Joint Special Interest Group meeting between euspen and ASPE
Advancing Precision in Additive Manufacturing
Inspire AG, St. Gallen, Switzerland 2021
www.euspen.eu

Exploratory research on the correlation of probability of detection and image quality during the tomographic characterisation of additive manufacturing defects

M. Sinico[1,2], S. Bellens[1,2,3], T. van Herreweghen[1], G.M. Probst[1,2], W. Dewulf[1]

[1]*Department of Mechanical Engineering, KU Leuven, 3001 Leuven, Belgium*
[2]*Member of Flanders Make - Core Lab MaPS, KU Leuven, 3001 Leuven, Belgium*
[3]*Materialise NV, 3001 Leuven, Belgium*

mirko.sinico@kuleuven.be

Abstract

This contribution reports on exploratory research to streamline the Probability of Detection (PoD) calculation for defect analysis with industrial X-ray Computed Tomography (XCT). Reliable, reproducible and universally applicable PoD information is becoming relevant for all non-destructive testing (NDT) methods, like XCT, where many acquisition parameters can affect the final inspection quality. Especially for complex shaped parts, where XCT is often the only suitable NDT inspection system, the a-priori determination of the inspection limit is paramount to confidently meet the required acceptance levels. In this work, three additively manufactured (AM) metal parts are employed as reference measurands to generate a set of PoD curves, using a binary response, for different XCT acquisition conditions. For this goal, an initial high resolution/quality XCT acquisition was considered as ground truth for measuring the size of the components' internal defects. Subsequently, the image quality of the different CT datasets was evaluated employing a set of quality metrics. The measured strong correlation between the PoD response and the image quality metrics suggests that an estimation of the PoD stemming directly from the quantification of CT image quality is viable and could shorten the XCT PoD assessment.

Computed Tomography, Additive Manufacturing, Probability of Detection, NDT, Image Quality, Porosity Analysis

1. Introduction

Probability of Detection (PoD) is used in various industry sectors to determine the capability of an inspection to detect flaws [1]. A PoD curve can be useful to value a-priori whether the inspection equipment is suitable for the inspection goal, and to estimate the likelihood of detecting a flaw in function of a characteristic parameter (e.g. size). However, PoD determination can be time consuming or even impractical for detection systems where a large number of variables contribute to the final quality of the measurement. For instance, the achievable scan quality of X-ray Computed Tomography (XCT) is influenced by numerous parameters, stemming from the CT system characteristics, the operator's choice of measurement strategy and scan settings, and the workpiece material, geometry and orientation [2]. Therefore, while calculating the PoD can be straightforward for one fixed CT inspection condition (if a ground truth is available for the measurand), this PoD will only be valid for the given set of parameters.

In the effort of gathering reproducible and universally applicable XCT PoD curves, the direct approach becomes the design and implementation of large experimental tests, where all the influencing factors are classified and correlated to the resulting achievable defect detection [3,4]. Another, more recent, approach is the automatic generation of arbitrarily large synthetic data set of realistic CT scans [5]. This data can be employed to train deep learning neural networks for e.g. improved defect segmentation [6], while at the same time it can be used to virtually estimate the PoD of a real CT machine as long as the realism of the simulation is validated. Additionally, the simulation model can be merged with the experimental results, through e.g. a Bayesian multilevel statistical model [7], with the goal of extending the measurement range while keeping XCT experimental costs low.

Nonetheless, all the aforementioned approaches rely on considerable amounts of data and analysis steps to finally derive a PoD for a specific XCT machine and parameter ranges. Yet, it is well known that defect detection performance can be just reduced as a direct result of the achieved XCT image quality after reconstruction, since image quality directly affects defects measurement [8]. Today, various image quality metrics (IQM) have been developed and implemented in different software packages [9], and can be employed to quantify the amount of noise, blur, perceptual quality, etc. Therefore, if an image can be correctly described with a selected combination of quality metrics, the same quantifiable metrics could be directly correlated to the achievable PoD. Undoubtedly, biases can be introduced, since the correlation will be highly dependent on the IQMs in use which might not describe completely the subtle variations between different XCT images. Nevertheless, achieving such goal would streamline a PoD estimation, enabling PoD assessment directly after XCT inspection. Furthermore, this estimate would be agnostic of the XCT equipment and capabilities, since it relies solely on the image metrics and the used defect detection algorithm/thresholding.

We therefore explore this correlation by experimentally determining PoD curves for three AM objects, employing a fixed 2D porosity analysis method, and subsequently studying how their PoD relate to different IQMs.

2. Materials and methods

Three reference metallic objects produced through the Laser Powder Bed Fusion (LPBF) technique were utilized in this study to assess the Probability of Detection of AM defects for different XCT acquisition parameters. The objects are presented in Figure 1, consisting of a manifold, a turbine and a cooler all fabricated in Ti-6Al-4V. The objects had been manufactured on a bygone LPBF system, and present various AM porosities such as gas, lack of fusion (LOF), and keyholing pores [10]. Each object was first inspected in a Nikon XTH 225ST machine to provide a high resolution, detailed scan of the object's overall dimensions and internal defects. Subsequently, the components were inspected again employing a Nikon XTH 450 machine and a diversified range of inspection parameters considering a Taguchi Orthogonal Array Design of Experiment (OA-DoE). Next, Probability of Detection curves were calculated using the hit and miss method, with a binary analysis that considers the true size of the defect (a) extrapolated from the high quality XCT inspection (i.e. our reference ground truth) against the signal response of the defect, from the inspections performed on the Nikon 450 system. Finally, for each XCT scan, multiple image quality metrics were computed through the MATLAB 2020a Image Processing Toolbox and correlated to the PoD values.

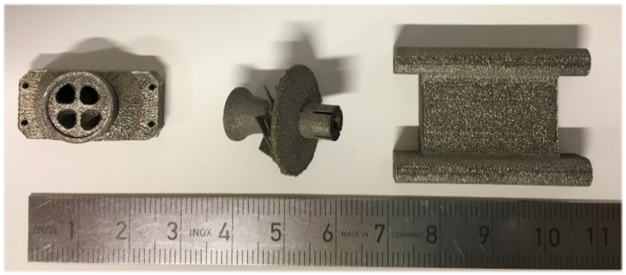

Figure 1. The three LPBF objects which are subjects of this work; from left to right an additively manufactured manifold, turbine and cooler.

2.1. XCT experiment design

Object's preparation prior to the tomographic characterisation consisted in fixing each component on one end of a ⌀= 16 mm Plexiglas tube, where also three reference ⌀= 2.5 mm stainless steel spheres were glued to the tube. The three spheres, together with the object itself, were used to align the different datasets after FDK reconstruction, in the VGSTUDIO MAX 3.4.3 software. The centre-to-centre distance between the spheres was moreover employed to avoid any voxel rescaling error between the distinct acquisitions. Following object's preparation, the ground truth scan on the 225 machine was performed according to the parameters listed in Table 1.

Table 1. XCT acquisition parameters for the 225 GT (ground truth) scan and the latter inspections on the 450 XCT machine.

Parameter	225 GT	450 L5F	450 S5	450 L2	450 S2F
Projections	3142	3142	3142	3142	3142
Magnification	11	5	5	2	2
Voxel size [µm]	18	40	40	100	100
Source spot size [µm]	<6	~80	~80	~80	~80
Exposure time [ms]	4000	2000	1415	2000	1415
Filter	Cu_1mm	Cu_1mm	-	-	Cu_1mm

The subsequent four separate acquisitions per component on the 450 machine were completed following a Taguchi Orthogonal Array design of experiment, where the variables filter, magnification and exposure time are studied at two levels. The X-ray source current was used as a free variable to adjust the brightness of the X-ray projections. The Taguchi Orthogonal Array design acquisition parameters are also summarized in Table 1. As expected, the inspections acquired with the 450 system resulted in lower quality datasets, considering both the lower spatial resolution achievable in the system, and the employed magnifications. A glimpse on the ground truth image quality compared to 450 data can be appreciated in Figure 2.

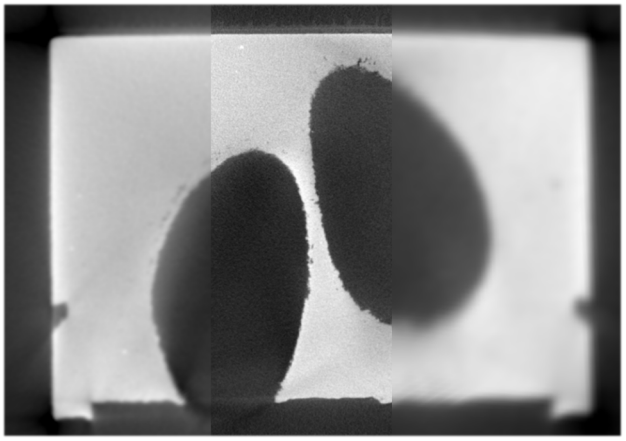

Figure 2. Reconstructed slices of the manifold object; from left to right, 450 S5 image, 225 GT image and 450 S2F image.

2.2. Probability of Detection calculation

In order to compute the PoD curves, every dataset was aligned and exported creating image stacks with a voxel resolution equal to the 225 GT scan (18 µm). Therefore, the image stacks from the 450 acquisitions were exported in interpolation mode, to compensate for the difference in voxel size. Since the subsequent porosity analysis was implemented on the exported 2D images, the interpolation of the 450 data was deemed necessary to accurately compare positions in the 225 GT data to corresponding positions in the 450 data. Next, for each dataset the list of defects was determined with MATLAB using first a global Otsu thresholding, for background removal, and second a local adaptive thresholding together with the function regionprops, for defects identification inside the material.

Subsequently, every defect list was binned considering the average of the defect's major and minor axis lengths ($\varnothing_{avg}$) with a bin size of 9 µm and, for each bin, a binary analysis was performed. The binary analysis considers the original defect found in the 225 GT scan, determines a cut-out region around the defect (adding a safety margin of 3 pixels) and computes a hit or miss response by looking at the same cut-out region in the 450 datasets. This is enough to finally calculate the PoD values; for a bin a with a total bin defect count n_a in the ground truth, all hits and misses are considered as:

$$PoD(\phi_{avg}) = \frac{\sum_{k=1}^{n_a} H_k}{n_a} \quad where \begin{cases} H = 1 \; if \; Hit \\ H = 0 \; if \; Miss \end{cases}$$

Figure 3 shows the obtained PoD curves for the cooler object.

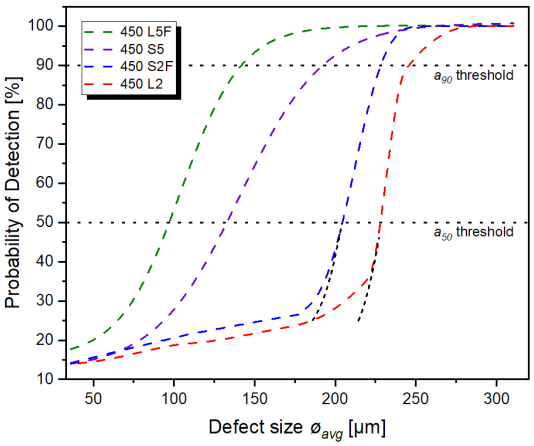

Figure 3. Calculated Probability of Detection curves for the cooler object. The curves are derived as sigmoidal fits over the raw data points.

2.3. Determination of the image quality metrics

From the 12 OA-DoE datasets, 10 % of the images in each image stack were selected randomly to compute the IQMs. All the metrics available in the MATLAB 2020a Image Processing Toolbox were determined [9]; in addition, a blur metric -which reflects what F. Crete et al. proposed in [11]- and the signal-to-noise ratio (SNR) were calculated. In this essay, only the mean-squared error (MSE), the structural similarity index (SSIM), the blind image spatial quality evaluator (BRISQUE) and the blur metric (BM) are discussed. MSE and SSIM are full reference (FR) metrics, whereas BRISQUE and BM are blind or no reference (NR). This is based on whether a reference image is required to calculate the score. For the FR metrics, the reference image was the corresponding image in the 225 GT dataset.

3. Results and discussion

3.1. Probability of Detection

Albeit the goal of this research is not the correlation of the PoD with the employed XCT settings, it is still valuable to briefly comment on the resulting insights. At first sight, looking for example at Figure 3 for the PoD of the cooler object, it is already clear how the magnification plays the major role on the final position of the PoD curve. The use of an X-ray filter and exposure time instead are less significant factors, with the exposure time being statistically non-significant. This of course is valid only for the DoE, objects and XCT machine in play. The statistical response is calculated with a linear regression model (R^2 of 93.1 %), where the PoD from all the objects and XCT measurement conditions is the response and the DoE parameters are the categorical predictors. Their interaction is as well accounted for, and a Pareto chart can be defined as in Figure 4.

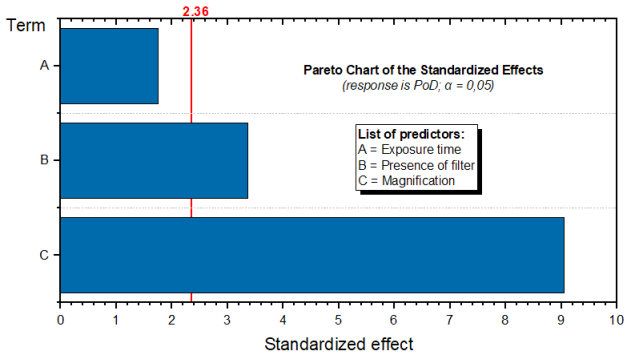

Figure 4. Pareto chart of standardized effects for the regression model calculated over the DoE results; PoD is the response.

It must be pointed out that the PoD curves for low magnification scans present an unexpected declining tail for the low values of PoD (below ~30 %), which is not consisted with the resolution limit and the conventional 'S' shape. The main reason for this behaviour can be traced back to the way the $\varnothing_{avg}$ is calculated for each bin size, considering the average of the defect's major and minor axis lengths. For non-spherical, elongated defects, such as cracks or lack of fusion, this method can introduce a bias in the PoD response, since the average of the defect's axes can be low but detectability is still achieved. As a consequence, the PoD is increased in the lower response region. Determining $\varnothing_{avg}$ with a different method, like the equivalent diameter, could fix this issue. Another less elegant approach would be performing the sigmoidal fit only in the centre region of the 'S' curve, obtaining more steep lower curves as the black dashed lines in Figure 3.

3.2. Image quality metrics

For exemplification, the IQMs for the manifold object are listed in the following Table 2.

Table 2. Image Quality Metrics for the manifold object.

Manifold datasets	Full reference		No reference	
	MSE·10^{-3}	SSIM	BRISQUE	BM
450 L5F	1.7	0.81	45.41	0.77
450 S5	7.4	0.79	45.90	0.80
450 S2F	4.1	0.76	58.13	0.93
450 L2	9.4	0.75	64.45	0.93

The general trend of the selected IQMs meets the expectations. The MSE values more the presence or not-presence of the filter, due the difference in between corresponding pixels of the DoE images and the GT images. This effect stems from the introduction of pronounced beam hardening artefact at the edge of the objects in the absence of filtering, and the possible shift in the global grey-values histogram. All the remaining IQMs follow the same trend of the PoD curves, as the scores worsen with decreased magnification or removal of filtering. It is worth stating that the BRISQUE IQM can be as well trained on sets of images with similar distortions as the images to be scored. This could further improve the reliability of the metric; however, for this exploratory research, the standard training LIVE IQA database was employed.

3.3. Correlation between IQM and PoD

Correlation or dependence can be studied by multiple means in statistics. A simple approach which can take into account multiple IQMs is, for instance, the implementation of a multi-criteria decision-making (MCDM) analysis where IQMs are first studied individually and then ranked according to their performance. A final MCDM score normally transfers this ranking to individual weights w_i for each criteria, and subdivides the criteria themselves in benefit or cost criteria.

The use of MCDM can be powerful when the effects to be measured are clearly defined, being for IQMs the single impact of e.g. noise, blur, contrast, luminance, etc. In the case of the four IQMs described in this paper, a combination of their IQMs can be made for each set as:

$$score_{set} = \sum_{i=0}^{4} \left(w_i \cdot IQM_{i,set} \right)$$

Specifically, in this following example, for the MSE metric we defined half the weight w_i (and hence importance) compared to the other IQMs, given the fact that MSE is skewed towards beam hardening ranking, a XCT artefact less troublesome for internal

defect detection. A linear min-max normalization is applied beforehand on the IQMs, to normalize their range between 0 and 1. The final MCDM score for each set can be directly correlated with the PoD response, using different *a* thresholds. In Figure 5, the dependence of the PoD a_{90} against the IQMs MCDM score can be appreciated.

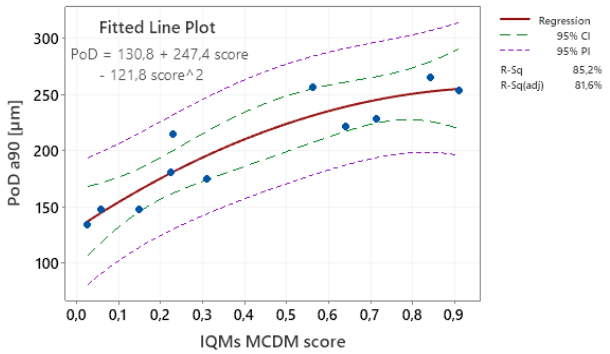

Figure 5. Regression analysis for PoD a_{90} against the IQMs MCDM score.

By all means, this simple MCDM analysis is far from being a usable tool for PoD estimation, since it relies on a small amount of datasets and utilizes only a small selection of IQMs for demonstration. However, the concept is ready to be extended to a more complex correlation analysis.

Besides, this MCMD approach still employs some full reference metrics, which might not be usable without a ground truth. For a more universally applicable correlation model, only NR IQMs should be selected. In the case of the two NR IQMs listed in section 3.2, the correlation with the PoD a_{90} can be visualized with a simple contour plot after a full quadratic response analysis, as in Figure 6.

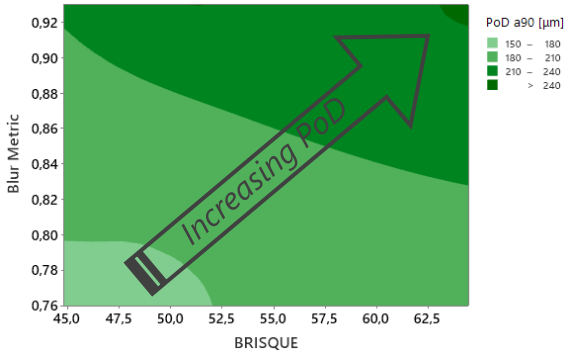

Figure 6. Contour plot of PoD against the BM and BRISQUE NF metrics; obtained R^2 for the full quadratic response analysis is 93.6 %.

Other means of correlation could rely only on a trained BRISQUE model, as described in section 3.2, even if, by employing a single metric, the risk is to lose the capability to later tune or rank your global score on individual image properties. Finally, deep learning algorithm could also be viable, as long as suitable training and validation data is accessible.

The most important take away from this exploratory research is that correlation between PoD and IQMs is possible and, being agnostic of XCT settings and/or machine type, should be applicable in every situation where the user has direct access to the raw image stack after reconstruction.

5. Conclusions

In industrial sectors like aeronautic or aerospace, the availability of Probability of Detection information for the employed NDT inspection systems is paramount. However, the determination of reliable PoD curves can be extremely time and resource intensive.

In this work, we proposed a novel approach where PoD estimation for industrial X-ray Computed Tomography is disjointed from the XCT settings and/or machine type in use. We suggest the possibility of deriving PoD information through the sole computation of image quality metrics on the reconstructed image stack after CT inspection. To model this correlation, an initial set of PoD curves was calculated for three metal AM objects, while at the same time IQMs were determined for each CT scanning condition. Both full reference (FR) and blind or no reference (NR) metrics were used for the scope. Afterwards, a simple correlation study was performed, connecting the PoD to a global MCDM score which took into account normalized IQMs and weight criteria.

The results of the study confirm the strong dependency of the PoD response with the selected IQMs. The results also suggest how an estimation of the PoD stemming directly from the quantification of blind-NF IQMs is viable, as long as the correlation is initially determined for a larger XCT dataset at different scanning conditions. This larger initial training data would be best determined through calibrated XCT artifacts like the one described in [12]. Finally, the resulting PoD response from IQM data, even if considered as an estimate, could still aid the initial screening of XCT conditions for a proper MIL-1823A [13] compliant PoD determination.

Acknowledgements

This research was funded in the framework of the VLAIO SIM-ICON project ALMA, grant agreement HBC.2018.0427. In addition, Simon Bellens acknowledges the Baekeland mandate grant HBC.2020.2280.

References

[1] European Commission - Joint Research Centre. *Institute for Energy. Probability of detection curves: statistical best practices*. LU: Publications Office; 2010.

[2] Carmignato S, Dewulf W, Leach R, editors. *Industrial X-Ray Computed Tomography*. Springer International Publishing; 2018.

[3] Amrhein S, Rauer M, Kaloudis M. Characterization of Computer Tomography Scanners Using the Probability of Detection Method. *J Nondestruct Eval* 2014;**33**:643–50.

[4] Kiefel D, Scius-Bertrand M, Stößel R, GmbH T. Computed Tomography of Additive Manufactured Components in Aeronautic Industry. *8th Conference on Industrial Computed Tomography (iCT) 2018*, 2018, p. 7.

[5] Fuchs P, Kröger T, Dierig T, Garbe CS. Generating Meaningful Synthetic Ground Truth for Pore Detection in Cast Aluminum Parts. *9th Conference on Industrial Computed Tomography (iCT) 2019*, 2019, p. 10.

[6] Bellens S, Vandewalle P, Dewulf W. Deep learning based porosity segmentation in X-ray CT measurements of polymer additive manufacturing parts. *Procedia CIRP* 2021;**96**:336–41.

[7] Kim FH, Pintar A, Obaton A-F, Fox J, Tarr J, Donmez A. Merging experiments and computer simulations in X-ray Computed Tomography probability of detection analysis of additive manufacturing flaws. *NDT & E International* 2021;**119**:102416.

[8] Pavan M, Craeghs T, Kruth J-P, Dewulf W. Investigating the influence of X-ray CT parameters on porosity measurement of laser sintered PA12 parts using a design-of-experiment approach. *Polymer Testing* 2018;**66**:203–12.

[9] Image Quality - MATLAB & Simulink - MathWorks Benelux n.d. https://nl.mathworks.com/help/images/image-quality.html?s_tid=CRUX_lftnav (accessed June 27, 2021).

[10] Sola A, Nouri A. Microstructural porosity in additive manufacturing: The formation and detection of pores in metal parts fabricated by powder bed fusion. *Jnl Adv Manuf & Process* 2019;**1**:10021.

[11] Crete F, Dolmiere T, Ladret P, Nicolas M. The blur effect: perception and estimation with a new no-reference perceptual blur metric. *Electronic Imaging 2007*, 2007, p. 64920I.

[12] Hermanek P, Carmignato S. Reference object for evaluating the accuracy of porosity measurements by X-ray computed tomography. *Case Studies in Nondestructive Testing and Evaluation* 2016;**6**:122–7.

[13] Virkkunen I, Koskinen T, Papula S, Sarikka T, Hänninen H. Comparison of â Versus a and Hit/Miss POD-Estimation Methods: A European Viewpoint. *J Nondestruct Eval*;**38**:89.