of the data set. The pore percentages were calculated using limits 70 μm for the pore length and 0.7 for the roundness for both types of pore.

where the point distance and exposure time replace the velocity of a continuous laser beam.<sup>23</sup>

Increasing energy density should increase the number of KH pores and decrease the number of LF pores. An example of this type of energy density plot is shown in Fig. 4.

On Fig. 4 two best fit lines have been added for LF and KH pores. The best fit lines were plotted using either a polynomial, logarithmic power law or exponential equation, whichever fit the data best. The polynomial fit goes to third order to allow a close fit without overfitting the data. The quality of the agreement between the best fit line and the data can be measured by finding the coefficient of determination ( $R^2$ ) for each line.

This process of plotting the data and finding the quality of the agreement can then be optimised using trial and error with different defined limits (the limits for the boxes in Fig. 3). The different limits will produce different levels of agreement between best fit lines and the data on the energy density plots (as shown in Fig. 4). The resultant  $R^2$  values can then be compared to find the defined limits that best describe the LF and KH pores.

**Classification Method: K-Means Clustering**

An alternative approach is to cluster the pore data to find the LF and KH pores rather than to define them using set limits. This avoids the problem of not knowing where to set the limits for pore classification. The clustering can be done using any number of variables, although adding extra variables will slow the calculations and potentially increase the noise on the data.

To ensure each variable contributes equally to the clustering the values are all scaled to a similar range. The new values,  $x_{new}$ , were calculated using the average value,  $x_{avg}$ , and the range of values, such that

$$x_{new} = \frac{x_i - x_{avg}}{x_{max} - x_{min}} \quad (6)$$

where  $x_i$  is the original value and the range is calculated using the minimum value,  $x_{min}$ , and maximum value,  $x_{max}$ , for each variable. This scaling ensures that all variables have a range of 1 and will therefore cluster the data with equal weight across all variables.

A K-means algorithm was used to cluster the data.<sup>18</sup> Clusters were given a random initialisation and optimised until the clusters were stable. This was repeated 500 times with the best solution (the one with the closest fitting groups) used. The variables used for grouping the 3D pore data were pore length, sphericity and ratio of height to width.

Even though there are three main types of pores (LF, KH and gas as described earlier) this does not necessarily mean that there are three clusters within the data. It may be that some types of pore vary in properties enough that they are best described using multiple clusters. For example, instead of a single cluster to describe gas pores there may need to be a cluster for typical gas pores and another cluster for the smallest gas pores.

The total number of clusters for the data was found by analysing the data using different numbers of clusters (two to ten). The distance between the pores in each cluster, the cost function from Eq. 3, describes how well the clusters describe the data. More clusters will always reduce the cost function. If an additional cluster produces a sudden drop in the cost function this usually indicates the correct number of clusters.<sup>18</sup>

## RESULTS

### Summary of Results

The main focus of this research is the development of the methods rather than investigating specific materials or process parameters. As such, materials were selected that best illustrate the strengths and weaknesses of the different methods.

The material selection was also based on the materials and data sets that were available. The details of the methods and materials are listed in Table I.

### Results: Defined Limits Method—3D Pore Data

Criteria limits were applied to 3D pore data (2664 pores) from a Ti-6Al-4V cuboid. KH pores were identified as having a length  $> 0.1 \text{ mm}^{21}$  and an aspect ratio (in the build direction)  $> 2^{14,15}$  as shown in Fig. 5a. LF pores were identified as being  $> 0.1 \text{ mm}^{20}$  with a sphericity  $< 0.6$ , which was judged to be the limit for being ‘irregular’ as shown in Fig. 5b.

These limits found there were 186 LF pores and just 1 KH pore. The most significant problem with the defined limits method is deciding what limits should be used. For criteria such as minimum sizes there are multiple sources with different values. For sphericity values there are no published values, and which value counts as ‘irregular’ is open to interpretation. As Fig. 5 shows, these limits have a noticeable effect on the results.

### Results: K-Means Clustering—3D Pore Data

The number of clusters selected was five, as the sudden drop in J (from Eq. 3) occurred then. The cluster results are plotted in Fig. 6.

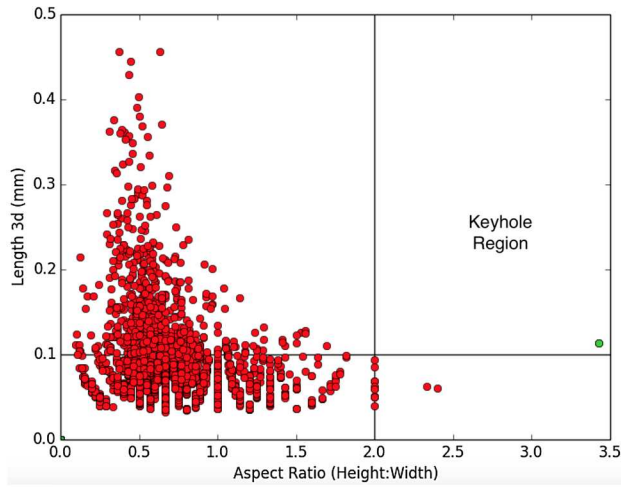
Figure 6a shows the LF pores (the largest and least spherical) and the gas pores (the smallest and most spherical). The KH pores (elongated and larger) can be found in Fig. 6b. The other clusters (orange and green) are unclear but are likely to be slightly larger and slightly more irregular gas pores, although further investigation would be required to confirm these clusters. The cluster sizes show 183 LF pores and 373 KH pores.

### Results: K-Means Clustering—2D Pore Data

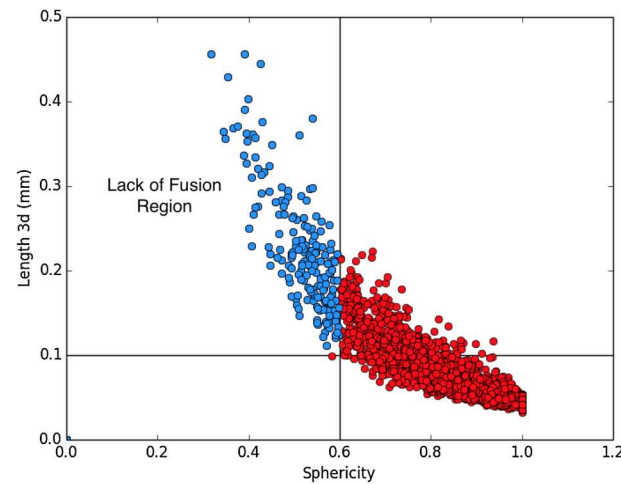
Clustering on the roundness and pore length was applied on 2D pore data from 81 Inconel 718 samples. The average number of pores was 21955 with the smallest data set having 2981 pores. Clustering results did not consistently identify pores. LF pores were occasionally found but KH pores were rarely identified. Some examples of the clustering results are shown in Fig. 7.

**Table I. Summary of the pore classification methods carried out**

Pore classification method	Data type	Material
Defined limits	3D pore data (XCT)	Ti-6Al-4V
Clustering	3D pore data (XCT)	Ti-6Al-4V
Clustering	2D pore data (micrographs)	Inconel 718
Defined limits with optimisation	Multiples sets of 2D pore data (micrographs)	Inconel 718, Haynes 282 and Ti-5553



(a)



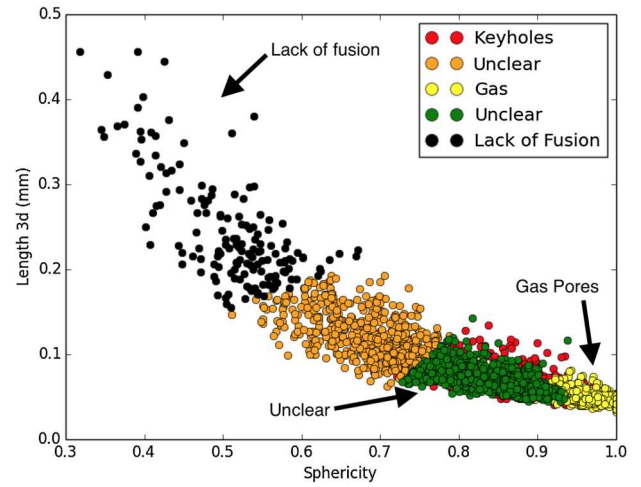
(b)

Fig. 5. The 3D pore sample data. All non-classified pores are shown in red. (a) The length and aspect ratio (in the build direction) of the pores. The KH region is labelled with the KH pore shown in green. (b) The length and sphericity of the pores. The LF region is labelled with the LF pores shown in blue (Color figure online).

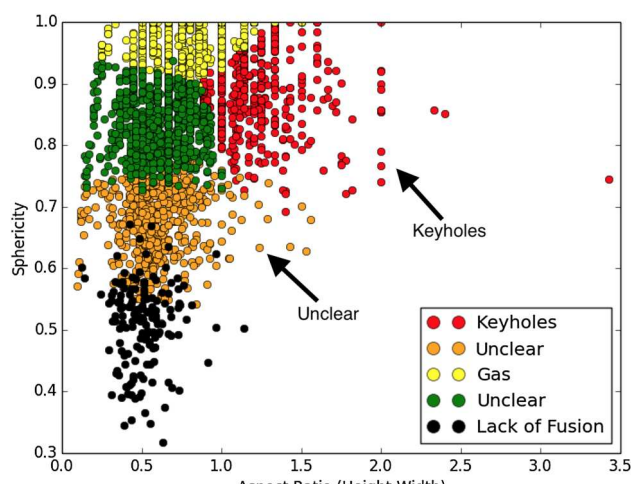
**Results: Defined Limits Method with Optimization—2D Pore Data**

Using defined limits on 2D pore data does not work well. The difficulty in selecting appropriate limits for 2D pore data is even more problematic than for 3D pore data. As a result, this optimisation approach is used as described in the Methods section.

Analysis was carried out on 81 Inconel 718 samples. The results showed that the best limits for classifying LF pores was  $> 31 \mu\text{m}$  and  $< 0.32$  roundness. The best limits for KH pores were  $> 61 \mu\text{m}$  and  $> 0.8$  roundness. The best fit lines on the energy density plot (similar to Fig. 4) produced  $R^2$  values of 0.44 for LF and 0.4 for KH.



(a)



(b)

Fig. 6. Clustered results for the 3D pore data. (a) Pore length and sphericity. (b) Sphericity and vertical aspect ratio.

For the 31 Haynes 282 samples the average number of pores per sample was 9750 with the smallest data set having 2621 pores. The limits for LF pores were  $10 \mu\text{m}$  and 0.48 roundness. The limits for KH pores were  $59 \mu\text{m}$  and 0.82 roundness. The limits generated  $R^2$  values of 0.29 for LF and 0.77 for KH. The weak relationship for the LF pores is explained by the reasonably large spread in values and relatively few data points.

For the 34 Ti-5553 samples there were an average of 348 pores. However, some of the samples had relatively few pores and so could provide anomalous results. The samples with  $< 100$  pores were excluded, leaving 27 samples with an average of 419 pores. The limits for LF pores were  $34 \mu\text{m}$  and 0.47 roundness. The limits for KH pores were  $73 \mu\text{m}$  and 0.98 roundness. The limits created  $R^2$  values of 0.47 for LF and 0.26 for KH.



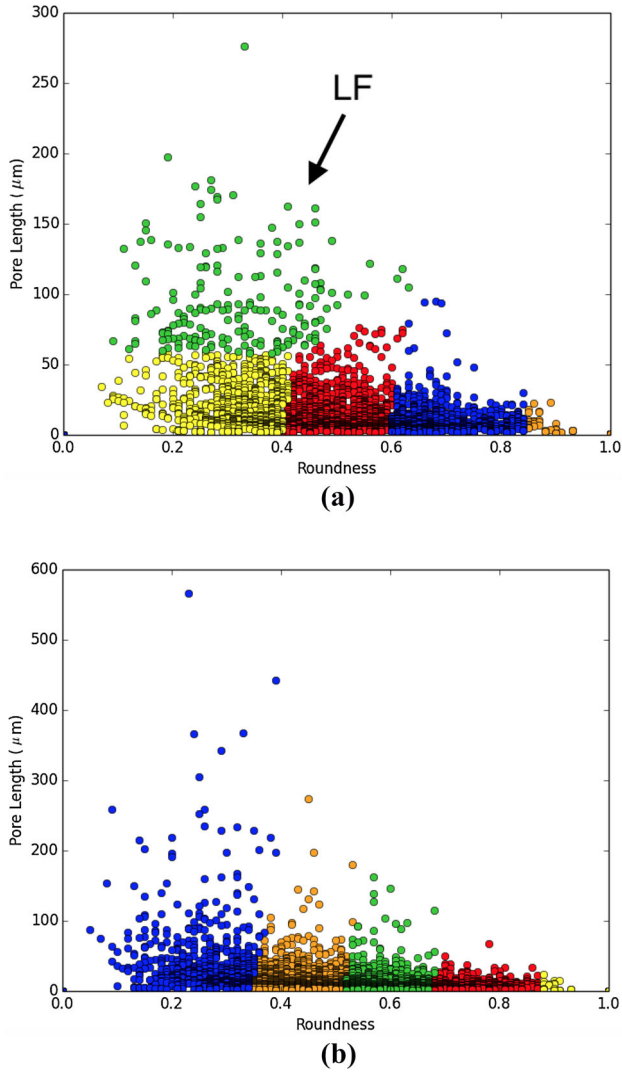


Fig. 7. Clustering examples for two sets of 2D pore data. (a) LF pores are identified (green) but KH pores could not be found. (b) Neither LF or KH pores could be located (Color figure online).

## DISCUSSION

### Three-dimensional Pore Data Results

For the 3D pore data either K-means clustering or defined limits may be suitable. For LF pores both methods had similar results, respectively predicting 183 and 186 pores out of 2664 pores. For KH pores there was a significant difference in results with 373 predicted by the K-means clustering and 1 using defined limits. The values for the defined limits though could vary significantly depending on the limits selected.

The veracity of the two approaches can be confirmed by using more data sets. Data sets on samples that have been made with purposefully different energy densities should show different fractions of LF and KH pores. This can be used to refine the two approaches and should reveal which approach is superior.

### Two-dimensional Pore Data Statistics

All the analysis on 2D pore data has assumed that the cross-sections of the pores are cutting across the middle of the pores at the widest part. This is not a true reflection though of the shape of pores, and there will be some pores that are cross-sectioned at a thinner part. This means that some number of the pores that appear small will actually be much larger pores. Consequently the predicted numbers of the larger pores, LF and KH, will be an underestimate.

From an optimisation standpoint this statistical bias is not important. All of the conclusions from this work, in terms of both pore classification and optimisation, use relative amounts of porosity. As pore data are only compared with other samples the bias on the overall amount should not affect the conclusions.

If the absolute numbers of pores were desired then they could be calculated by considering using the probabilities of where pores have been cross-sectioned. This is a feasible task but would require 3D pore data to train the statistical corrections. These 3D pore data would require detailed mapping of pores to confirm what the 2D slices of each pore type look like along the full height of each pore.

### Two-dimensional Pore Data Results

The 2D pore data could not be analysed using K-means clustering. It may be possible to classify the pores using a different machine-learning method, but there will always be the same data problems: too few variables and a continuous spectrum of values that are difficult to consistently classify.

Using defined limits with optimisation on 2D pore data was more successful. The limits were not predetermined, as was the case in the 3D pore data, but were found by fitting the pore classification results to the energy densities of the samples. This approach allowed defined limits to be found that could be used to classify LF and KH pores.

For the Inconel 718 samples, the LF and KH limits were generated with  $R^2$  values of 0.44 for LF and 0.4 for KH. These are relatively modest  $R^2$  values indicating only a moderate relationship at best. This is to be expected given the values that are being used in the  $R^2$  calculations. Describing the porosity formation with just energy density values is unlikely to be perfect and so the relatively low  $R^2$  values are understandable. A better model may be possible if energy density is replaced with a more complex model of energy input and material conductivity.

The other results showed similar trends and conclusions as the data from Inconel 718. The values for pore length and roundness for the Haynes 282 and Ti-5553 seem reasonable. There are large differences in where the defined limits form but this is expected. It reflects the fact that pore shapes and sizes vary between materials and classifying pores requires adaptable values rather than a set limit for all materials.

## CONCLUSION

A K-means clustering and a defined limit method were trialled on both 2D (micrographs) and 3D pore data (XCT). The different methods had different degrees of success and not all would be suitable for future use.

1. Three-dimensional Pore Data Defined Limits was a reasonable method but there is difficulty selecting suitable limits. Some of the limits had no sources and even those with sources often had values that varied significantly.
2. Three-dimensional Pore Data Clustering was a good approach with it being possible to identify LF, KH and gas pores. A benefit of this approach though was the removal of user input in deciding any defined limits.
3. Two-dimensional Pore Data Clustering did not work well. Although available data established a fair starting point for 2D pore data clustering, more information needs to be stored to reliably classify pores.
4. Two-dimensional Pore Data Defined Limits did work but only after an optimisation process. The arbitrary nature of the defined limits was removed by calculating the limits using the data across many data sets. This approach was used successfully on three different alloys.

## ACKNOWLEDGEMENTS

The authors acknowledge EPSRC funding from the Advanced Metallic Systems CDT (EP/L016273/1), Cummins Turbo Technologies, the Manufacture using Advanced Powder Processes hub (EP/P006566/1) and the Henry Moseley X-ray Imaging Facility (EP/F007906/1, EP/F001452/1 and EP/M010619/1). We also acknowledge the EPSRC for funding the Henry Moseley X-ray Imaging Facility, which has been made available through the Royce Institute for Advanced Materials through grants (EP/F007906/1, EP/F001452/1, EP/I02249X, EP/M010619/1, EP/F028431/1, EP/M022498/1 and EP/R00661X/1).

## CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

## OPEN ACCESS

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## ELECTRONIC SUPPLEMENTARY MATERIAL

The online version of this article (<https://doi.org/10.1007/s11837-019-03761-9>) contains supplementary material, which is available to authorized users.

## REFERENCES

1. I. Gibson, D. Rosen, and B. Stucker, *Additive Manufacturing Technologies: 3D Printing, Rapid Prototyping, and Direct Digital Manufacturing* (Springer, New York, 2014).
2. C. Achillas, D. Aidonis, E. Iakovou, M. Thymianidis, and D. Tzetzis, *J. Manuf. Syst.* 37, 328 (2015).
3. A.J. Sterling, B. Torries, N. Shamsaei, S.M. Thompson, and D.W. Seely, *Mater. Sci. Eng. A* 655, 100 (2016).
4. A. Uriondo, M. Esperon-Miguez, and S. Perinpanayagam, in *Proceedings of the Institution of Mechanical Engineers. Part G: Journal of Aerospace Engineering* 229, p. 2132 (2015).
5. M.C. Kayacan, Y.B. Baykal, T. Karaaslan, K. Özsoy, İ. Alaca, B. Duman, and Y.E. Delikanlı, *J. Appl. Biomater. Funct. Mater.* 16, 68 (2018).
6. S. Bland, and N.T. Aboulkhair, *Met. Powder Rep.* 70, 79 (2015).
7. E. Hernández-Nava, P. Mahoney, C.J. Smith, J. Donoghue, I. Todd, and S. Tammas-Williams, *Sci. Rep.* 9, 4070 (2019).
8. H. Bikas, P. Stavropoulos, and G. Chryssoulouris, *Int. J. Adv. Manuf. Technol.* 83, 389 (2016).
9. L. Sheridan, O.E. Scott-Emuakpor, T. George, and J.E. Gockel, *Mater. Sci. Eng. A* 727, 170 (2018).
10. S. Gangireddy, E.J. Faierson, and R.S. Mishra, *J. Dyn. Behav. Mater.* 4, 441 (2018).
11. J.A. Choren, S.M. Heinrich, and M.B. Silver-Thorn, *J. Mater. Sci.* 48, 5103 (2013).
12. R. Biswal, X. Zhang, A.K. Syed, M. Awd, J. Ding, F. Walther, and S. Williams, *Int. J. Fatigue* 122, 208 (2019).
13. F. Kim, S. Moylan, E. Garboczi, J. Slotwinski, *Addit. Manuf.* 17, 23 (2017).
14. H. Gong, H. Gu, K. Zeng, J. Dilip, D. Pal, B. Stucker, D. Christiansen, J. Beuth, and J. Lewandowski, *25th Annual International Solid Freeform Fabrication Symposium* (2014). [https://www.researchgate.net/publication/280114580\\_Melt\\_Pool\\_Characterization\\_for\\_Selective\\_Laser\\_Melting\\_of\\_Ti-6Al-4V\\_Pre-alloyed\\_Powder](https://www.researchgate.net/publication/280114580_Melt_Pool_Characterization_for_Selective_Laser_Melting_of_Ti-6Al-4V_Pre-alloyed_Powder).
15. W.E. King, H.D. Barth, V.M. Castillo, G.F. Gallegos, J.W. Gibbs, D.E. Hahn, C. Kamath, and A.M. Rubenchik, *J. Mater. Process. Technol.* 214, 2915 (2014).
16. C. Boig, The application of additive manufacturing to nickel-base superalloys for turbocharger applications, Ph.D. thesis, University of Sheffield (2018).
17. H.-P. Kriegel, P. Kröger, J. Sander, and A. Zimek, *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* 1, 231 (2011).
18. R. Banik, *Hands-On Recommendation Systems with Python: Start Building Powerful and Personalized Recommendation Engines with Python* (Packt Publishing, 2018).
19. H. Qi, M. Azer, and A. Ritter, *Metall. Mater. Trans. A* 40, 2410 (2009).
20. T. Vilaro, C. Colin, and J.D. Bartout, *Metall. Mater. Trans. A* 42, 3190 (2011).
21. A. du Plessis, I. Yadroitsev, I. Yadroitsava, and S.G. Le Roux, *3D Print. Addit. Manuf.* 5, 227 (2018).
22. D. Clymer, J. Cagan, and J. Beuth, *J. Mech. Des.* 139, 100907 (2017).
23. P. Karimi, T. Raza, J. Andersson, and L.-E. Svensson, *Int. J. Adv. Manuf. Technol.* 94, 2199 (2018).

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.