Wilkinson, T., Rendell, P., Barba, D. et al. A spatial statistics framework for detection of build defects in laser powder bed fusion using on-axis photodiode sensors. Prog Addit Manuf (2025). https://doi.org/10.1007/s40964-025-01095-4

A spatial statistics framework for detection of build defects in laser powder bed fusion using on-axis photodiode sensors

Abstract

In-process monitoring techniques provide critical insights into components manufactured using laser powder bed fusion. However, one of the most significant challenges is handling the immense volume of data generated during a print, particularly for large components and high-volume production runs. While current state-of-the-art methods in academia are focused on the scan track level or for small parts, industrial applications require scalability. This work presents a new approach to analysing on-axis photodiode data by transforming it into a voxelised spatial structure. It uses statistical reduction functions to consolidate multiple sensor readings into a representative value, exposing characteristic thermal and melt pool stability behaviour. This reduces complexity while preserving key patterns. We demonstrate the framework's effectiveness with two case studies: first, we validate the approach using a benchmark dataset from an overheating test part. Second, we investigate the correlation between photodiode data and spatter-induced porosity showing how the statistical features of the dataset can be used to highlight the location of defects. We found that using the mean of the photodiode response had a moderate negative correlation with porosity, up to -0.6, and the standard deviation had a moderate positive correlation, up to 0.45. Finally, we show how this correlation is improved by combining multiple

statistical features into a single indicator, improving the correlation strength to up to 0.68.

Keywords: laser powder bed fusion, in-process monitoring, defect detection, spatter

1 Introduction

In-process monitoring (IPM) might be the catalyst needed to make the next technological leap in metal additive manufacturing (AM). Current quality control methods rely heavily on post-build inspection methods such as computed tomography (CT) scanners or destructive testing - both of which can add significant time and cost to a production run. A comprehensive suite of sensors coupled with powerful analysis tools could remove this dependency, reduce costs, and speed up quality control while providing engineers with suggestions for optimising the manufacturing route.

During laser powder bed fusion (LPBF) manufacturing, variations in the powder feedstock [1, 2], powder spreading artefacts [3, 4], unfavourable thermal conditions in the melt pool [5], and spatter particles [6] can all cause defects to appear in the final printed component. An IPM system should be capable of identifying some (or all) of these defects, alerting the engineer to their presence.

Among the many examples of IPM sensors - including high-speed [7, 8] and thermal cameras [9], and photodiodes [5, 10–15] - photodiode sensors offer a promising approach to defect detection. While cameras record spatial information, the data storage requirements are immense, especially for large components. Photodiodes offer much higher temporal resolution and lower storage demand at the cost of spatial information. Many commercial LPBF machine manufacturers, such as Trumpf and Renishaw, include photodiode sensors in their process monitoring systems. These systems often feature photodiodes made from different sensor materials to cover a wider

wavelength range, enabling accurate monitoring of various materials with different melting points.

The currently available research on using photodiode sensors for defect detection can be divided into two categories: single track and 3D sample studies. The single-track studies typically use a design of experiments to print single tracks and use either a high-speed camera or CT scanner to generate ground truth data for the defect detection workflow. Within these two categories, the studies are further split into photodiode-only and sensor fusion techniques, where the photodiodes are used in conjunction with other sensor data.

Forien et al. [5] studied single tracks in their work, aiming to link *in-situ* pyrometer measurement data with *ex-situ* X-ray radiography scans. They showed that the measured pyrometer signal correlated with the normalised enthalpy of the input laser parameters and could, therefore, be used to determine the melting regime of the track. They finish their work by showing that the slope change in the pyrometer signal and melt pool dimensions with normalised enthalpy define a threshold for the conduction to keyhole melting regime. They use this information to define three zones that describe the likelihood of pore formation based on the pyrometer signal (low, $<5\%$, high, $>95\%$, and a transition zone in between). Their work shows that the pyrometer signal is a useful tool for identifying where pores are more or less likely to form for single tracks printed on a base plate, but it does not show how these observations extend to 3D.

Harbig et al. [16] extend the work of Forien et al. [5] using a sensor fusion technique to combine photodiodes with high-speed and thermographic cameras for the identification of balling defects in single tracks. The authors used a height map of the track profile to identify balling defects, with significant height differences being classed as defective. They used three different filters for anomaly detection for the in-process data: absolute fluctuations (AF), dynamic, and short fluctuations (SF). The

full details can be found in the original work, but as a summary, the AF filter identifies points that exceed set thresholds, the dynamic filter identifies points where the value has changed more than a set amount from the previous sample, and the SF filter defines up and lower bounds based on the moving average of the signal shifted up or down by a set amount. They applied these filters to different features from their monitoring sensors, such as melt pool width from the high-speed camera. They showed that using data fusion techniques improves the sensitivity of defect detection over using the individual sensors by up to 20% in some cases. This work was also not extended to 3D.

One of the main issues that will be faced when extending the single track approaches into 3D is that while a defect may be correctly detected at a certain instance in time, it has been shown that these defects can be healed during subsequent layers with appropriate parameter selection. Feng et al. [17] demonstrate this in their work, observing that high volumetric energy density (VED) causes keyhole pores several layers below the current printing layer, low (VED) causes lack-of-fusion pores in the current layer, and medium (VED) can heal lack of fusion pores in the layers below. This phenomenon is unique to printing multiple layers, and as such, the algorithms developed for the single tracks will not take it into account. This is also demonstrated by Chen et al. [18], where they demonstrate the potential for lack of fusion (LOF) defects to be healed by subsequent layers with an appropriate scanning strategy. Both of these works show the importance of multi-layer defect detection algorithms to capture the complex formation mechanisms moving forward.

Moving into 3D methods, Jayasinghe et al. [12] used several machine-learning methods to cluster high and low-density sample cubes, as well as directly predict the density. They used singular value decomposition (SVD) to combine three photodiode sensors and extract features to use for machine learning. Their unsupervised clustering technique could cluster high and low-density samples (greater than or less than 99%

density) with 93.5% accuracy. They also attempted to use the same feature set to predict the sample cubes densities using a Gaussian process model, achieving a root mean squared (RMS) error of 3.7%. Their clustering algorithm shows good accuracy and is potentially useful in a quality control setting where a binary pass or fail result based on density is required. The main drawback of this study is that the proposed models cannot localise the porosity in space and instead focus on predicting the density of the entire component. A defect detection system that can localise areas of high and low porosity would provide a more use case-specific quality assessment, with more weight on porosity found in highly-stressed regions, for example. This would prevent part rejection due to porosity found in non-critical regions. This method may also suffer when scaling to larger samples, as computing the SVD of very large matrices requires a large amount of memory.

Gronle et al. [10] also developed a method for process monitoring in 3D using a photodiode sensor to detect overheating defects in printed parts. Their method uses a control chart on the layerwise mean of the photodiode signal to identify overheating layers. This will scale well to larger components but will only indicate in which layer an overheating defect may occur. This method will also likely miss defect events that occur on a shorter time scale compared to the overall layer time, as they will be diluted by the ”normal” signals in the remainder of the layer. The layerwise mean is also likely to be sensitive to the cross-sectional area of the part, so it may not be a reliable indicator for more realistic components.

In 2023, the ASTM International Additive Manufacturing Centre of Excellence released a report covering the technological readiness of *in-situ* monitoring for AM [19]. In this comprehensive report, participants from an international consortium of AM experts were tasked with determining the technological readiness of *in-situ* monitoring for AM. Some of the key insights from this report are:

- Some AM-produced parts are beginning to exceed the current limits of *ex-situ* inspection methods, such as CT scanning.
- Detecting sub-surface flaws is a key challenge given that post-processing and subsequent layers can heal defects.
- The scale of the data generated by *in-situ* monitoring is a significant challenge, and data reduction techniques are proposed as a potential solution.
- The currently available *in-situ* monitoring methods are suitable for defect detection, but more work is needed to develop algorithms capable of extracting key information from the data.

Another important aspect highlighted in the ASTM report is the need for a standardised approach to data collection and analysis. They detail that each machine manufacturer has its own data collection methods, file formats, and data processing methods, making it difficult to compare results across different machines. This point is also relevant when considering the potential impact of machine learning and artificial intelligence on the field, something which participants in the report were optimistic about. Having a standardised approach to data collection and analysis would allow for the development of machine learning algorithms that could be applied across different machines, and increase the amount of data available for training.

This review of the literature shows that significant challenges remain in predicting build defects at both the single track and 3D levels. The single track methods are capable of predicting the appearance of a defect with high levels of accuracy, but they may not scale well into 3D due to the algorithms involved, or they do not take into account the complexities of defect formation in 3D, where pores may appear in other layers, or be healed by subsequent layers. At the same time, the currently proposed 3D methods fall short also in terms of scalability or localisation of defects beyond a layer. Hence, we have identified the following gaps in the literature:

- The current state-of-the-art methods are not scalable to larger components, either due to the algorithms used or the data processing methods.
- The current state-of-the-art methods do not localise defects in space, only indicating which layer a defect may occur in.
- The current state-of-the-art methods do not take into account the complexities of defect formation in 3D, where pores may appear in other layers, or be healed by subsequent layers.
- The current state-of-the-art methods do not use a standardised approach to data collection and analysis, making it difficult to compare results across different machines.

In this work, our proposed framework directly targets these research gaps in the following ways:

- We propose a method for reducing the scale of the dataset using statistical reduction functions and voxelisation
- This voxel structure takes multiple layers into account, allowing for the detection of defects that may appear below the current printing layer or be healed by subsequent layers.
- The voxel structure visualises the data in 3D, allowing for the localisation of defects in space.
- The voxel structure is machine-agnostic, allowing for the development of algorithms that can be applied across different machines.
- We propose a method for transforming porosity data from CT scans into the voxel representation using kernel density estimation (KDE), allowing for direct comparison with the photodiode data in 3D, further standardising the approach.

This manuscript is organised into three sections. Section 2.1 describes the experimental setup of the two case studies in this work, designed to induce defects we wish to detect using the in-process monitoring equipment. Section 2.2 talks about

our proposed solution for detecting the induced defects, including how the raw data is converted into a spatially organised voxel structure, as well as the spatial statistics calculations and how the ground truth porosity data is transformed into a voxel representation also. Finally, Section 3 presents the results for both case studies, highlighting the successes of the proposed solution and the routes to improvement. Three signal processing methods are used for comparison with examples of current analysis methods.

2 Methodology

2.1 Experimental Setup

2.1.1 Case study 1: On the detection of overheating defects

The Open Data Challenge is a free, open dataset released by Gronle et al. [10] to aid researchers in developing novel solutions to in-process monitoring with LPBF. The data is taken from an on-axis indium gallium arsenide (InGaAs) photodiode sensor while manufacturing a part containing intentional overheating defects. Each overheating defect is induced by leaving unexposed powder regions within the part. Leaving more unexposed powder should cause more significant overheating in the solid layers above due to the lower thermal conductivity of the powder material. The part design is given in Figure 1. There are ten overheating defect blocks with an increasing number of unexposed layers, increasing the severity with the build height. Block 1 has one unexposed layer, block 2 has two unexposed layers, and so on for the rest of the part. The part was printed on a TRUMPF TruPrint 5000 laser powder bed fusion system in Al-Si-10Mg, using 1500mm/s scan speed, 480W laser power, and a spot diameter of 100µm. The InGaAs photodiode sensor has a sampling rate of 100kHz and a wavelength sensitivity of 1100-1700nm, covering the short infrared range. The

sampling rate in the published dataset was reduced to 50kHz. Full details of this dataset are given in the original work by Gronle et al. [10].

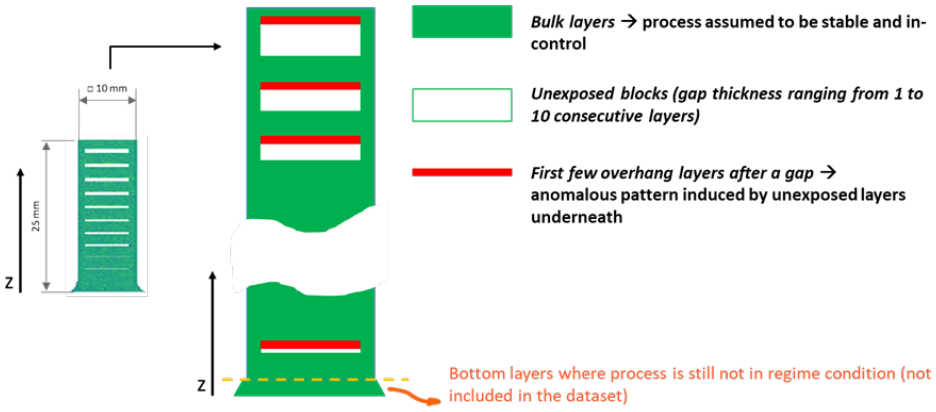


Fig. 1 The printed test part for the Open Data Challenge, reproduced with permission [10].

2.1.2 Case study 2: On the detection of spatter-induced porosity

For the second case study, six parts (cubes with an edge length of 12.7mm) were manufactured in AlSi10Mg using a Renishaw AM500Q LPBF system, as shown in Figure 2. Each of the four optical assemblies in the machine has three photodiodes: one for observing the laser output and two for observing the melt pool region. These two melt pool photodiodes have sensitivity ranges of 700-1040nm and 1070-1700nm with a sampling rate of 100kHz. Only one laser was used for this build. The six test parts were built next to a large solid block of material with parameters tuned to cause significant spattering, such that the ejected particulates are carried by the gas flow towards the parts. Spatter refers to particulates being ejected from the melt pool zone during printing [20]. The printing order was set to print the spatter producer first, depositing significant spatter onto the parts and contaminating the powder before printing them in order from 1 to 6. The spatter producer was only half the height of the part, meaning only the lower half had significant spatter. The parts were printed

with a laser power of 310W, 25µm point distance, 120µm hatch spacing and 50µm layer height, proven to produce dense parts by a previous internal study. A meandering scan strategy was used with a random layer rotation angle that also avoided scanning directly into the gas flow such that any spatter or other emissions would not interfere with the laser. Please refer to the published dataset to see the exact scan strategy. The spatter producer block was printed using 270W laser power, 17µm point distance and 60µm hatch spacing, using the same scanning strategy as the parts. These parameters were internally validated to ensure that high levels of visible spatter were produced, exceeding that from normal printing. The gas flow rate was set to 190m^3/h.

2.2 Data processing framework

This section discusses the data processing pipeline in detail, going from raw data to the correlation with CT defects. The entire process takes place across two parallel branches. The first branch concerns the *in-situ* monitoring data, which is voxelised using mean and standard deviation as reduction functions, then normalised to remove any optical artefacts and positional variation, and finally the spatial correlation is calculated to identify significant clusters of high and low values. The second branch concerns the *ex-situ* CT data, which is voxelised using kernel density estimation to create a continuous porosity field. This output porosity field and the processed *in-situ* data are then compared, and the correlation is calculated on a voxel-by-voxel basis. This process is summarised graphically in Figure 3.

2.2.1 Preprocessing

Several preprocessing steps are applied to the raw data before it can be voxelised. Firstly, since the two machines used in this study export data in different file formats, they are both transformed into a standard .csv format read by the voxelisation software. Each row in the .csv file contains the timestamp, coordinates, and photodiode reading of each sample as a minimum. The data from the Trumpf is filtered to remove

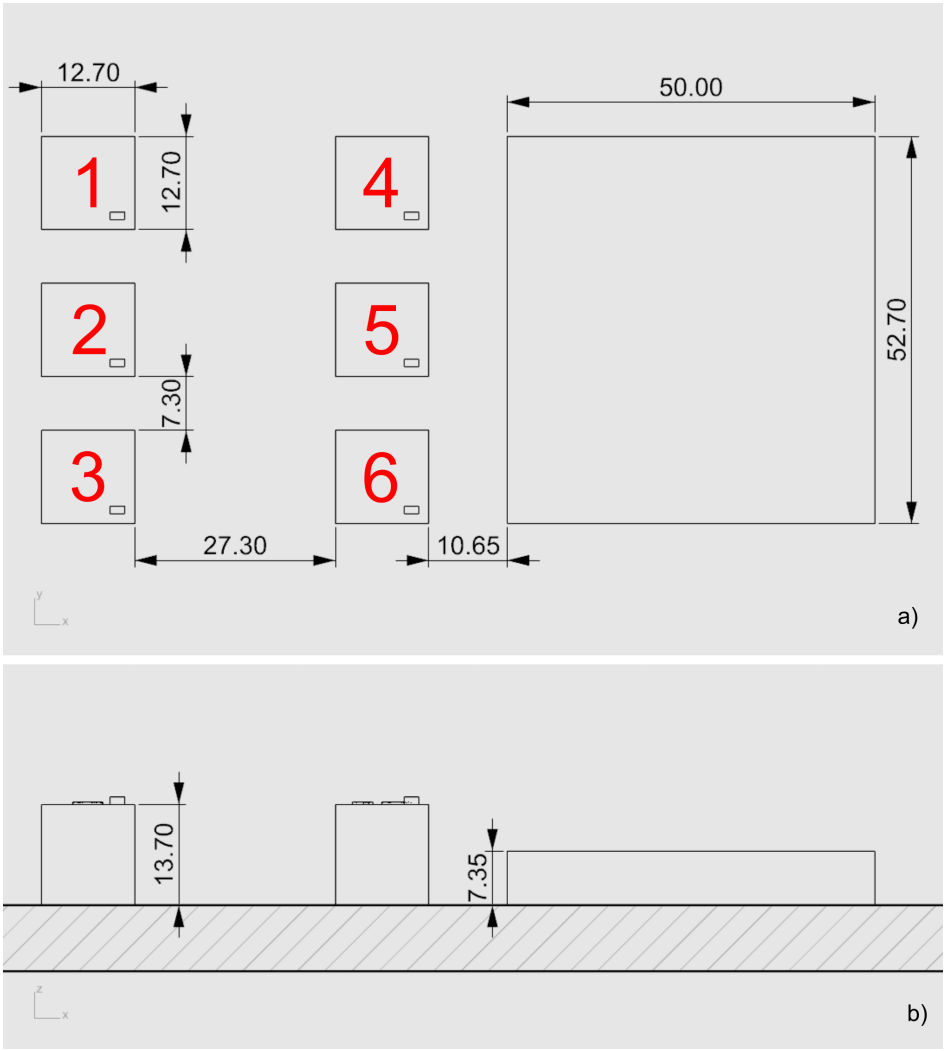


Fig. 2 The a) top and b) front views of the build layout for case study 2. The gas flow direction is from right to left.

samples where the photodiode value is lower than 1000 units to ensure only samples where the machine is printing are left in the dataset. The data from the Renishaw AM500Q is also filtered to remove samples where the laser is either deactivated or below a minimum power threshold to ensure that only samples where the machine is printing are kept. Further details on the filtering and the effect of the filtering on the data is detailed in the appendices, Section A.

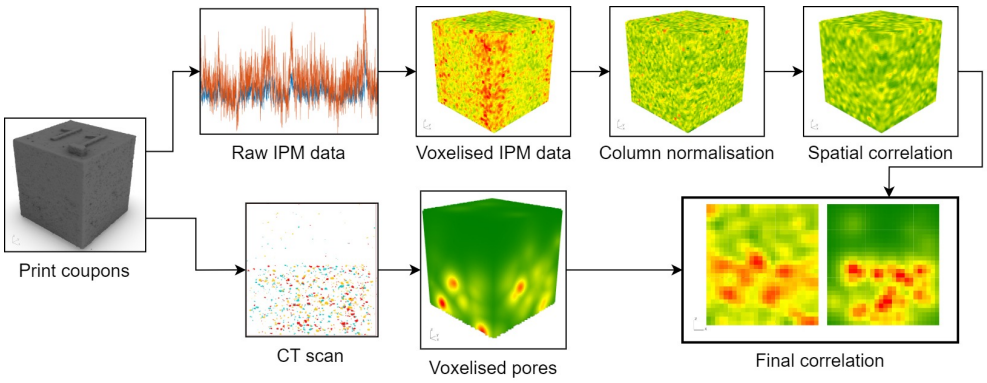


Fig. 3 A graphical overview of the voxelisation, spatial statistics, and defect correlation processes. The raw IPM data is voxelised so that a series of normalisation and correlation operations can happen before it is compared with a voxelised representation of the porosity extracted from the part using high-resolution CT. Data is voxelised using 200µm voxels.

2.2.2 Voxelisation

The voxelisation process was implemented using C# (.NET framework 4.8). A set of components for Rhino 8 / Grasshopper (Robert McNeel & Associates, 2024) was written to interface between Grasshopper and the voxel data structure for visualisation in Rhino, (url removed for anonymity). The voxel structure is stored as a dictionary in the class, with a 3D tuple of integers as the key and a custom data structure as the value. This structure stores the data points within each voxel and a scalar for the final voxel value. The dictionary structure was chosen for its fast lookup speed, and the integer keys make it simple to query neighbouring voxels.

The CSV files contain each sample point's x and y laser coordinates and photodiode response. To add a point to the voxel structure, the relevant voxel key in the dictionary is first calculated using the world to voxel space transform:

$$\mathbf{v} = \lfloor \frac{\mathbf{w} - \mathbf{o}}{s} \rfloor \tag{1}$$

where $\mathbf{v}$ is the integer voxel position, $\mathbf{w}$ is the world-space coordinate of the voxel centre in millimetres, $\mathbf{o}$ is an offset in millimetres, s is the voxel size in millimetres,

and $\lfloor\ \rfloor$ is the integer floor operator. Much like finite element analysis, the size of the discretisations will affect the final analysis results, and a sensitivity analysis is performed later on in Section 3.2. The voxel structure also includes a voxel-to-world-space transformation to visualise the structure and register it with input CAD and CT scanned data, Equation (2).

$$\mathbf{w}_i = s(\mathbf{v}_i + 0.5) + \mathbf{o}_i \tag{2}$$

This voxelisation process results in the raw data being transformed into a 3D grid of voxels, with each being represented in index notation as $V(i, j, k)$, where i, j, and k are the voxel indices in the x, y, and z directions respectively.

2.2.3 Voxel statistics

Once the input data is voxelised, statistical reduction function combine sample points within each voxel and its immediate neighbours into a single number. The statistical reduction functions are selected based on further of the literature on photodiode *in-situ* monitoring, and the types of patterns expected. For example, Forien et al. [5] showed that the mean photodiode signal correlates with the normalised enthalpy of the laser, and hence the melting regime. Gronle et al. [10], the source of the first dataset in this manuscript, showed that the mean photodiode signal can be used to detect overheating defects in a part, maintaining the laser parameters constant. Berumen et al. [21] showed that both the mean and standard deviation can indicate powder layer inconsistencies, noting both a higher mean and standard deviation for thicker powder layers. Montazeri et al. [22] investigated the photodiode response of printing with Inconel powder that was artificially contaminated with tungsten and aluminium particulates. They noted that scan tracks that covered contaminated regions showed clear peaks in the photodiode response. For this work, we anticipate that the spatter-contaminated regions will show a similar trend, based on the work by Simonelli et

al. [23]. They showed that spatter particulates from Al-Si-10Mg powder often have a thick magnesium oxide layer, which will have a higher melting point than the feedstock material. We anticipate that this will cause a higher standard deviation in the photodiode response signal due to the elevated peaks in the signal.

By using each voxel's neighbours in the reduction, we capture the same volume of data as a voxel three times the size while having high spatial resolution from a smaller voxel size. As highlighted in the work by Jayasinghe et al. [12], defects can appear at a distance of 0.4-0.5mm away from the in-process signal, so using the neighbouring voxels takes this into account for smaller voxel sizes. This process is shown graphically in 2D in Figure 4.

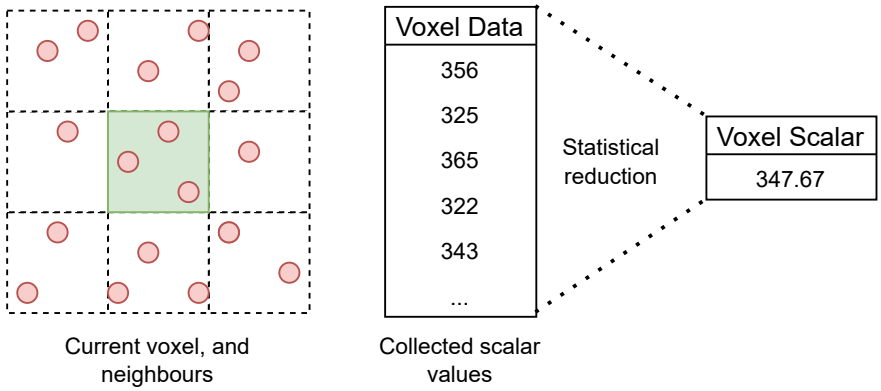


Fig. 4 A schematic showing how the raw sample points from each voxel are combined into a single number. Depending on the analysis requirements, this statistical reduction could be any appropriate function, such as the mean, standard deviation or skewness.

Mathematically, let $V(i, j, k)$ represent a voxel at position (i, j, k) in the 3D voxel grid. Each voxel contains a set of photodiode samples $S(i, j, k) = \{s_1, s_2, ..., s_n\}$. The neighborhood $N(i, j, k)$ of a voxel includes all 26 adjacent voxels that share a face, edge, or vertex with $V(i, j, k)$:

$$N(i,j,k) = \{V(i+\alpha, j+\beta, k+\gamma) \mid \alpha,\beta,\gamma \in \{-1,0,1\}, (\alpha,\beta,\gamma) \neq (0,0,0)\} \quad (3)$$

For a voxel $V(i,j,k)$, we compute its reduced value $R(i,j,k)$ using all samples from both the voxel itself and its neighbours. X is the union of all sample sets in the neighborhood:

$$X = S(i,j,k) \cup \left(\bigcup_{V' \in N(i,j,k)} S(V') \right) \quad (4)$$

Hence, the mean reduction function is:

$$R(i,j,k) = \frac{1}{n} \sum_{x \in X} s_x \quad (5)$$

For the standard deviation reduction function:

$$R(i,j,k) = \sqrt{\frac{1}{n} \sum_{x \in X} (s_x - \mu)^2} \quad (6)$$

where μ is the mean calculated above and n is the total number of samples in X.

2.2.4 In-column voxel normalisation

After the initial voxelisation of the raw in-process data, we noticed that the in-process data collected from the Renishaw machine varies significantly depending on the XY position of the voxels. For example, cubes 4, 5 and 6 have generally higher mean photodiode values than cubes 1, 2 and 3, despite using the same laser parameters. This is discussed later in the results section. Besides this large-scale positional effect, we also noticed a localised optical artefact, shown in Figure 5. The data suggests that this artefact is caused by the laser incidence angle being close to vertical since the centroid of the artefact aligns with the laser home position. We correct this using a

normalisation process that normalises voxels based on the mean and standard deviation of other voxels that share XY coordinates. This normalisation process is made possible by the efficient organisation of the voxel data structure, $V(i, j, k)$, where i, j and k are the voxel indices in the x, y and z directions respectively. Given our voxel structure $V(i, j, k)$ with reduced values $R(i, j, k)$ calculated as described in Section 2.2.3, the in-column mean μ_{ij} and standard deviation σ_{ij} are calculated:

$$\mu_{ij} = \frac{1}{N_k} \sum_{k=1}^{N_k} R(i, j, k) \tag{7}$$

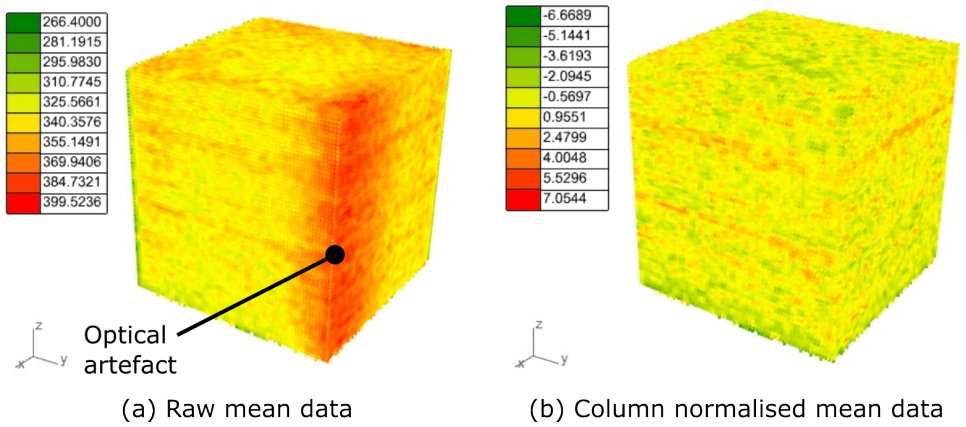
$$\sigma_{ij} = \sqrt{\frac{1}{N_k} \sum_{k=1}^{N_k} (R(i, j, k) - \mu_{ij})^2} \tag{8}$$

where N_k is the voxel count in the z-direction for column ij. Finally, the normalised voxel structure, $R_{norm}(i, j, k)$, is calculated using μ_{ij} and σ_{ij} as follows:

$$R_{norm}(i, j, k) = \frac{R(i, j, k) - \mu_{ij}}{\sigma_{ij}} \tag{9}$$

In other words, we calculate the mean and standard deviation for each column of voxels in the Z direction, using them to normalise the voxels in those columns.

It is important to emphasise that this column normalisation process is implemented to compensate for positional variation observed in the *in-situ* photodiode data when reduced using the mean. This phenomenon has been identified by other authors using the same machine and process monitoring setup from Renishaw. Carter et al. [24] recently studied and applied a calibration process to correct for a positional trend in their data. In future work, special attention must be paid to fully quantifying and correcting for this trend. It should be noted that this positional trend is not present when the standard deviation is used as a statistical reduction function.



(a) Raw mean data (b) Column normalised mean data

Fig. 5 A perspective view of cube 3, showing (a) the raw mean data with the optical artefact identified and (b) the column normalised data. Column normalisation removes the in-plane artefacts by normalising the voxel values based on vertical columns of data. Data is voxelised using 200µm voxels.

2.2.5 Spatial statistical correlation

In this work, the goal of the analysis is to use photodiode data to predict the location of defects in a printed component. Some of these defects will be caused by stochastic events in the melt pool, which are either extremely difficult or impossible to predict. This, combined with noise in the recorded photodiode data, results in a 3D voxel structure with both random and significant variation. These significant variations in data are more likely to cause predictable, systematic defects. In this work, we separate the random and significant spatial variation using the Getis-Ord metric [25, 26], which calculates how well a particular data sample correlates with its neighbours. More positive Getis-Ord values mean the sample is a significant cluster of above-average samples, more negative values show highly correlated, below-average values, and zero represents random variation.

Given a voxel $V(i, j, k)$ with neighborhood $N(i, j, k)$ and neighbour count $m(i, j, k)$, the Getis-Ord metric is then calculated:

$$G(i,j,k) = \frac{\sum_{V' \in N(i,j,k)} R(V') - m(i,j,k)\overline{R}}{S\sqrt{\frac{nm(i,j,k) - m(i,j,k)^2}{n-1}}} \tag{10}$$

where $\overline{R}$ and S are the global mean and standard deviation of all reduced values and n is the total number of voxels. The Getis-Ord metric does not include the centre sample in the calculation. This work assumes a spatial weighting of 1 between all neighbouring voxels. By scaling the local statistics using the global mean and standard deviation, the final Getis-Ord metric represents how statistically significant the hot and cold spots in the data are, and thresholds can be used to reveal statistically significant clusters within the part. This makes the Getis-Ord metric effectively a z-score. Highly positive values represent clusters of high values, and highly negative values represent clusters of low values. Figure 6 shows the effect of spatial correlation on both raw and column-normalised data.

2.2.6 CT Mesh Processing

The printed samples from case study 2 were scanned using a Phoenix V|tome|X M300 (Waygate Technologies) CT scanner to extract the ground truth porosity. The scanner used a 25µm voxel size, a 140kV voltage, 370µA current, and 500ms exposure time. The resulting meshes are transformed to the same voxel representation as the IPM data for direct voxel-by-voxel comparison. There are many possible ways to perform this transformation, and in this work, we choose to represent the porosity as a continuous density field using kernel density estimation (KDE). By representing the porosity as a continuous field using KDE, we are making the assumption that for repeated prints of the same build using the same parameters, the pore locations will differ, but the general porosity distribution will be similar. Each pore is treated as a point source whose influence is weighted by its volume and decays exponentially with distance. For a voxel $V(i,j,k)$, its porosity value $W(i,j,k)$ is calculated using the n nearest pores:

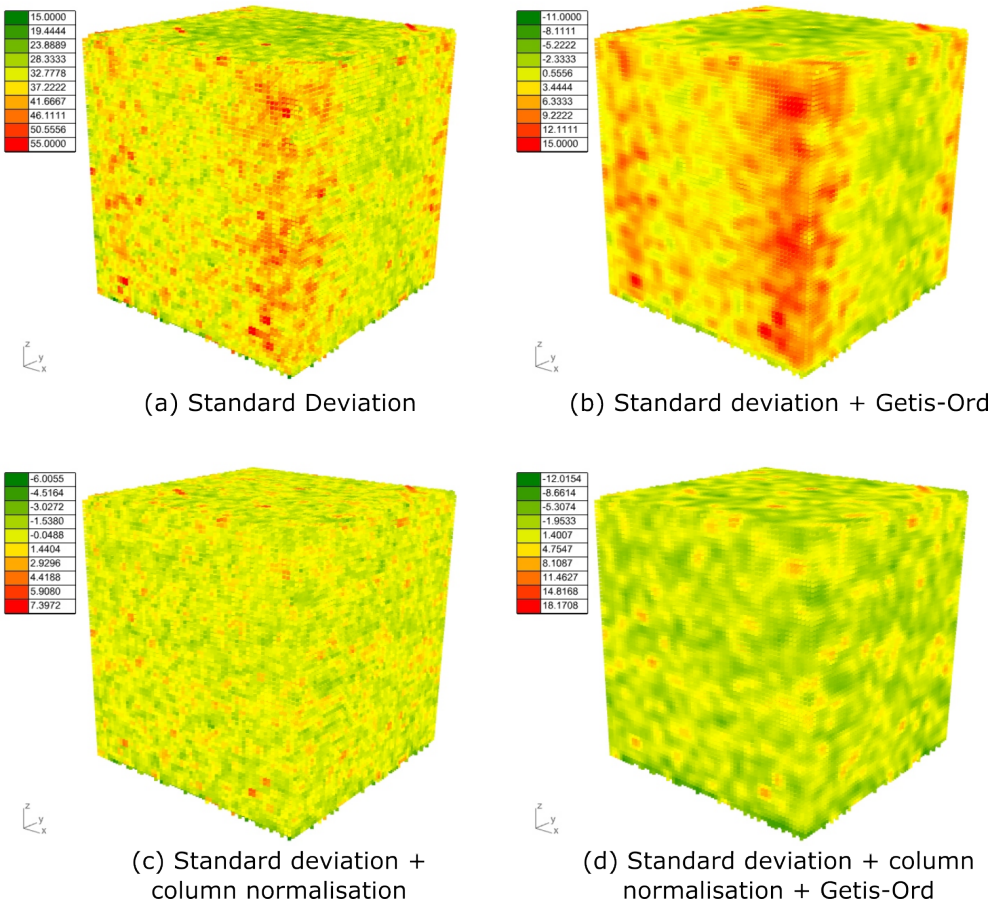


Fig. 6 Point cloud views of (a) the raw standard deviation, (b) the spatial correlation of the standard deviation, (c) the column-normalised standard deviation data and (d) the spatially correlated, column-normalised standard deviation. The data shown is from cube 3. Data is voxelised using 200µm voxels.

$$W(i,j,k) = \sum_{p=1}^{n} v_p K(d_p) \tag{11}$$

where $K(d) = e^{-d}$ is an exponential kernel function, v_p is the volume of the pth nearest pore, and d_p is the Euclidean distance between the voxel center, $\mathbf{x}_{ijk}$ and the pth pore centroid, $\mathbf{x}_p$:

$$d_p = \|\mathbf{x}_{ijk} - \mathbf{x}_p\| \tag{12}$$

The parameter n effectively controls the bandwidth of the kernel density estimate, balancing spatial resolution and field smoothness. A small value of n produces a highly localized density field with distinct maxima at individual pores, while larger values create a smoother distribution that better represents regional porosity. We use an R-Tree spatial indexing structure [27] for efficient nearest-neighbour querying during the calculations. In this work, we chose a value of $n = 10$ as a balance between a smooth porosity field and computational efficiency. This process is shown graphically in Figure 7.

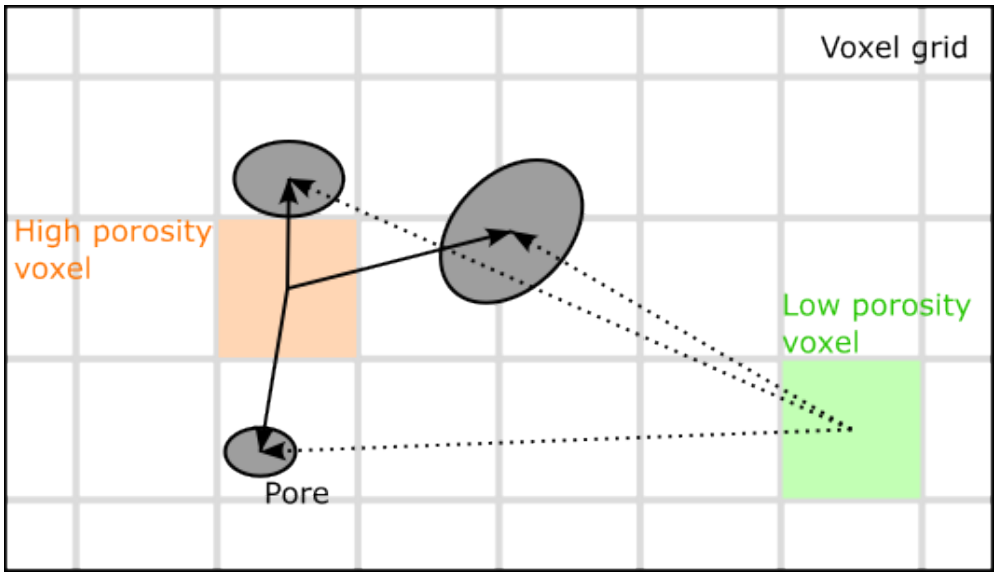


Fig. 7 A schematic view of calculating the porosity for a given voxel. The voxel grid is shown in grey, and the arrows show the distance between the voxel and pore centroids. The orange box represents a high porosity voxel since it is closer to the three pores, and the green box represents a low porosity voxel.

This voxelisation process is performed for all six parts in the study, and Figure 9 plots the average porosity in each voxel layer. What is clear is that for all six cubes, the porosity is higher in the first half before reducing in the second half. This shows that inducing spatter causes increased porosity as intended. As expected, cubes 1-3 also show significantly less porosity than cubes 4-6. Being further away from the

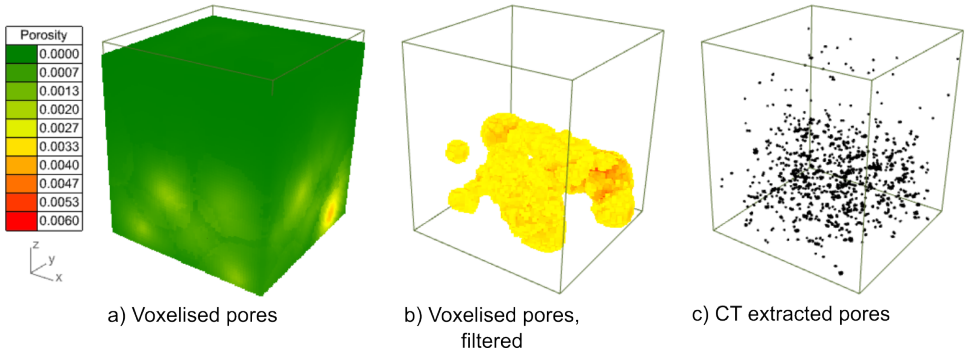


Fig. 8 (a) An example of voxelised porosity, with (b) the same data but filtered to show the more porous regions. Data is voxelised using 200µm voxels.

spatter producer means the powder for cubes 1-3 is less contaminated, resulting in lower porosity.

2.3 Analysis Methods

2.3.1 Statistical Thresholding

One method to identify potential defect locations using the voxelised data is to perform statistical outlier testing using an appropriate significance level (α). This test assumes the data follows a normal distribution and sets a threshold for the likelihood of a particular point being a random variation or a significant outlier. The choice of α will vary depending on the specific analysis and the tolerance of errors—using a higher significance level (e.g. 0.1 instead of 0.05) would result in more false positives or falsely detecting defect regions, and using a lower significance level (e.g. 0.01) would increase the number of false negatives, or fail to identify defect regions. The optimal choice would depend on the particular application. For example, it would be better to use a higher significance level and over-predict porosity in a safety-critical application. However, in a system where a false alarm is costly, such as stopping production unnecessarily, it may be wise to decrease the significance level. In this work, the voxel data are normalised using the Getis-Ord metric, converting each value to a z-score.

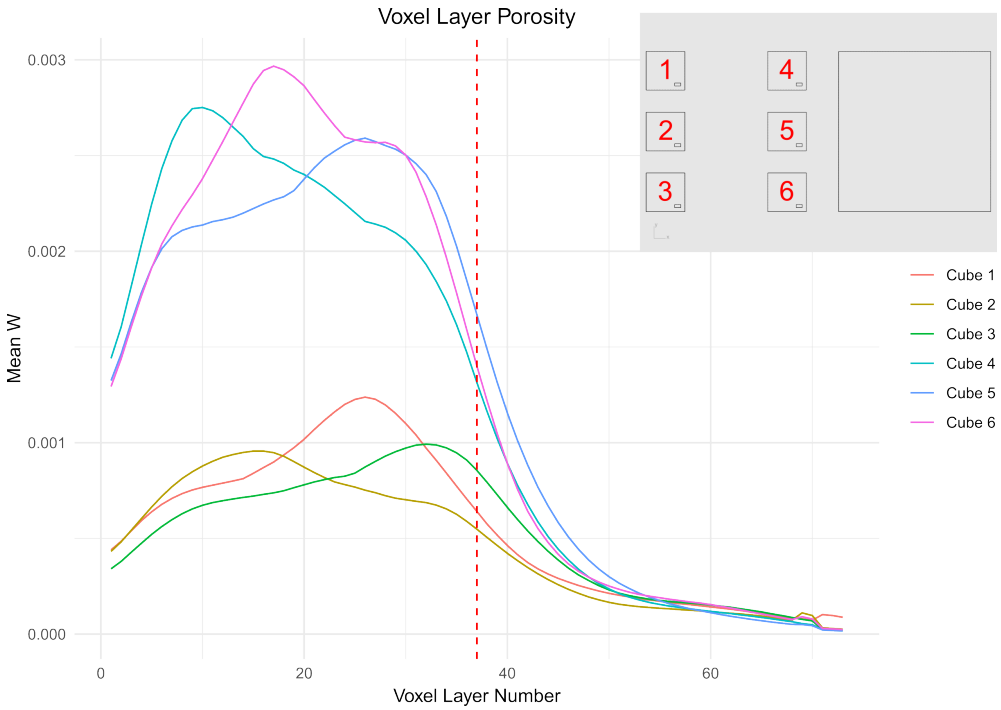


Fig. 9 The average porosity W per layer for each of the 6 parts. The red dotted line indicates where the spatter block finished printing. The build layout is shown in the upper right for reference. Data is voxelised using 200µm voxels.

We then select a commonly used α of 0.05 to determine outlier points. In the case of overheating for example, one might use a one-tailed test to only threshold significantly high mean voxel values. This converts to a z-score threshold of 1.645, where points exceeding this are determined to be outliers, or potential overheating defects. For a two-tailed test where both overheating and undermelting defects are to be detected, at an α value of 0.05, this converts to z-score thresholds of ± 1.96. We use a one-tailed test to identify significant overheating regions in case study 1, presenting the results in Section 3.1.

2.3.2 Ground Truth Correlation

When ground truth data is available, such as in case study 2, we can take the analysis a step further and directly link the IPM data with pores. To do this, both the IPM

and porosity data are voxelised as described in Section 2.2. We also normalise the voxelised in-process data using the Getis-Ord metric. Now that both the sensor and porosity data are in a common format, they are compared voxel by voxel to determine their correlation. This work uses Spearman's Rank as the correlation measure due to the non-linear relationship between the IPM data and the porosity. In this case, we expect Spearman's Rank to be a more representative correlation measure than the alternative Pearson correlation because the porosity data is heavily skewed, with many small pores and fewer large pores. The results of this analysis with case study 2 are discussed in Section 3.2.

3 Results

3.1 Case study 1

In this section, the dataset from the Open Data Challenge [10] is voxelised with a voxel size of 150µm. This voxel size was chosen for a high spatial resolution. The mean and standard deviation of the data are calculated as described in Section 2.2.3, and the spatial correlation is calculated as described in Section 2.2.5. No anomalous in-plane variation was observed in this data, so it was not column normalised.

3.1.1 Mean response

Figure 10 shows the raw calculated mean and the corresponding spatial correlation values. In Figure 10(a), the overheating effect for the defects above the unexposed layers is visible, with the first voxel layer after each overheating defect having higher mean values compared to the rest of the part. Figure 10(b) shows this more clearly, with the hotspots of high mean showing up in red above the unexposed powder layers. Figure 10(b) also reveals the less severe overheating defect layers as hotspots, whereas these lower layers are hidden in the raw mean data.

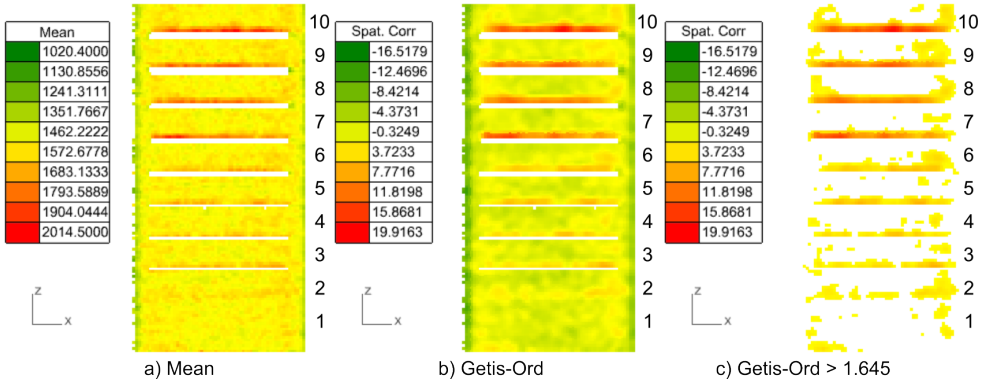


Fig. 10 Slices through the (a) raw mean data, (b) the Getis-Ord G_i spatial correlation of the mean data, and (c) the thresholded Getis-Ord data. Image (c) only shows voxels with a Getis-Ord value greater than or equal to 1.645. (b) and (c) have the same colour scale. Data is voxelised using 150µm voxels.

Now that the data is normalised, we can apply the statistical testing discussed in Section 2.3.1 isolate outlier voxels. In this case, we use a one-tailed statistical test with $\alpha = 0.05$ to identify only overheating voxels, matching the analysis in the original work by Gronle et al. [10].

Thresholding the data to only show values with a G_i metric greater than 1.645 (a significance level of 0.05 for a one-tailed test) reveals the statistically significant hotspot regions within the sample, shown in Figure 10(c). The overheating effect for nine out of the ten blocks is visible, with only the first block being undetected. Blocks 7-10 have a much higher G_i metric, making them more statistically significant than blocks 3-6. Block 2 has statistically significant voxels in certain areas, and block 1 is not considered to be overheating. This correlates well with the results obtained in the work by Gronle et al., where they also concluded that block 1 was not a significant overheating defect and blocks 6-10 were the most severe. What is clear is the positive correlation between the number of unexposed powder layers and the severity of the overheating defects.

Gronle et al. did not publish any porosity information for this sample, so it is impossible to link the sample's porosity with the observed overheating defects. We

would expect porosity to correlate with the overheating zones since higher surface temperatures can lower the energy density threshold for keyhole mode melting, as observed by Zheng et al. [28]. This means that keyhole mode melting is more likely to occur in overheating zones, even with constant laser parameters, resulting in trapped gas porosity [5, 29].

Gronle et al. focus on overheating defects in their study, so in our analysis, we use a one-tailed statistical test. A two-tailed test would reveal voxels significantly below the mean, likely indicating undermelting regions within the part.

3.1.2 Standard deviation response

The standard deviation of a voxel is used to indicate the melting process's stability in that region of space. Figure 11(a) and (b) show the raw and spatially correlated standard deviation values for the data in case study 1. Figure 11(c) shows the statistically significant data thresholded with a one-tailed statistical test with a significance level of 0.05. This analysis reveals that blocks 7-10 are the most significant outliers. Blocks 5 and 6 contain statistically significant voxels, but these only represent a small portion of each block. Blocks 1-4 are not considered to be statistically significant outliers in this analysis. The interpretation of this result is that in regions where overheating exists, the standard deviation of the photodiode response is also higher. It is not clear from this data if these two effects are independent or if the overheating causes the increased standard deviation.

3.1.3 Case study 1 summary

The results obtained from case study 1 show that the voxelisation framework is capable of identifying statistically significant overheating defects in a sample, with the mean being a more reliable indicator than the standard deviation. Our analysis of the mean identified blocks 3 to 10 as significant overheating zones and blocks 7 to 10 as being much more overheated compared to blocks 3 to 6. This is almost the same conclusion

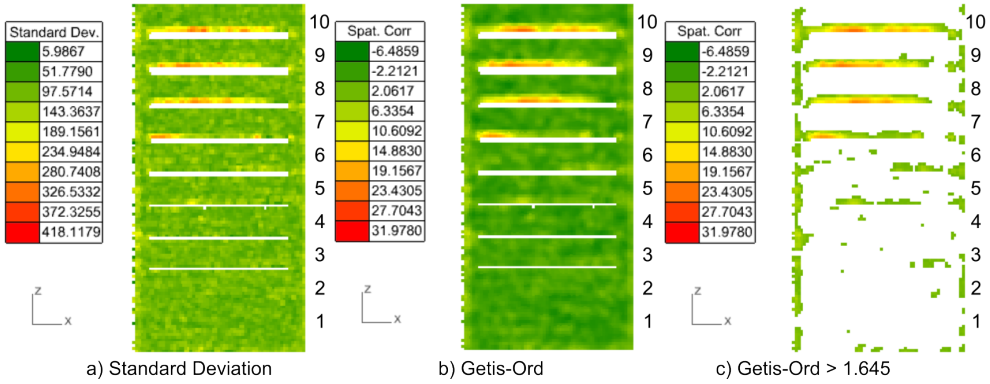


Fig. 11 Slices through the (a) raw standard deviation data, (b) the Getis-Ord G_i spatial correlation of the standard deviation data, and (c) the thresholded Getis-Ord data. The colour map is from green (low) to yellow (medium) and red (high) and is scaled to the maximum and minimum of each dataset. Image (c) only shows voxels with a Getis-Ord value greater than or equal to 1.645. (b) and (c) have the same colour scale. Data is voxelised using 150µm voxels.

as was made by Gronle et al. in their analysis. We also determined that the standard deviation is only useful in the most extreme overheating cases, with blocks 7 to 10 being statistically significant. Blocks 1 to 4 were not identified as outliers in the standard deviation analysis, and blocks 5 and 6 only have small areas of overheating voxels. In both cases, block 1 was not identified as a significant defect.

3.2 Case study 2

In the second case study, we investigate the effect of spatter-induced porosity on the photodiode system's response. During spattering, particulates of various sizes and morphologies are ejected from the melting zone, contaminating the exposed powder layer downwind of the gas flow. The laser then scans across this contaminated powder, disturbing the melting process from its nominal behaviour. This section will show how the proposed statistical approach can detect this change in melting behaviour and be used to correlate with porosity observed using high-resolution CT. We present the results of two statistical analyses, one using the mean to quantify the thermal effects and another using the standard deviation to quantify the melt pool stability.

In both cases, we vary the voxel size from 200µm to 500µm in increments of 100µm as a sensitivity analysis.

We only analyse the data collected from photodiode 1 in the Renishaw AM500Q, the sensor with the 700-1040nm range. We found that this sensor correlated better with the spatter defects observed during testing. We hypothesise that this enhanced correlation stems from the chosen processing parameters generating higher melt pool temperature and increased plasma production, to which photodiode 1 is more sensitive. The input data was downsampled to a 25kHz sampling rate for additional gains in storage space and execution speed, which had a negligible impact on the final result, validated during testing.

3.2.1 Mean response

For case study 2, the printed parts are cubes with uniform laser parameters, eliminating geometrically induced overheating regions such as overhangs. Therefore, any significant thermal differences would be due to the scanning strategy or spatter contamination.

Figure 12(a) shows the raw data for each voxel layer of the parts in case study 2 and Figure 12(b) shows the 3D view of the raw data from cube 6. Here, the mean data has a clear periodic pattern, repeating approximately every ten layers. It is not immediately clear why this pattern emerges, but one likely explanation is that the scan rotation strategy causes a systematic variation in the photodiode response depending on the direction of laser travel. There is also a clear positional dependence, with each part having a different vertical offset in Figure 12(a). Figure 12(c) plots the average layer data from each cube sample after column normalisation and spatial correlation calculation. The column normalisation and spatial correlation remove the spatial dependence from the data, allowing the trends to be compared side by side. Figure 12(d) shows this in 3D for cube 6. There is no obvious change from printing with contaminated vs clean powder in this view of the data.

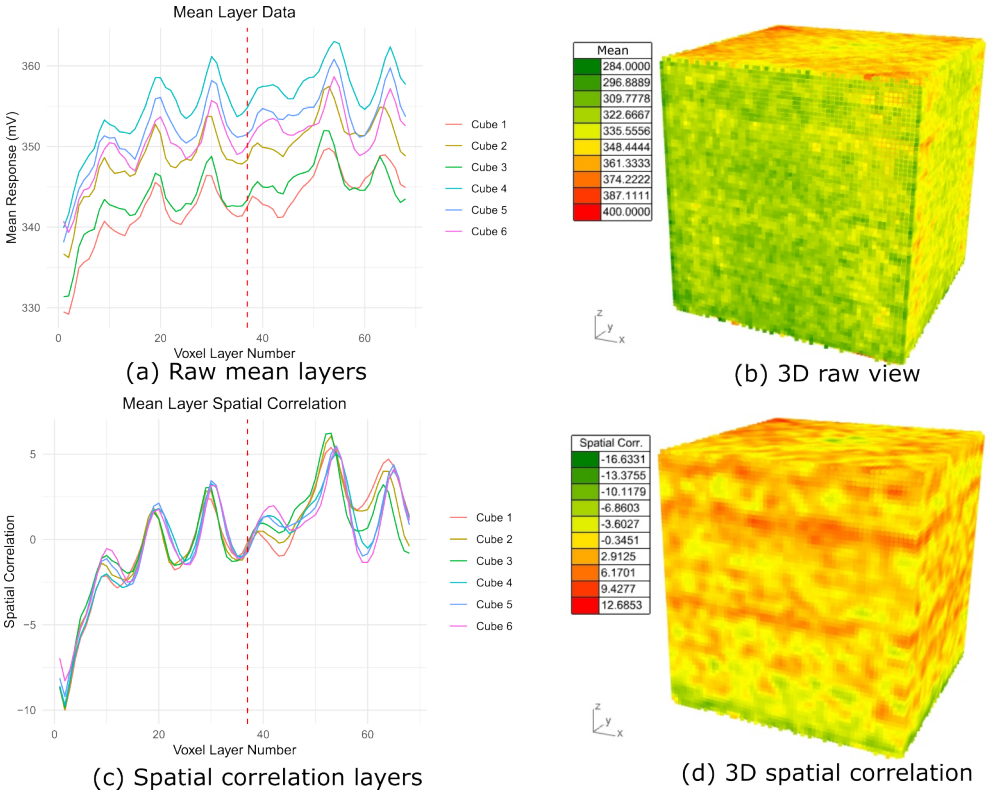


Fig. 12 (a) The raw mean data for each voxel layer in the spatter dataset and (b) a 3D view of the data for cube 6. (c) The voxel layer mean statistics for the parts in case study 2 after column normalisation and spatial correlation calculations, and (d) a perspective view of the normalised data for cube 6. The scan layer rotation effect can be seen. The vertical dashed line shows the point at which the spatter stops. Data is voxelised using 200µm voxels.

We compared this voxelised dataset to the voxelised porosity generated in Section 2.2.6 using the Spearman Rank correlation coefficient to see how well porosity can be predicted in 3D. The results for the mean structure for each voxel size are given in Table 1.

The data in Table 1 shows a consistent negative correlation with porosity at all voxel sizes, with the correlation being stronger at larger voxel sizes. The coefficients vary from around -0.37 for the 200µm voxel size up to -0.53 for the 500µm voxel size, a weak to moderate correlation. The larger voxel sizes likely have a stronger correlation because of more data averaging and reduced noise. Cube 3 has a consistently

Part	Voxel size (µm)			
	200	300	400	500
1	-0.48	-0.55	-0.57	-0.60
2	-0.42	-0.50	-0.52	-0.56
3	-0.28	-0.35	-0.37	-0.42
4	-0.39	-0.46	-0.50	-0.54
5	-0.33	-0.41	-0.44	-0.50
6	-0.31	-0.39	-0.44	-0.50

Table 1 Spearman's Rank correlation coefficients for the mean voxel data and the voxelised pore meshes for the six parts in case study 2 at four different voxel sizes.

weaker correlation than the other five, which may be caused by the presence of the optical artefact identified, shown in Figure 5. While this was removed using column normalisation, it is not clear if this has contributed to the weaker correlation or if there are other factors.

The analysis of the mean data aligns with the conclusions made by Gronle et al., where the photodiode response is used as a temperature proxy. The negative correlation suggests that lower melt pool temperatures cause the porosity. However, the fact that only a moderate correlation is seen even at the largest voxel size also indicates that the mean alone may not be sufficient to predict porosity reliably.

3.2.2 Standard Deviation Response

Figure 13 summarises the results for the six parts with the standard deviation as the reduction function, both with the raw data (a, b) and the column-normalised, spatially-correlated data (c, d). What is immediately apparent is that there are differences in magnitude between all six parts, with cubes 1, 2 and 3 demonstrating a lower standard deviation than cubes 4, 5 and 6. What is not clear is why the standard deviations of the 6 parts differ after the spattering has stopped. One possible explanation for this is position-dependent behaviour, where printing samples at different locations on the build plate changes how emitted light from the melt pool is recieved by the

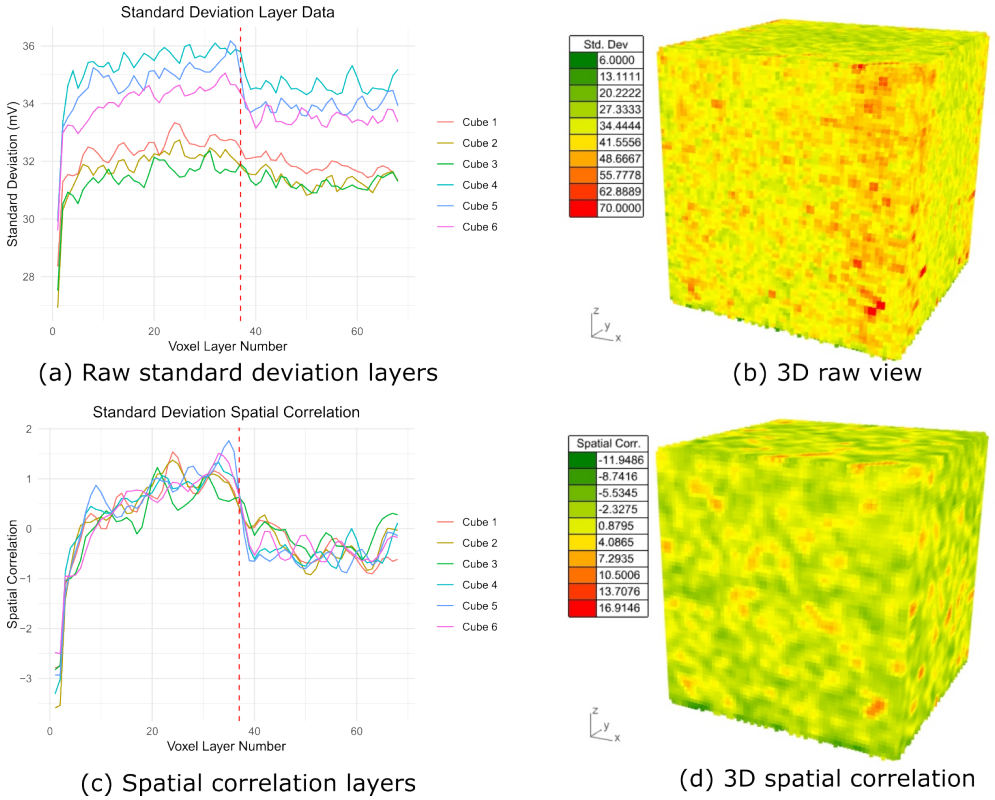


Fig. 13 (a) The raw voxel layer standard deviation statistics for all six parts, and (b) the raw data for cube 6 in 3D. Also shown are (c) the column normalised and spatially correlated data, with (d) the processed data from cube 6. The vertical dashed line shows the point at which the spatter stops. Data is voxelised using 200µm voxels.

optical system and photodiodes. Further experiments would be needed to quantify position-dependent effects to correct the raw data.

Figure 13(c) and (d) show the same data after column normalisation and spatial correlation. This process removes the absolute differences in magnitude between the curves so their trends can be directly compared. All six parts exhibit the same trend: a gradual increase towards a peak in standard deviation in the first half before sharply reducing and remaining constant for the second half. This is expected because the parts were printed with contaminated powder for the first half and clean powder for the second, and is similar to that seen in the porosity data from Figure 9.

The Spearman's Rank coefficient is calculated to determine how well the standard deviation predicts the porosity in 3D. It is calculated for four voxel sizes, shown in Table 2. These coefficients in Table 2 reveal several important insights. Firstly, the correlations are positive for all samples and voxel sizes but weak to moderate in magnitude. This suggests that while standard deviation indicates potential porosity, it may not be enough to detect it reliably. The second key insight is that as the voxel size increases, the correlation with porosity increases. This, similar to the mean results, is possibly due to greater averaging at larger voxel sizes, which removes noise from the input data. Thirdly, the same four cubes have the highest correlation coefficient at all voxel sizes: cubes 1, 2, 5 and 6. Cube 4 has a slightly lower correlation, and cube 3 has the lowest correlation. This might be caused by the optical artefact previously identified having a more significant effect than previously considered. While the severe artefact was removed using column normalisation, further preprocessing may be required to correct these.

part	Voxel size (µm)			
	200	300	400	500
1	0.17	0.28	0.34	0.42
2	0.19	0.30	0.38	0.45
3	0.10	0.17	0.20	0.23
4	0.17	0.29	0.33	0.40
5	0.20	0.31	0.37	0.43
6	0.17	0.27	0.31	0.35

Table 2 Spearman's Rank correlation coefficients for the standard deviation voxel data and the voxelised pore meshes for the six parts in case study 2 at four different voxel sizes.

3.2.3 Combined Indicator

Section 3.2.1 and Section 3.2.2 show that after voxelisation and processing, the mean, M and standard deviation S have moderate correlation with the porosity extracted

from the CT scans. To attempt to improve on these results, we devised a method to combine these two metrics into a single indicator, P, using a linear combination with a weight w:

$$P_i = -w * M_i + (1 - w) * S_i \tag{13}$$

P_i, M_i and S_i are the porosity indicator, the preprocessed mean and standard deviation data for the ith voxel, respectively. We found that a w of 0.5 yielded the best correlation with porosity across the different voxel sizes, and these results are presented below in Table 3.

Part	Voxel size (µm)			
	200	300	400	500
1	0.52	0.58	0.62	0.64
2	0.48	0.58	0.61	0.65
3	0.31	0.41	0.42	0.46
4	0.48	0.58	0.63	0.68
5	0.43	0.55	0.59	0.65
6	0.40	0.52	0.58	0.64

Table 3 Spearman's Rank correlation coefficients for the combined indicator metric P and the voxelised pore meshes for the six parts in case study 2 at four different voxel sizes. w is set to 0.5 for these results.

Table 3 shows that the proposed combination of the two variables improves the correlation between the IPM data and porosity across all parts and voxel sizes. The data exhibits similar trends to those seen when the two metrics were analysed separately: higher voxel sizes result in the strongest correlation, and cube 3 has the weakest correlation with the IPM data. A cross-section through the voxel structure calculated using the combined metric is displayed along with cross-sections from the mean, standard deviation, and porosity structures in Figure 14.

This improved correlation suggests that looking simultaneously at multiple process variables is the correct path forward in defect detection for LPBF. Combining the

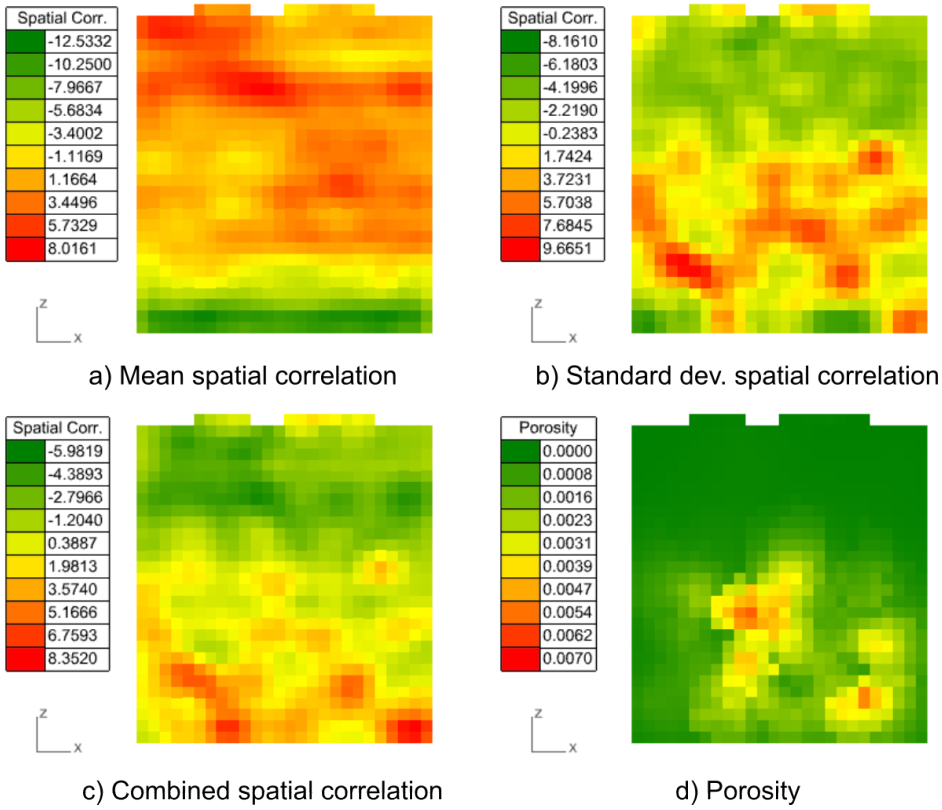


Fig. 14 Cross-sections of four different voxel representations of the data collected from cube 6: a) the spatially-correlated mean, b) spatially-correlated standard deviation, c) spatially correlated combined indicator, and d) the voxel representation of the porosity extracted by CT. Data is voxelised using 500µm voxels.

mean and standard deviation into a single indicator means the analysis can consider thermal and melt pool stability effects when predicting porosity, leading to improved correlation.

3.2.4 Case study 2 summary

For case study 2, we compared both the mean and standard deviation as reduction functions when constructing the voxel grid, as well as a linear combination of these. The analysis showed that the mean data negatively correlates with porosity according to Spearman's Rank, and the strength of this correlation depends on the voxel size,

with larger voxels having a stronger correlation. The standard deviation data had a positive correlation with the porosity and, in general, was weaker than the mean. The strength of this correlation was also dependent on the voxel size, with larger voxels showing a stronger correlation. The results show that the standard deviation is a useful indicator of powder bed imperfections, as was previously highlighted by Berumen et al. [21], and similar results were seen in the powder contamination work of Montazeri et al. [22]. It is likely that these powder bed imperfections change the melt pool dimensions, which can be observed as a change in the standard deviation of the photodiode response.

It is important to note that the column normalisation implemented in this paper is not a perfect solution to the problem of position-dependent behaviour. For a start, it requires that the part be completed and all *in-situ* build data collected before it can be calculated. This means that the data cannot be used to detect defects in real-time, and the analysis must be performed offline. Second, this normalisation process does not account for deviation caused by the geometry of the printed part. In this experiment, the parts had a uniform cross section, so this was not a factor. In future work, a more sophisticated normalisation process will be required to account for the geometry of the part, as well as positional dependence.

The final piece of the analysis combined these two statistical features into a single indicator. When using $w = 0.5$, this indicator showed an improved correlation with observed porosity for all parts and voxel sizes, highlighting that porosity has multiple contributing factors.

The effects of spatter contamination on part porosity are well documented in the literature. Doğu et al. [30] showed that adding 10wt% of spatter powder to virgin powder increased the porosity of Inconel 939 parts by 0.3% and surface roughness by 80.8%. They also commented that the presence of spatter particles is likely to disturb the melting process from its nominal behaviour, since in their work, samples

printed using spatter-contaminated powder exhibited larger grain sizes in the vertical direction. This is likely due to an increased melt pool temperature, which, given that the laser parameters are constant, would be caused by an increased absorptivity at the powder bed. This increase in absorptivity could be due to the presence of an oxide layer on the spatter particles, a phenomenon observed in other materials [31]. It is not clear what effect this would have on the *in-situ* data, given that the higher laser absorptivity would likely result in higher melt pool temperatures, and a higher photodiode response, whereas our results show a lower mean response in the contaminated regions.

Other works discuss how the chosen scanning strategy can impact the degree of spatter contamination. For example, Liu et al. [32] showed that changing the angle of the scan vector relative to the gas flow direction can cause more or less spatter to be ejected from the melt pool. In our work, we were not able to use a camera to observe the spatter, so we cannot comment on the spatter distribution for each layer in our sample. Linking spatter observed from a camera with *in-situ* photodiode data could make for an interesting future study, assuming the data storage and processing challenges that come with the use of camera systems have been overcome.

3.3 Comparative Analysis

In this section, we show the results for three methods used to detect signal outliers applied to the in-process monitoring data. This will show how the voxel-based approach is superior in visualisation and analysis capabilities.

We implemented several methods from the literature discussed in the introduction to provide a direct comparison.

3.3.1 Signal Outlier Methods

First, we present the signal outlier detection filters discussed by Harbig et al. [16] applied to an example cube from the spatter dataset. The authors used three methods

to detect outliers: absolute fluctuations (AF), dynamic, and short fluctuations (SF). The AF method imposes fixed upper and lower limits on the photodiode response, and any points that fall outside of those limits are classed as outliers. The SF algorithm calculates the moving average of the signal using the specified window size and then uses this as upper and lower bounds for thresholding the original signal. The SD algorithm calculates and thresholds the change in the signal from one sample to the next. Figure 15 shows the three algorithms applied to a layer of data from cube 4 of the spatter dataset, along with the porosity, calculated with a similar method to that shown in Section 2.2.6, calculating a porosity value for each sample point using KDE.

In this comparison, the upper and lower limits are set to 1.96 standard deviations, or a p-score of 0.05, on either side of the mean response for each layer, the same statistical limits used in the previous analysis in this manuscript, since [16] make no recommendation in their work. A section of data is shown in Figure 15(a). The figure shows the raw data sampled at 100kHz, with normal samples coloured in yellow, high outliers coloured in red, and low outliers coloured in blue. Pore meshes from the CT scan are coloured in pink. Only pores up to 0.5mm below the current layer are shown. The chosen layer is one where the laser was scanning spatter-contaminated powder. Figure 15(a) shows little to no correlation between the photodiode data and the porosity from the CT scan. We would expect the thresholded high and low sample points to align with the pores, but this is not the case.

Figure 15(b) shows the SF algorithm applied to the same section of in-process data using a moving average window size of 50. In our tests, the moving average signal was shifted by 1.96 standard deviations on either side of the original signal. A range of moving average window sizes from 2 to 150 were tested, but there was no clear link between the thresholded samples and the CT pores.

The final signal processing method attempted was the SD algorithm. In our tests, the mean and standard deviation of the sample-to-sample changes were calculated,

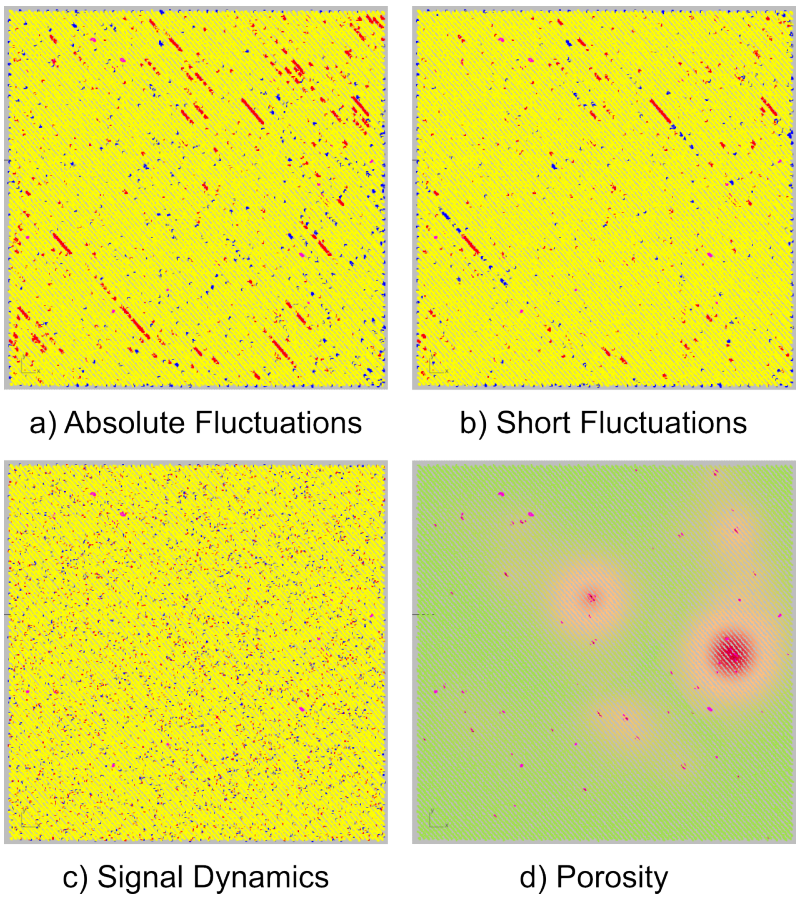


Fig. 15 Layer 50 from cube 4 of the spatter dataset, showing (a) the absolute fluctuations, (b) short fluctuations, and (c) signal dynamics algorithms from Harbig et al. [16], along with the porosity of each point, calculated using the KDE method introduced in Section 2.2.6.

and the same limit of ± 1.96 was used to threshold significant increases and decreases in the signal value. Figure 15(c) shows the results for the same section of in-process data as the previous two analyses. Once again, there is no correlation between the in-process data and the pores using the SD algorithm.

In summary, the three signal outlier methods discussed in this section were not able to correlate with 3D porosity for our dataset. The main complication is that the ground truth data available in the present work is different to that obtained in the studies from the literature. For example, in the work of Harbig et al. [16], they were

able to label balling defects using the height map of single tracks. Forien et al. [5] used radiographic imaging to determine exactly where along a single track a pore had formed, and the pyrometer signal could be overlayed. Since this only occurred in one dimension, it was easier to determine when a pore formed in the pyrometer signal. In the current work, the pore data comes from CT scans post-fabrication, so it is impossible to know when a particular pore is formed. As highlighted by Feng et al. [17] and Chen et al. [18], pores often form in layers below the current printing layer or can be healed by subsequent layers, making their detection more challenging. This means calculating the prediction accuracy for these other methods using our CT data is impossible. Nevertheless, the presented analysis provides a useful discussion about the difficulties in extending defect detection methods into 3D. One particularly complex part of the analysis with these methods is that they apply layer by layer, requiring manual inspection of every layer for every part. This would add significant time for little benefit for large datasets with many components and layers. In comparison, the 3D voxel method proposed in this manuscript provides a faster way to inspect the data but also correlates better with the ground truth pore data.

4 Conclusions

This work presents a novel framework for processing and analysing in-process monitoring data from laser powder bed fusion systems and correlating this data with ground truth porosity using an efficient voxel structure. The proposed framework offers three primary contributions:

- Spatial statistical analysis: the framework discretises the data in 3D space and provides statistical analysis of the discrete regions, revealing previously unseen patterns in the data. The Getis-Ord metric highlights statistically significant high and low clusters in the dataset in 3D space, quickly showing areas where more focused analysis should occur.

- Ground truth validation: we utilise KDE to transform porosity data from a high-resolution CT scanner into a voxel grid that is compatible in shape to the *in-situ* data, allowing for voxel-by-voxel comparision and correlation.
- Machine-Agnostic Data reduction and organisation: the voxel structure reduces the immense IPM dataset to a manageable size while enriching the space with local statistics, and providing an organised structure for querying results in 3D space. This format is machine agnostic, proven to work with data from two different machines, and can be extended to other sensors and machines.

The proposed method is conceptually simple and computationally efficient, making it suitable for large datasets and industrial applications. The choice of reduction function, such as the mean and standard deviation, allows the flexibility to analyse different aspects of the data which reveal patterns that correlate to ground truth porosity. This simultaneously addresses the challenges of defect detection as well as data standardisation.

The method offers immense flexibility and future development potential. Some of the primary areas of focus moving forward should be investigating other reduction functions that quantify other aspects of the data. For example, further investigation of the frequency response in localised regions may reveal additional defect formation mechanisms. This work has also revealed the necessity to understand and compensate for variation in the photodiode response as a result of printing location. This effect is separate from geometrically induced variations and requires a more sophisticated mapping between the printing location and the expected photodiode response.

The presented method represents a new way to analyse time-series data from in-process monitoring sensors, offering a practical solution to the challenges of defect detection and big data in AM. As the industry moves forward, such frameworks will minimise the time and labour costs associated with quality control, ensuring consistently high part quality.

Author Contributions.

Funding. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 956401.

Availability of data, code and material. The voxelisation library code is freely available on Github, (link removed for anonymity). The dataset for case study 1 is available at the IC labs page, https://www.ic.polimi.it/open-data-challenge/. The dataset for case study 2 has been made publicly available at the following link: (link removed for anonymity).

Declarations

Conflict of Interest. The authors declare that there is no conflict of interest regarding the publication of this article.

Financial Declarations.

Appendix A Preprocessing

A.1 Photodiode data filtering

As mentioned in Section 2.2.1, the photodiode signal data from both the Trumpf and Renishaw machines were filtered before voxelisation to remove data where the laser was not actively scanning the powder bed. For the Trumpf machine, data points with a photodiode value below 1000 were removed. Raw data from a layer is shown in

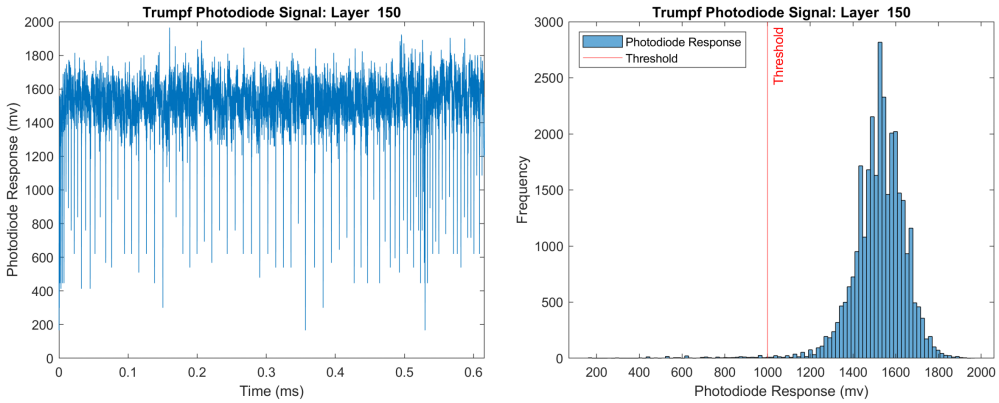


Fig. A1 Example data showing a) the raw photodiode data extracted from a layer of the Dataset Challenge test part as well as b) the histogram of the data. The vertical dashed line show the threshold values used to filter the data.

Figure A1(a), and the histogram of this layer is shown in Figure A1(b). The vertical dashed line shows the threshold value used to filter the data. This thresholding translates to a very minimal amount of data being removed from the dataset.

For the Renishaw machine, the laser modulation and power monitor channels were used to filter out samples where the laser was deactivated, or at a low power level, $<$ 30000 units. The power monitor channel records the power draw by the laser module, and the modulation channel is a boolean flag that records if the laser module is activated. The analysis is intended to focus on the "normal" steady-state melting, so removing low power samples also removes where the laser is ramping up or down the power level at the start and ends of each scan vector. The power monitor channel is more reliable for filtering the data because the separation between laser off, power ramping, and constant power is more defined. A section of the filtered data from the Renishaw machine is shown in Figure A2. The lower bound on the power monitor was chosen to exclude extremely low power levels, while maintaining all tracks in the dataset.

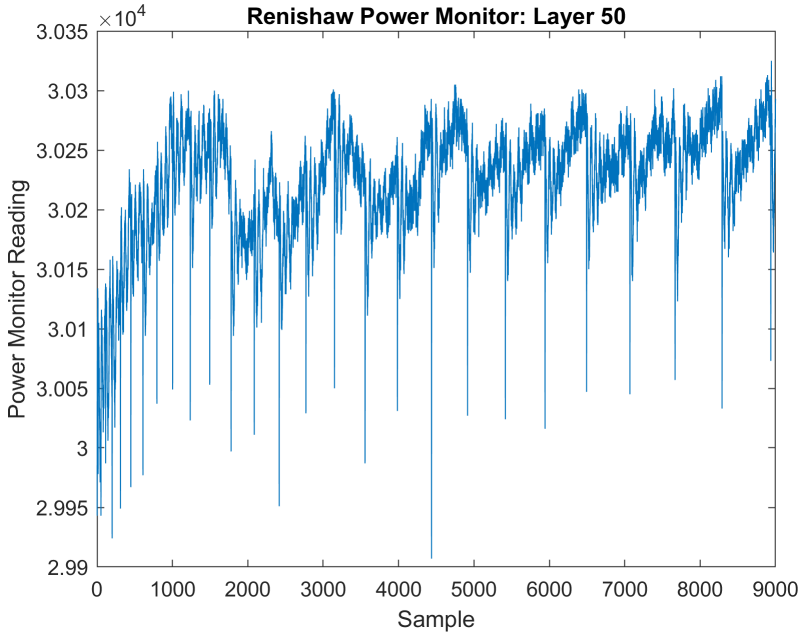


Fig. A2 The filtered power monitor channel from the Renishaw machine for an example layer

It is worth mentioning at this point that even though there is a constant demand laser power level, the power monitor shows some variation which is linked to the temperature of the laser module. At lower temperatures, the laser operates more efficiently, hence a lower power draw is required to reach the demand laser power. This is seen at the beginning of layer in Figure A2, where the first few scan vectors have a lower power draw compared to the later scan vectors.

References

[1] Riener, K., Albrecht, N., Ziegelmeier, S., Ramakrishnan, R., Haferkamp, L., Spierings, A.B., Leichtfried, G.J.: Influence of particle size distribution and morphology on the properties of the powder feedstock as well as of alsi10mg parts produced by laser powder bed fusion (lpbf). Additive Manufacturing **34**, 101286 (2020) https://doi.org/10.1016/j.addma.2020.101286

[2] Ziri, S., Hor, A., Mabru, C.: Combined effect of powder properties and process parameters on the density of 316l stainless steel obtained by laser powder bed fusion. International Journal of Advanced Manufacturing Technology **120**, 6187–6204 (2022) https://doi.org/10.1007/s00170-022-09160-w

[3] Scime, L., Beuth, J.: Anomaly detection and classification in a laser powder bed additive manufacturing process using a trained computer vision algorithm. Additive Manufacturing **19**, 114–126 (2018) https://doi.org/10.1016/j.addma.2017.11.009

[4] Zhang, B., Ziegert, J., Farahi, F., Davies, A.: In situ surface topography of laser powder bed fusion using fringe projection. Additive Manufacturing **12**, 100–107 (2016) https://doi.org/10.1016/j.addma.2016.08.001

[5] Forien, J.-B., Calta, N.P., DePond, P.J., Guss, G.M., Roehling, T.T., Matthews, M.J.: Detecting keyhole pore defects and monitoring process signatures during laser powder bed fusion: A correlation between in situ pyrometry and ex situ x-ray radiography. Additive Manufacturing **35**, 101336 (2020) https://doi.org/10.1016/j.addma.2020.101336

[6] Yang, L., Lo, L., Ding, S., Özel, T.: Monitoring and detection of meltpool and spatter regions in laser powder bed fusion of super alloy inconel 625. Progress in Additive Manufacturing **5**, 367–378 (2020) https://doi.org/10.1007/s40964-020-00140-8

[7] Ma, H., Mao, Z., Feng, W., Yang, Y., Hao, C., Zhou, J., Liu, S., Xie, H., Guo, G., Liu, Z.: Online in-situ monitoring of melt pool characteristic based on a single high-speed camera in laser powder bed fusion process. Applied Thermal Engineering **211**, 118515 (2022) https://doi.org/10.1016/j.applthermaleng.2022.118515

[8] Barrett, C., Carradero, C., Harris, E., Rogers, K., MacDonald, E., Conner, B.: Statistical analysis of spatter velocity with high-speed stereovision in laser powder bed fusion. Progress in Additive Manufacturing **4**, 423–430 (2019) https://doi.org/10.1007/s40964-019-00094-6

[9] Lough, C.S., Wang, X., Smith, C.C., Landers, R.G., Bristow, D.A., Drallmeier, J.A., Brown, B., Kinzel, E.C.: Correlation of swir imaging with lpbf 304l stainless steel part properties. Additive Manufacturing **35**, 101359 (2020) https://doi.org/10.1016/j.addma.2020.101359

[10] Gronle, M., Grasso, M., Granito, E., Schaal, F., Colosimo, B.M.: Open data for open science in industry 4.0: In-situ monitoring of quality in additive manufacturing. Journal of Quality Technology **55**, 253–265 (2023) https://doi.org/10.1080/00224065.2022.2106910

[11] Lapointe, S., Guss, G., Reese, Z., Strantza, M., Matthews, M.J., Druzgalski, C.L.: Photodiode-based machine learning for optimization of laser powder bed fusion parameters in complex geometries. Additive Manufacturing **53**, 102687 (2022) https://doi.org/10.1016/j.addma.2022.102687

[12] Jayasinghe, S., Paoletti, P., Sutcliffe, C., Dardis, J., Jones, N., Green, P.L.: Automatic quality assessments of laser powder bed fusion builds from photodiode sensor measurements. Progress in Additive Manufacturing **7**, 143–160 (2022) https://doi.org/10.1007/s40964-021-00219-w

[13] Power, J.J., Humphreys, O., Hartnett, M., Egan, D., Dowling, D.P.: Application of in situ process monitoring to optimise laser parameters during laser powder bed fusion printing of ti-6al-4v parts with overhang structures. International Journal of Advanced Manufacturing Technology **130**, 2297–2311 (2024) https://doi.org/10.1007/s00170-023-12794-z

[14] Cao, L., Hu, W., Zhou, T., Yu, L., Huang, X.: Monitoring of single-track melting states based on photodiode signal during laser powder bed fusion. Sensors **23**, 9793 (2023) https://doi.org/10.3390/s23249793

[15] Mao, Z., Feng, W., Ma, H., Yang, Y., Zhou, J., Liu, S., Liu, Y., Hu, P., Zhao, K., Xie, H., Guo, G., Liu, Z.: Continuous online flaws detection with photodiode signal and melt pool temperature based on deep learning in laser powder bed fusion. Optics & Laser Technology **158**, 108877 (2023) https://doi.org/10.1016/j.optlastec.2022.108877

[16] Harbig, J., Wenzler, D.L., Baehr, S., Kick, M.K., Merschroth, H., Wimmer, A., Weigold, M., Zaeh, M.F.: Methodology to determine melt pool anomalies in powder bed fusion of metals using a laser beam by means of process monitoring and sensor data fusion. Materials **15**, 1265 (2022) https://doi.org/10.3390/ma15031265

[17] Feng, S., Chen, Z., Bircher, B., Ji, Z., Nyborg, L., Bigot, S.: Predicting laser powder bed fusion defects through in-process monitoring data and machine learning. Materials & Design **222**, 111115 (2022) https://doi.org/10.1016/j.matdes.2022.111115

[18] Chen, X., Liao, W., Yue, J., Liu, T., Zhang, K., Li, J., Yang, T., Liu, H., Wei, H.: Unveiling the layer-wise dynamics of defect evolution in laser powder bed fusion: Insights for in-situ monitoring and control. Additive Manufacturing, 104414 (2024) https://doi.org/10.1016/j.addma.2024.104414

[19] Bay, C.: Strategic Guide: Additive Manufacturing In-Situ Monitoring Technology Readiness (2023). https://doi.org/10.1520/amcoe-guide-in-situ-tech-readiness . https://amcoe.org/in-situtechnologyreadiness/

[20] Young, Z.A., Guo, Q., Parab, N.D., Zhao, C., Qu, M., Escano, L.I., Fezzaa, K., Everhart, W., Sun, T., Chen, L.: Types of spatter and their features and formation mechanisms in laser powder bed fusion additive manufacturing process. Additive Manufacturing **36**, 101438 (2020) https://doi.org/10.1016/j.addma.2020.101438

[21] Berumen, S., Bechmann, F., Lindner, S., Kruth, J.P., Craeghs, T.: Quality control of laser- and powder bed-based additive manufacturing (am) technologies. In: Physics Procedia, vol. 5, pp. 617–622. Elsevier B.V., ??? (2010). https://doi.org/10.1016/j.phpro.2010.08.089

[22] Montazeri, M., Yavari, R., Rao, P., Boulware, P.: In-process monitoring of material cross-contamination defects in laser powder bed fusion. Journal of Manufacturing Science and Engineering, Transactions of the ASME **140** (2018) https://doi.org/10.1115/1.4040543

[23] Simonelli, M., Tuck, C., Aboulkhair, N.T., Maskery, I., Ashcroft, I., Wildman, R.D., Hague, R.: A study on the laser spatter and the oxidation reactions during selective laser melting of 316l stainless steel, al-si10-mg, and ti-6al-4v. Metallurgical and Materials Transactions A: Physical Metallurgy and Materials Science **46**, 3842–3851 (2015) https://doi.org/10.1007/s11661-015-2882-8

[24] Carter, L.N., Villapún, V.M., Andrews, J., Grandjean, T.R.B., Dardis, J., Cox, S.C.: Modelling process monitoring data in laser powder bed fusion: A pragmatic route to additive manufacturing quality assurance. Additive Manufacturing Letters **11** (2024) https://doi.org/10.1016/j.addlet.2024.100252

[25] Getis, A., Ord, J.K.: The analysis of spatial association by use of distance statistics. Geographical Analysis **24**, 189–206 (1992) https://doi.org/10.1111/j.1538-4632.1992.tb00261.x

[26] Ord, J.K., Getis, A.: Local spatial autocorrelation statistics: Distributional issues and an application. Geographical Analysis **27**, 286–306 (1995) https://doi.org/10.1111/j.1538-4632.1995.tb00912.x

[27] Guttman, A.: R-trees. ACM SIGMOD Record **14**, 47–57 (1984) https://doi.org/10.1145/971697.602266

[28] Zheng, H., Wang, Y., Xie, Y., Yang, S., Hou, R., Ge, Y., Lang, L., Gong, S., Li, H.: Observation of vapor plume behavior and process stability at single-track and multi-track levels in laser powder bed fusion regime. Metals **11** (2021) https://doi.org/10.3390/met11060937

[29] King, W.E., Barth, H.D., Castillo, V.M., Gallegos, G.F., Gibbs, J.W., Hahn, D.E., Kamath, C., Rubenchik, A.M.: Observation of keyhole-mode laser melting in laser powder-bed fusion additive manufacturing. Journal of Materials Processing Technology **214**, 2915–2925 (2014) https://doi.org/10.1016/j.jmatprotec.2014.06.005

[30] Doğu, M.N., Mussatto, A., Yalçın, M.A., Ozer, S., Davut, K., Obeidi, M.A., Kumar, A., Hudson, S., O'Neill, D., O'Connor, R., Gu, H., Brabazon, D.: A comprehensive characterization of the effect of spatter powder on in939 parts fabricated by laser powder bed fusion. Materials and Design **235** (2023) https://doi.org/10.1016/j.matdes.2023.112406

[31] Yan, J., Zhou, Y., Gu, R., Zhang, X., Quach, W.M., Yan, M.: A comprehensive study of steel powders (316l, h13, p20 and 18ni300) for their selective laser melting additive manufacturing. Metals **9** (2019) https://doi.org/10.3390/met9010086

[32] Liu, Z., Yang, Y., Wang, D., Chen, Z., Yan, W.: Correlation between the scan strategy, residing spatter distribution, and parts quality in laser powder bed

fusion. Materials and Design **234** (2023) https://doi.org/10.1016/j.matdes.2023.112317