NDT&E International 162 (2026) 103769

Contents lists available at ScienceDirect

NDT and E International

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

Research Paper

Probability of detection based evaluation of XCT resolution limits and voxel sizing errors in defect detection

Miroslav Yosifov [a,b,*], Jonathan Glinz [a], Lorenzo Rusnati [c], Bernhard Plank [a], Reinhard Hubmann [d], Johann Kastner [a], Sascha Senck [a]

[a] CT Research Group, University of Applied Sciences Upper Austria, Stelzhamerstrasse 23, Wels, 4600, Austria
[b] imec-Vision Lab, Department of Physics, University of Antwerp, Universiteitsplein 1, Antwerpen, 2610, Belgium
[c] Department of Mechanical Engineering, Politecnico di Milano, via La Masa 1, Milano, 20156, Italy
[d] Fotec GmbH, Wiener Neustadt, Lower Austria, 2700, Austria

ARTICLE INFO

Keywords:
X-ray computed tomography
Probability of detection
Voxel size error
Additive manufacturing
Image resolution

ABSTRACT

This study investigates sizing error and the impact of scan resolution on defect detectability and porosity analysis in aluminum alloy samples, using a probability of detection framework. X-ray computed tomography (XCT) scans were generated at resolutions of 10 µm, 20 µm, 35 µm, and 70 µm. The highest-resolution XCT scan (10 µm) was established as the ground truth, serving as a reference to quantify detection performance, sizing errors, and segmentation accuracy (via dice coefficient and Jaccard index) across lower-resolution scans. Defects were segmented, matched across resolutions using centroid-based registration with a 60 µm tolerance radius, and evaluated via POD curves generated from probit regression on hit/miss data. Results show that detectability declines markedly with lower voxel size: the defect size achieving 90% POD at 95% confidence for equivalent diameter rises from 100 µm at 20 µm voxel size to 235 µm at 35 µm voxel size. Lower resolutions systematically overestimate defect size due to partial volume effects, inflate sphericity (from ≈ 0.80 to ≈ 0.87), and reduce segmentation overlap (Dice coefficient drops from > 0.99 to ≈ 0.97). Our study shows that optimizing scanning resolution is essential, as critical defects can be missed or mis-sized at lower resolutions, directly impacting downstream analyses such as finite element or fatigue simulations. These sizing errors propagate into such analyses, potentially leading to over- or underestimation of defect size and safety margins. The results and generated POD curves offer practical guidance for selecting scan parameters that effectively balance imaging fidelity and sizing error against processing time in nondestructive evaluation.

1. Introduction

X-ray Computed Tomography (XCT) has emerged as a cornerstone in the field of non-destructive testing (NDT), providing advanced capabilities for the three-dimensional (3D) visualization and analysis of internal material structures without the need for invasive physical sectioning [1]. Originating from medical imaging applications in the 1970s, XCT has since been adapted for industrial and materials science purposes, leveraging the differential attenuation of X-rays through materials to reconstruct detailed volumetric images [1,2]. This technique is particularly valuable for industrial materials, including metallic alloys such as aluminum and titanium, as well as lightweight materials such as carbon and glass fibers [3,4]. These materials are extensively used in high-stakes industries — such as aerospace, automotive, and energy — due to their favorable strength-to-weight ratios, corrosion resistance, and machinability [5].

The growing adoption of metal additive manufacturing (AM) across various industries can be attributed to its ability to enable complex designs, consolidate multiple parts into a single component, streamline production processes, lower component weight, and minimize material waste. In metal components, manufacturing processes like casting, forging, and AM often introduce defects such as voids, inclusions, cracks, and porosity [6,7]. The laser-powder bed fusion (PBF-LB/M) process most commonly produces the following types of defects: (i) pores caused by trapped gas, (ii) incomplete fusion areas, (iii) residual or trapped powder, (iv) keyhole-induced voids, and (v) surface roughness [8,9]. These imperfections can significantly compromise structural integrity, leading to premature failure under operational stresses [10]. Porosity, in particular, arises from gas entrapment during solidification in casting or from keyhole instabilities in AM processes, resulting in

* Corresponding author at: CT Research Group, University of Applied Sciences Upper Austria, Stelzhamerstrasse 23, Wels, 4600, Austria.
E-mail address: miroslav.yosifov@fh-wels.at (M. Yosifov).

https://doi.org/10.1016/j.ndteint.2026.103769
Received 23 January 2026; Received in revised form 4 April 2026; Accepted 5 May 2026
Available online 12 May 2026

reduced mechanical properties like fatigue strength and ductility. Traditional NDT methods, such as radiography or ultrasonic testing, offer limited insights into the 3D nature and distribution of these defects, often requiring destructive validation. XCT addresses these limitations by enabling quantitative assessment of defect morphology, size, and spatial distribution, thereby facilitating improved quality control and process optimization.

The efficacy of XCT in defect detection is heavily influenced by scan parameters, with resolution being a critical factor. Resolution, typically defined by voxel size, determines the smallest detectable feature and directly impacts the trade-off between imaging fidelity and practical constraints such as scan time, radiation dose, and computational resources [11]. Higher resolutions (smaller voxel sizes) improve the visibility of defects but escalate data volume and processing demands, potentially introducing artifacts like noise or beam hardening that can obscure true features [12]. Conversely, lower resolutions may overlook small defects or lead to oversegmentation, where adjacent voids are erroneously merged or noise is misidentified as defects. However, each manufacturing process has its own criteria and acceptable defect thresholds [13,14]. For instance, in the automotive sector, larger flaws such as 1 mm pores may be considered non-critical for high-volume cast components where cost-efficiency is paramount [15]. Conversely, in the aerospace industry, the tolerance for defects is significantly lower; flaws as small as 0.5 mm or even 0.2 mm are often classified as critical due to the risk of fatigue-driven catastrophic failure [14,16]. Recent studies in additive manufacturing emphasize that even sub-millimeter pores (<0.5 mm) can act as primary sites for crack initiation in high-performance alloys, necessitating higher resolution nondestructive evaluation (NDE) despite the increased processing time [17].

The influence of acquisition parameters on image quality and defect detectability has been systematically examined in additive manufacturing specimen. In particular, Senck et al. [18] proposed a simulation-based methodology for CT inspection of topology-optimized aerospace brackets, improving dimensional accuracy and quality assessment of additively manufactured components. Experimental designs employing fractional factorial methods have ranked parameters like tube voltage, current, and voxel size in terms of their impact on noise levels and image quality, showing that optimal settings can improve detection confidence limits for defects [11]. Rapid XCT characterization techniques using sparse projections and deep learning-based reconstruction have been proposed to reduce scan times while maintaining defect detection rates above 78% for flaws exceeding 50 μm, leveraging CAD models to mitigate artifacts [19].

The ability of an NDE method or system to identify defects can be expressed quantitatively using the probability of detection (POD) [20–22]. POD analysis has become an essential tool for evaluating the performance of defect detection techniques applied to XCT data [23]. POD offers a numerical measure of the likelihood that a detection algorithm will accurately identify and characterize flaws of specific sizes, thus aiding in the assessment of imaging systems, processes, and detection methodologies [24]. Specifically, POD frameworks applied to XCT have evolved from simulation-only approaches toward hybrid experimental–numerical validations that better reflect industrial conditions. Yosifov et al. introduced a superimposition method for inserting synthetic defects into real XCT volumes, revealing that realistic background noise and artifacts increase detection limits by factors of 2.5–3.3× compared with purely simulated data—directly supporting the superiority of the present experimental ground-truth methodology over earlier simulation-based POD studies [25]. Ziabari et al. combined CAD-guided deep-learning MBIR reconstruction with sparse projections and achieved $> 78\%$ detection of flaws $> 50\,\mu m$ and 100% detection of flaws $> 90\,\mu m$, even at reduced scan times, underscoring the potential of DL-assisted reconstruction to mitigate resolution constraints [26]. For example, Lee et al. presented a model-assisted POD framework that provides more reliable quantification of inspection performance. The reliability of defect detection in XCT is increasingly defined by the transition from traditional binary models to more sophisticated probabilistic approaches. The shift from traditional “hit/miss” models to sophisticated signal-response ($\hat{a}$ vs. a) models allows for a more nuanced understanding of defect sensitivity in complex geometries [27]. While binary hit/miss data only records the presence or absence of a defect, the $\hat{a}$ vs. a approach correlates the actual physical size of the defect (a) with the recorded signal response ($\hat{a}$), such as the gray-value contrast or reconstructed diameter in XCT. By leveraging Model-Assisted Probability of Detection (MAPOD) and digital twins, researchers can now simulate thousands of inspection scenarios, effectively bridging the gap between theoretical resolution and practical detection limits [28,29].

The observed detectability limits are in good agreement with previously reported studies on XCT resolution effects. Sietins et al. demonstrated that reliable detection of pores typically requires a defect diameter of at least 3–5 voxels, depending on noise level and segmentation strategy [30]. In the present study, the $a_{90/95}$ thresholds correspond to approximately 5 voxels at $20\,\mu m$ resolution and 6–7 voxels at $35\,\mu m$ resolution, confirming these empirical observations. Similarly, Galvez-Hernandez et al. [31] demonstrated that coarser scan resolutions lead to segmentation degradation and systematic overestimation of porosity due to pronounced partial volume effects, particularly at voxel sizes above 25 μm. Comparable observations were reported by Cooper et al. who showed that increasing voxel size significantly biases 3D micro-CT-based pore quantification, with partial volume effects becoming the dominant source of error as voxel dimensions approach the characteristic pore size [32].

Our study systematically examines the influence of XCT scan resolution on defect detectability and porosity characterization in aluminum components, with a particular focus on quantitative performance assessment using POD analysis. All other scanning parameters were optimized empirically and through XCT simulations [18]. XCT datasets were acquired at voxel sizes of $10\,\mu m$, $20\,\mu m$, $35\,\mu m$, and $70\,\mu m$, enabling a controlled evaluation of resolution-dependent effects. Defect characteristics — including volume, equivalent diameter, sphericity, and spatial distribution — were extracted and analyzed across all resolutions, alongside global porosity metrics. By correlating these morphological descriptors with POD outcomes, this work provides insight into how resolution-driven imaging limitations affect sizing error, defect detection reliability and establishes practical guidelines for selecting XCT acquisition parameters that balance inspection fidelity with efficiency in industrial NDE. While previous studies have investigated the influence of XCT resolution on defect detectability and sizing accuracy, most works rely on qualitative comparisons or deterministic thresholds, without explicitly quantifying detection reliability. In contrast, the present study introduces a POD-based framework to systematically evaluate the relationship and trade-off between voxel size, defect detectability, and sizing error using a high-resolution XCT dataset as reference. In contrast to previous studies, this work:

1. quantifies detectability using statistically defined metrics on real CT data rather than X-ray simulations (e.g., POD),
2. directly relates voxel size to both detection probability and image similarity across resolutions, and
3. provides application-oriented guidelines for selecting the optimal voxel size based on defect criticality.

Understanding the relationship between minimum detectability and voxel size can help develop strategies to reduce inspection time and effort while improving the overall accuracy of XCT-based critical defect detection. The findings of this study may inform future improvements in POD methodology and guide industrial XCT inspection practices.

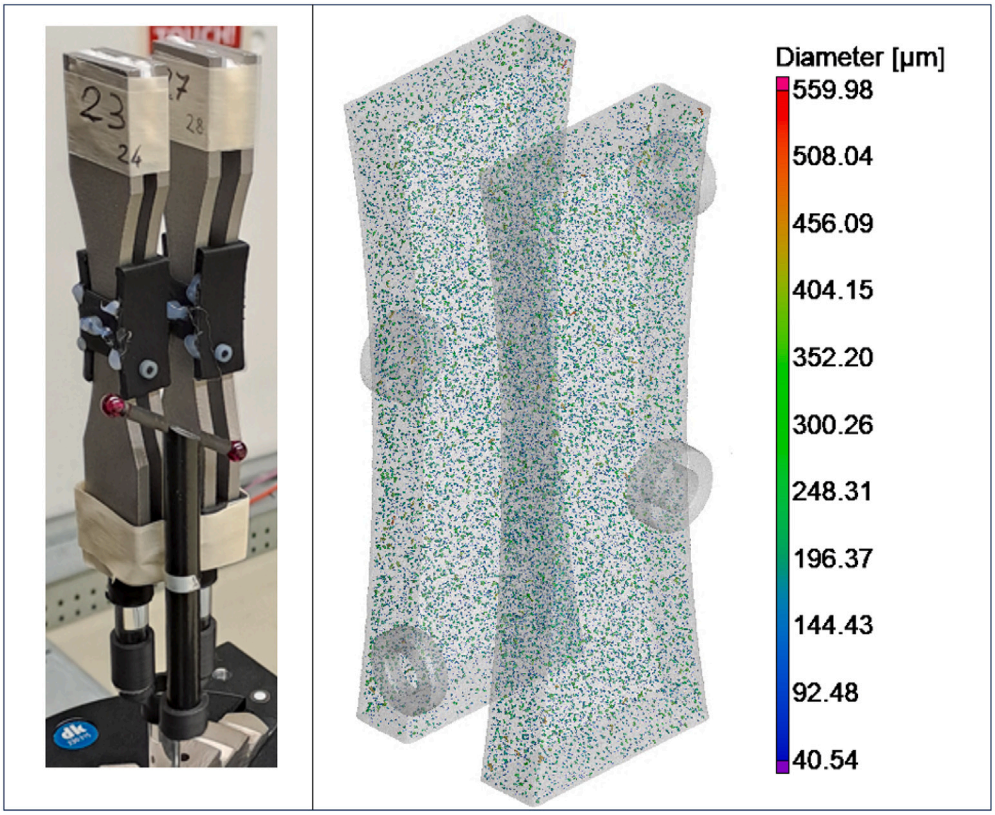


Fig. 1. Four physical test specimens mounted in a fixture for XCT scanning (left), and the resulting composite 3D model visualizing their internal porosity (right).

2. Materials and methods

In this section, we describe the methodologies used, including sample preparation, Micro-CT scanning, data analysis, probability of defect detection, image similarity assessment, and the overall processing pipeline.

2.1. Sample preparation and micro-CT scans

In this study, a series of continuous radius flat fatigue specimens labeled S20 to S30 was printed on a selective laser melting system (EOS M400). Each specimen was manufactured from an aluminum alloy and designed in a dog-bone geometry following standard tensile specimen dimensions. The specimen dimensions were as follows: length 13 cm, width ranging from 1 cm to 2.4 cm, and thickness of 0.4 cm. Fig. 1 shows a quality inspection setup where four physical specimens are prepared for a XCT scan to analyze their internal quality and porosity levels. The 3D model on the right presents a combined visualization from two specimen, mapping the internal pores and color-coding them by diameter to provide a comprehensive analysis of defect size and distribution. Additionally, four ring-shaped markers were added to each specimen pair to improve the registration and alignment process. Then, all specimens were scanned using two different XCT systems at four voxel resolutions: 10, 20, 35, and 70 µm. These resolutions were selected based on the typical scanning settings used for real components [33]. High-resolution scans at 10 µm were performed using a Phoenix Nanotom 180NF system, while the remaining scans were carried out with a ProCon 240 ALPHA XCT system. To reduce scanning time, the latter measurements were performed in batches of four specimens. To achieve higher resolution and establish a reliable ground truth, the Nanotom 180NF was employed. For comparison of detectability, all measurements were performed on the ProCon 240 Alpha at voxel sizes of 20, 35, and 70 µm to ensure a fair evaluation. Although each device has distinct hardware and software characteristics that may influence image quality and defect visibility, both CT scanners were carefully calibrated, and all measurements followed standardized protocols to minimize systematic differences. Minor instrument-specific effects may still exist, but they are negligible and do not significantly affect the results. Any remaining variability was further mitigated by using calibrated scans and applying consistent software algorithms across all data.

Table 1 summarizes the XCT scanning parameters, porosity evaluation settings, and scan details for the aluminum fatigue specimens. It includes the voxel size, tube voltage and current, detector integration time, number of averaged projections, total scan time, number of projections, and threshold values used for porosity segmentation. These settings were selected to balance image quality, resolution, and acquisition time across different scan conditions. The applied parameters were 150 kV tube voltage, 150 µA current, 1500 projections, an integration time of 600 ms, and a 0.1 mm Cu pre-filter for 10 µm. Additionally, specimens were scanned at lower resolutions of 20 µm, 35 µm, and 70 µm using a ProCon 240 ALPHA XCT system at 220 kV with tube currents of 100 , 170 , and 330 µA, integration times of 1700 ms, 1400 ms, and 600 ms, numbers of projections of 1260, 1040, and 780, and threshold values of 30%, 30%, and 20%, respectively.

The X-ray tube current and detector integration time were selected to ensure sufficient signal-to-noise ratio and contrast for reliable defect detection under low-contrast conditions. Due to the limited attenuation differences between pores and the surrounding material, higher photon statistics were required to reduce noise and enhance feature visibility. The final parameters were determined through preliminary trials, balancing image quality, acquisition time, and avoidance of detector saturation. Current and integration time were adjusted empirically to achieve clear images with acceptable noise while keeping scan duration as short as possible.

Table 1
XCT scanning parameters, porosity evaluation settings, and scan details for aluminum fatigue specimens.

XCT System	Voxel size [μm]	Voltage [kV]	Current [μA]	Int. Time [ms]	Avg	Scan time [min]	Projections	Threshold [%]
Nanotom 180NF	10	150	150	600	5	75	1500	44
ProCon 240 ALPHA	20	220	100	1700	1	36	1260	30
ProCon 240 ALPHA	35	220	170	1400	1	24	1040	30
ProCon 240 ALPHA	70	220	330	600	1	8	780	20

2.2. Data analysis

Porosity analysis was evaluated from the reconstructed, unfiltered volume data using the porosity/inclusion analysis module in VGStudioMax 2024.1 (Volume Graphics). The EasyPore module was applied with a relative threshold of 44% for the 10 μm scans. For the 20 μm and 35 μm scans, a threshold of 30% was used, while for the 70 μm scans, a threshold of 20% was applied. No image filters were used, and the minimum pore volume was defined as 27 voxels [34,35].

Thresholds were determined through a combination of visual inspection and analysis of image contrast. In EasyPore, void detection is controlled by the "Contrast" parameter, which defines the minimum gray-value difference between a void and the surrounding material [36]. For instance, a contrast value of 20% indicates that a voxel is classified as part of a void only if it is at least 20% darker than the neighboring material. Threshold optimization is performed interactively, with real-time visualization of detected pores allowing adjustments for under- or over-segmentation. If too few defects are detected or small pores are missed, the contrast value is decreased, whereas if large artifacts or excessive indications appear, the value is increased. The optimal threshold is chosen to balance sensitivity and specificity, typically corresponding to the transition region between material and void peaks in the gray-value histogram.

The following metrics were extracted from the reconstructed volumes: pore volume V, diameter D, equivalent diameter D_{eq}, and voxel number of pores. The equivalent diameter, which corresponds to the diameter of a sphere with the same volume as the pore, is defined as

$$D_{\mathrm{eq}} = 2\left(\frac{3V}{4\pi}\right)^{1/3} \tag{1}$$

2.3. Probability of detection

The probability of detection (POD) of porosity defects was evaluated using a probit regression model [37,38]. Defect size was characterized by one of three metrics: the diameter D, the equivalent diameter D_{eq}, or the pore volume V.

The binary detection outcome, Y_i, was defined as:

$$Y_i = POD \begin{cases} 1, & \text{if defect } i \text{ is detected,} \\ 0, & \text{if defect } i \text{ is not detected.} \end{cases} \tag{2}$$

The Probit model relates the detection probability to the chosen defect size metric via:

$$\Phi^{-1}\left[\mathrm{POD}(S)\right] = \beta_0 + \beta_1 S, \tag{3}$$

where S represents the defect size (D, D_{eq}, or V), Φ^{-1} is the inverse standard normal cumulative distribution function, and β_0 and β_1 are regression coefficients estimated from the dataset. The model provides the estimated detection probability for any defect size and allows computation of characteristic sizes at specified confidence levels, such as a_{90}, the defect size detected with 90% probability. Maximum likelihood estimation was used to fit the model, and the fit quality was assessed using confidence intervals and visual inspection of the POD curves. This approach enables quantitative evaluation of defect detectability across different size metrics and XCT voxel resolutions.

In conventional POD studies, defect detectability is often assessed using XCT simulations to establish a controlled ground truth. To establish correspondence between high- and low-resolution scans, we identified matching defects. For each defect, we defined an imaginary square with a radius of 60 μm, referred to as the tolerance value. Additionally, we ensured that no overlaps occurred, so that defects were not matched more than once. The tolerance value was determined such that overlaps between non-matching defects were excluded.

2.4. Image similarity

To quantitatively assess the similarity between reconstructed XCT volumes, two approaches were employed: comparison of 3D image intensities and evaluation of segmentation overlap.

3D Image Similarity: The structural similarity between volumes was quantified using the multi-scale structural similarity index (MS-SSIM) [39,40] and peak signal-to-noise ratio (PSNR) [41,42]. These metrics provide complementary information: MS-SSIM captures perceptual structural differences, while PSNR measures absolute voxel intensity deviations.

Segmentation Similarity: The agreement between segmented pores in different resolutions was evaluated using the Jaccard index and the Dice coefficient. For two sets of segmented voxels, A and B, the Jaccard index is defined as

$$J = \frac{|A \cap B|}{|A \cup B|}, \tag{4}$$

and the Dice coefficient as

$$\mathrm{Dice} = \frac{2|A \cap B|}{|A| + |B|} \tag{5}$$

For all comparisons, the high-resolution 10 μm scans were used as the ground truth XCT volumes. Lower-resolution scans were registered to this reference to evaluate image fidelity and segmentation accuracy across different voxel sizes.

2.5. Processing pipeline

Fig. 2 illustrates the proposed processing pipeline for defect detection and validation using dual-resolution XCT scans. The workflow employs a high-resolution voxel size of 10 μm as ground truth for accurate defect segmentation, and a low-resolution XCT scan for efficient large-scale screening, enabling a balance between accuracy and computational efficiency. The pipeline begins with *volume registration and alignment* of the high-resolution (HR) and low-resolution (LR) scans to ensure spatial correspondence between the datasets. For alignment, a best-fit algorithm was applied using the fiducial (ring) markers in VGStudio Max. Following alignment, *porosity analysis and defect segmentation* is performed exclusively on both the HR and LR scan. This step identifies potential defects (e.g., pores or voids) using image processing techniques such as thresholding, morphological operations, and possibly machine learning-based segmentation to delineate defect boundaries accurately. For each segmented defect in the HR scan, the *defect centroid* is computed as a representative point (e.g., the geometric center of mass). This centroid is then *mapped* to the corresponding location in the LR scan using the established registration transformation. To validate the defect's presence in the coarser LR scan, a localized *search within a tolerance radius* R_{tol} is conducted around the mapped centroid. This radius accounts for potential misalignment errors or resolution-induced discrepancies between the scans. The search evaluates whether

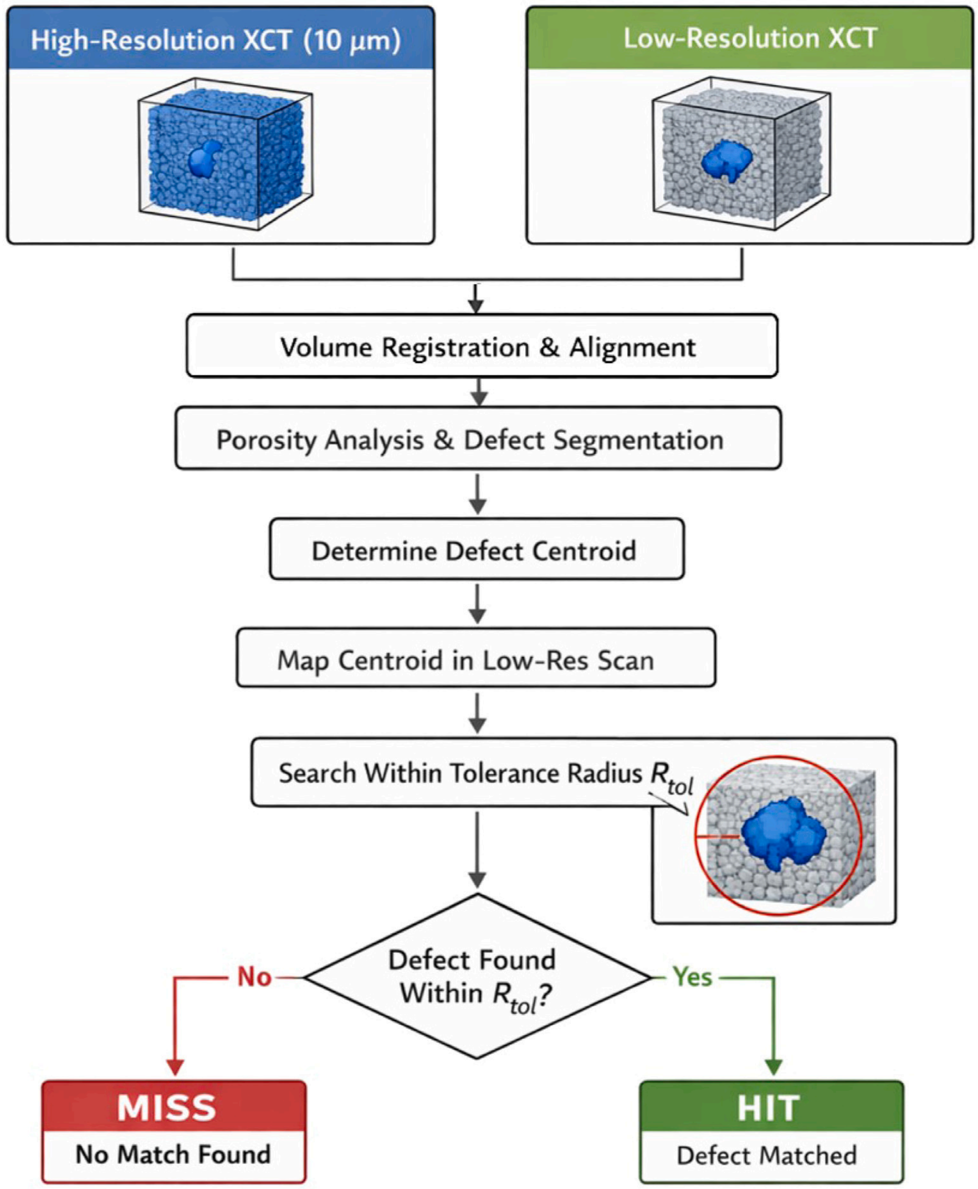


Fig. 2. The proposed processing pipeline for defect detection and validation using dual-resolution X-ray computed tomography (XCT) scans. The workflow integrates high-resolution (HR) scanning for precise defect segmentation with low-resolution (LR) scanning for efficient validation, ensuring spatial alignment and tolerance-based matching.

a defect-like anomaly (e.g., a low-density region indicative of porosity) exists within R_{tol}.

A tolerance radius of 60 μm was employed in this study to establish correspondences between defects across datasets with differing spatial resolutions (10 μm vs. 20 μm and 10 μm vs. 35 μm). This value was not arbitrarily selected but determined through an iterative optimization procedure aimed at ensuring robust one-to-one defect matching. Specifically, the tolerance radius was systematically varied, and the resulting defect correspondences were evaluated based on matching consistency. At small tolerance values, true corresponding defects were frequently missed due to slight spatial deviations caused by resolution differences, registration inaccuracies, and segmentation variability. Conversely, larger tolerance values increased the likelihood of multiple defects being incorrectly assigned to a single reference defect, leading to ambiguous (one-to-many) matches. The optimal tolerance radius was therefore defined as the smallest value that maximized unique one-to-one correspondences, i.e., each defect in one dataset was matched to at most one defect in the reference dataset. This was achieved through an iterative process in which the tolerance was gradually increased until duplicate matches were minimized while maintaining a high number of valid correspondences. The selected value of 60 μm reflects a balance between spatial uncertainty (arising from voxel size, partial volume effects, and segmentation variability) and matching specificity. This choice directly impacts matching accuracy: a smaller tolerance would underestimate detection performance by missing valid matches, whereas a larger tolerance would artificially inflate performance by introducing false correspondences. Therefore, the adopted value represents a conservative and physically meaningful compromise.

In the final step, each reference defect is evaluated within the defined tolerance radius. If a segmented defect is found within this tolerance region in the low-resolution dataset, the defect is classified as a hit; otherwise, it is classified as a miss. This hit–miss classification forms the basis for subsequent POD evaluation.

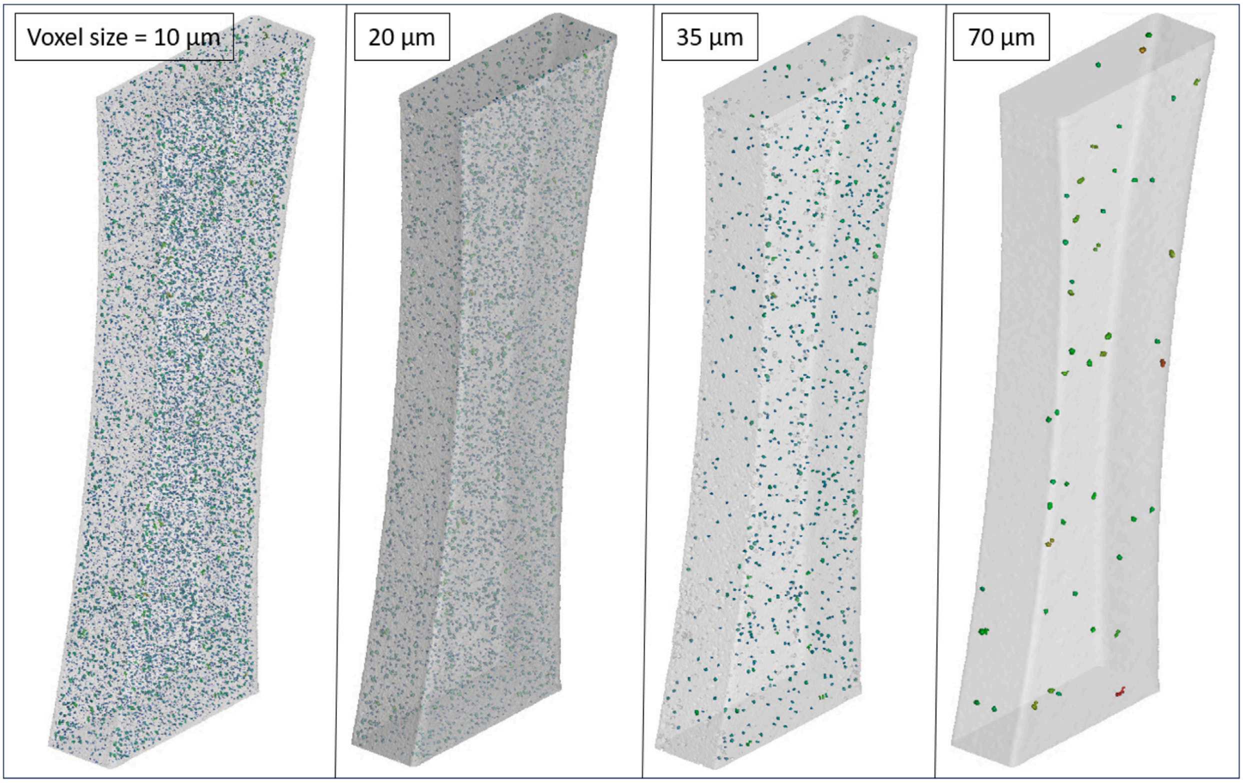


Fig. 3. Each subpanel shows a three-dimensional rendering of the reconstructed XCT volumes (sample S26) at voxel sizes of 10, 20, 35, and 70 µm, with automatically detected pores highlighted to illustrate their spatial distribution within the material.

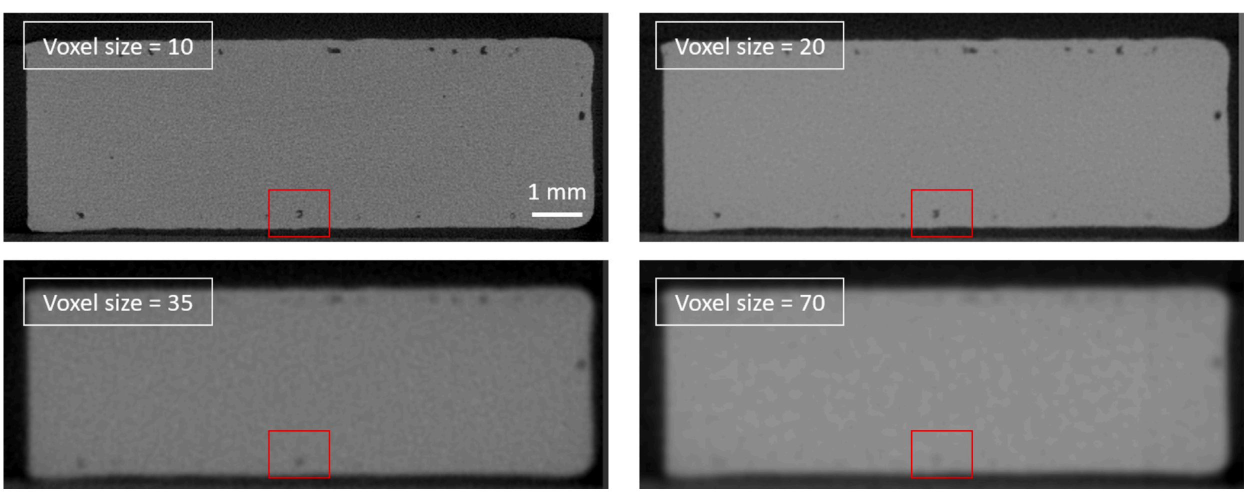


Fig. 4. Image slices from 3D reconstructed data showing porosity defects in an aluminum specimen at different voxel sizes. From left to right: 10 µm, 20 µm, 35 µm, and 70 µm. The number and clarity of detected defects decrease with lower voxel resolution. The red square highlights a representative internal pore.

3. Results

This section presents a comprehensive visualization of the defect detection results, including POD outcomes, statistical analysis of defect characteristics, segmentation results, and quantitative similarity analyses.

3.1. Visualization of defect detection

Fig. 3 presents the visualization of porosity defects in a representative aluminum specimen (S26) scanned at four different voxel sizes. Each subpanel shows reconstructed XCT slices extracted from the 3D volume, with one pore highlighted, illustrating the effect of voxel resolution on defect detectability.

At high resolution (10 µm), the majority of defects are clearly resolved, while lower resolutions (20 µm, 35 µm and 70 µm) show a reduction in the number of detected pores and increased partial volume effects. This visualization highlights the importance of voxel size in accurately capturing small-scale defects and supports the quantitative analysis presented in subsequent sections. In this step, it is also important to define the critical defect size (threshold for OK and NOT OK), which can vary depending on the type of component. Establishing these definitions early can save considerable time and effort in subsequent evaluation and inspection processes. More specifically, once the optimal resolution and minimum detectability thresholds are

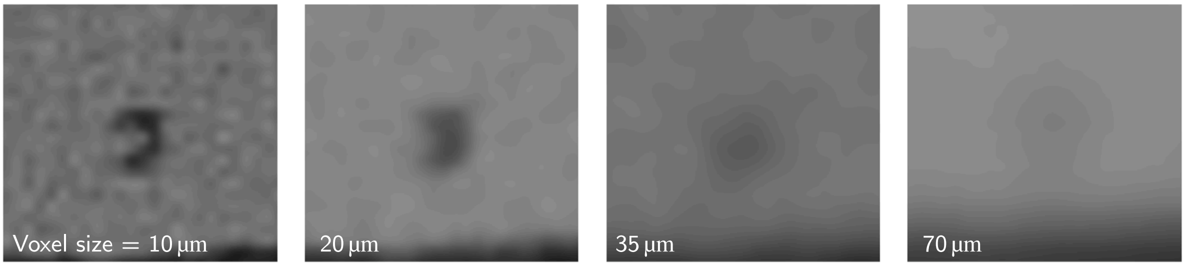


Fig. 5. Close-up of a representative internal defect extracted from the XCT scans at decreasing spatial resolutions (left to right, corresponding to increasing voxel sizes). Detection is only possible at voxel sizes of 10 and 20 μm for the highlighted defect.

determined, there is no need to process high-resolution volumes or perform unnecessarily long scans.

Fig. 4 shows XCT cross-sections of a material sample acquired using four progressively larger voxel sizes (10, 20, 35, and 70 μm). As mention before, these resolutions were chosen based on a previous study of real components [33]. As the voxel size increases, the spatial resolution decreases dramatically, which is evident by the increased blurring and loss of fine textural details in the background material. The red box highlights a small internal defect, illustrating that its sharp definition at low voxel sizes is progressively lost until it is barely discernible at the coarsest resolution of 70 μm. In Fig. 5, the zoomed view of the highlighted defect demonstrates that defect visibility diminishes with decreasing resolution. Notably, defects in the 70 μm entirely undetectable at the lower resolution.

3.2. POD analysis

In this section, we present the POD results together with the corresponding minimum detectability values. The ground-truth dataset is obtained from a 10 μm resolution XCT scan segmented using VG EasyPore. Based on this reference, POD results are subsequently determined for reconstructions at 20 and 35 μm voxel resolutions. Fig. 6 shows that the minimum detectable equivalent diameter pore size ($a_{90/95}$) is 100 μm for the 20 μm resolution case, while it increases to roughly 235 μm for the 35 μm case. The minimum detectable diameter pore size ($a_{90/95}$) is 180 μm for the 20 μm resolution case, while it increases to roughly 550 μm for the 35 μm case (see Fig. 7). In other words, any pore with an equivalent diameter of 100 μm must be detected with 90% probability at a 95% confidence level using a 20 μm voxel resolution. However, smaller defects can still be detected, with the minimum observed defect size being 59 μm (see Table 2).

If we compare the equivalent diameter and diameter results, it is evident that both metrics exhibit similar trends with respect to resolution, with high resolution scans consistently enabling the detection of smaller defects as expected. However, the equivalent diameter generally provides a more robust representation of irregularly shaped pores, capturing variations in volume and morphology that a single linear diameter measurement may underestimate. This indicates that using equivalent diameter as a descriptor can improve defect characterization, particularly for complex geometries commonly found in additive-manufactured components.

The tables (Tables 2 and 3) summarize the estimated detectability limits ($a_{90/95}$), defined as the defect size corresponding to 90% POD, for both measured defect diameters and equivalent defect diameters across specimens S21–S30. For each specimen, the minimum and maximum diameters of detected defects are reported alongside the $a_{90/95}$ values, highlighting the influence of voxel size on detection performance. At 20 μm resolution, smaller defects are consistently detected compared to 35 μm, yielding lower $a_{90/95}$ values and a wider detectable range, as expected from improved sensitivity at finer resolutions.

The mean ± standard deviation values across specimens further quantify this effect. At 20 μm resolution, the mean minimum detected

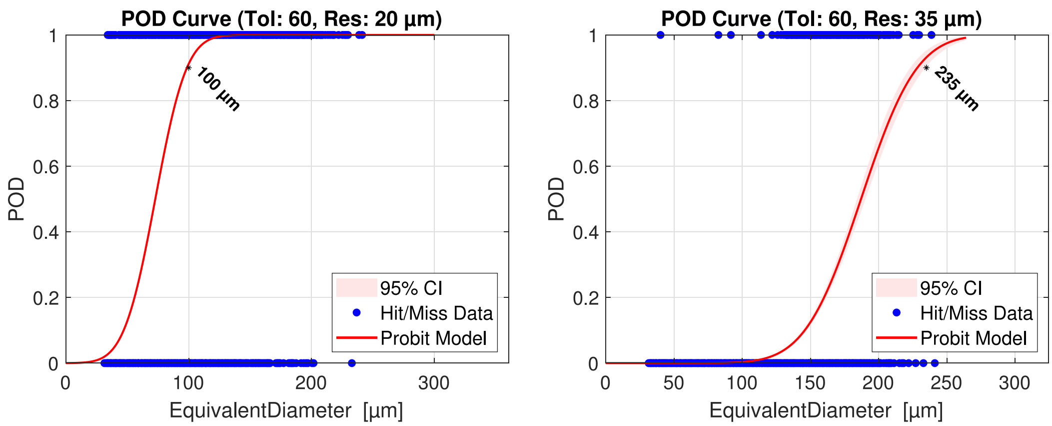
Table 2
Estimated minimum detectability limit values for **defect diameter ($\mathbf{a_{90/95}}$)**, with minimum and maximum sizes of detected defects for each specimen at 20 μm and 35 μm voxel sizes.

Specimen	Resolution (μm)	Min. detected Defect (μm)	Max. Detected Defect (μm)	$a_{90/95}$ (μm)
S21–22	20	59.13	651.91	180
	35	99.87	508.84	556
S23–24	20	69.27	614.22	186
	35	113.09	441.07	521
S25	20	71.18	564.21	233
	35	122.33	485.05	428
S26	20	81.11	551.49	211
	35	148.24	490.58	401
S27–28	20	79.47	546.13	309
	35	142.27	488.88	476
S29–30	20	80.65	549.04	203
	35	139.17	417.40	497
Mean ± Std	20	73.96 ± 8.73	577.20 ± 38.49	223.7 ± 52.1
	35	128.59 ± 21.06	466.57 ± 34.47	476.5 ± 57.0

Table 3
Estimated minimum detectability limit values for **equivalent defect diameter ($\mathbf{a_{90/95}}$)**, with minimum and maximum sizes of detected defects for each specimen at 20 μm and 35 μm voxel sizes.

Specimen	Resolution (μm)	Min. detected Defect (μm)	Max. detected Defect (μm)	$a_{90/95}$ (μm)
S21–22	20	38.97	290.42	100
	35	66.20	254.56	235
S23–24	20	51.44	301.28	103
	35	73.30	229.34	227
S25	20	54.70	248.70	126
	35	92.67	265.85	191
S26	20	60.93	250.48	116
	35	112.25	272.15	187
S27–28	20	64.71	262.03	157
	35	111.36	251.90	208
S29–30	20	61.51	286.99	117
	35	111.38	236.74	221
Mean ± Std	20	55.88 ± 9.17	273.83 ± 14.28	119.83 ± 22.37
	35	94.03 ± 19.00	251.76 ± 14.92	211.50 ± 31.80

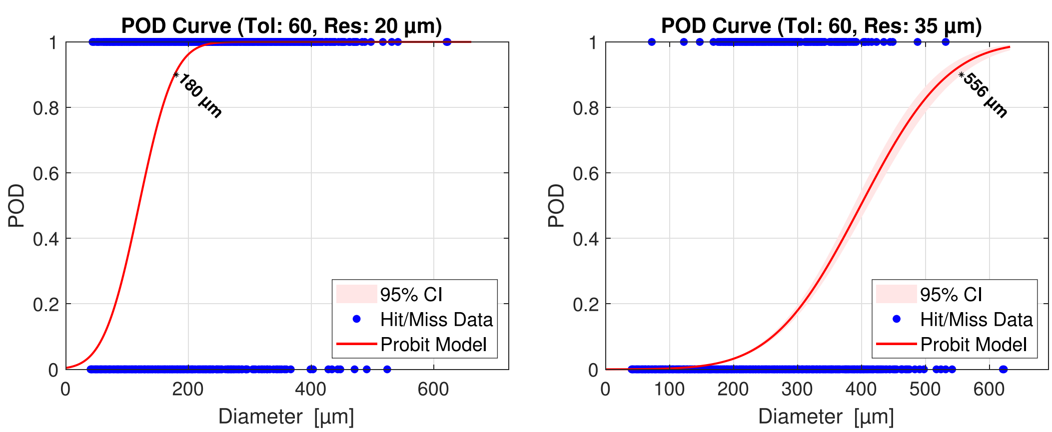
diameter is 73.96 ± 8.73 μm (measured) versus 55.88 ± 9.17 μm (equivalent). The mean maximum detected diameter is 577.20 ± 38.49 μm (measured) versus 273.83 ± 14.28 μm (equivalent), with corresponding $a_{90/95}$ values of 223.7 ± 52.1 μm and 119.83 ± 22.37 μm. At 35 μm, these metrics increase: minimum detected diameters to 128.59 ± 21.06 μm (measured) versus 94.03 ± 19.00 μm (equivalent); maximum detected diameters to 466.57 ± 34.47 μm versus 251.76 ± 14.92 μm; and $a_{90/95}$ values to 476.5 ± 57.0 μm versus 211.50 ± 31.80 μm. At lower voxel resolutions,



(a) POD results at resolution 20 µm

(b) POD results at resolution 35 µm

Fig. 6. POD curves for samples S21–S22, illustrating pore detectability as a function of equivalent pore diameter. The red curves correspond to POD with 95% confidence bounds, while the blue markers represent experimental hit/miss data. The annotated $a_{90/95}$ values indicate the pore diameters detected with a 90% probability at a 95% confidence level. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



(a) Resolution 10 µm vs. 20 µm

(b) Resolution 10 µm vs. 35 µm

Fig. 7. POD curves for samples S21–S22, illustrating pore detectability as a function of pore diameter. The red curves correspond to POD with 95% confidence bounds, while the blue markers represent experimental hit/miss data. The annotated $a_{90/95}$ values indicate the pore diameters detected with a 90% probability at a 95% confidence level. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the minimum detected defect appears larger while the maximum detected defect appears smaller, indicating that small pores are generally overestimated, whereas large pores tend to be underestimated.

Furthermore, paired t-tests were conducted to evaluate the effect of voxel size (20 µm vs. 35 µm) on the estimated minimum detectability limit $a_{90/95}$. For the defect diameter, a statistically significant increase in $a_{90/95}$ was observed at 35 µm resolution ($t(5) = 7.29$, $p = 0.000760$, $p < 0.001$), with a mean difference of 259.50 ± 87.19 µm and a 95% confidence interval ranging from 168.00 to 351.00 µm. Similarly, for the equivalent defect diameter, the detectability limit was significantly higher at 35 µm resolution ($t(5) = 6.56$, $p = 0.001239$, $p < 0.01$), with a mean difference of 91.67 ± 34.26 µm and a 95% confidence interval ranging from 55.72 to 127.62 µm. These results show that voxel size has a statistically significant and practically substantial impact on defect detectability, with consistently higher $a_{90/95}$ values observed at lower spatial resolution.

Comparing the tables reveals systematic differences between measured and equivalent diameters (EqDiam), the latter standardizing defect size for cross-specimen comparability. Equivalent diameters yield consistently lower values across all metrics and resolutions, reflecting their normalization to a spherical equivalent and providing a more conservative assessment of detectability. These findings emphasize the importance of imaging resolution in defect evaluation and offer benchmarks for establishing POD thresholds in quality control and structural integrity applications.

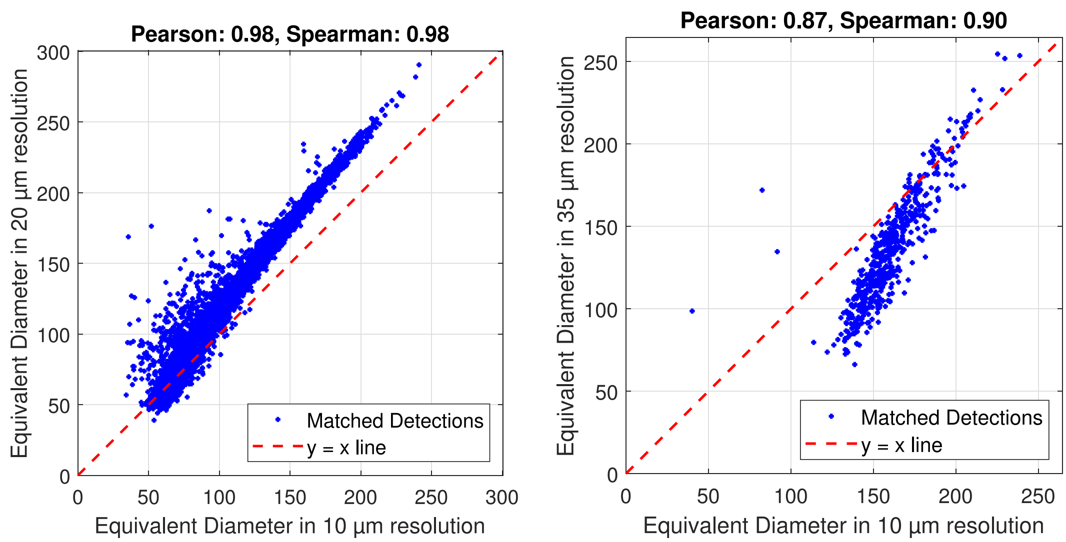
3.3. Impact of voxel resolution on defect morphology

This subsection presents a statistical characterization of detected defects across specimens S21–S30, focusing on their geometric properties and the influence of voxel resolution. A summary of these properties is provided in Table 4.

The analysis shows a significant dependency between imaging resolution and the calculated morphology of the defects. As resolution decreases (lower voxel sizes), there is a systematic overestimation of defect size. The mean equivalent diameter increases from 78.17 ± 5.38 µm at 10 µm resolution to 123.83 ± 4.84 µm and 147.69 ± 7.59 µm at 20 µm and 35 µm, respectively. This sizing inflation is further visualized in the scatter plots in Fig. 8, where the deviation from the $y = x$ line highlights the impact of voxelization on volumetric data. Alongside this, the mean sphericity of detected defects increases from 0.7950 ± 0.0020 at the highest resolution to 0.8703 ± 0.0106 at 20 µm and 0.8730 ± 0.0129 at 35 µm. This trend suggests a "smoothing" effect; as the voxel size increases, the ability to resolve irregular surface features and complex

Table 4
Average defect properties (mean ± standard deviation) for different XCT resolutions for S21 to S30.

Resolution	Sphericity	Equiv. diameter [μm]	Diameter [μm]	Voxels
10 μm	0.7950 ± 0.0020	78.17 ± 5.38	125.04 ± 7.71	457.63 ± 68.62
20 μm	0.8703 ± 0.0106	123.83 ± 4.84	177.15 ± 6.86	168.32 ± 20.17
35 μm	0.8730 ± 0.0129	147.69 ± 7.59	204.08 ± 8.83	53.01 ± 6.51



(a) 10 vs. 20 μm resolution
(b) 10 vs. 35 μm resolution

Fig. 8. Scatter plots of matched defect equivalent diameters between ground-truth 10 μm scans and coarser resolutions. The trend lines ($y = x$) show size inflation due to voxelization, with increased scatter at 35 μm indicating higher matching uncertainty.

branch-like geometries is lost, causing the segmented defects to appear more artificially spherical.

The spatial sampling density also drops dramatically across the tested resolutions. The average voxel count per defect falls from 457.63 at 10 μm to only 53.01 at 35 μm. This nearly nine-fold reduction in voxels per defect explains the increased scatter observed in Table 4, as there is insufficient information to accurately define defect boundaries at lower scales. These statistical trends demonstrate that defect morphology metrics are highly sensitive to imaging resolution, providing essential context for the POD results presented in subsequent sections.

3.4. Defect sizing error analysis

To evaluate the reliability of defect quantification at lower resolutions, a comparative analysis was performed between the high-resolution ground-truth scans (10 μm) and low resolutions (20 μm and 35 μm). The sizing accuracy was assessed using two metrics: Equivalent Diameter and Diameter.

Fig. 8 presents the scatter plots for the equivalent diameter. A strong correlation is observed between the 10 μm and 20 μm resolutions (Fig. 8a), though a systematic positive bias is evident where the data points lie largely above the $y = x$ reference line. This indicates a general overestimation of defect volume in the lower scans, a phenomenon likely attributed to the partial volume effect and voxelization artifacts, where border voxels artificially inflate the total volume of small defects. At 35 μm (Fig. 8b), while the positive correlation remains, the scatter significantly increases, reflecting a higher degree of uncertainty in volumetric estimation as the defect size approaches the resolution limit.

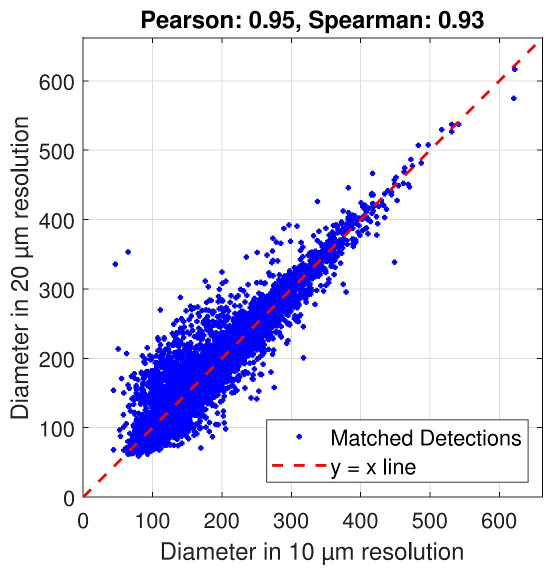
Fig. 9 illustrates the comparison for the diameter metric. Consistent with the volumetric data, the 20 μm resolution maintains reasonable linearity with the ground truth. However, the 35 μm resolution (Fig. 9b) exhibits severe scattering. Unlike the equivalent diameter, which integrates volume, the maximum diameter is highly sensitive to the geometric representation of the defect shape. The discrete nature of the larger voxels at 35 μm leads to substantial deviations, with defects being both significantly over- and under-sized compared to the ground truth.

To provide a quantitative assessment of these observations, Pearson correlation [43] and Spearman correlation [44] coefficients were computed to evaluate the agreement between defect size measurements across different voxel resolutions. For the equivalent diameter metric, the comparison between 10 μm and 20 μm scans shows very strong correlation (Pearson = 0.98, Spearman = 0.98), indicating highly consistent defect size estimation across these resolutions. In contrast, the comparison between 10 μm and 35 μm scans exhibits lower correlation (Pearson = 0.87, Spearman = 0.90), reflecting increased variability and reduced measurement reliability at coarser voxel resolutions. A similar trend is observed for the diameter metric. The 10 μm and 20 μm comparison still demonstrates strong agreement (Pearson = 0.95, Spearman = 0.93), whereas the 10 μm and 35 μm comparison shows substantially reduced correlation (Pearson = 0.68, Spearman = 0.64). Sizing error between resolutions is further quantified by the scatter around the $y = x$ line in Figs. 8 and 9, as well as by the Pearson and Spearman correlation coefficients, confirming the systematic partial-volume overestimation while providing a direct measure of variability.

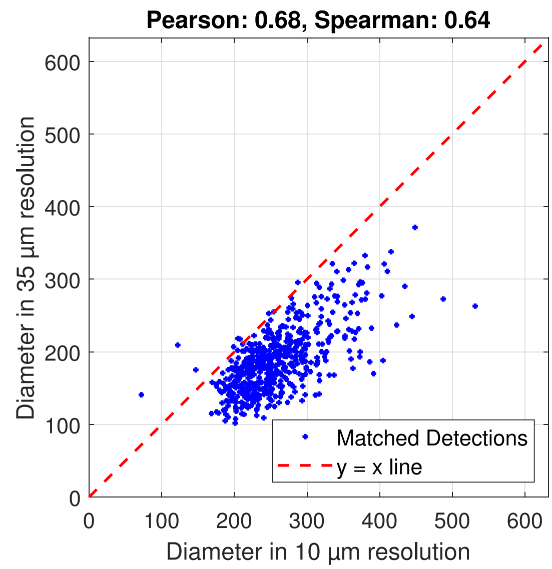
Overall, the data suggests that while defect detection may be feasible at lower resolutions, accurate sizing — particularly using raw diameter measurements — degrades rapidly as the voxel size increases relative to the defect scale. The equivalent diameter appears to be a more robust metric than maximum diameter for characterizing defects in lower-resolution XCT data.

3.5. Segmentation and image similarity

The impact of resolution degradation on both image fidelity and segmentation accuracy was evaluated using a combination of pixel-based metrics (PSNR, MS-SSIM) and volumetric overlap metrics (Dice



(a) 10 vs. 20 μm resolution



(b) 10 vs. 35 μm resolution

Fig. 9. Scatter plots of matched defect diameters between ground-truth 10 μm scans and coarser resolutions. The trend lines ($y = x$) show size inflation due to voxelization, with increased scatter at 35 μm indicating higher matching uncertainty.

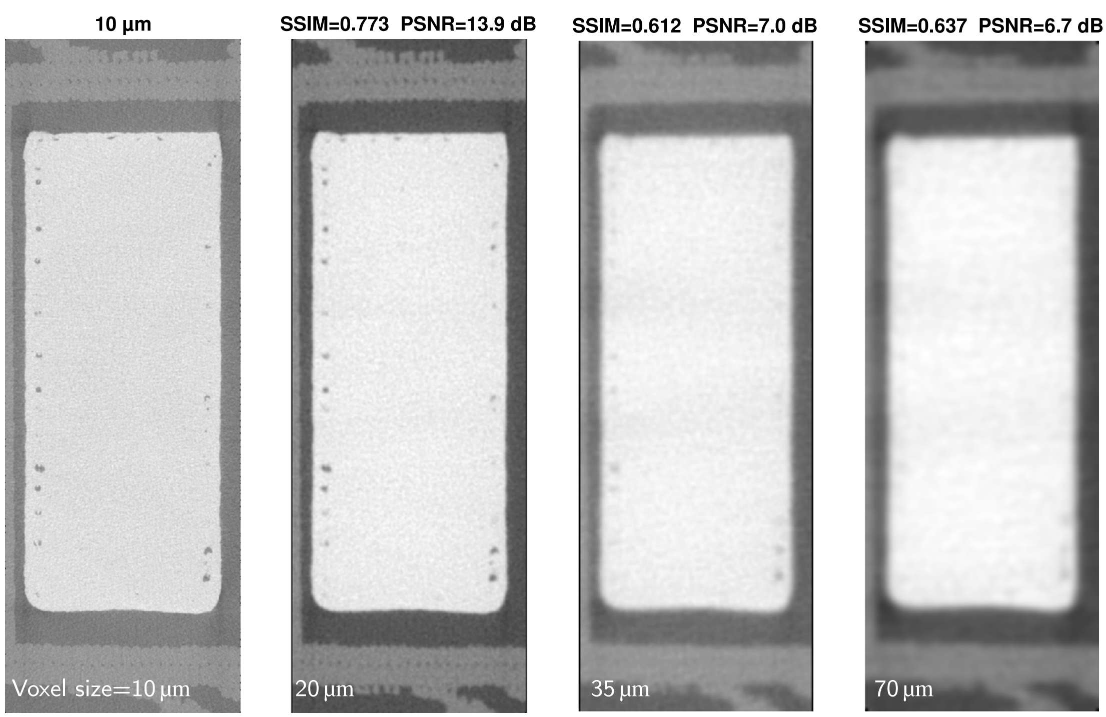


Fig. 10. Volume slices from XCT reconstructions at 10, 20, 35, and 70 μm resolutions, where the 10 μm reconstruction is treated as the ground truth (GT). Structural similarity index (SSIM) and peak signal-to-noise ratio (PSNR) are annotated for each comparison to GT, highlighting progressive degradation in image fidelity and defect resolvability due to partial volume effects.

coefficient, Jaccard index). Fig. 10 illustrates the progressive loss of structural detail due to partial volume effects. While the 20 μm scan maintains high fidelity compared to the ground truth (GT) with a PSNR of 13.9 dB and an SSIM of 0.773, a sharp decline is observed at lower resolutions. At 35 μm resolution, the PSNR drops significantly to 7.0 dB, and fine internal features — specifically the smaller micropores — become indistinguishable from the background noise. This blurring effect confirms that as the voxel size exceeds the scale of the smallest defects, the gray-value contrast necessary for accurate thresholding is severely compromised.

This resolution effect directly impacts the segmentation quality shown in Fig. 11. The 20 μm resolution yields a segmentation mask nearly identical to the ground truth, achieving a Dice coefficient of 0.993. However, at 35 μm and 70 μm, the binary masks exhibit noticeable edge roughening and a loss of finer defect structures. Although the global shape of the specimen is preserved — reflected in Dice scores remaining above 0.97 even at 70 μm — the drop in the Jaccard index from 0.985 (20 μm) to 0.943 (70 μm) highlights increasing discrepancies at the object boundaries and defect interfaces. Fig. 12 summarizes the statistical distribution of these metrics across the dataset. The

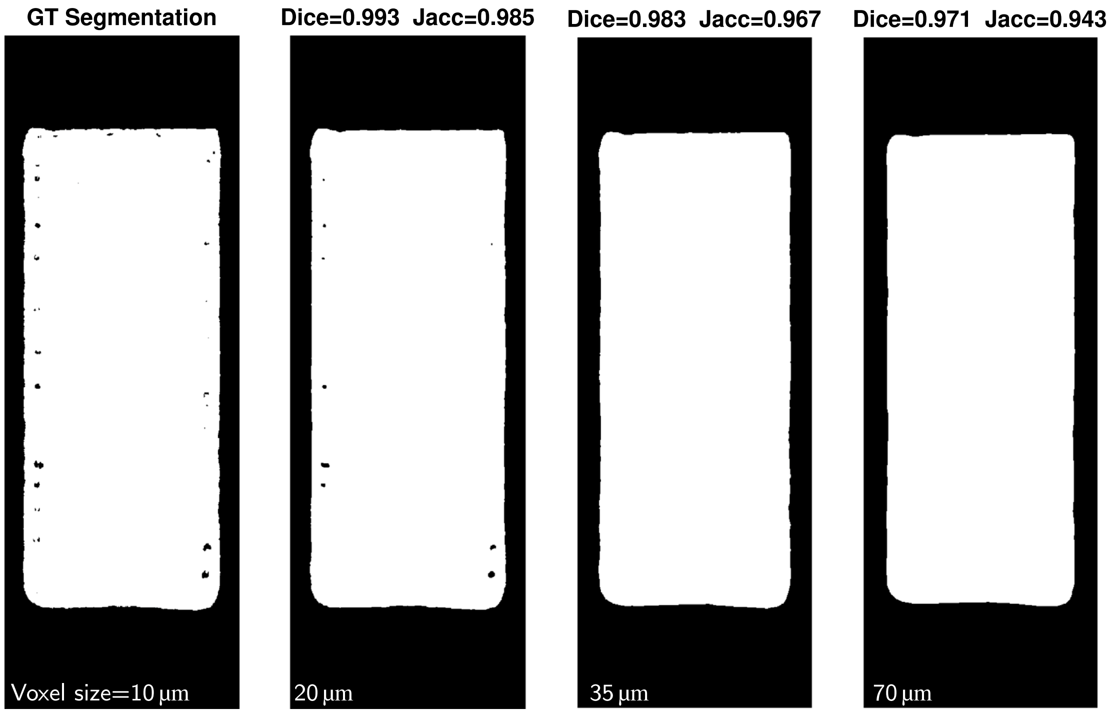


Fig. 11. Segmented slices comparing ground-truth (GT, left) to outputs at 20, 35, and 70 µm resolutions. Dice coefficients and Jaccard index are annotated for each resolution.

boxplots reveal a consistent trend: image quality metrics (MS-SSIM and PSNR) exhibit higher variance and a steeper decline than segmentation metrics. Notably, the PSNR shows a distinct step-change between 20 µm and 35 µm, suggesting a critical threshold for image quality loss. Conversely, the high median values for Dice (>0.98) and Jaccard (>0.96) across all resolutions suggest that the specimen geometry is well-captured even at 70 µm; however, these metrics mask the fact that defects were not segmented at all.

4. Discussion

In this section, we discuss the influence of voxel size on defect detectability, the use of diameter versus equivalent diameter as POD descriptors, and comparisons with existing literature. We also address the practical implications for industrial XCT inspections, including trade-offs between voxel resolution, scan efficiency, and defect detectability. Finally, we outline the study's limitations and directions for future work.

4.1. Influence of voxel size on defect detectability

The presented results clearly show that voxel size is the dominant factor governing defect detectability in XCT-based porosity analysis. When using the 10 µm reconstruction as ground truth, a systematic increase in the minimum detectable defect size is observed as voxel size increases to 20 µm and 35 µm. This trend is consistent across all investigated specimens and defect descriptors (diameter, equivalent diameter, and volume), and is in agreement with established XCT resolution studies [34,45].

From a physical standpoint, this behavior is primarily driven by partial volume effects and reduced spatial sampling. As voxel size increases, small pores occupy fewer voxels and increasingly blend with surrounding material, leading to under-segmentation or complete loss of detectability [46]. The POD results quantify this effect explicitly: the $a_{90/95}$ equivalent diameter increases from approximately 100 µm at 20 µm voxel size to about 235 µm at 35 µm voxel size. This corresponds to detectability thresholds of roughly 5–6 voxels across a defect diameter, which aligns well with commonly accepted empirical rules in XCT metrology and defect analysis [30,45]. Importantly, these findings emphasize that nominal voxel size alone is not a sufficient indicator of detectability. Instead, voxel size must be interpreted relative to the expected defect size distribution and the acceptable POD for a given application, as required in reliability-based NDT qualification frameworks [33,47]. As previously discussed, a voxel resolution of 20 µm is generally suitable for detecting defects larger than 100 µm. However, for safety-critical applications, such as in the aerospace industry, there remains a low but non-negligible risk of missing critical defects due to factors including scan image quality, material properties, and geometric complexity. Therefore, the optimal resolution threshold must be determined on a case by case basis under carefully controlled conditions to ensure reliable defect detection and compliance with industry safety standards.

Building on these resolution-dependent limitations, super-resolution techniques have demonstrated strong performance in improving image quality, effectively transforming low-resolution scans toward high-resolution detail [48–50]. However, there are limited studies addressing the improvement of industrial defect detection through the use of super-resolution techniques. For example, recent work by Klos et al. [51] demonstrates that deep learning-based super-resolution XCT can improve image quality and enable accurate 3D quantification of defects in lattice structures, significantly reducing scan time while improving porosity and surface roughness analysis compared to conventional low-resolution data. These approaches have the potential to improve defect detection rates from low-resolution XCT data.

4.2. Diameter versus equivalent diameter as POD descriptors

A key observation of this study is the systematic difference between POD results obtained using the measured defect diameter and those based on equivalent diameter. While both descriptors exhibit similar trends with respect to resolution, the equivalent diameter consistently yields lower $a_{90/95}$ values and reduced inter-specimen variability, indicating improved robustness.

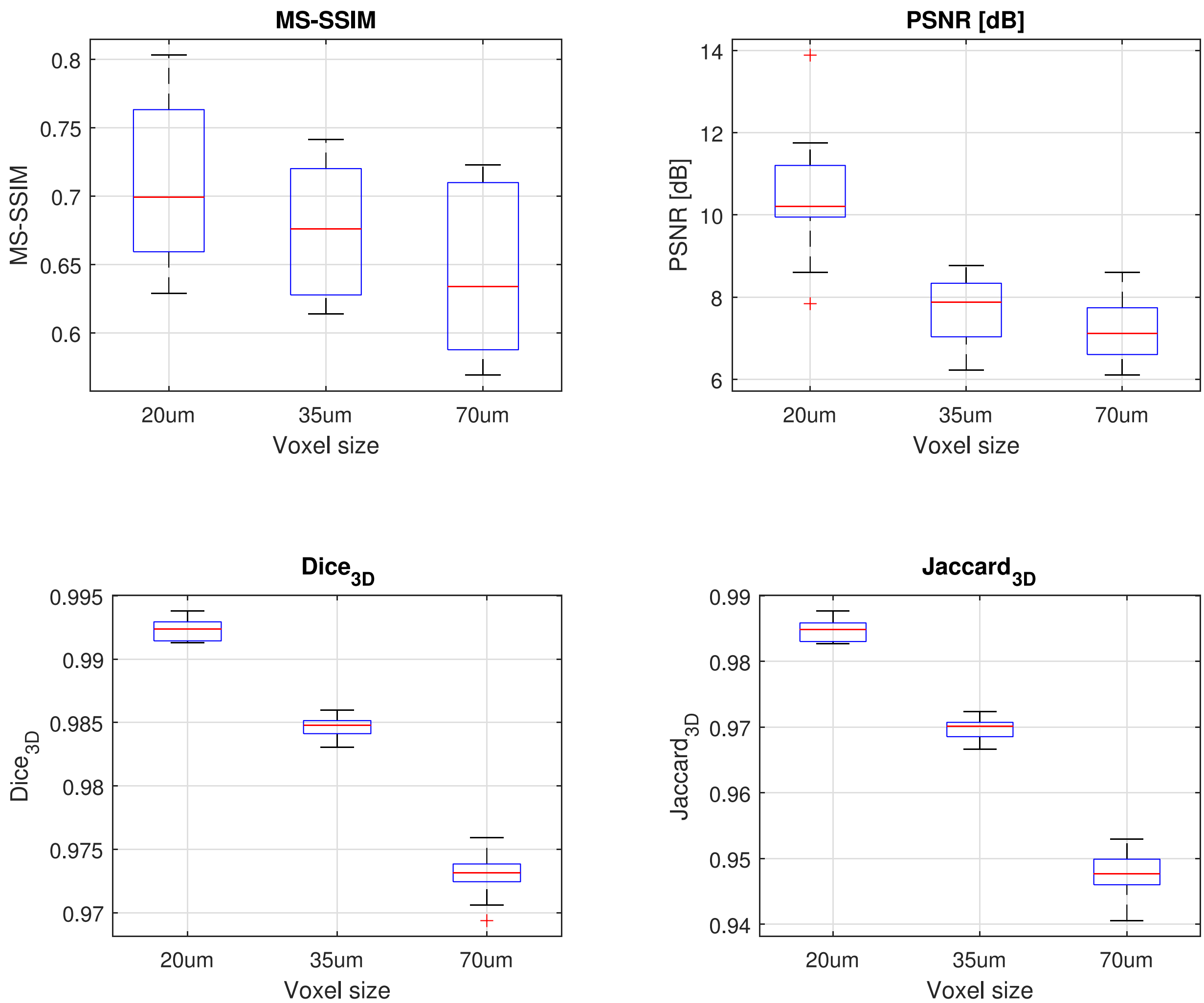


Fig. 12. Statistical distribution of image similarity (MS-SSIM, PSNR) and segmentation overlap ($Dice_{3D}$, $Jaccard_{3D}$) metrics across resolutions for S21 to S30. While volumetric overlap remains numerically high (>0.97), the sharp drop in PSNR and MS-SSIM at 35 μm.

This behavior can be attributed to the fact that real porosity defects — particularly those originating from additive manufacturing or casting — rarely exhibit ideal spherical or cylindrical shapes [45,46]. Linear diameter measurements are therefore sensitive to orientation, segmentation noise, and boundary smoothing effects at lower resolutions. In contrast, the equivalent diameter integrates volumetric information and provides a more robust and resolution-tolerant representation of defect size [52].

These results support the use of equivalent diameter as the preferred descriptor for POD analysis in XCT-based porosity studies, particularly when irregular or partially merged defects are present. This conclusion is consistent with prior XCT-based defect characterization studies, where volumetric descriptors were shown to correlate more reliably with mechanical performance and fatigue behavior than single linear dimensions [46].

4.3. Comparison with existing literature

Recent metrological studies have further quantified the influence of voxel size on resolution-driven biases in additive manufacturing (AM) components. Holgado et al. conducted a detailed metrological evaluation and classification of porosity in metal AM using XCT, confirming that voxel size directly governs both detection limits and sizing accuracy. They showed that partial-volume-induced overestimation becomes significant above approximately 25–30 μm voxel sizes, which aligns closely with the ~5–6 voxel empirical rule observed in this work [36]. Similarly, Jones et al. validated XCT detection limits for stochastic flaws in additively manufactured Ti-6Al-4V and reported comparable voxel-size thresholds (typically 4–6 voxels for reliable probability of detection, POD). Their findings highlight that real stochastic defect populations amplify the impact of resolution limits compared with idealized synthetic defects [52].

Advanced segmentation techniques, particularly deep-learning methods, have emerged as a powerful complement to traditional thresholding approaches for improving POD at lower resolutions. Perghem et al. systematically compared Otsu, Triangle, and U-Net segmentation methods across varying voxel sizes in L-PBF AlSi10Mg, demonstrating that U-Net — especially when combined with machine learning-based interaction and 3D convex-hull reconstruction — significantly improves POD curves and reduces sizing errors compared with Otsu, which tends to underestimate defects at larger voxel sizes [53]. Ledwaba et al. showed that a 2.5D U-Net trained on a limited number of calibrated slices outperforms global Otsu and manual thresholding on Ti-6Al-4V datasets, while remaining robust to beam-hardening artifacts and varying porosity levels [54].

Era et al. proposed an unsupervised, promptable porosity segmentation framework based on the Segment Anything Model (SAM), enabling scalable, label-free analysis of XCT volumes where annotated data are scarce, offering a promising pathway toward automated, resolution-agnostic defect characterization [55]. Wilbig et al. and Villarraga-Gómez et al. further demonstrated that deep learning-driven multi-scale

XCT/XRM workflows (e.g., DeepRecon and DeepScout) can reduce scan times while preserving defect detectability, highlighting that segmentation accuracy — not only voxel size — is a key limiting factor in industrial porosity analysis [56,57].

Additionally, recent benchmarking studies comparing porosity analysis algorithms such as VGDefX, OnlyThreshold, and VGEasyPore have shown that VGDefX achieves the highest repeatability, with subvoxel errors and low measurement uncertainty for features larger than five voxels [36]. In the present work, the VGEasyPore method was used for segmentation and defect detection. While it enables reliable detection of small pores, the results suggest that POD performance and sizing accuracy could be further improved by adopting deep-learning-based segmentation techniques.

Duba-Sullivan et al. explored 2.5D super-resolution networks to overcome inherent XCT resolution limitations in AM components, achieving improved flaw detectability without increasing scan time or radiation dose [58]. Collectively, these studies demonstrate that although voxel size remains the dominant physical constraint on detectability — as also quantified in the present work — modern deep learning-based segmentation and reconstruction approaches can partially compensate for coarser resolutions, enabling more efficient XCT inspection workflows.

In contrast to many prior POD studies that rely on simulated XCT data, the present work is based on experimentally acquired high-resolution XCT data used as ground truth. This approach avoids uncertainties associated with simplified defect geometries and idealized noise models, thereby providing POD estimates that are more representative of real industrial inspection scenarios.

4.4. Practical implications for industrial XCT inspection

From an application perspective, the results offer concrete guidance for selecting XCT scan resolution and managing sizing errors to optimize the POD in industrial of AM parts that are characterized by process induced internal imperfections. For aluminum components with critical defect sizes below approximately 150 μm, voxel sizes of 20 μm or finer are required to achieve a POD of at least 90%, consistent with reliability-based inspection requirements in safety-critical applications [47]. Conversely, when only larger defects are of concern, lower resolutions may be sufficient and can significantly reduce scan time, data volume, and computational cost. The results further demonstrate that once detectability thresholds are established using POD analysis, it is unnecessary to routinely perform high-resolution scans for all components. Instead, resolution can be tailored to the specific inspection task, enabling more efficient and cost-effective XCT-based inspection workflows [33].

Furthermore, the POD-based framework developed in this study also establishes a direct connection between XCT detectability and mechanical performance, particularly in the context of fatigue-critical defect assessment. In metallic materials, and especially in additively manufactured aluminum alloys, fatigue behavior is strongly governed by defect populations, where pores act as local stress concentrators and preferential crack initiation sites. This relationship has been extensively demonstrated in defect-based fatigue approaches, such as those proposed by Murakami et al. where fatigue strength is correlated to defect size. In particular, three-dimensional analyses have shown that both internal and surface-connected defects can dominate fatigue life, with crack initiation strongly influenced by defect size, morphology, and location [59,60]. In this context, XCT-derived defect populations are increasingly used as direct input for finite element (FE) and fatigue simulations, enabling physically informed predictions of crack initiation and propagation. Studies such as those by Tammas-Williams et al. [61] and Romano et al. [62] demonstrate that incorporating realistic defect geometries obtained from XCT into computational models enables accurate prediction of local stress concentration, fatigue strength, and crack initiation behavior. These works show that explicitly resolving defect morphology and spatial distribution is essential for capturing failure-critical locations and reproducing experimentally observed fatigue performance. More recent work, investigating the tailoring of porosity and mechanical properties in wire-based directed energy-deposited molybdenum alloys via hot isostatic pressing (HIP), demonstrates that significant reductions in porosity achieved through optimized HIP treatments are directly correlated with substantial improvements in tensile strength and ductility [63]. These findings underscore the critical relationship between porosity mitigation and enhanced mechanical performance.

Within this framework, the present POD analysis provides quantitative detectability metrics (e.g., $a_{90/95}$) that define the range of defects reliably captured at a given voxel size. This can enable a consistent alignment between XCT acquisition parameters and simulation requirements, ensuring that the defect populations used in FE or fatigue models are representative of the features governing crack initiation. Consequently, voxel size can be interpreted as a mechanics-informed parameter, where resolution is selected not only to achieve sufficient image quality, but also to capture the defect size regime controlling structural performance. Overall, the integration of POD-based detectability with defect characterization provides a robust basis for linking NDE to performance-driven modeling, supporting improved fatigue life prediction and more reliable defect-tolerant design strategies.

Fig. 13 highlights the trade-off between voxel size, scan efficiency, and defect resolution. As voxel size increases, both scan time and computational cost decrease, but smaller defects become less detectable. Moreover, both the mean defect size and the $a_{90/95}$ minimum detectability metrics increase with voxel size, indicating reduced sensitivity to fine features. This emphasizes that selecting a voxel size requires balancing efficiency with the need to detect small defects: smaller voxels are favored for high-resolution inspections, while larger voxels are more suitable for rapid, coarse assessments. In our experiment, the optimal voxel size was determined to be 20 μm, enabling reliable detection of defects approximately 100 μm in size. However, this optimal value may vary depending on the specific application and material conditions. For example, in the automotive industry, a voxel size of 35 μm may be sufficient for certain components, whereas in aerospace applications, defects smaller than 100 μm can be critical and require higher-resolution scans.

4.5. Limitations and future work

The proposed approach offers notable advantages, particularly through the implementation of a POD-based framework for evaluating sizing errors and determining the minimum detectability thresholds for real CT data at specific resolutions; however, it remains subject to certain constraints and limitations in its applicability. Although the 10 μm CT scans were considered as reference data, even higher-resolution imaging could reveal additional micro-porosity, as noted in prior XCT capability studies [45]. Consequently, features smaller than the voxel size may be partially averaged or undetected, potentially affecting defect characterization and sizing accuracy. Recent comparative studies have demonstrated that XCT scans at this scale may fail to reach a 90% detection rate for the "true" flaw population until the defects are approximately 7 to 17 times larger than the voxel size [52]. Furthermore, the reliance on laboratory-scale micro-CT can lead to the underestimation of complex, non-spherical features — such as thin lack-of-fusion "webbings" — which are often only fully resolvable through higher-fidelity synchrotron-based imaging or destructive serial sectioning [64,65]. Consequently, using such scans as a baseline may introduce a systematic bias by neglecting the smallest population of defects that significantly contribute to the material's fatigue life.

In this study, CT scans of aluminum alloy specimens were acquired under standard industrial conditions. Although the reconstruction process included standard corrections, residual beam hardening

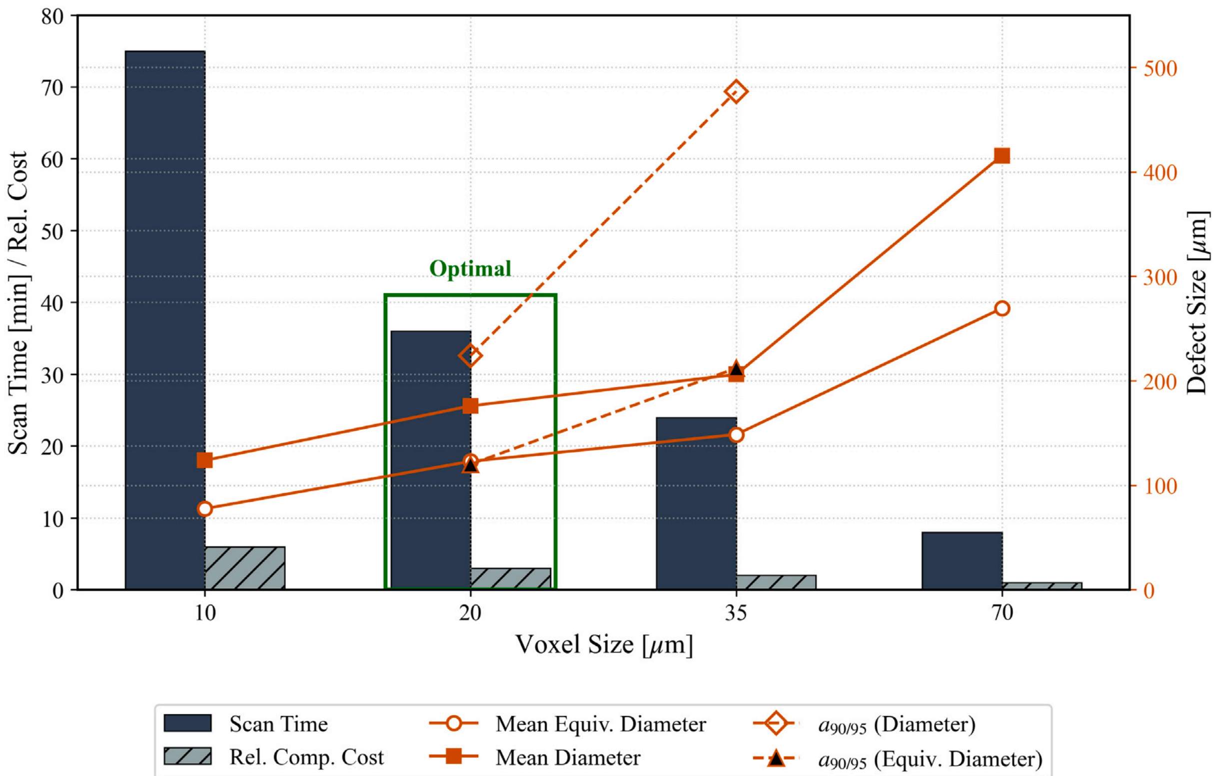


Fig. 13. Resolution trade-offs for defect analysis. Scan time and relative computational cost decrease with increasing voxel size (left y-axis, bar plot), while defect detectability metrics — including mean diameter, mean equivalent diameter, and $a_{90/95}$ values (right y-axis, line plot) — show that coarser voxel resolutions reduce the ability to detect smaller defects. This figure highlights the practical trade-offs between voxel resolution, computational effort, and defect detectability.

or scatter effects may locally alter intensity values, particularly near dense regions, thereby impacting measured defect dimensions and their detectability. Beam hardening and scatter artifacts can introduce systematic distortions in reconstructed CT volumes, leading to potential under- or overestimation of defect sizes. Beam hardening occurs when lower-energy X-rays are preferentially absorbed as they pass through dense regions, resulting in nonlinear attenuation profiles and artificially bright or dark areas near dense features [66]. Scatter effects, caused by X-rays deflecting from their original paths, can generate diffuse intensity patterns, streaks, or shading in the reconstructed volume [67]. Together, these artifacts can locally alter voxel intensities, blur defect boundaries, and create apparent variations in defect morphology, particularly for small pores or thin features. Consequently, these distortions may cause segmentation algorithms to misclassify regions, underestimate defect volumes, or even fail to detect subtle defects, directly influencing the accuracy of sizing and the estimated probability of detection.

The analysis was performed on aluminum fatigue specimens with naturally occurring porosity, produced from a single material and processed under a specific manufacturing route. Consequently, the observed defect detectability may not directly translate to other materials, manufacturing methods, or defect morphologies [46]. Materials with higher X-ray attenuation, such as steels or titanium alloys, may exhibit stronger beam hardening and scatter effects, which could exacerbate sizing errors and reduce detectability of small defects. Likewise, defects arising from different mechanisms — such as gas pores, shrinkage cavities, or cracks in additive-manufactured or cast components — may present distinct shapes, densities, or contrast levels, influencing segmentation accuracy and probability of detection estimates. Variations in surface finish, internal microstructure, and density heterogeneity across materials may further alter CT image quality and the efficacy of threshold-based segmentation. Therefore, while the current results provide valuable insights for aluminum fatigue porosity, additional validation studies are necessary to confirm the applicability of the proposed approach to other materials, defect types, and processing conditions.

AI-based segmentation methods, particularly the U-Net convolutional neural network [68], have proven highly effective in medical and materials science applications. In materials science, Yosifov et al. [24] showed that CNN-based approaches outperform traditional methods such as k-means, watershed, and Otsu thresholding by 5%–20% in defect detection on simulated XCT data. Employing deep learning-based segmentation approaches could provide more accurate defect detection and potentially improve the resulting POD estimates.

Building on the present study, several promising avenues exist for further research. First, the integration of advanced machine learning approaches for defect segmentation could improve detection accuracy and enable automated analysis of large datasets. Second, the methodology could be extended to other materials, such as titanium alloys or composite structures, broadening its applicability in industrial quality control. Finally, incorporating super-resolution reconstruction techniques may enhance defect visibility in low-resolution CT scans, potentially improving the POD for small or subtle defects.

In addition, integrating mechanical testing results with POD-based detectability thresholds would enable direct linkage between XCT inspection performance and structural integrity requirements, further strengthening the relevance of XCT-based POD analysis for industrial NDT.

5. Conclusion

This study successfully quantified the impact of sizing error on defect detectability in additively manufactured aluminum alloys using a POD framework. By utilizing 10 μm high-resolution scans as a ground truth, we demonstrated a significant shift in detection thresholds across varied voxel sizes. The mean $a_{90/95}$ values demonstrate that reducing the voxel resolution from 20 μm to 35 μm nearly doubles the minimum reliably detectable defect size, highlighting the strong dependence of detection performance on spatial resolution. Specifically, the $a_{90/95}$ equivalent diameter limit increased from approximately 100 μm at a 20 μm resolution to 235 μm at a 35 μm resolution. Lower resolutions systematically overestimate defect size due to partial volume effects, inflate apparent sphericity (from ≈ 0.80 to ≈ 0.87), and reduce segmentation accuracy, while small defects (smaller than 50 μm) become undetectable. The results provide actionable guidance for industrial XCT inspection: 20 μm voxels offer an effective compromise for applications tolerating defects over 100 μm, whereas safety-critical components demanding detection of smaller defects require 10 μm or higher resolution. By leveraging real multi-resolution experimental data rather than simulations, this work delivers representative POD benchmarks, enabling tailored scan protocols that balance detection reliability, scan time, and computational resources.

These findings confirm that while equivalent diameter serves as a more robust descriptor for irregular pores, scan resolution remains the dominant factor in capturing critical sub-millimeter defects. From an application perspective, the results provide concrete guidance for selecting scan parameters that balance imaging fidelity and sizing error with processing time in NDE. This work provides NDE practitioners with a data-driven basis for optimizing XCT inspection workflows to ensure the structural integrity of safety-critical components in industries such as aerospace and automotive manufacturing. Building on these findings, practical recommendations are provided to guide NDT and NDE practitioners in optimizing inspection workflows:

1. For high-criticality applications (e.g., aerospace), voxel sizes of ≤ 20 μm are recommended to reliably detect defects of approximately 100 μm equivalent diameter ($a_{90/95}$).
2. For moderate-criticality applications (e.g., automotive), voxel sizes of 20–35 μm provide a balance between detectability (∼100–235 μm) and inspection efficiency.
3. For low-criticality or high-throughput screening, voxel sizes ≥ 35 μm are acceptable, with detectability limited to defects larger than 200 μm.

To balance resolution, scan time, and computational cost, a practical trade-off strategy can be adopted. As shown in Fig. 13, higher-resolution scans (smaller voxel sizes) increase both scan time and computational cost, while improving the detectability of smaller defects. Conversely, lower voxel sizes reduce scan time and computational demand but compromise the detection of fine defects. An intermediate voxel size (e.g., 20 μm) provides an optimal balance, maintaining reliable detection of critical defects while significantly reducing scanning and processing effort. Equivalent diameter is identified as a robust descriptor for POD analysis, exhibiting reduced sensitivity to voxelization effects. Overall, the proposed POD-based framework enables quantitative and application-specific selection of XCT parameters aligned with defect criticality and inspection efficiency requirements.

Building on the established POD framework, future research will focus to automate the dual-resolution registration pipeline using machine learning-based segmentation to reduce manual processing time. Another significant avenue of investigation involves applying this methodology to different samples and materials, such as carbon fibers and other type of metals, where beam hardening and scatter artifacts may introduce different sizing error characteristics. Finally, integrating these POD curves into a real-time adaptive XCT control system could allow for dynamic resolution adjustment based on the local criticality of the component's geometry.

CRediT authorship contribution statement

Miroslav Yosifov: Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Jonathan Glinz:** Writing – review & editing, Data curation. **Lorenzo Rusnati:** Writing – review & editing, Data curation. **Bernhard Plank:** Writing – review & editing, Funding acquisition. **Reinhard Hubmann:** Resources. **Johann Kastner:** Writing – review & editing, Supervision. **Sascha Senck:** Writing – review & editing, Funding acquisition, Data curation, Conceptualization.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the author(s) used ChatGPT for language editing purposes only. After using this tool, the author(s) thoroughly reviewed and edited the content as needed and take full responsibility for all aspects of the published article.

Funding

We acknowledge financial support for the project sustaiNDT (grant nr. 909801) funded within the framework "Production & Materials" of Austrian Research Promotion Agency (FFG). The work was also supported by the project IBW/EFRE & JTF 2021–2027 – HyperMAT.

Declaration of competing interest

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

Appendix

The tables in Appendix summarize key defect properties extracted from XCT scans of specimens S21–S30 at voxel sizes of 10 μm, 20 μm, and 35 μm. Tables A.1–A.3 summarize the average defect properties for each specimen pair at voxel sizes of 10 μm, 20 μm, and 35 μm, respectively. At 10 μm resolution, defects exhibit relatively small equivalent diameters and voxel counts, with mean sphericity values close to 0.8, indicating moderately irregular defect shapes. As the voxel size increases to 20 μm and 35 μm, both equivalent diameter and measured diameter increase systematically for all specimens. This trend is primarily attributed to partial volume effects and the merging of fine defect features at lower resolutions.

Table A.1
Average defect properties from 10 μm scans for Specimens S21–S30.

Specimens	Sphericity	Equiv. diam. [μm]	Diameter [μm]	Voxels
S21–22	0.79527	74.50	120.23	419.68
S23–24	0.79775	77.52	125.14	441.04
S25	0.79594	87.59	137.96	584.35
S26	0.79181	73.56	118.41	403.23
S27–28	0.79385	81.14	129.79	485.18
S29–30	0.79564	74.70	118.69	412.30

Table A.2
Average defect properties from 20 μm scans for Specimens S21–S30.

Specimens	Sphericity	Equiv. diam. [μm]	Diameter [μm]	Voxels
S21–22	0.86018	120.46	173.78	168.73
S23–24	0.86286	130.78	188.80	198.13
S25	0.88562	126.68	178.76	170.76
S26	0.87398	117.51	168.90	141.74
S27–28	0.87855	121.63	173.46	150.87
S29–30	0.86089	125.93	179.19	179.68

At the finest resolution (10 μm, Table A.1), defects generally have small equivalent diameters and voxel counts, with mean sphericity

Table A.3
Average defect properties from 35 µm scans for Specimens S21–S30.

Specimens	Sphericity	Equiv. diam. [µm]	Diameter [µm]	Voxels
S21–22	0.85616	136.53	191.49	45.49
S23–24	0.86702	146.44	203.86	50.50
S25	0.89434	158.18	214.64	63.50
S26	0.87479	152.27	212.11	56.37
S27–28	0.87803	150.05	205.61	54.27
S29–30	0.86752	142.65	196.79	47.92

Table A.4
Defect properties summary (S25, S26, S27–28, S29–S30 combined).

Dataset	EqDiam (µm) Mean ± Std [Min–Max]	Diam (µm) Mean ± Std [Min–Max]
S21–22		
S21–22 10 µm	74.50 ± 36.62 (31.22–241.13)	120.23 ± 67.97 (40.45–622.23)
S21–22 20 µm	120.46 ± 42.90 (38.97–290.42)	173.78 ± 70.61 (59.13–651.91)
S21–22 35 µm	136.53 ± 32.80 (66.20–254.56)	191.49 ± 49.72 (99.87–508.84)
S23–24		
S23–24 10 µm	77.52 ± 35.60 (31.94–234.14)	125.14 ± 67.27 (41.22–552.31)
S23–24 20 µm	130.78 ± 39.78 (51.44–301.28)	188.80 ± 67.59 (69.27–614.22)
S23–24 35 µm	133.61 ± 27.88 (73.30–229.34)	187.56 ± 41.53 (113.09–441.07)
S25–26		
S25–26 10 µm	80.58 ± 36.17 (26.44–232.36)	128.19 ± 67.77 (35.65–573.06)
S25–26 20 µm	122.10 ± 32.92 (57.82–249.59)	173.83 ± 56.91 (76.64–557.85)
S25–26 35 µm	155.23 ± 29.28 (102.46–268.00)	213.38 ± 38.85 (135.29–487.81)
S26 70 µm	269.26 ± 29.88 (227.71–335.25)	415.48 ± 102.99 (286.46–755.91)
S27–28		
S27–28 10 µm	81.14 ± 35.83 (31.79–251.22)	129.79 ± 67.17 (40.54–559.98)
S27–28 20 µm	121.63 ± 29.60 (64.71–262.03)	173.46 ± 53.35 (79.47–546.13)
S27–28 35 µm	150.05 ± 24.84 (111.36–251.90)	205.61 ± 44.25 (142.27–488.88)
S29–S30		
S29–S30 10 µm	74.70 ± 35.92 (21.42–234.92)	118.69 ± 63.71 (39.76–577.49)
S29–S30 20 µm	125.93 ± 38.84 (61.51–286.99)	179.19 ± 61.83 (80.65–549.04)
S29–S30 35 µm	142.65 ± 23.53 (111.38–236.74)	196.79 ± 38.03 (139.17–417.40)

Dataset	Sphericity Mean ± Std [Min–Max]	Voxels Mean ± Std [Min–Max]
S21–22		
S21–22 10 µm	0.7953 ± 0.0935 (0.4500–0.9600)	419.68 ± 636.47 (27–7474)
S21–22 20 µm	0.8602 ± 0.0580 (0.4900–0.9600)	168.73 ± 166.43 (27–1615)
S21–22 35 µm	0.8562 ± 0.0436 (0.6600–0.9400)	45.49 ± 30.05 (27–215)
S23–24		
S23–24 10 µm	0.7977 ± 0.0965 (0.4500–0.9600)	441.04 ± 610.59 (27–6826)
S23–24 20 µm	0.8629 ± 0.0573 (0.5500–0.9600)	198.13 ± 172.80 (27–1779)
S23–24 35 µm	0.8551 ± 0.0399 (0.6800–0.9400)	41.60 ± 22.91 (27–163)
S25–26		
S25–26 10 µm	0.7939 ± 0.0957 (0.4450–0.9550)	493.79 ± 653.17 (27–6716)
S25–26 20 µm	0.8798 ± 0.0573 (0.5900–0.9700)	156.25 ± 123.70 (27–1024)
S25–26 35 µm	0.8846 ± 0.0406 (0.6150–0.9600)	59.94 ± 33.23 (27–256)
S26 70 µm	0.8370 ± 0.0855 (0.5700–0.9300)	39.58 ± 11.29 (27–68)
S27–28		
S27–28 10 µm	0.7938 ± 0.0958 (0.4500–0.9600)	485.18 ± 640.36 (27–8421)
S27–28 20 µm	0.8786 ± 0.0511 (0.6200–0.9600)	150.87 ± 113.96 (27–1204)
S27–28 35 µm	0.8780 ± 0.0361 (0.6700–0.9500)	54.27 ± 27.04 (27–211)
S29–S30		
S29–S30 10 µm	0.7956 ± 0.0945 (0.2700–0.9600)	412.30 ± 588.56 (27–6905)
S29–S30 20 µm	0.8609 ± 0.0600 (0.4400–0.9600)	179.68 ± 158.29 (27–1566)
S29–S30 35 µm	0.8675 ± 0.0407 (0.6200–0.9300)	47.92 ± 22.93 (27–178)

values around 0.79–0.80, indicating moderately irregular shapes. When the voxel size increases to 20 µm (Table A.2), both equivalent and linear diameters increase systematically, sphericity slightly improves, and voxel counts decrease due to coarser spatial sampling.

Table A.4 provides a detailed overview of equivalent diameter, linear diameter, sphericity, and voxel counts for all combined datasets while Tables A.1–A.3 present the average values per specimen pair for each resolution.

Data availability

Data will be made available on request.

References

[1] Carmignato S, Dewulf W, Leach R, editors. Industrial X-Ray computed tomography. first ed.. Cham: Springer; 2018, http://dx.doi.org/10.1007/978-3-319-59573-3.

[2] Thompson A, Maskery I, Leach R. X-ray computed tomography for additive manufacturing: a review. Meas Sci Technol 2016;27. http://dx.doi.org/10.1088/0957-0233/27/7/072001.

[3] Garcea S, Wang Y, Withers P. X-ray computed tomography of polymer composites. Compos Sci Technol 2017;156:305–19. http://dx.doi.org/10.1016/J.COMPSCITECH.2017.10.023.

[4] Wang Z, Georgarakis K, Nakayama K, Li Y, Tsarkov AA, Xie G, Dudina D, Dudina D, Louzguine-Luzgin DV, Yavari A. Microstructure and mechanical behavior of metallic glass fiber-reinforced Al alloy matrix composites. Sci Rep 2016;6. http://dx.doi.org/10.1038/srep24384.

[5] Ostapiuk M, Bieniaś J. Corrosion resistance in NaCl environment of fiber metal laminates based on aluminum and titanium alloys with carbon and glass fibers. Adv Eng Mater 2020;23. http://dx.doi.org/10.1002/adem.202001030.

[6] Wycisk E, Solbach A, Siddique S, Herzog D, Walther F, Emmelmann C. Effects of defects in laser additive manufactured Ti-6Al-4V on fatigue properties. Phys Procedia 2014;56:371–8. http://dx.doi.org/10.1016/j.phpro.2014.08.120.

[7] Yadollahi A, Shamsaei N. Additive manufacturing of fatigue resistant materials: Challenges and opportunities. Int J Fatigue 2017;98:14–31. http://dx.doi.org/10.1016/J.IJFATIGUE.2017.01.001.

[8] Dutton B, Vesga W, Waller J, James S, Seifi M. Metal Additive Manufacturing Defect Formation and Nondestructive Evaluation Detectability. ASTM Spec Tech Publ 2020;STP 1620:1–50. http://dx.doi.org/10.1520/STP162020180136.

[9] Kotadia H, Gibbons G, Das A, Howes P. A review of Laser Powder Bed Fusion Additive Manufacturing of aluminium alloys: Microstructure and properties. Addit Manuf 2021;46:102155. http://dx.doi.org/10.1016/j.addma.2021.102155.

[10] Nourian-Avval A, Fatemi A. Characterization and analysis of porosities in high pressure die cast aluminum by using metallography, X-ray radiography, and micro-computed tomography. Materials 2020;13(14). http://dx.doi.org/10.3390/ma13143068.

[11] Kim F, Pintar A, Moylan S, Garboczi E. The influence of X-Ray computed tomography acquisition parameters on image quality and probability of detection of additive manufacturing defects. J Manuf Sci Eng 2019;141(11):111012. http://dx.doi.org/10.1115/1.4044515.

[12] Fuchs P, Kröger T, Garbe CS. Defect detection in CT scans of cast aluminum parts: A machine vision perspective. Neurocomputing 2021;453:85–96. http://dx.doi.org/10.1016/j.neucom.2021.04.094.

[13] Senck S, Happl M, Scheerer M, Glinz J, Reiter T, Kastner J. Quantification of surface-near porosity in additively manufactured aluminum brackets using X-ray microcomputed tomography. 2021, http://dx.doi.org/10.2514/6.2021-0536.

[14] Galleguillos C, Senck S, Vilches C, Heupl S, Lasagni F. Comparison of defect detectability between computed tomography inspection and CT simulation using a calibrated defect phantom. E-Journal Nondestruct Test 2022;27. http://dx.doi.org/10.58286/26579.

[15] Matzkanin G, Yolken H. Probability of detection (POD) for nondestructive evaluation (NDE). 2001, p. 65.

[16] Aeronautics N, Administration S. Nondestructive evaluation requirements for fracture-critical metallic components. Tech. rep. NASA-STD-5009C, NASA; 2023, URL https://standards.nasa.gov/standard/NASA/NASA-STD-5009.

[17] Nudelis N, Mayr P. A novel classification method for pores in laser powder bed fusion. Metals 2021;11(12):1912. http://dx.doi.org/10.3390/met11121912.

[18] Senck S, Heupl S, Kastner J, Ehrenfellner P, Happl M, Reiter M. Simulation-based optimization of microcomputed tomography inspection parameters for topology-optimized aerospace brackets. 2022, http://dx.doi.org/10.2514/6.2022-0798.

[19] Ziabari A, Venkatakrishnan SV, Snow Z, Lisovich A, Sprayberry M, Brackman P, Plotkowski A. Enabling rapid X-ray CT characterisation for additive manufacturing using CAD models and deep learning-based reconstruction. Npj Comput Mater 2023;9:91. http://dx.doi.org/10.1038/s41524-023-01032-5.

[20] Berens AP. NDE reliability data analysis. In: ASM handbook, volume 17 - nondestructive evaluation and quality control. Materials Park, Ohio, US: ASM International; 1989, p. 689–701.

[21] Amrhein S, Rauer M, Kaloudis M. Characterization of computer tomography scanners using the probability of detection method. J Nondestruct Eval 2014;33:643–50. http://dx.doi.org/10.1007/s10921-014-0258-4.

[22] USDepartment of Defense. In: Handbook Non-Destructive Evaluation System Reliability Assessment, vol. MIL-HDBK-, Washington, D.C.: U.S. Department of Defense; 1999, p. 1823.

[23] Yosifov M, Reiter M, Heupl S, Gusenbauer C, Fröhler B, Fernández-Gutiérrez R, De Beenhouwer J, Sijbers J, Kastner J, Heinzl C. Probability of detection applied to X-ray inspection using numerical simulations. Nondestruct Test Eval 2022;37(5):536–51. http://dx.doi.org/10.1080/10589759.2022.2071892.

[24] Yosifov M, Weinberger P, Reiter M, Fröhler B, De Beenhouwer J, Sijbers J, Kastner J, Heinzl C. Defect Detectability Analysis via Probability of Defect detection Between Traditional and Deep Learning Methods in Numerical Simulations. E-Journal Nondestruct Test 2023;28(3). http://dx.doi.org/10.58286/27716.

[25] Yosifov M, Fröhler B, Sijbers J, Beenhouwer J, Kastner J, Heinzl C. Superimposing synthetic defects into real XCT data and segmentation-based comparison for advanced probability of detection evaluation. J Nondestruct Eval 2025;44. http://dx.doi.org/10.1007/s10921-025-01262-1.

[26] Ziabari A, Venkatakrishnan SV, Snow Z, et al. Enabling rapid X-ray CT characterisation for additive manufacturing using CAD models and deep learning-based reconstruction. Npj Comput Mater 2023;9:91. http://dx.doi.org/10.1038/s41524-023-01032-5.

[27] Tai JL, Sultan MTH, Shahar FS. Analysis of phased array corrosion mapping data using probabilistic detection (POD) method. Pertanika J Sci Technol 2025;33(4):1887–906. http://dx.doi.org/10.47836/pjst.33.4.10.

[28] MDPI Editorial. Advanced non-destructive testing simulation and modeling approaches for fiber-reinforced polymer pipes. Materials 2024;18(11):2466. http://dx.doi.org/10.3390/ma18112466.

[29] Tai JL, Sultan MTH, ukaszewicz AL, Józwik J, Oksiuta Z, Shahar FS. Recent trends in non-destructive testing approaches for composite materials: A review of successful implementations. Materials 2025;18(13):3146. http://dx.doi.org/10.3390/ma18133146.

[30] Sietins JM, Green WH, Jones JS. Material evaluation using X-ray computed tomography. In: Comprehensive structural integrity. second ed.. Elsevier; 2022, URL https://api.semanticscholar.org/CorpusID:261460815.

[31] Galvez-Hernandez P, Smith R, Gaska K, Mavrogordato M, Sinclair I, Kratz J. The effect of X-ray computed tomography scan parameters on porosity assessment of carbon fibre reinfored plastics laminates. J Compos Mater 2023;57:4535–48. http://dx.doi.org/10.1177/00219983231209383.

[32] Cooper D, Turinsky AL, Sensen C, Hallgrímsson B. Effect of voxel size on 3D micro-CT analysis of cortical bone porosity. Calcif Tissue Int 2007;80:211–9. http://dx.doi.org/10.1007/s00223-005-0274-6.

[33] Senck S, Happl M, Reiter M, Scheerer M, Kendel M, Glinz J, Kastner J. Additive manufacturing and non-destructive testing of topology-optimised aluminium components. Nondestruct Test Eval 2020;35(3):315–27. http://dx.doi.org/10.1080/10589759.2020.1774582.

[34] du Plessis A, Yadroitsev I, Yadroitsava I, le Roux SG. X-ray microcomputed tomography in additive manufacturing: A review of the current technology and applications. In: 3D printing and additive manufacturing, vol. 5, (3):2018, p. 227–47. http://dx.doi.org/10.1089/3dp.2018.0060.

[35] Withers PJ, Bouman C, Carmignato S, Cnudde V, Grimaldi D, Hagen CK, Maire E, Manley M, Du Plessis A, Stock SR. X-ray computed tomography. Nat Rev Methods Prim 2021;1(1):18. http://dx.doi.org/10.1038/s43586-021-00015-4.

[36] Holgado I, Ortega N, Yagüe-Fabra JA, Plaza S, Villarraga-Gómez H. Metrological evaluation and classification of porosity in metal additive manufacturing using X-ray computed tomography. Mater Des 2025;254:114057. http://dx.doi.org/10.1016/j.matdes.2025.114057.

[37] Spiess M, Nagl W, Hamerle A. Probit models: Regression parameter estimation using the ML principle despite misspecification of the correlation structure. 67, 1997, http://dx.doi.org/10.5282/ubm/epub.1461.

[38] Demidenko E. Computational aspects of probit model. In: 2, Math Commun In: 2, 1997;6:233–47.URL https://hrcak.srce.hr/829,

[39] Wang Z, Simoncelli E, Bovik A. Multiscale structural similarity for image quality assessment. In: Multiscale structural similarity for image quality assessment, vol. 2, 2003, p. 1398 – 1402 Vol.2. http://dx.doi.org/10.1109/ACSSC.2003.1292216.

[40] Bakurov I, Buzzelli M, Schettini R, Castelli M, Vanneschi L. Structural similarity index (SSIM) revisited: A data-driven approach. Expert Syst Appl 2022;189:116087. http://dx.doi.org/10.1016/j.eswa.2021.116087.

[41] Sheikh H, Sabir M, Bovik A. A statistical evaluation of recent full reference image quality assessment algorithms. IEEE Trans Image Process 2006;15(11):3440–51. http://dx.doi.org/10.1109/TIP.2006.881959.

[42] Behammer L, Weinberger P, Nepelius L, Fröhler B, Yosifov M, Kastner J, Senck S. Correlation of image quality metrics with expert perception for industrial computed tomography. E-Journal Nondestruct Test 2025;30. http://dx.doi.org/10.58286/30753.

[43] Pearson K. VII. Mathematical contributions to the theory of evolution.—III. Regression, heredity, and panmixia. Philos Trans R Soc Lond Ser A: Contain Pap A Math Or Phys Character 1896;(187):253–318. http://dx.doi.org/10.1098/rsta.1896.0007.

[44] Spearman C. The proof and measurement of association between two things. Am J Psychol 1904;15(1):72–101. http://dx.doi.org/10.2307/1412159.

[45] Snell R, Tammas-Williams S, Chechik L, Prangnell PB. Methods for rapid pore classification in metal additive manufacturing. JOM 2020;72:101–9. http://dx.doi.org/10.1007/s11837-019-03761-9.

[46] Shah S, Gray F, Crawshaw J, Boek E. Micro-computed tomography pore-scale study of flow in porous media: Effect of voxel resolution. Adv Water Resour 2016;95:276–87. http://dx.doi.org/10.1016/j.advwatres.2015.07.012.

[47] MIL-HDBK-1823a: Nondestructive evaluation system reliability assessment. 2009.

[48] Senck S, Weinberger P, Nepelius L, Haghofer A, Woegerer B, Glinz J, Yosifov M, Behammer L, Kastner J, Trieb K, Kranioti E, Winkler S. Optimizing μ CT resolution in tarsal bones: A comparative study of super-resolution models for trabecular bone analysis. Tomogr Mater Struct 2025;8:100063. http://dx.doi.org/10.1016/j.tmater.2025.100063.

[49] Klos A, Salvo L, Lhuissier P. Effective super-resolution X-ray tomography using MSDnet for nondestructive testing of metallic lattices: Analysis of training dynamics and strategies. J Nondestruct Eval 2025;44:142. http://dx.doi.org/10.1007/s10921-025-01273-y.

[50] Li J, Gsaxner C, Pepe A, Schmalstieg D, Kleesiek J, Egger J. Sparse convolutional neural network for high-resolution skull shape completion and shape super-resolution. Sci Rep 2023;13(1):20229. http://dx.doi.org/10.1038/s41598-023-47437-6.

[51] Klos A, Salvo L, Lhuissier P. Super-resolution X-ray tomography using deep learning applied to the 3D quantification of defects in lattice structures. Sci Rep 2025;15. http://dx.doi.org/10.1038/s41598-025-20372-4.

[52] Jones G, Sundar V, Reed R, Stecko M, Keist J. Validation of X-ray computed tomography detection limits for stochastic flaws in additively manufactured Ti-6Al-4 V. J Mater Eng Perform 2024;34. http://dx.doi.org/10.1007/s11665-024-10383-z.

[53] Perghem D, Salehnasab B, Beretta S, Shao S, Shamsaei N. ML-based detection of critical defects in additively manufactured parts via X-ray computed tomography. Mater Des 2025;260:116053. http://dx.doi.org/10.1016/j.matdes.2025.115184.

[54] Ledwaba T, Steenkamp C, Moore K, Kouprianoff D, du Plessis A. X-ray computed tomography-based porosity segmentation in additive manufacturing: comparison of traditional and AI approaches. In: MATEC web of conferences, vol. 406, 2024, p. 05011. http://dx.doi.org/10.1051/matecconf/202440605011.

[55] Era IZ, Ahmed I, Liu Z, Das S. An unsupervised approach towards promptable porosity segmentation in X-ray computed tomography of laser powder bed fusion parts using the segment anything model. Npj Adv Manuf 2025;2:10. http://dx.doi.org/10.1038/s44334-025-00021-4.

[56] Wilbig J, Wilson-Heid AE, Bernard L, Baptista J, Obaton A-F. Comparison of porosity analysis based on X-ray computed tomography on metal parts produced by additive manufacturing. Appl Sci 2025;15(18):9876. http://dx.doi.org/10.3390/app15189876.

[57] Villarraga-Gómez H, Brackman P, Ziabari A, et al. Multiscale characterization of additive manufacturing components using X-ray computed tomography and X-ray microscopy. J Nondestruct Eval 2025. http://dx.doi.org/10.1007/s10921-025-01231-8.

[58] Duba-Sullivan H, Rahman O, Venkatakrishnan S, Ziabari A. 2.5D super-resolution approaches for X-ray computed tomography of additively manufactured parts. 2024, arXiv preprint arXiv:2412.04525. URL https://arxiv.org/abs/2412.04525.

[59] Murakami Y, Endo M. Effects of defects, inclusions and inhomogeneities on fatigue strength. Int J Fatigue 1994;16(3):163–82. http://dx.doi.org/10.1016/0142-1123(94)90001-9.

[60] Murakami Y. Metal fatigue: effects of small defects and nonmetallic inclusions. Academic Press; 2019, http://dx.doi.org/10.1016/C2016-0-05272-5.

[61] Tammas-Williams S, Withers PJ, Todd I, Prangnell PB. The influence of porosity on fatigue crack initiation in additively manufactured titanium components. Sci Rep 2017;7:7308. http://dx.doi.org/10.1038/s41598-017-06504-5.

[62] Romano S, Brückner-Foit A, Brandão A, Gumpinger J, Ghidini T, Beretta S. Fatigue properties of AlSi10Mg obtained by additive manufacturing: Defect-based modelling and prediction of fatigue strength. Eng Fract Mech 2018;187:165–89. http://dx.doi.org/10.1016/j.engfracmech.2017.11.002, SI: 50th Anniversary Issue.

[63] Karim MA, Tanvir G, Jadhav S, Islam S, Kim Y-M, Villarraga-Gomez H, Lee H-J, Jeon Y, Kim DB. Tailoring porosity and mechanical properties of wire-based directed energy deposited molybdenum alloys through hot isostatic pressing. Appl Mater Today 2025;42:102618. http://dx.doi.org/10.1016/j.apmt.2025.102618.

[64] Chuirazzi W, Kancharla RR, Morankar S. Laboratory-based micro- X-ray Computed Tomography of Energy Materials at Idaho National Laboratory. JOM 2025;77(3):1337–53. http://dx.doi.org/10.1007/s11837-024-06970-z.

[65] Zhao R, Shmatok A, Fischer R, Prorok BC. Investigation of causal relationships between printing parameters, pore properties and porosity in laser powder bed fusion. Metals 2023;13(2). http://dx.doi.org/10.3390/met13020330.

[66] Brooks RA, Chiro GD. Beam hardening in X-ray reconstructive tomography.. Phys Med Biol 1976;21 3:8–19. http://dx.doi.org/10.1088/0031-9155/21/3/004.

[67] Joseph PM, Spital R. The effects of scatter in X-ray computed tomography.. Med Phys 1982;9 4:464–72. http://dx.doi.org/10.1118/1.595111.

[68] Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-Net: Learning dense volumetric segmentation from sparse annotation. In: Medical image computing and computer-assisted intervention – MICCAI.. Lecture notes in computer science, vol. 9901, Springer, Cham; 2016, p. 424–32. http://dx.doi.org/10.1007/978-3-319-46723-8_49.