Theoretical and Applied Fracture Mechanics 147 (2026) 105875

Contents lists available at ScienceDirect

Theoretical and Applied Fracture Mechanics

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

Fatigue of AM TiAl6V4: Benchmark of defect criticality models

Sebastian Mansky [a,*], Dirk Herzog [a,b], Ingomar Kelbassa [a,b]

[a] Fraunhofer Research Institution for Additive Manufacturing Technologies IAPT, Am Schleusengraben 14, Hamburg, 21029, Germany
[b] Institute for Industrialization of Smart Materials ISM, TU Hamburg, Harburger Schloßstraße 28 (Channel 4), Hamburg, 21079, Germany

ARTICLE INFO

Keywords:
Additive manufacturing
X-ray computed tomography
Defects
Fatigue
Defect criticality model

ABSTRACT

Internal defects in laser powder bed fusion (PBF-LB/M) components are the primary drivers of fatigue failure, yet no systematic comparison of defect criticality models exists for ranking detected defects by severity. This work addresses two complementary research questions: (i) *which defect is most critical?* and (ii) *what fatigue-life does the critical defect imply?* For the first question, nine defect criticality ranking models, grouped into three geometry-based approaches (size ranking, Murakami $\sqrt{\text{Area}}$ model, Tammas-Williams model) and six simulation-based approaches (stress–strain concentration factor, averaged strain energy density, relative stress gradient, Smith–Watson–Topper parameter, equivalent plastic strain, Theory of Critical Distances), are benchmarked on two series of PBF-LB/M TiAl6V4 specimens (93 fatigue-tested; 84 of these inspected by X-ray CT prior to testing; 73 fractographically analysed after testing). The simulation-based models are evaluated with both a linear-elastic and an elastic–plastic material model. A matching procedure links CT defect data to fractographically identified fracture origins (21 of 84 CT-inspected specimens matched). For the second question, three crack growth models (Hartman-Schijve, Murakami–Endo, and a novel hybrid formulation) are compared for fatigue-life prediction based on the identified critical defect. Results show that geometry-based models achieve comparable ranking accuracy to computationally expensive simulation-based approaches, and that distinguishing between internal and surface crack growth environments is essential for accurate life prediction.

1. Introduction

Laser powder bed fusion of metals (PBF-LB/M) has matured from a prototyping technique to a viable production process for structural components in aerospace, medical and energy applications [1,2]. The layer-wise manufacturing principle enables geometries that are difficult or impossible to achieve with conventional subtractive or formative processes, but it also introduces process-inherent defects such as gas porosity, lack-of-fusion (LOF) voids and keyhole pores [3]. These defects act as stress concentrators and are widely recognised as the primary driver limiting the fatigue performance of PBF-LB/M components [4,5]. For safety-critical applications, regulatory frameworks increasingly demand a quantitative understanding of how individual defects affect structural integrity, shifting the design philosophy from a safe-life towards a damage-tolerant approach [5,6].

The fatigue strength of PBF-LB/M components is governed less by the bulk material properties than by the size, morphology, location and local stress state of the most critical defect in the loaded volume [3,7]. Surface roughness and residual stresses constitute additional influencing factors [8,9]; however, when the surface is machined and a stress-relief heat treatment is applied, as is common practice for fatigue-critical parts, internal defects remain as the dominant life-limiting feature [4,10]. The characterisation of the internal defect population is therefore a prerequisite for any defect-tolerant fatigue assessment.

Non-destructive testing (NDT) by X-ray computed tomography (X-ray CT) provides three-dimensional information on the internal defect population of a part prior to service [11,12]. With volumetric defect data available, the natural question arises: *which of the detected defects is most likely to initiate a fatigue crack?* Answering this question requires a defect criticality model that ranks the detected defects according to their expected severity. Such models vary widely in complexity and in the type of input data they require: from simple geometric measures extracted directly from the X-ray CT output, through semi-analytical fracture mechanics estimates, to individual finite element simulations that resolve the full three-dimensional defect morphology.

Among the geometry-based approaches, the $\sqrt{\text{Area}}$ parameter introduced by Muarakami [13] relates the projected defect area to an equivalent stress intensity factor and has become the most widely

* Corresponding author.
E-mail address: sebastian.mansky@iapt-extern.fraunhofer.de (S. Mansky).

https://doi.org/10.1016/j.tafmec.2026.105875
Received 5 June 2026; Received in revised form 12 August 2026; Accepted 13 August 2026
Available online 29 August 2026

Nomenclature

$\bar{\eta}$	Mean logarithmic error (bias)
$\bar{R}$	Mean normalised rank
$\bar{W}$	Averaged strain energy density
χ^*	Relative stress gradient
ΔK	Stress intensity factor range
ΔK_{thr}	HS fitting threshold
ΔK_t	Stress concentration factor
ΔK_{th}	Threshold stress intensity factor range
$\Delta\sigma$	Applied stress range
$\Delta\sigma_0$	Fatigue strength range
η_i	Logarithmic life ratio
$\mathbf{d}$	Defect parameter set
$\mathbb{1}(\cdot)$	Indicator function
$\mathrm{d}a/\mathrm{d}N$	Crack growth rate
ν	Poisson's ratio
ρ	Relative mean density
σ_1	First principal stress
σ_∞	Nominal stress
σ_{net}	Net-section stress
$\sigma_{\mathrm{supported}}$	Support-effect corrected stress
σ_{UTS}	Ultimate tensile strength
σ_w	Crack-size-dependent fatigue limit
σ_y	Yield strength
σ_{max}	Maximum principal stress
σ_{TCD}	TCD ranking parameter
$\sqrt{\mathrm{Area}_{\mathrm{eff}}}$	Effective Murakami parameter
$\sqrt{\mathrm{Area}}$	Murakami parameter, square root of the defect area normal to the maximum stress
ε_∞	Nominal strain
$\varepsilon_{\mathrm{pl,eq}}$	Equivalent plastic strain
$\varepsilon_{\mathrm{max}}$	Maximum principal strain
a	Crack size (in terms of $\sqrt{\mathrm{Area}}$)
a^*	Material-dependent support factor
a_0	Initial crack size
A_{crack}	Crack area
A_{HS}	Cyclic fracture toughness
A_{total}	Specimen cross section area
a_c	Critical crack length at failure
C^*, m^*, n^*	ME crack growth coefficient, stress exponent, crack size exponent
D, p	HS crack growth coefficient, exponent
E	Young's modulus
H	Hit rate
HV	Vickers hardness
K	Stress intensity factor
K_{max}	Maximum stress intensity factor
k_g	Stress–strain concentration factor
k_σ	Stress factor
k_ε	Strain factor
K_{Ic}	Static fracture toughness
L	Critical distance (TCD)
l_n	Distance from defect to surface
$l_{n,\mathrm{thr}}$	Ligament threshold for surface/internal classification
N_f	Number of cycles to failure
n_i	Number of defects in specimen i
N_m	Number of matched specimens per series
N_{air}	Crack growth cycles in air
N_{int}	Internal crack growth cycles in self-contained environment
$N_{f,\mathrm{exp}}$	Total experimental life
$N_{f,\mathrm{pred}}$	Total predicted life
P_{SWT}	SWT fatigue indicator parameter
R	Stress ratio
R_c	SED control radius
r_i	Rank assigned to the fracture-initiating defect
Ra	Arithmetic mean roughness
V_c	Control volume (SED)
V_P	Volume of all detected defects
V_S	X-ray CT scanned volume
$W(\mathbf{x})$	Local strain energy density
Y	Geometry factor
AM	Additive Manufacturing
AUC	Area Under the ROC Curve
CoV	Coefficient of Variation
EP	Elastic–plastic
FE	Finite element
HS	Hartman-Schijve
HSV	Highly stressed volume
LE	Linear-elastic
LOF	Lack-of-fusion
ME	Murakami–Endo
NDT	Non-destructive testing
PBF-EB/M	Powder Bed Fusion - Electron Beam/Melting
PBF-LB/M	Powder Bed Fusion - Laser Beam/Melting
$\mathrm{RMSE}_{\log}$	Root-mean-square error of logarithmic life ratio
SED	Strain Energy Density
SWT	Smith–Watson–Topper
TCD	Theory of Critical Distances
X-ray CT	X-ray computed tomography

adopted metric for small defects and inclusions. Its application to additively manufactured materials has been demonstrated for TiAl6V4 [7] and AlSi10Mg [14], often in combination with extreme-value statistics to predict the largest expected defect in a given volume [15,16]. Tammas-Williams et al. [17] extended the $\sqrt{\mathrm{Area}}$ concept by incorporating stress concentration factors derived from a-priori finite element simulations of idealised spherical defects, thereby accounting for surface proximity, defect interaction and aspect ratio effects. While this approach captures first-order geometric interactions, it retains the idealisation of defects to spherical or ellipsoidal shapes, which may not adequately represent the tortuous LOF morphologies typical of PBF-LB/M processes [18].

The limitations of geometric idealisations have motivated a growing body of work on simulation-based defect assessment, as reviewed by Torries et al. [19] and Awd et al. [20]. These approaches range from macroscopic continuum finite element (FE) models to crystal-plasticity frameworks that resolve the microstructure around individual defects. At the continuum level, classical elastic or elastic–plastic (EP) FE analyses can quantify stress and strain concentrations at realistic, X-ray CT-derived defect geometries without idealising the defect shape. The choice of constitutive model has a direct influence on the predicted local fields: linear-elastic (LE) models tend to overestimate peak stresses at sharp notch roots, while EP formulations with

isotropic, kinematic or combined hardening capture stress redistribution through local yielding [21]. At a finer scale, crystal-plasticity finite element models have been used to study fatigue crack nucleation around pores in AlSi10Mg [22] and the combined effect of microstructure and defect morphology on fatigue damage accumulation in dual-phase TiAl6V4 [23]. Mikkola et al. [24] demonstrated mesoscale modelling of crack nucleation from defects in steel, linking microstructural shear stress and plastic strain to critical configurations even near the endurance limit. While crystal-plasticity finite element models offer the highest physical fidelity, their computational cost currently limits their application to small numbers of defects or representative volume elements rather than to the ranking of entire defect populations comprising hundreds of pores. For population-level ranking, continuum-scale FE sub-models therefore remain the most practical approach.

Within the continuum FE framework, several defect criticality ranking strategies have been proposed. Li et al. [25] and Gao et al. [26] introduced a ranking based on a combined stress–strain concentration factor k_g extracted from individual defect simulations, originally applied to cast aluminium alloys. This methodology was adopted by Tammas-Williams et al. [17] for PBF-EB/M (Powder Bed Fusion - Electron Beam/Melting) TiAl6V4, by Afroz et al. [27] for PBF-LB/M AlSi10Mg with a LE material model, and by Li et al. [28] for PBF-LB/M TiAl6V4 with nonlinear isotropic-kinematic hardening. Beyond the k_g factor, other fatigue-relevant metrics can be extracted from the same FE sub-model results, including the averaged strain energy density [29,30], relative stress gradients [31], and the Smith–Watson–Topper (SWT) fatigue indicator parameter [32]. However, these alternative metrics have not been systematically compared as defect criticality ranking criteria. An alternative framework is provided by the Theory of Critical Distances (TCD), which treats defects as notches with finite tip radii rather than idealised cracks [33–35]. The TCD has been applied to notched metallic components and, more recently, to additively manufactured surfaces by Barricelli et al. [36] and to CT-detected internal defects by Gillham et al. [37]. However, its systematic use for ranking entire defect populations within a part has not yet been demonstrated. Furthermore, data-driven approaches combining defect features with machine-learning models have shown promise for fatigue-life prediction [20,38], yet they require large training datasets and do not directly provide a physically interpretable defect criticality ranking.

Despite the growing number of individual studies, a systematic comparison of defect criticality models, spanning from simple geometric metrics to full FE-based approaches, on the *same* defect population and validated against experimentally identified fracture origins is lacking. Previous benchmarks, such as the work of Tammas-Williams et al. [17] on PBF-EB/M TiAl6V4, were limited to a small number of models, a single specimen geometry, and did not investigate the influence of the constitutive model used in the FE simulations. The latter aspect is particularly relevant, as the choice between LE, EP and cyclic-plasticity formulations affects both the magnitude and distribution of the predicted stress concentrations [19,20]. Furthermore, the role of the specimen geometry, specifically the highly stressed volume (HSV), on the predictive capability of the models has not been addressed. A meaningful benchmark requires (i) a well-characterised and representative defect population obtained by X-ray CT, (ii) fatigue tests to failure, and (iii) a reliable matching between the fracture-initiating defect identified on the fracture surface and the corresponding defect in the CT data. This matching is non-trivial and depends on the X-ray CT resolution relative to the defect size [12,39].

This work addresses two complementary aspects of defect-tolerant fatigue assessment for PBF-LB/M components. The **first part** benchmarks nine defect criticality models on two series of PBF-LB/M TiAl6V4 fatigue specimens that differ in gauge section volume and, consequently, in the competition of defect-controlled failure. All specimens are inspected by X-ray CT before testing, and a matching procedure links the CT defect data to the fracture surfaces after testing. The defect criticality models are grouped into two categories: three *geometry-based* models (defect size ranking, Murakami $\sqrt{\text{Area}}$ model [13], and Tammas-Williams model [17]) and six *simulation-based* models (k_g factor following Li et al. [25], averaged strain energy density, relative stress gradient, SWT parameter, maximum equivalent plastic strain, and TCD [33,35]). For the simulation-based models, a semi-automated workflow is developed that isolates, meshes and simulates each defect from the CT reconstruction as a finite element sub-model, from which multiple ranking metrics are extracted. Two material model variants, LE and EP, are investigated to quantify the influence of the constitutive model on the ranking outcome.

The **second part** addresses the subsequent research question: given the identified critical defect, what fatigue-life does it imply? Three crack propagation models of increasing complexity are compared: (i) the Hartman-Schijve (HS) equation with dual-environment parameters (surface-in-air vs. internal) [40,41], (ii) the parameter-free Murakami–Endo (ME) scS-N model [42,43], and (iii) a novel hybrid formulation that combines the Murakami-derived, crack-size-dependent threshold with the HS framework and its environmental sensitivity. A key hypothesis tested is that the distinction between surface crack growth in air and internal crack growth in self-contained environment conditions is essential for accurate fatigue-life prediction of PBF-LB/M components.

The remainder of this paper is structured as follows. Section 2 describes specimen manufacturing, X-ray CT inspection, fatigue testing and defect matching, followed by the defect criticality models (Part A) and the crack growth models (Part B). Section 3 presents the ranking benchmark results and the fatigue-life predictions. Section 4 summarises the findings and discusses implications for practical defect-tolerant design of PBF-LB/M components.

2. Methodology

The overall methodology of this work is structured in two main parts, preceded by a common experimental basis. The experimental basis comprises manufacturing, X-ray CT inspection, fatigue testing, and defect matching (Sections 2.1–2.4).

2.1. Manufacturing

To assess the influence of fatigue specimen shapes on fatigue results, two different series of TiAl6V4 (Grade 5) fatigue test samples are manufactured with a PBF-LB/M process. To reflect the current possibilities of industrial additive manufacturing (AM), the samples are manufactured by an industrial supplier on a TruPrint 5000 (TRUMPF, Ditzingen, Germany) machine. Due to external manufacturing, the process parameters and powder properties cannot be disclosed. The production parameters are designed to produce high-quality, dense parts. A subset of the first build job specimens are manufactured horizontally, the following results focus exclusively on the vertical specimens. In the second series, all samples are manufactured vertically. All samples are manufactured as cylindrical blanks and machined down for fatigue testing. This is done to ensure that the surface finish meets the standard requirement of $Ra < 0.8\,\mu\text{m}$ [44] and to completely remove subsurface porosity, both of which are shown to be detrimental to fatigue-life [4,45]. Since machining removes the as-built surface and any subsurface porosity, all *CT-detected* defects lie within the bulk material. Defects that intersect the final machined surface are nevertheless classified as surface-connected in the fractographic analysis (Section 2.4). The estimated fatigue-life properties therefore reflect the hatch-region defect population of the manufacturing process. The manufacturing of cylindrical blanks ensures a constant interlayer time for the relevant areas. The build platform is exclusively used to manufacture these blanks, no other parts are added to the platform. Thermal variations due to changes in the scan area per layer are minimised. Even though the fatigue specimens from both series are manufactured in different

build jobs, a comparison is considered valid because the production parameters are left unchanged. Since the production parameters target high-quality, dense parts with standard industrial settings, the resulting defect population is considered representative of current PBF-LB/M practice, ensuring that the subsequent criticality metrics are not biased towards a specific defect type or morphology.

After manufacturing, the samples are stress-relieved below β-transus at 710 °C for 130 min under vacuum, transforming the initial martensitic α' structure into a mixture of α and β phases within prior-β grains [46]. The resulting microstructure is consistent with literature reports [46–48] and is documented in Supplementary Material S1.

The shape and size of the specimens are varied between both testing series. To avoid confusion, in the following sections the first series will be referred to as 'small' and the second series will be named 'large'. The series 'small' is manufactured according to DIN EN 6072 FCE A12.5 [44]. This specimen geometry creates a concentrated HSV, allowing a fatigue property assessment with a reduced impact of internal defects. The HSV is defined as the volume in which the local maximum principal stress $\sigma_1(x)$ is equal to or larger than 90 % of the nominal stress σ_∞. The specimen geometry 'large' is chosen to increase the effect of defects on the fatigue-life, by increasing the volume and HSV, which in turn increases the competition between defects. Series 'large' consists of a cylindrical gauge section with blended radii at both ends for a constant HSV which is approximately 13 times larger. The two series serve complementary purposes: the 'small' geometry concentrates a small highly stressed volume (HSV) to characterise the baseline fatigue strength with minimal defect competition, whereas the 'large' geometry deliberately increases the HSV (≈13×) to promote defect-controlled failure and thus enable the CT-to-fracture matching required for the ranking benchmark.

The specimen geometries of both series and their respective HSV are shown in Fig. 1. Series 'small' follows DIN EN 6072 FCE A12.5 [44] (gauge length 18 mm, minimal gauge diameter 3.99 mm, transition radius 28 mm); series 'large' has a gauge length of 43 mm, a gauge diameter of 6 mm and a compound transition radius (R25.15 + R50 + R300). A total of 50 specimens were tested in series 'small' and 43 in series 'large', of which 41 and 43, respectively, were inspected by X-ray CT prior to testing (see Table 5 for an overview of all subsequent subsets).

2.2. X-ray computed tomography

After machining to the final geometry, all parts undergo non-destructive testing via X-ray CT. Series 'small' is scanned in full, including the threaded ends, which limits the achievable voxel size to 25.6 µm; for series 'large' only the gauge section is scanned, permitting 14 µm. The CT acquisition parameters are kept identical for both series (Table 1). At 25.6 µm, defects near the detection limit cannot be reliably detected or accurately reconstructed, which is the primary reason for the lower matching rate in series 'small' (Section 2.4). To avoid confounding this resolution effect, the ranking metrics are evaluated *within* each series and are not pooled across resolutions. The resulting defect populations are characterised in Section 3.2.

Fig. 2 shows two representative CT-reconstructed defects spanning the observed size range: a smaller defect from series 'small' ($\sqrt{\text{Area}_{\text{eff}}}$ = 51.39 µm) and a larger defect from series 'large' ($\sqrt{\text{Area}_{\text{eff}}}$ = 206.46 µm), each shown before and after the smoothing step. Defects represented by only a small number of voxels are more susceptible to voxel-induced staircase artefacts that reduce the morphological accuracy of the reconstructed surface. For larger defects, these artefacts are present but do not significantly alter the reconstructed defect shape. The rescaling and smoothing workflow described in Section 2.7.2 and in more detail in [39] reduces these artefacts and improves the geometric accuracy of the defect surface for FE sub-modelling.

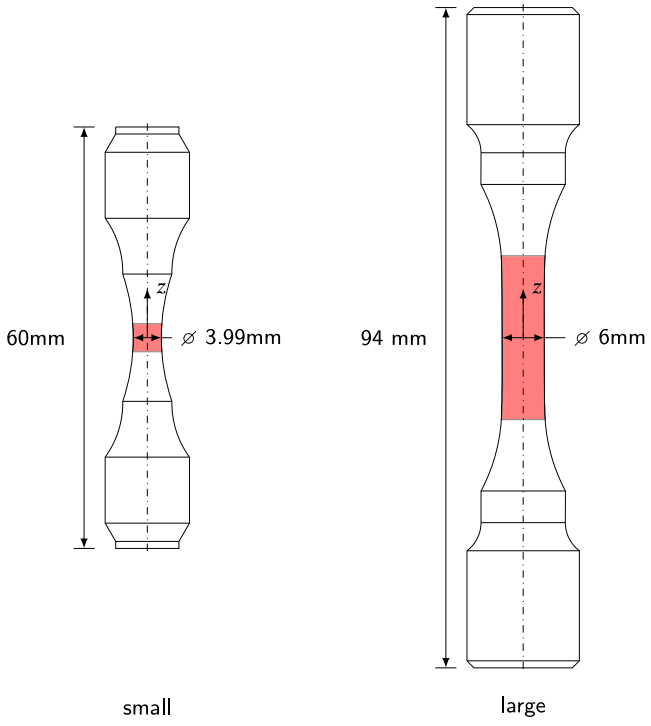


Fig. 1. Specimen geometries from both testing series. The red areas mark the highly stressed volume, where the maximum principal stress $\sigma_1(z)$ is at least 90 % of the nominal stress σ_∞. The origin is set at the centre of symmetry, with z aligned to the build direction per DIN EN ISO 17295 [49].

Table 1
X-ray CT acquisition parameters (identical for both series).

Parameter	Acc. voltage	Current	Exposure	Projections	Integrations
Value	260 kV	100 µA	300 ms	1001	4

2.3. Fatigue testing

Fatigue testing is performed at MPA Stuttgart on a Testronic 50 kN (Rumul, Switzerland) resonance test machine in accordance with DIN 50100 [50] at room temperature under axial tension-compression loading in force-controlled mode with a stress ratio of $R = -1$ at a test frequency of 70 Hz to 90 Hz. The applied waveform is sinusoidal. The runout criterion is set to $N = 3 \times 10^6$ cycles.

The two specimen series serve different experimental purposes. Series 'small' is used to establish the baseline fatigue behaviour of the machined and stress-relieved PBF-LB/M TiAl6V4 material with reduced defect competition due to the smaller HSV. Series 'large' is then tested at stress levels selected relative to this baseline in order to promote defect-controlled failures in a larger HSV and thereby increase the probability of CT/fractography matching.

For series 'small', two complementary fatigue test procedures are used. First, 26 specimens are tested at multiple stress amplitudes using a pearl-string approach to establish the finite-life P-S-N behaviour. Second, 24 specimens are tested by the staircase method to determine the fatigue strength at 3×10^6 cycles. Of these staircase specimens, 12 runout specimens are subsequently retested to failure or until 1×10^7 cycles. Of these retested specimens, 2 remain runouts at 1×10^7 cycles, while the remaining 10 fail before reaching this limit. The retests are shown in Fig. 4 but are not included in the fatigue-strength evaluation. Consecutive staircase stress levels differ by a factor of 1.059.

For series 'large', two stress amplitudes are selected relative to this reference fatigue strength: one below, σ_a = 530 MPa, and one above, σ_a = 600 MPa. In total, 34 specimens are tested at 530 MPa and 9

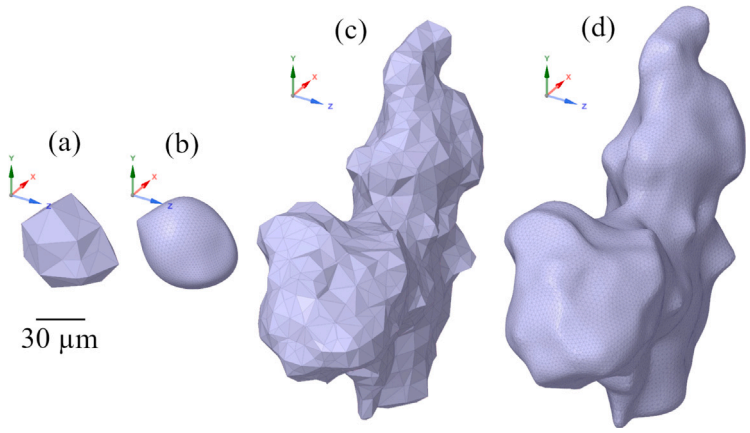


Fig. 2. Representative CT-reconstructed defects before and after the smoothing and rescaling step applied prior to FE sub-modelling. (a) Raw voxel-based reconstruction, series 'small' ($\sqrt{\text{Area}_{\text{eff}}}$ = 51.39 µm); (b) same defect after smoothing and rescaling; (c) Raw reconstruction, series 'large' ($\sqrt{\text{Area}_{\text{eff}}}$ = 206.46 µm); (d) same defect after smoothing and rescaling. Voxel-induced staircase artefacts visible in (a) and (c) are substantially reduced by the smoothing procedure.

specimens at 600 MPa. This test design is chosen to probe whether the larger HSV shifts failure towards defect-dominated behaviour even at a nominal stress amplitude below the fatigue strength derived from series 'small'.

In total, 93 specimens are fatigue-tested in the present study: 50 in series 'small' and 43 in series 'large'. The resulting fatigue data and the corresponding P-S-N representation are presented in Section 3.1. The complete specimen-level dataset, including stress amplitudes, cycles to failure, fractographic measurements ($\sqrt{\text{Area}}$, l_n), CT-derived defect data, and the classification of each specimen for the crack growth models, is provided in Supplementary Material S5.

2.4. Defect matching

After fatigue testing, the fracture surfaces are analysed by digital optical microscopy (VHX5000, Keyence, Japan). Optical microscopy is selected because it provides a sufficiently large field of view to document the complete fracture origin region and both fracture halves efficiently, while still allowing geometric measurement of the initiating defect. Higher-resolution SEM imaging is not required for the present purpose, since the matching procedure is based primarily on defect position, projected defect geometry, and ligament distance rather than on sub-micrometre fractographic features.

For each specimen, the fracture-initiating defect is identified on the fracture surface. From the optical micrographs, the projected defect area and the nearest distance to the specimen free surface (l_n) are measured using ImageJ [51]. The reported values of $\sqrt{\text{Area}}$ and l_n represent the mean of four manual measurements obtained from two viewing angles and from both fracture halves. The corresponding minimum and maximum values are used in the sensitivity study (Section 3.5.3). The fracture origin is further classified as surface-connected or internal based on its ligament to the free surface.

To assess the ranking accuracy of the defect criticality models, the fracture-initiating defect identified by fractography is matched to the pre-test X-ray CT defect population. For this purpose, the fracture image and the segmented CT defects are imported into a CAD environment. The fracture surface image is aligned with the specimen coordinate system, scaled to the specimen diameter, and translated along the specimen axis to the experimentally documented fracture location. Candidate CT defects are then evaluated based on their axial position, radial distance from the specimen axis, projected shape, and relative position with respect to the specimen surface. A match is accepted only when these criteria are jointly consistent. Fig. 3 illustrates the procedure for a representative specimen from series 'large': the full CT-detected defect population (a) and the optical fracture surface micrograph (b) are aligned in the CAD environment, and the confirmed match between the CT defect and the fractographic fracture origin is shown in (c).

For the comparison between fractography and CT, the CT defect size is defined by the projected area of the segmented three-dimensional defect normal to the loading direction, expressed as the Murakami parameter $\sqrt{\text{Area}}$. This definition ensures consistency with the fractographic measurement, which is likewise based on the projected initiation area on the fracture surface. A perfect quantitative agreement between both measurements is not expected, since they rely on fundamentally different measurement principles: CT-based segmentation involves greyscale thresholding and is subject to partial-volume effects, whereas fractographic measurement is a direct projection of the fracture surface. Systematic deviations between CT-based and microscopy-based defect size measurements are well documented in the literature and are attributed to CT scan parameters, segmentation threshold choice and partial-volume effects at the defect boundary [12,39]. A quantitative comparison of CT-derived and fractographically measured defect sizes for the 21 matched specimens is presented in Section 3.2 (Fig. 6).

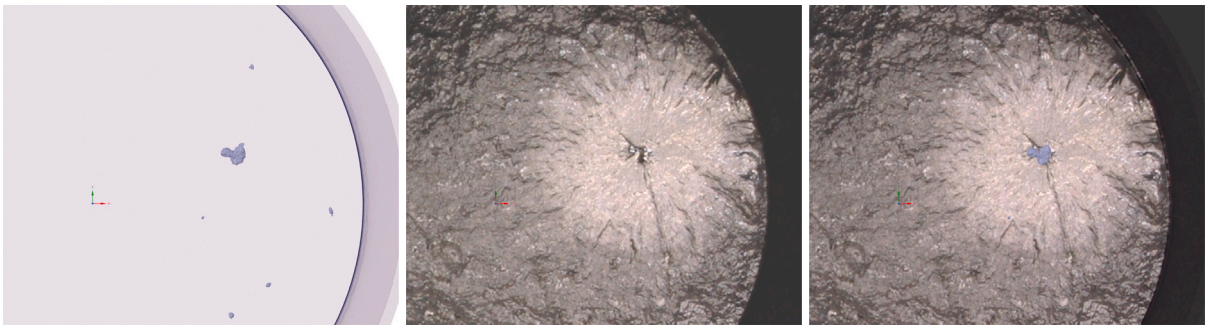
Of the 93 fatigue tested specimens, 84 specimens were available for the matching procedure. The remaining 9 specimens (all from series 'small') were not inspected by X-ray CT prior to fatigue testing and therefore could not contribute to the matching procedure. The matching rate itself is considered an experimental result rather than only a methodological limitation. In series 'small', only 5 of 41 specimens could be matched, whereas in series 'large' 16 of 43 specimens are matched successfully. This difference reflects not only the finer voxel size in series 'large', but also the larger HSV, which shifts the fracture-initiating defects towards larger and therefore more reliably detectable defects. The present dataset therefore highlights the practical difficulty of closing the chain from non-destructive inspection to fracture-origin validation, especially when the expected critical defects are close to the CT detection limit. A more detailed discussion of the limiting factors for successful matching is provided in Supplementary Material S2.

2.5. Evaluation metrics

The performance of the models is assessed separately for the two parts of this study.

2.5.1. Part A: Ranking accuracy

The purpose of the ranking analysis is to evaluate how well each defect criticality model identifies the defect that actually initiated fatigue failure. For each matched specimen, the experimentally observed fracture-initiating defect is known from fractography and is located within the pre-test CT defect population by the matching procedure described in Section 2.4. Each defect criticality model is then applied



(a) CT reconstruction: defect population (b) Optical micrograph: fracture surface (c) Overlay: matched CT defect and fracture origin

Fig. 3. Representative example of the CT-to-fractography matching procedure (series 'large'). (a) X-ray CT reconstruction showing all segmented defects detected in the gauge section of the specimen. (b) Optical micrograph of the fracture surface after fatigue testing; the fracture-initiating defect is visible at the centre of the fracture pattern. (c) Matched configuration: the CT-detected defect is aligned with the fracture origin identified in (b). Agreement in axial position, radial distance from the specimen axis, projected defect shape and surface-proximity confirms a successful match.

to all CT-detected defects of that specimen and assigns an ordered list from the most critical to the least critical defect.

Let r_i denote the rank assigned by a given model to the experimentally identified fracture-initiating defect in specimen i. A value of $r_i = 1$ means that the model correctly assigns the highest criticality to the actual fracture origin. If $r_i > 1$, the model predicts one or more other defects to be more critical than the true fracture initiator. The larger the value of r_i, the larger the ranking error. Since the number of detected defects varies between specimens, the rank must be interpreted relative to the defect population of the respective specimen.

To compare the models, two complementary ranking metrics are used in this work.

Hit rate. The hit rate H measures how often a model correctly places the fracture-initiating defect at rank 1:

$$H = \frac{1}{N_m} \sum_{i=1}^{N_m} \mathbb{1}(r_i = 1)\,, \tag{1}$$

where N_m is the number of matched specimens and $\mathbb{1}(\cdot)$ is the indicator function. A value of $H = 1$ indicates perfect ranking, i.e. the model identifies the correct fracture origin in every matched specimen. The hit rate is therefore the most direct measure of practical ranking success.

Mean normalised rank. The hit rate alone does not distinguish between a near miss and a severe mis-ranking. To capture how far the true fracture origin lies from the top of the predicted ranking when it is not ranked first, the mean normalised rank $\bar{R}$ is additionally computed as

$$\bar{R} = \frac{1}{N_m} \sum_{i=1}^{N_m} R_i\,, \qquad R_i = \begin{cases} \dfrac{r_i - 1}{n_i - 1}\,, & n_i > 1 \\ 0\,, & n_i = 1\,. \end{cases} \tag{2}$$

where n_i is the total number of defects in specimen i. This normalisation maps the rank to the interval $[0, 1]$, allowing comparison between specimens with different defect counts. A value of $\bar{R} = 0$ corresponds to perfect ranking in all specimens, whereas $\bar{R} = 1$ corresponds to the worst possible ranking in all specimens.

Together, the two metrics provide complementary information: the hit rate quantifies how often a model identifies the correct defect directly. The mean normalised rank quantifies the severity of the ranking error when the critical defect is not correctly identified. While both metrics build on established rank-based measures, they are here specialised to the present benchmark to compare models consistently across specimens with different defect populations. Specifically, the hit rate corresponds to the rank-1 accuracy commonly used in information retrieval and recognition tasks [52–54], while the mean normalised rank is adapted from normalised rank measures used in retrieval evaluation [55,56]. Because the ranking benchmark is based on only 21 matched specimens, small differences between models should be interpreted cautiously; the metrics are therefore used primarily to identify robust performance trends rather than to claim strict superiority based on marginal differences.

For the present benchmark, the hit rate and mean normalised rank are evaluated separately for each series on the respective matched subset, comprising $N_m = 5$ specimens for series 'small' and $N_m = 16$ specimens for series 'large' (21 in total, see Table 5).

2.5.2. Part B: Fatigue-life prediction accuracy

For the fatigue-life predictions, the logarithmic life ratio is used as the primary accuracy measure:

$$\eta_i = \log_{10}\left(\frac{N_{f,\mathrm{pred},i}}{N_{f,\mathrm{exp},i}}\right)\,. \tag{3}$$

A value of $\eta_i = 0$ indicates perfect prediction. The overall model accuracy is characterised by:

- The **mean logarithmic error** (bias): $\bar{\eta} = \frac{1}{n}\sum_i \eta_i$. A positive value indicates systematic overprediction.
- The **root-mean-square logarithmic error**: $\mathrm{RMSE}_{\log} = \sqrt{\frac{1}{n}\sum_i \eta_i^2}$.
- The **percentage of predictions within factor-k scatter bands** ($k = 3, 6$), i.e. $|\eta_i| \le \log_{10}(k)$.

Predictions within a scatter band of factor 3 are commonly considered acceptable for fatigue-life models [18]. The results are presented in predicted-vs-experimental life diagrams with scatter bands.

2.6. Material properties

Table 2 summarises the material parameters used in this study. The TCD parameters ΔK_{th} and $\Delta\sigma_0$ are taken from the engineering estimation of Rigon and Meneghetti [57] and the Murakami hardness-fatigue strength correlation [13]. The lower part of the table lists derived characteristic quantities that are not independent material properties but follow directly from the primary parameters. These quantities define the length scales and thresholds used in the individual ranking and life prediction models.

The elastic constants E and ν are well-established for stress-relieved PBF-LB/M TiAl6V4 and show only minor process-to-process variation [28]. The yield strength σ_y is taken from a directly comparable PBF-LB/M process variant [28]; its influence on the defect ranking is evaluated implicitly through the comparison of LE and EP model variants in Section 2.7. A separate calibration of FE material parameters to the present material is not pursued.

Table 2
Material properties for TiAl6V4 (PBF-LB/M, stress relieved) and derived characteristic quantities.

Parameter	Symbol	Value	Unit	Source
Primary material properties				
Young's modulus	E_1	103.7	GPa	[28]
	E_2, E_3	95.238	GPa	[28]
Poisson's ratio	$\nu_{12}, \nu_{13}, \nu_{23}$	0.342	–	[28]
Shear modulus	G_{12}, G_{13}	35.966	GPa	[28]
	G_{23}	39.161	GPa	[28]
Yield strength	σ_y	1018	MPa	[28]
Ultimate tensile strength	σ_{UTS}	1165	MPa	[58,59]
Hardness	HV	380	$\mathrm{kgf/mm^2}$	[59]
Static fracture toughness	K_{Ic}	49	$\mathrm{MPa}\sqrt{\mathrm{m}}$	[60]
Threshold stress intensity factor range[a]	ΔK_{th}	6.363	$\mathrm{MPa}\sqrt{\mathrm{m}}$	[57]
Fatigue strength range[b]	$\Delta\sigma_0$	1216	MPa	[13]
Derived characteristic quantities				
Critical distance (TCD)	L	8.72	µm	Eq. (22)
SED control radius	R_c	6.62	µm	Eq. (12)
Support length	a^*	0.07	mm	Eq. (16)
ASTM threshold (air)	$\Delta K_{th,\mathrm{ASTM}}$	2.63	$\mathrm{MPa}\sqrt{\mathrm{m}}$	Eq. (33)

Notes: L and R_c are computed from ΔK_{th}, $\Delta\sigma_0$, and ν. The ASTM threshold is the long-crack threshold at $\mathrm{d}a/\mathrm{d}N = 10^{-10}$ m/cycle implied by the HS surface-in-air parameters (Eq. (33)). The ME fatigue limit σ_w (Eq. (29)) depends on the current crack size a and is therefore not listed as a single value.

[a] Estimated from the engineering correlation of Rigon and Meneghetti [57] for PBF-LB/M TiAl6V4 (stress relieved). Not directly measured on the present material; value corresponds to $R = -1$.

[b] Derived from the Murakami hardness-fatigue strength correlation [13]: $\Delta\sigma_0 = 2{\times}1.6\,HV$, representing the intrinsic defect-free fatigue strength range.

2.7. Defect criticality models

Defect criticality models estimate the effect of defects on the fatigue-life and try to generate a ranking of the detected defects inside of a part. These models vary in complexity and input data requirements. Tammas-Williams et al. [17] compared four different models, including their own model on defect data from PBF-EB/M. In this paper, nine defect criticality models, including the four models of Tammas-Williams et al. [17], are benchmarked and the results are compared. The benchmark also considers data sources, as well as computational cost.

2.7.1. Geometry-based criticality models

The following three models assess defect criticality based on geometric features extracted directly from the X-ray CT data. They do not require individual finite element simulations of each defect and are therefore computationally inexpensive. The models differ in the number of geometric and material parameters considered, ranging from a simple size ranking to semi-analytical stress concentration estimates.

Defect size ranking. One of the simplest metrics to rank the defects is the defect size. The size of all detected defects can be gathered directly from the NDT data, with the option for size ranking already built into CT software. In this model, the defect size, either by volume or surface area, governs the defect criticality. The largest defect is deemed to be the fracture critical defect. The ranking is defined in the following way. For a part with n defects $\mathbf{d} = d_1, d_2, \ldots, d_n$ the critical defect is the defect with the largest projected area perpendicular to the principal stress direction. To establish comparability with subsequent models, the square root of this value, defined as $\sqrt{\mathrm{Area}}$ is considered:

$$d_{\mathrm{crit}} = \max \sqrt{\mathrm{Area}}\,. \tag{4}$$

This model may be inaccurate in a non-uniform stress field. Larger defects at locations with a low stress can reduce the accuracy of the model since the stress is not considered in this model. A potential solution lies in the definition of thresholds, where a defect might become critical. This leads to the definition of HSVs, in which the stress is close to the maximum stress. Common thresholds for the HSV are 80 % to 90 % of the maximum nominal stress. The resulting volumes are likely locations for fracture critical defects. However, the arbitrary definition of these volumes can cause issues with rare, large defects that can be fracture critical at significantly lower stresses. The HSV approach was not used in the benchmark; defect size was the only metric in this specific model. While the applied stress is not part of the ranking, it is necessary to determine the stress field to determine the normal direction for the area calculation.

Murakami model. The Murakami model [13] can be seen as an expansion of the defect size model. It uses the square root of the defect area perpendicular to the load direction as a criticality estimator. Larger areas are deemed more harmful and more likely to initiate a fatigue fracture. The model also includes the local stress σ in the ranking. The approach of Murakami et al. calculates a stress intensity factor K based on the square root of the maximum defect area normal to the stress direction $\sqrt{\mathrm{Area}}$, the local stress σ and a factor based on the proximity to the surface Y. The critical defect is then described as the defect with the maximum stress intensity factor:

$$K = Y\sigma\sqrt{\pi\sqrt{\mathrm{Area}}}, \tag{5}$$

$$d_{\mathrm{crit}} = \max K_1, K_2, \ldots, K_n, \tag{6}$$

for n defects in a part. Multiple expansions of the original Murakami model were developed to account for stress increasing factors like surface proximity or defect interaction due to clustering. The square root of the defect area normal to the maximum stress $\sqrt{\mathrm{Area}}$ is increased once the defect distance to other defects or the surface falls below a size specific threshold. A detailed explanation of the different cases and their effect on $\sqrt{\mathrm{Area}}$ can be found in Masuo et al. [7]. The expanded model uses the term $\sqrt{\mathrm{Area}_{\mathrm{eff}}}$ to indicate the usage of the effective area, which may be larger than the area of the initial defect. The Murakami model can be used in combination with a Hardness correlation and a local stress field to estimate the fatigue strength of a part with defects based on empirical data. This allows the definition of fatigue strength based on defect sizes, similar to other fracture mechanical models developed for cracks, like the El-Haddad–Smith–Topper [61] model.

Tammas-Williams model. The Tammas-Williams model [17] uses a priori finite element simulations on idealised, spherical defects. The simulations are used to model the effects of surface proximity, defect interaction and pore aspect ratio on the tensile stress as a ratio of the nominal stress σ_∞. In combination with the $\sqrt{\mathrm{Area}}$ parameter, a relative tensile stress concentration factor ΔK_t is calculated, which can be used to rank the defects in a part. The stress concentration factors of

a defect are calculated by comparison of the defect parameters with the previous simulations. The product of all stress concentration factors, the local nominal stress σ_∞ and the square root area $\sqrt{\text{Area}}$ equals the relative stress concentration factor used for the ranking, as shown below:

$$\Delta K_t = \sigma_\infty \, K_t(\text{surface}) \, K_t(\text{aspect ratio}) \, K_t(\text{proximity}) \sqrt{\sqrt{\text{Area}}}, \tag{7}$$

$$d_{\text{crit}} = \max \Delta K_{t_1}, \Delta K_{t_2}, \dots, \Delta K_{t_n}, \tag{8}$$

for n defects in a part. The stress concentration factors are determined by interpolation from pre-computed FE results for idealised spherical and oblate spheroidal pores: $K_t(\text{surface})$ is a function of the dimensionless depth ratio d/D (pore centroid depth to pore diameter); $K_t(\text{aspect ratio})$ depends on the ratio of pore axes; and $K_t(\text{proximity})$ accounts for the interaction between two pores with face-to-face separation $s < D$, following Tammas-Williams et al. [17]. The model was developed for tensile load cases. For other load cases, further simulation would be necessary to determine the accurate stress intensity factors. Tammas-Williams et al. [17] report an increase in accuracy on their dataset compared to the Murakami model and the defect size ranking. However, the idealisation of defects to spherical, or ellipsoidal objects might reduce the accuracy when applied to PBF-LB/M due to the existence of tortuous LOF defects.

2.7.2. Simulation-based criticality models

In contrast to the geometry-based models, the following defect criticality models require individual finite element simulations for each detected defect. By resolving the actual three-dimensional defect morphology in the FE model, these approaches capture stress and strain concentrations that arise from the tortuous, irregular shapes typical of PBF-LB/M defects. This comes at a significantly higher computational cost, but avoids the idealisation to spherical or ellipsoidal shapes inherent in the geometry-based models. All simulation-based criticality models operate on the same FE sub-model results; they share a common defect processing workflow, which is described in the following before the individual ranking criteria are introduced.

For the simulation-based criticality models, a semi-automated defect processing workflow is developed to isolate, mesh and simulate each detected defect individually. The workflow operates on the X-ray CT reconstruction data and consists of five steps:

1. **Defect isolation:** A cuboidal bounding region (5× defect size) is defined around each defect. Neighbouring defects and the specimen surface are included in the sub-model if they intersect this region.
2. **Geometry preparation:** Surface subdivision, smoothing and rescaling are applied to reduce voxel-induced staircase artefacts and compensate for systematic CT size underestimation [12,39].
3. **Volume meshing:** Graded tetrahedral mesh with fine elements at the defect surface and coarser elements towards the boundaries.
4. **Boundary conditions:** Displacement boundary conditions interpolated from a defect-free global model onto the sub-model cut faces.
5. **FE simulation:** Solution with either a LE or EP material model (see below).

The sub-modelling approach decouples global and local length scales and enables parallelisation. For the present dataset (472 + 605 defects), the workflow was executed using a scripted pipeline in *Python 3.10* with the *PyMAPDL* interface.

The material model used in the sub-model simulation influences both the magnitude and the distribution of the local stress and strain concentrations and, consequently, the ranking outcome. To quantify this influence, two material model variants are investigated:

1. **LE**: Young's modulus E and Poisson's ratio ν only. This is the computationally cheapest option and serves as a baseline. Because no yielding is captured, the model may overestimate peak stresses at sharp notch roots, as used by Afroz et al. [27].
2. **EP**: An elastic-perfectly-plastic model with yield strength σ_y. Stress redistribution around the defect is captured, which limits the peak stress and shifts the ranking towards defects that produce larger plastic zones, as used by Li et al. [25].

Both variants are applied to the same set of defect sub-models, so that the effect of the material model on the ranking can be isolated. The following paragraphs describe the individual ranking metrics extracted from these simulation results. All simulation-based metrics are extracted from the same simulation results using a post-processing script based on Ansys DPF (Data Processing Framework).

Stress–strain concentration factor. Following the work of Li et al. [25] and Gao et al. [26], the stress–strain concentration factor k_g is defined as the geometric mean of the local stress concentration $k_\sigma = \sigma_{\max}/\sigma_\infty$ and strain concentration $k_\varepsilon = \varepsilon_{\max}/\varepsilon_\infty$, with the maximum principal stress $\sigma_{\max}$ and strain $\varepsilon_{\max}$ extracted from the defect sub-model and the far-field stress σ_∞ and strain ε_∞ obtained from the defect-free global model:

$$k_g = \sqrt{k_\sigma \, k_\varepsilon}, \tag{9}$$

$$d_{\text{crit}} = \max \left(k_{g,1}, k_{g,2}, \dots, k_{g,n} \right) . \tag{10}$$

For purely LE simulations, $k_\sigma = k_\varepsilon$ and thus $k_g = k_\sigma$. For EP simulations, stress redistribution reduces k_σ while increasing k_ε, so that k_g captures the combined severity of the local stress–strain state. Tammas-Williams et al. [17] adopted this factor from Li et al. [25] for their comparison of defect rankings on PBF-EB/M data.

Averaged strain energy density (SED). The averaged strain energy density method, originally proposed by Lazzarin et al. [29,30] for notch fatigue assessment, evaluates the mean elastic (and, if applicable, plastic) strain energy density $\bar{W}$ averaged over a material-dependent control volume surrounding the stress hotspot:

$$\bar{W} = \frac{1}{V_c} \int_{V_c} W(\mathbf{x}) \, \mathrm{d}V \, , \tag{11}$$

where V_c is a spherical control volume of radius R_c centred at the point of maximum principal stress on the defect surface, and $W(\mathbf{x}) = \frac{1}{2}\sigma_{ij}\varepsilon_{ij}$ is the local strain energy density. For elastic–plastic simulations, the plastic work contribution is included: $W = W_e + W_p$, with $W_p = \sigma_{ij}\varepsilon_{ij}^{pl}$.

The control radius R_c is a material property derived from the threshold stress intensity factor range and the defect free fatigue limit [30]:

$$R_c = \frac{(1+\nu)(5-8\nu)}{4\pi} \left(\frac{\Delta K_{th}}{\Delta \sigma_0} \right)^2 . \tag{12}$$

The defect ranking is based on the maximum averaged SED:

$$d_{\text{crit}} = \max \left(\bar{W}_1, \bar{W}_2, \dots, \bar{W}_n \right) . \tag{13}$$

Compared to the peak-value-based k_g metric, the SED method averages over a finite volume and is therefore less sensitive to mesh refinement and local stress singularities at sharp geometric features of the defect surface. The control radius R_c was originally derived by Lazzarin and Berto [29,30] for blunt V-notches in metallic components under fatigue loading, where it defines the radius of the process zone over which crack initiation is governed. In the present work, it is applied to the defect surface as an equivalent notch root, following the interpretation that CT-resolved pores and LOF defects feature finite tip radii rather than idealised crack tips.

Relative stress gradient. The relative stress gradient concept, originally introduced by Siebel and Stieler [31] to characterise the notch support effect on fatigue strength, quantifies how rapidly the stress field decays from the defect surface into the bulk material. The approach has since been widely adopted in fatigue assessment of notched components [62–64]:

$$\chi^* = \frac{1}{\sigma_{\max}} \left| \frac{\mathrm{d}\sigma_1}{\mathrm{d}r} \right|_{r=0}, \tag{14}$$

where σ_1 is the first principal stress, r is the distance from the hotspot perpendicular to the defect surface (into the material), and $\sigma_{\max}$ is the peak stress at the surface. The gradient is determined by sampling the stress field along a line from the hotspot into the material and performing a linear fit over a distance of $2R_c$.

The support length a^* was originally calibrated by Mei et al. [63] for notched metallic components under uniaxial fatigue loading. Its application here assumes that the local stress gradient at a CT-detected defect can be treated analogously to a macro-notch gradient, which is a simplification that may introduce uncertainty for irregular LOF morphologies.

A high χ^* indicates a steep stress decay, implying that the stress concentration is highly localised, characteristic of sharp notch-like features. A low χ^* indicates a more gradual stress field extending deeper into the material, which is more damaging from a fatigue perspective since a larger material volume is subjected to near-peak stresses, promoting crack initiation and early propagation [31,62]. This observation forms the basis of the support factor concept: steep gradients allow load redistribution to less-stressed subsurface material, effectively increasing the tolerable peak stress [31,63]. Following the modified effective stress approach of Mei et al. [63], defects are ranked by a support-effect corrected stress:

$$\sigma_{\text{supported}} = \frac{\sigma_{\max}}{1 + \chi^* a^*}, \tag{15}$$

where a^* is a material-dependent support length that governs the sensitivity to stress gradients. The formulation corresponds to a first-order linearisation of the stress decay over the characteristic length scale [63,64]. The parameter a^* is estimated from the empirical correlation [63]:

$$a^* = \left(\frac{270}{\sigma_{\text{UTS}}} \right)^{1.8}, \tag{16}$$

where σ_{UTS} is the ultimate tensile strength in MPa. The ranking criterion is:

$$d_{\text{crit}} = \max \left(\sigma_{\text{supported},1}, \sigma_{\text{supported},2}, \dots, \sigma_{\text{supported},n} \right). \tag{17}$$

Smith-Watson-Topper (SWT) parameter. The Smith–Watson–Topper [32] fatigue indicator parameter P_{SWT} combines the maximum stress and the strain range into a single damage metric that accounts for mean stress effects:

$$P_{\text{SWT}} = \sigma_{\max} \frac{\Delta\varepsilon}{2}, \tag{18}$$

where $\sigma_{\max}$ is the maximum principal stress at the defect surface and $\Delta\varepsilon$ is the total strain range (elastic plus plastic) at the same location. For fully reversed loading ($R = -1$), the strain range equals twice the maximum strain: $\Delta\varepsilon = 2(\varepsilon_{\text{el,max}} + \varepsilon_{\text{pl,max}})$. The SWT parameter has physical units of stress and can be interpreted as a fatigue damage intensity. Defects are ranked by:

$$d_{\text{crit}} = \max \left(P_{\text{SWT},1}, P_{\text{SWT},2}, \dots, P_{\text{SWT},n} \right). \tag{19}$$

The SWT parameter is particularly relevant for elastic–plastic simulations, where stress redistribution limits the peak stress but increases the local plastic strain, capturing a trade-off that neither k_σ nor k_ε alone reflects.

Table 3
TCD method variants used for defect ranking.

Method	Evaluation domain	Criterion
Point Method (PM)	Stress at $r = L/2$	$\sigma(L/2) \geq \Delta\sigma_0$
Line Method (LM)	Average over line $[0, 2L]$	$\frac{1}{2L}\int_0^{2L} \sigma(r)\,\mathrm{d}r \geq \Delta\sigma_0$
Area Method (AM)	Average over hemisphere, radius L	$\frac{2}{\pi L^2}\iint_A \sigma \mathrm{d}A \geq \Delta\sigma_0$
Volume Method (VM)	Average over sphere, radius L	$\frac{3}{2\pi L^3}\iiint_V \sigma\,\mathrm{d}V \geq \Delta\sigma_0$

Maximum equivalent plastic strain. For elastic–plastic simulations, the maximum equivalent plastic strain at the defect surface provides a direct measure of local plastic damage accumulation:

$$\varepsilon_{\text{pl,eq}} = \sqrt{\frac{2}{3}\varepsilon_{ij}^{\text{pl}}\varepsilon_{ij}^{\text{pl}}}. \tag{20}$$

This metric is only available for the elastic–plastic model variant; for LE simulations it is identically zero. Defects that produce larger plastic zones or higher peak plastic strains are considered more critical, since plastic strain localisation is a precursor to fatigue crack nucleation. The ranking criterion is:

$$d_{\text{crit}} = \max \left(\varepsilon_{\text{pl,eq},1}, \varepsilon_{\text{pl,eq},2}, \dots, \varepsilon_{\text{pl,eq},n} \right). \tag{21}$$

Theory of critical distance (TCD). The TCD treats the existing defects as notches, rather than cracks [35]. This is arguably more realistic for additive manufacturing defects, since the three-dimensional shape of porosity and LOF defects features notch roots with a finite tip radius, as opposed to the idealised infinitely sharp crack tip assumed in LE fracture mechanics-based approaches [33,34,41]. TCD is a family of methods that use a material length parameter (referred to as the critical distance, L), together with material parameters to establish the critical fracture conditions [35]. The length parameter L is defined as

$$L = \frac{1}{\pi}\left(\frac{\Delta K_{th}}{\Delta\sigma_0} \right)^2, \tag{22}$$

with the threshold stress intensity factor ΔK_{th} and the theoretical, defect-free fatigue limit $\Delta\sigma_0$. The parameter L is a material constant and does not depend on defect geometry. The TCD is evaluated in four variants that differ in how the stress field is sampled relative to the notch tip (Table 3). All variants use the same critical distance L (Eq. (22)) and predict failure when the respective averaged stress range reaches $\Delta\sigma_0$. For the purpose of a defect ranking, the TCD is combined with the FE sub-model results from the shared defect simulation workflow. The stress field around each defect, already computed for the other simulation-based metrics, is reused and the TCD criterion is evaluated without additional simulations. A ranking parameter σ_{TCD} can be defined for each defect. Using the PM as an example:

$$\sigma_{TCD,i} = \sigma_i \left(r = \frac{L}{2} \right) \tag{23}$$

The fracture-critical defect is then:

$$d_{crit} = \max \left(\sigma_{TCD,1}, \sigma_{TCD,2}, \dots, \sigma_{TCD,n} \right) \tag{24}$$

for n defects in a part. Compared to the direct FE ranking based on k_g, TCD avoids the strong mesh-size sensitivity of maximum stress and strain values at the defect surface, because the evaluation point or averaging domain lies away from the singularity-prone surface. This makes the result less dependent on mesh refinement. However, TCD requires the additional material parameters ΔK_{th} and $\Delta\sigma_0$, increasing the experimental effort. The TCD is applicable to a range of loading conditions (uniaxial, multiaxial, static, cyclic) and has been validated for notched metallic components [33,34], though its application to AM defect populations is still limited [36].

2.8. Fatigue-life prediction models

While Part A of this study addresses the question of *which* defect is most critical, this section (Part B) addresses the complementary question: *what fatigue-life does the critical defect imply?* Once the fracture-initiating defect is identified, either by the defect criticality models (predicted) or by fractographic analysis (ground truth), a crack growth model translates the defect characteristics (size, location, morphology) into a predicted number of cycles to failure.

A key aspect of defect-controlled fatigue in PBF-LB/M components is the distinction between surface and internal crack growth. Internal cracks propagate in a self-contained environment, shielded from the ambient atmosphere by the surrounding material. It has been demonstrated that the crack growth rate in vacuum can be orders of magnitude lower than that in air, particularly at low stress intensity factor ranges ΔK close to the threshold [65–67]. This difference has a profound effect on the fatigue-life of specimens failing from internal defects: models that do not account for the reduced internal crack growth rate systematically underestimate the fatigue-life of internally initiated failures [41]. Conversely, once an internal crack breaches the free surface, it transitions to surface-crack growth in air, and the growth rate increases accordingly [41].

To quantify the influence of the crack growth model and the environmental distinction on the fatigue-life prediction, three crack propagation models of increasing complexity are benchmarked:

1. **HS** with dual-environment parameters (Section 2.8.2),
2. **ME** scS-N model without environmental distinction (Section 2.8.3),
3. **Hybrid** combining the Murakami-derived evolving threshold with the HS dual-environment framework (Section 2.8.3).

For consistency and to isolate the effect of the crack growth model from the defect criticality model, the initial crack size for all life predictions is taken from fractographic analysis of the actual fracture-initiating defect, independent of the criticality models from Part A. The crack propagation models are described in the following sections.

2.8.1. Initial crack size

Each ranking model provides a different representation of the defect severity, which must be translated into an equivalent initial crack size a_0 for the crack propagation calculation. Following Murakami [13], the initial crack size is taken directly as the $\sqrt{\text{Area}}$ parameter of the dominant defect identified on the fracture surface via optical microscopy

$$a_0 = \sqrt{\text{Area}}\,, \tag{25}$$

and the stress intensity factor is computed as

$$\Delta K = Y \cdot \Delta\sigma \cdot \sqrt{\pi a}\,, \tag{26}$$

with the geometry factor $Y = 0.5$ for internal defects and $Y = 0.65$ for surface-connected defects, the crack size a and the applied stress range $\Delta\sigma$.

The defect location relative to the free surface governs whether the crack is classified as a surface crack (growing in air from the first cycle) or an internal crack (initially growing in a self-contained environment). For internal cracks, the transition to air-environment growth occurs when the crack front reaches the specimen surface, i.e. when the crack depth equals the distance between the defect and the free surface l_n [41]. At a ligament threshold of $l_n \leq 20\,\mu\text{m}$ the crack is defined as surface-connected and the surface crack growth regime applies.

Table 4
Crack growth model parameters.

Model	Environment	Parameter	Value	Source
HS	Air (surface)	D_{air}	$2.791 \times 10^{-10}\,\text{MPa}\sqrt{\text{m}}$	[40]
		p_{air}	2.12	[40]
		ΔK_{thr}	$2.0\,\text{MPa}\sqrt{\text{m}}$	[40]
		A_{air}	$90\,\text{MPa}\sqrt{\text{m}}$	[41]
	Self-contained (internal)	D_{int} [a]	$2.897 \times 10^{-13}\,\text{MPa}\sqrt{\text{m}}$	[41]
		p_{int} [a]	4.18	[41]
		ΔK_{thr} [a]	$2.0\,\text{MPa}\sqrt{\text{m}}$	[41]
		A_{int}	$90\,\text{MPa}\sqrt{\text{m}}$	[41]
ME	(single)	C^*	10^{-4}	[42]
		m^*	2.0	[42]
		n^*	1.0	[42]
Hybrid	Air/self-contained	D, p	As HS	[40,41]
		$\Delta K_{\text{thr}}(a)$	From Eq. (30)	This work

[a] Parameters are refitted to the present PBF-LB/M TiAl6V4 dataset in Phase 2 (Section 3.5.2); all other values are taken from the cited literature without modification.

2.8.2. Crack growth model I: HS (dual environment)

The HS equation [40] explicitly includes the crack growth threshold ΔK_{thr} and the cyclic fracture toughness A:

$$\frac{\mathrm{d}a}{\mathrm{d}N} = D\left(\frac{\Delta K - \Delta K_{\text{thr}}}{\sqrt{1 - K_{\max}/A_{\text{HS}}}}\right)^{p}, \tag{27}$$

with the material constants D and p, the stress intensity factor range ΔK computed from Eq. (26), the maximum stress intensity factor $K_{\max}$ and the threshold ΔK_{thr}. Following Han et al. [41], separate parameter sets are used for surface crack growth in air and internal crack growth:

- **Surface crack in air:** D_{air}, p_{air}, $\Delta K_{\text{thr,air}}$, A_{air} (Table 4).
- **Internal crack (self-contained):** D_{int}, p_{int}, $\Delta K_{\text{thr,int}}$, A_{int} (Table 4).

The transition from internal to surface growth occurs when the crack depth reaches the ligament distance ($a \geq a_0 + l_n$), as described in Section 2.8.4. Compared to the Paris–Erdogan [68] law, the HS equation captures the near-threshold behaviour more accurately, which is particularly relevant for high-cycle fatigue where the majority of cycles are consumed at low ΔK values close to the threshold [40,41]. Furthermore, the explicit threshold term prevents non-physical crack growth below ΔK_{thr}, which is important for internal defects where the initial ΔK may be close to or below the threshold.

2.8.3. Crack growth model II: ME scS-N model

In addition to the fracture-mechanics-based crack growth equations described above, a second model is included in the benchmark that does not require the determination of crack growth parameters from fatigue crack growth testing. The scS-N prediction model proposed by Murakami and Endo [42] predicts fatigue-life based solely on the Vickers hardness HV and the initial defect size, without curve-fitting of crack growth constants to experimental data. The model is based on the small crack growth equation:

$$\frac{\mathrm{d}a}{\mathrm{d}N} = C^*\left(\frac{\sigma}{\sigma_w} - 1\right)^{m^*} a^{n^*}, \tag{28}$$

where a is the crack size in terms of $\sqrt{\text{Area}}$ in µm, σ is the applied stress amplitude, and σ_w is the fatigue limit, which decreases cycle by cycle as the crack grows. The constants $C^* = 10^{-4}, m^* = 2$ and $n^* = 1$ are claimed to be material-independent [42] and were validated for steels, cast iron, and AM TiAl6V4 as well as AM Inconel 718 [42,43]. The fatigue limit σ_w at any crack size a is given by the $\sqrt{\text{Area}}$ parameter model:

$$\sigma_{w,R=-1} = \frac{C(HV + 120)}{a^{1/6}} \tag{29}$$

where $C = 1.43$ for surface defects and $C = 1.56$ for subsurface defects, HV is the Vickers hardness in $\mathrm{kgf/mm^2}$, and $a = \sqrt{\mathrm{Area}}$ in µm. For materials with $HV > 450$, Murakami and Endo recommend using 80 % of the value given by Eq. (29).

The key conceptual difference between this model and the HS formulation is that the threshold behaviour is embedded implicitly: when $\sigma \leq \sigma_w$, the crack growth rate is zero by construction, and σ_w evolves with crack size. This eliminates the need for a fixed ΔK_{thr} parameter and inherently captures the small-crack effect, since smaller cracks have a higher σ_w (and hence a lower effective threshold in ΔK space) than larger cracks.

It is noted that the conversion between the tension-only stress range $\Delta\sigma$ used in the HS formulation and the stress amplitude σ used in the ME model depends on the stress ratio. For $R = -1$, the tension-only stress range equals the maximum stress: $\Delta\sigma = \sigma_{\max} = \sigma_{\mathrm{amp}}$, since only the tensile half-cycle contributes to crack growth. This convention is adopted consistently across all models.

A limitation of the ME model as originally formulated is that it does not distinguish between surface crack growth in air and internal crack growth environments (though the geometric location factor C in Eq. (29) differs between surface and subsurface defects). To address this within the present benchmark, two variants are considered:

1. **ME (single environment)**: The model is applied as published, with a single set of constants (C^*, m^*, n^*) regardless of defect location. This serves as a test of whether the model's simplicity is sufficient for the present dataset.
2. **Hybrid model**: The Murakami-derived, crack-size-dependent threshold is incorporated into the HS framework. At each integration step, the fixed ΔK_{thr} in Eq. (27) is replaced by

$$\Delta K_{\mathrm{thr}}(a) = Y \cdot f(R) \cdot \sigma_w(a) \cdot \sqrt{\pi a}\,, \tag{30}$$

where $f(R) = 1$ for $R < 0$ and $f(R) = 2$ for $R \geq 0$, and $\sigma_w(a)$ is computed from Eq. (29) at the current crack size a. The dual-environment distinction (surface-in-air vs. internal parameters for D and p) is retained from the HS model. This hybrid formulation combines the physically motivated, evolving threshold of the ME approach with the environmental sensitivity of the HS equation, eliminating ΔK_{thr} as a free parameter while preserving the distinction between surface and internal crack growth rates.

The integration procedure for both variants follows the same scheme as described in Section 2.8.4, with the physical failure criterion (Section 2.8.4) applied to all three models. Murakami and Endo [42] originally used a fixed $a_c = 1000\,\mu\mathrm{m}$; the sensitivity study (Supplementary Material S3) confirms that the choice of failure criterion has negligible influence on predicted life in the present HCF regime.

2.8.4. Integration and life calculation

For all three crack growth models, the fatigue-life is obtained by numerical integration of the respective crack growth equation from the initial crack size a_0 to the critical crack length a_c:

$$N_{f,\mathrm{pred}} = \int_{a_0}^{a_c} \frac{\mathrm{d}a}{(\mathrm{d}a/\mathrm{d}N)}\,. \tag{31}$$

The critical crack length a_c is determined by the lesser of the fracture toughness criterion ($K_{\max} = K_{Ic}$) or the net-section failure of the specimen cross-section. For the present high-cycle fatigue conditions, the predicted life is insensitive to the exact choice of a_c, since the majority of cycles are consumed at small crack lengths near a_0.

For internal defects, the integration is split at $a = a_0 + l_n$: the first segment uses self-contained/internal parameters, and the second segment uses air/surface parameters. The total predicted life is

$$N_{f,\mathrm{pred}} = N_{\mathrm{int}} + N_{\mathrm{air}} = \int_{a_0}^{a_0+l_n} \frac{\mathrm{d}a}{(\mathrm{d}a/\mathrm{d}N)_{\mathrm{int}}} + \int_{a_0+l_n}^{a_c} \frac{\mathrm{d}a}{(\mathrm{d}a/\mathrm{d}N)_{\mathrm{air}}}\,. \tag{32}$$

This two-stage formulation follows Han et al. [41] and accounts for the orders-of-magnitude difference in crack growth rate between the internal (self-contained) and surface (air) environments.

An adaptive integration scheme is employed: at each step, the cycle increment is chosen such that the crack growth does not exceed 1 % of the current crack length, ensuring numerical accuracy in the near-threshold regime where growth rates are extremely low. A maximum cycle count of 10^{10} is imposed as a runout criterion.

Physical failure criterion. In addition to the geometric crack size limit, two physical failure criteria are evaluated at each integration step:

1. **Unstable fracture**: Failure occurs when $K_{\max} \geq K_{Ic}$, with the static fracture toughness K_{Ic} (see Table 2).
2. **Net-section collapse**: Failure occurs when the net-section stress σ_{net} exceeds the ultimate tensile strength σ_{UTS}, approximated for the circular specimen cross-section as $\sigma_{\mathrm{net}} = \sigma_{\max} \cdot A_{\mathrm{total}}/(A_{\mathrm{total}} - A_{\mathrm{crack}})$.

The integration terminates at the first occurrence of either criterion. For the present HCF conditions, unstable fracture is reached before net-section collapse in all cases, and the predicted life is insensitive to the exact failure criterion since the vast majority of cycles are consumed at crack sizes where $K_{\max} \ll K_{Ic}$.

It is noted that the HS equation uses a fitting parameter ΔK_{thr} that differs from the ASTM E647 [69] threshold ΔK_{th} defined at $\mathrm{d}a/\mathrm{d}N = 10^{-10}$ m/cycle. Following Jones et al. [40], the relationship is:

$$\Delta K_{\mathrm{th}} \approx \Delta K_{\mathrm{thr}} + \left(\frac{10^{-10}}{D}\right)^{1/p} \tag{33}$$

For the surface-crack-in-air parameters ($D_{\mathrm{air}} = 2.791 \times 10^{-10}\,\mathrm{MPa}\sqrt{\mathrm{m}}$, $p_{\mathrm{air}} = 2.12$), this yields $\Delta K_{\mathrm{th}} \approx \Delta K_{\mathrm{thr}} + 0.62\,\mathrm{MPa}\sqrt{\mathrm{m}}$. This distinction ensures that the equation returns a small but non-zero growth rate at the ASTM-defined threshold, rather than exactly zero.

2.8.5. Fitting of internal crack growth parameters

The surface-crack-in-air parameters from Jones et al. [40] have been shown to be valid across multiple AM TiAl6V4 process variants and are therefore adopted without modification. The internal crack growth parameters, however, may differ between the wrought TiAl6V4 data used by Han et al. [41] and the present PBF-LB/M material due to differences in microstructure and defect morphology. A fitting procedure is therefore applied exclusively to the internal parameters.

For the HS model, the parameters D_{int}, p_{int}, and $\Delta K_{\mathrm{thr,int}}$ are optimised simultaneously, with $A_{\mathrm{int}} = 90\,\mathrm{MPa}\sqrt{\mathrm{m}}$ held fixed. For the hybrid model, only D_{int} and p_{int} are optimised, since ΔK_{thr} is determined by the Murakami formulation at each integration step. The objective function is the mean squared logarithmic life ratio over the subset of specimens with internal dominant defects (Eq. (34)).

The optimisation employs a two-stage strategy: first, a grid search over physically meaningful parameter ranges ($\log_{10}(D_{\mathrm{int}}) \in [-13, -8]$ $\mathrm{MPa}\sqrt{\mathrm{m}}$, $p_{\mathrm{int}} \in [2, 6]$, $\Delta K_{\mathrm{thr,int}} \in [0.5, 3.0]$ $\mathrm{MPa}\sqrt{\mathrm{m}}$) identifies the approximate optimum; second, a Nelder–Mead simplex refinement [70,71] converges to the final parameter values. This approach avoids convergence to local minima while remaining computationally tractable.

The defect classification into surface and internal is governed by the ligament threshold: defects with $l_n \leq l_{n,\mathrm{thr}}$ are classified as surface-connected and assigned air-environment parameters from the first cycle. The default threshold of $l_{n,\mathrm{thr}} = 20\,\mu\mathrm{m}$ is adopted, and its influence is investigated in the sensitivity study (Section 3.5.3).

2.8.6. Summary of crack growth parameters

Table 4 summarises the material parameters used for each crack growth model variant. For HS and Hybrid, two parameter sets are listed: one for surface crack growth in air and one for internal crack growth. The parameters for surface crack growth in air are taken from Jones et al. [40], who demonstrated their validity across multiple AM

Table 5
Specimen subsets used in each stage of the analysis.

Subset	small	large	Total
Fatigue-tested (Section 2.3)	50	43	93
Inspected by X-ray CT (Section 2.2)	41	43	84
Fractographically analysed (Section 2.4)	30	43	73
Matched CT/fractography (Section 2.4)	5	16	21

TiAl6V4 process variants. The internal crack growth parameters are adopted from Han et al. [41], who fitted them to literature data of crack growth in vacuum and synchrotron-observed internal cracks in wrought TiAl6V4.

3. Results

The results are presented in two parts, corresponding to the two main objectives of this work. Section 3.4 presents the ranking benchmark (Part A), comparing the ability of each criticality model to identify the fracture-initiating defect. Section 3.5 presents the fatigue-life predictions (Part B), comparing the three crack growth models against the experimentally observed cycles to failure.

Not all fracture-initiating defects detected in the fracture surface analysis could be matched with NDT data. A direct comparison of the defect model criticality prediction to the fracture-initiating defect is therefore only possible on a subset of specimens (21 out of 84 CT scanned specimens, see Table 5). Of all 93 fatigue-tested specimens, 73 contribute to the life prediction benchmark, where the fracture-initiating defect is identified from fractography independent of the CT data. The remaining 20 specimens were not available for fractography and therefore only contribute to the fatigue testing. Table 5 summarises the specimen subsets underlying each stage of the analysis, from the total fatigue-tested population down to the subset used for the defect-ranking benchmark.

3.1. Fatigue test results

Fig. 4 shows the fatigue test results of both specimen series in a P-S-N representation. For series ‘small’, the combined finite-life and staircase programme yielded an experimental fatigue strength of σ_F = 559.5 MPa at 3×10^6 cycles. In contrast, all specimens of series ‘large’ failed before reaching 3×10^6 cycles, including those tested at 530 MPa, i.e. below the fatigue strength reference derived from series ‘small’. This confirms that the increased HSV of series ‘large’ promotes defect-controlled failure and reduces the fatigue performance relative to the baseline established on series ‘small’.

The scatter in fatigue-life is substantially larger in series ‘large’, which is consistent with the stronger competition between defects in the enlarged highly stressed volume. This increased competition is one of the main reasons why series ‘large’ is particularly informative for benchmarking defect criticality models.

3.2. Defect characterisation

Table 6 summarises the CT-detected defect populations for both series. Due to the higher resolution (finer voxel size) in series ‘large’, a higher amount of defects is detected, even though the scanned volume V_S is lower. This results in a lower mean relative density, though both series exceed a relative density of 99.99 %, indicating a stable and successful build process. It is noted that these values represent an upper bound, as defects smaller than approximately 2–3× the voxel size cannot be reliably resolved by X-ray CT, causing the CT-derived density to tend towards 100 %. Complementary metallographic analysis of 15 cross-sections yields a minimum relative density of 99.84 % and a mean of 99.90 %, confirming that the true porosity level is higher than the CT data suggest but remains well within acceptable limits for structural applications. Fig. 5 shows a histogram of defect size for both series. The finer resolution in series ‘large’ results in a higher number of detected defects and a shift of the peak to smaller defect sizes. Fig. 5 therefore serves as an indirect resolution comparison: the shift in detected defect size distribution and the increase in high-sphericity defects at 14 µm resolution compared to 25.6 µm reflect the improved detection capability rather than a true change in the underlying defect population.

The practical impact of the resolution difference on the downstream analysis is quantified by two complementary outcomes. First, the CT/fractography matching rate increases from 12 % (5/41 specimens at 25.6 µm) to 37 % (16/43 specimens at 14 µm), reflecting the improved detectability of fracture-initiating defects above the resolution-dependent detection limit. Second, for the 21 matched specimens, the mean absolute deviation between CT-derived and fractographically measured $\sqrt{\text{Area}}$ is comparable across both series (Fig. 6), indicating that once a defect is detectable, the sizing accuracy is similar at both resolutions. The CT resolution therefore primarily determines *which* defects can be detected and matched to fracture origins, rather than the accuracy with which detected defects are sized.

Fig. 5 shows that both series are dominated by near-spherical gas pores ($\Psi > 0.7$), with a minor population of irregular LOF-type defects at lower sphericity values. This morphological distribution is consistent with the high relative density and the use of standard industrial process parameters targeting dense parts. The predominance of gas pores implies that the Murakami $\sqrt{\text{Area}}$ model, which assumes a compact defect projection, is an appropriate approximation for the present defect population. The scarcity of highly elongated LOF defects means that the potential advantage of simulation-based models capturing tortuous morphologies (Section 2.7.2) is limited for the present dataset.

Fig. 6 provides a quantitative consistency check of the CT characterisation by comparing CT-derived and fractographically measured $\sqrt{\text{Area}}$ values for the 21 matched specimens; the measurement definitions and underlying uncertainty sources are discussed in Section 2.4. In 16 of the 21 matched specimens the CT-derived $\sqrt{\text{Area}}$ exceeds the fractographically measured value, with a mean difference of approximately 6 µm. The five cases of CT underestimation are concentrated at CT values below 55 µm, consistent with partial-volume effects near the detection limit (3×14 µm ≈ 42 µm for series ‘large’). These deviations are within the expected range for the respective resolutions and do not compromise the validity of the CT-based defect population for the ranking benchmark.

3.3. Fracture-origin classification

Post-mortem fractographic analysis of all 93 tested specimens identifies the fracture-initiating defect type and location for 73 specimens; the remaining 20 specimens were not available for fractographic examination (retested run-outs, see Supplementary Material S5). Fig. 7 shows representative examples of the two observed fracture origin types: internal defects embedded in the bulk material (Fig. 7(a)) and surface-connected defects intersecting the machined specimen surface (Fig. 7(b)). Fig. 7(c) shows an internal fracture origin with a comparatively small ligament distance, representative of the range observed within the internal subset.

Table 7 summarises the distribution of fracture origin types for all specimens with available fractographic data. The binary classification at $l_{n,\text{thr}} = 20$ µm is adopted consistently with the crack growth models (Section 2.8), where defects with $l_n \leq 20$ µm are assigned surface-in-air parameters from the first cycle and all others are treated as internal. The defect criticality models (Section 2.7) employ their own proximity-based classification schemes: the Murakami model applies a size-dependent area correction for defects within a threshold distance of the free surface [7], and the Tammas-Williams model uses a $K_t(\text{surface})$ factor that increases with decreasing depth ratio [17]. However, no CT-detected defect in either series satisfied the proximity threshold

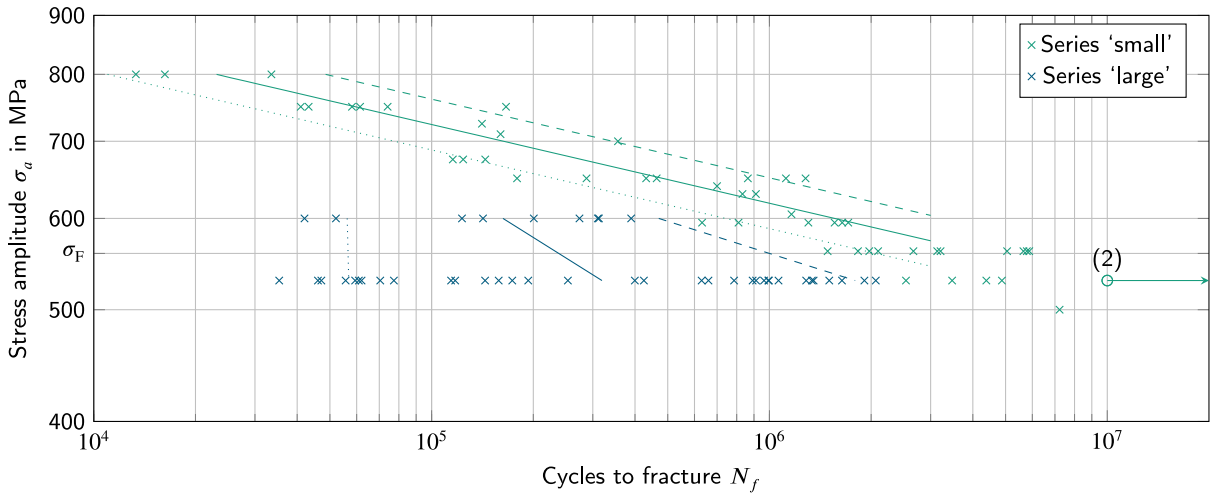


Fig. 4. Fatigue test results of both specimen series in a P-S-N representation. Series 'small' combines finite-life tests and staircase testing to determine the fatigue strength at 3×10^6 cycles. Series 'large' was tested at two stress amplitudes selected relative to this reference fatigue strength. The figure highlights the earlier failure of the large-volume specimens and the increased scatter associated with defect-controlled fatigue.

Table 6
Summary of defect data for both series.

Series	Specimen number	Resolution in µm	Detected defects	Scanned volume V_S in mm³	Defect volume V_P in mm³	Mean density $\rho = \frac{V_S - V_P}{V_S}$ in %
small	41	25.6	472	120 690.30	0.0688	99.99994
large	43	14	605	39 384.60	0.0507	99.99987

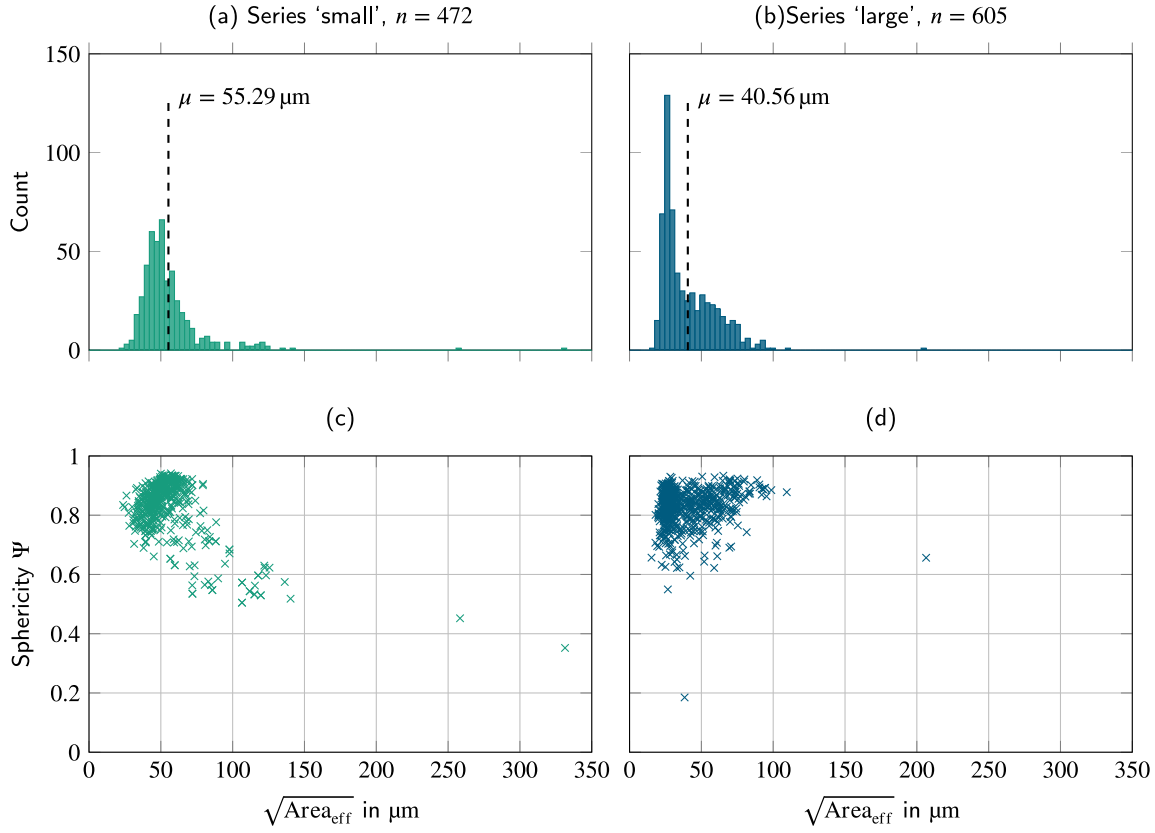


Fig. 5. CT-detected defect population for both series. (a,b) Histograms of defect size $\sqrt{\text{Area}_{\text{eff}}}$ for series 'small' ($n = 472$, 25.6 µm resolution) and series 'large' ($n = 605$, 14 µm resolution). The finer resolution in series 'large' results in detection of smaller defects, shifting the mean to lower values. Dashed lines indicate the respective mean values. (c,d) Sphericity $\Psi = \pi^{1/3}(6V)^{2/3}/A_s$ as a function of defect size. Values near unity indicate near-equiaxed gas pores; lower values indicate irregular LOF morphologies. Both series are dominated by high-sphericity pores. The differences between (a,b) and between (c,d) are attributable primarily to the difference in voxel size (25.6 µm vs. 14 µm) rather than to a different defect population, since both series are produced under identical process parameters.

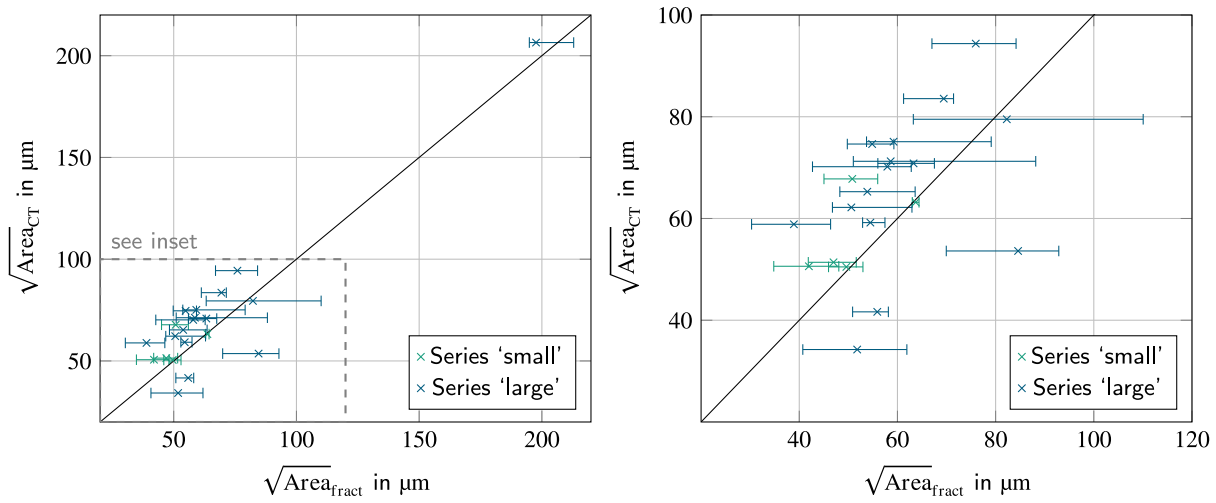


Fig. 6. Comparison of defect sizes measured by X-ray CT ($\sqrt{\text{Area}_{CT}}$, projected area of the segmented 3D volume normal to the loading direction) and fractography ($\sqrt{\text{Area}_{fract}}$, 2D projected area on the fracture surface) for the 21 matched specimens (5 from series 'small', 16 from series 'large'). Error bars on the horizontal axis indicate the range of the four individual fractographic measurements per specimen. The solid line represents the 1:1 correspondence. In 16 of 21 cases the CT-derived value exceeds the fractographic measurement (mean: ≈ 6 µm). The remaining five specimens show CT underestimation, with the largest relative deviations at CT values below 55 µm, consistent with partial-volume effects near the detection limit. Systematic differences between CT and microscopy-based defect size measurements are reported in the literature and depend on CT resolution, scan parameters and defect type [12,39]. The inset shows a zoomed view of the region $\sqrt{\text{Area}} \leq 120$ µm, where 20 of the 21 matched defects are located.

required to trigger these corrections; the effective classification in the criticality models is therefore identical to the binary scheme applied in the crack growth models. In series 'small', fracture is predominantly initiated from internal defects (93 %), reflecting the concentrated HSV that limits defect competition to the most critical, typically largest defect and reduces the influence of the surface area. In series 'large', a substantially higher fraction of surface-initiated failures is observed (40 %), consistent with the larger gauge section that exposes a greater surface area to the applied stress. The complete specimen-level data, including $\sqrt{\text{Area}}$, l_n, stress amplitude and cycles to failure, are provided in Supplementary Material S5. Fig. 8 visualises the complete fractographic dataset in a $\sqrt{\text{Area}}$-l_n plane. A logarithmic scale is used for the ligament distance to accommodate the large dynamic range of nearly three decades spanned by the internal fracture origins; surface-connected defects with $l_n = 0$ are plotted at $l_n = 5$ µm for visibility on this scale, as indicated in the figure caption. The surface-connected failures (solid circles) cluster below the classification threshold of $l_{n,\text{thr}} = 20$ µm and are substantially more frequent in series 'large' than in series 'small' (17 vs. 2 specimens), consistent with the larger exposed gauge surface area. The internal fracture origins (crosses) are broadly distributed in the depth direction, ranging from just above the threshold to $l_n \approx 2700$ µm, with no apparent correlation between defect size and ligament distance. The defect sizes of internal fracture origins in series 'large' tend to be larger than those in series 'small', reflecting the larger HSV and the increased probability of encountering a large defect at depth. This wide range of ligament distances directly determines the relative contribution of the internal crack growth phase: shallow internal defects transition to air-environment growth early, whereas deeply embedded defects accumulate the majority of their fatigue-life in the self-contained environment (Section 2.8).

3.4. Part A: Defect criticality ranking benchmark

21 specimens are matched between NDT and fracture surface data. Table 8 summarises the ranking accuracy of all investigated models, evaluated on the 21 matched specimens (5 from series 'small', 16 from series 'large'). Two metrics are reported: the hit rate H (fraction of specimens where the model correctly identifies the fracture-initiating defect as rank 1) and the mean normalised rank $\bar{R}$ (lower values indicate better performance). The TCD model is evaluated in four variants

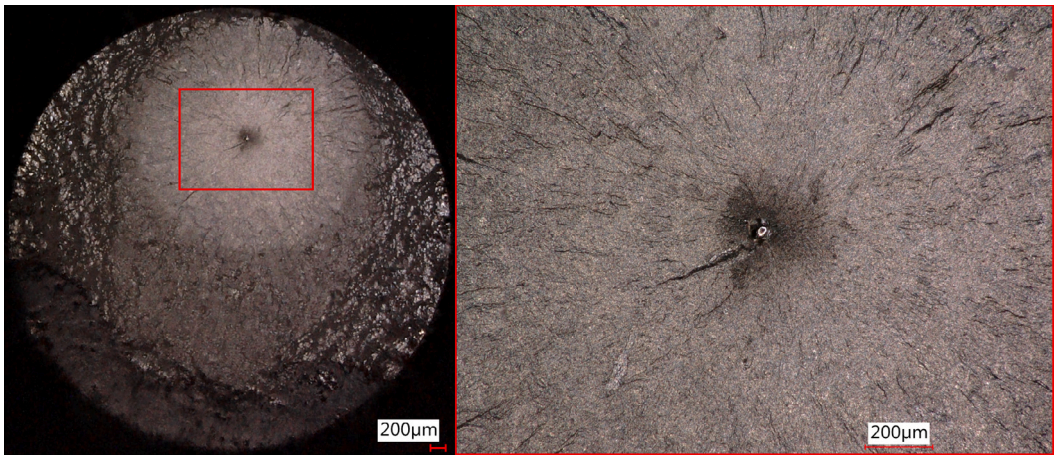
Table 7
Summary of fracture-origin classification for all specimens with available fractographic data.

Series	n_{total}	Surface-connected ($l_n \leq 20$ µm)	Internal ($l_n > 20$ µm)
'small'	30	2	28
'large'	43	17	26
Total	73	19	54

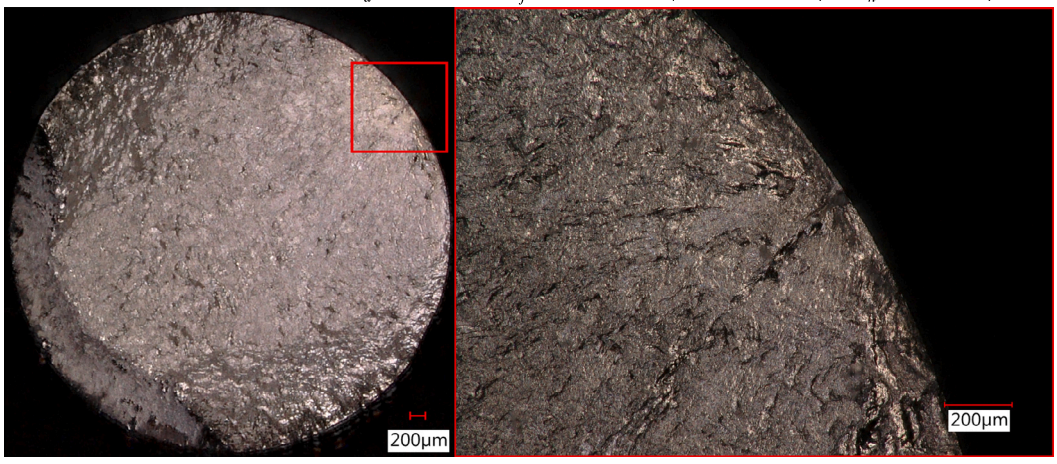
(PM, LM, AM, VM) as described in . The ε_{pl} metric is only applicable to EP simulations and therefore not listed under the LE results.

For series 'small', all models except the pure size ranking achieve identical hit rates of 3/5, with mean normalised ranks between 0.068 and 0.097. The size ranking fails to identify any fracture-initiating defect as the most critical ($H = 0/5$, $\bar{R} = 0.437$), indicating that defect size alone is an insufficient criticality criterion when the stress field is non-uniform. The negligible differences among the remaining models suggest that, for a small HSV with few competing defects, even simple geometry-based approaches perform as well as computationally expensive simulation-based methods.

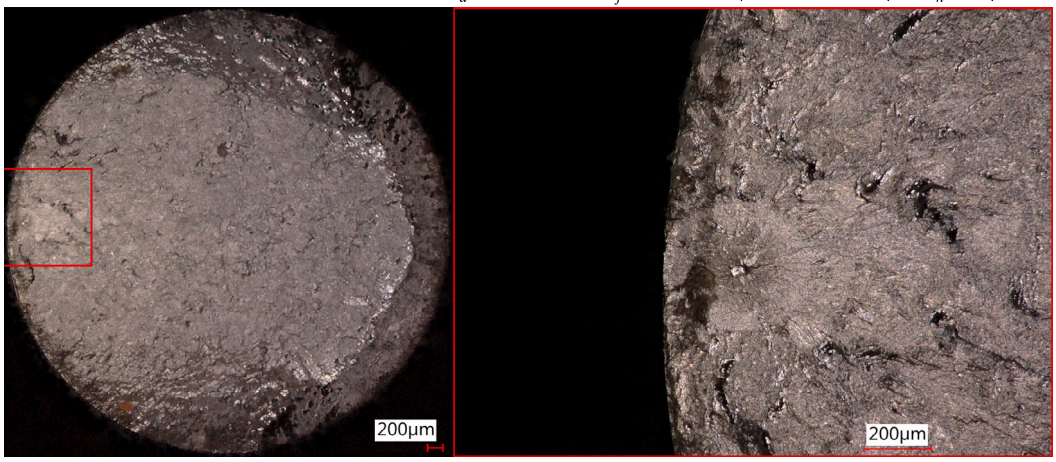
For series 'large', where a larger HSV increases the number of competing defects, greater differentiation between models emerges. Among the geometry-based models, the Tammas-Williams model achieves the highest hit rate ($H = 9/16$, $\bar{R} = 0.171$), outperforming both the Murakami model ($H = 6/16$) and the size ranking ($H = 5/16$). Among the simulation-based models with a LE material model, the averaged strain energy density $\bar{W}$ ($H = 8/16$, $\bar{R} = 0.136$), the relative stress gradient $\sigma_{\text{supported}}$ ($H = 8/16$, $\bar{R} = 0.132$), and the TCD line method ($H = 8/16$, $\bar{R} = 0.133$) achieve the highest hit rates. Notably, the k_g factor and the SWT parameter perform poorly in the LE variant ($H = 2/16$), suggesting that peak-value-based metrics are sensitive to mesh-dependent stress concentrations at sharp defect features. The TCD variants show a spread from $H = 5/16$ (σ_{AM}) to $H = 8/16$ (σ_{LM}), indicating that the choice of averaging domain influences the ranking outcome.



(a) Fracture from internal defect, $\sigma_a = 530\,\text{MPa}$, $N_f = 1\,910\,310$, $\sqrt{\text{Area}} = 69.43\,\mu\text{m}$, $l_n = 1744.14\,\mu\text{m}$



(b) Fracture from surface-connected defect, $\sigma_a = 530\,\text{MPa}$, $N_f = 35\,392$, $\sqrt{\text{Area}} = 42.34\,\mu\text{m}$, $l_n = 0\,\mu\text{m}$



(c) Fracture from internal defect with small ligament distance, $\sigma_a = 530\,\text{MPa}$, $N_f = 2\,064\,407$, $\sqrt{\text{Area}} = 40.95\,\mu\text{m}$, $l_n = 200.45\,\mu\text{m}$

Fig. 7. Representative examples of fracture origins observed in the present dataset. (a) Internal defect showing a fish-eye fracture pattern typical of subsurface crack initiation in a self-contained environment. (b) Surface-connected defect directly intersecting the machined specimen surface. (c) Near-surface defect with a measurable ligament to the free surface ($l_n \approx 200\,\mu\text{m}$).

The poor ranking performance of peak-value-based metrics (k_g, P_{SWT}) in the LE variant ($H = 2/16$) is attributed to the sensitivity of local peak stresses to mesh density and to near-singular stress concentrations at sharp features of the CT-segmented defect surface. At the scale of CT-resolved defects (10 µm to 200 µm), the continuum stress field at the defect boundary is strongly influenced by voxel-induced geometry artefacts, segmentation smoothing, and the local mesh resolution, all of which are inherent to the CT-to-FE workflow and are not representative of the true crack-tip field. Metrics that integrate the stress field over a finite, material-relevant length scale, such as $\bar{W}$ ($R_c = 6.62\,\mu\text{m}$), $\sigma_{\text{supported}}$, and the TCD variants ($L = 8.72\,\mu\text{m}$), are more robust in this respect, as the averaging domain partially filters mesh-scale

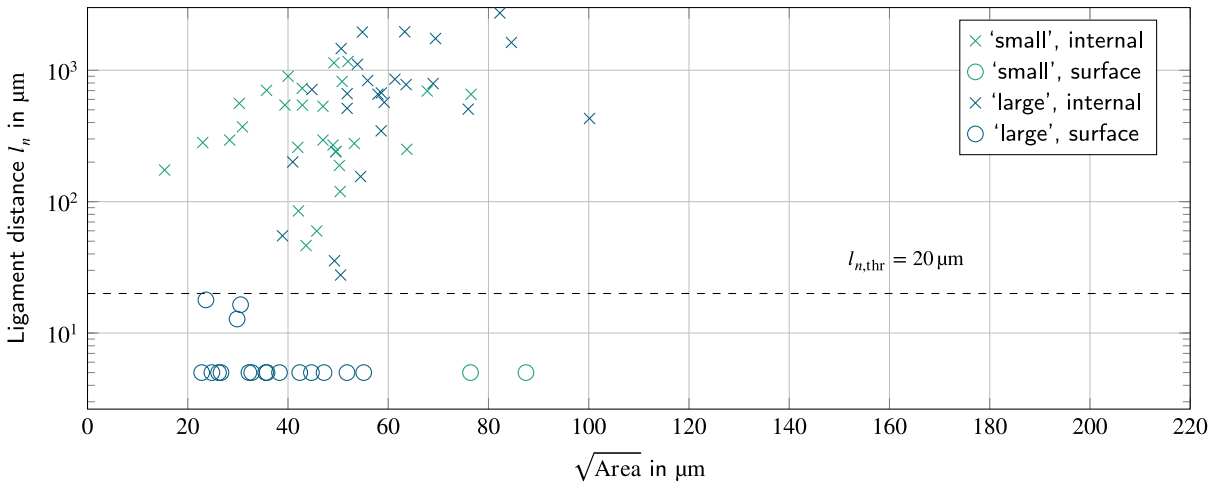


Fig. 8. Fractographically measured defect size $\sqrt{\mathrm{Area}}$ and ligament distance l_n for all 73 specimens with available fracture surface data. Filled circles: surface-connected fracture origins ($l_n = 0$, plotted at $l_n = 5$ µm for visibility); crosses: internal fracture origins. The dashed line marks the surface/internal classification threshold $l_{n,\mathrm{thr}} = 20\,\mu\mathrm{m}$ used in the crack growth models (Section 2.8). The complete numerical data are provided in Supplementary Material S5.

Table 8
Model comparison by ranking accuracy.

Series		Model	Hit rate H	Mean normalised rank $\bar{R}$
small	Geometry-based	Size	0/5	0.437
		Murakami	3/5	0.068
		Tammas-Williams	3/5	0.068
	Simulation-based LE	k_g	3/5	0.091
		$\bar{W}$	3/5	0.091
		P_{SWT}	3/5	0.091
		$\sigma_{\mathrm{supported}}$	3/5	0.097
	σ_{TCD}	PM	3/5	0.097
		LM	3/5	0.097
		AM	3/5	0.097
		VM	3/5	0.097
	EP	k_g	3/5	0.091
		$\bar{W}$	3/5	0.097
		P_{SWT}	3/5	0.091
		ϵ_{pl}	3/5	0.075
large	Geometry-based	Size	5/16	0.218
		Murakami	6/16	0.166
		Tammas-Williams	9/16	0.171
	Simulation-based LE	k_g	2/16	0.275
		$\bar{W}$	8/16	0.136
		P_{SWT}	2/16	0.269
		$\sigma_{\mathrm{supported}}$	8/16	0.132
	σ_{TCD}	σ_{PM}	6/16	0.193
		σ_{LM}	8/16	0.133
		σ_{AM}	5/16	0.220
		σ_{VM}	7/16	0.191
	EP	k_g	5/16	0.258
		$\bar{W}$	5/16	0.289
		P_{SWT}	5/16	0.300
		ϵ_{pl}	4/16	0.298

artefacts. This behaviour is consistent with the observation by Tammas-Williams et al. [17], who found that FE-based ranking on CT-segmented pores did not improve over geometry-based approaches, attributing this to CT resolution limitations and segmentation uncertainties.

The EP material model does not improve the ranking accuracy compared to the best LE results. All EP metrics (k_g, $\bar{W}$, P_{SWT}, ϵ_{pl}) achieve hit rates of 4/16–5/16 with mean normalised ranks between 0.258 and 0.300, which is consistently worse than the best-performing LE and geometry-based models. This indicates that stress redistribution through local yielding reduces the discriminative power of peak-based ranking metrics by homogenising the stress–strain response across defects of different severity.

The mean normalised rank $\bar{R}$ provides complementary information to the hit rate by quantifying the severity of ranking errors in specimens where the fracture-initiating defect is not assigned rank 1. For the best-performing LE simulation-based metrics, $\sigma_{\mathrm{supported}}$ and σ_{LM} achieve $\bar{R} = 0.132$ and $\bar{R} = 0.133$, respectively, indicating that on average the true fracture origin is ranked within the top 13 % of all detected defects even when the model assigns rank 1 to a different defect. The Tammas-Williams model achieves the highest hit rate ($H = 9/16$) at a somewhat higher mean normalised rank ($\bar{R} = 0.171$), reflecting a trade-off: it identifies the critical defect directly more often but, when incorrect, places the true fracture initiator slightly deeper in the ranking than the best LE simulation-based models. For the EP metrics, the simultaneously elevated $\bar{R} \approx 0.26$–0.30 indicates that even in specimens

Table 9
Input requirement matrix for the investigated models.

Data group	Parameter	Geometry-based			Simulation-based	
		Size	Murakami	Tammas-Williams	LE	EP
3D	CT defect data	•	•	•	•	•
	Specimen surface data	○	•	•	•	•
Mesh data	Mesh of defect proximity	○	○	○	•	•
Simulation	Defect free stress field	◐[a]	•	•	•	•
	Defect free strain field	○	○	○	◐[b]	•
	A-priori FE database	○	○	•	○	○
	Individual FE simulation	○	○	○	•	•
Defect	Distance to other defects	○	•	•	○	○
	$\sqrt{\text{Area}}$	○	•	•	○	○
	Aspect ratio	○	○	•	○	○
Material	Hardness HV	○	•	○	○	○
	Young's modulus E	○	○	•	•	•
	Poisson ratio ν	○	○	•	•	•
	Yield strength σ_y	○	○	•	○	•

[a] Stress field only required to assess projected area orientation.
[b] Strain field is function of stress field.

where the correct defect is missed, the fracture origin is on average only ranked within the top 25 %–30 % of the detected population. In practical terms, the mean normalised rank indicates how many defects an analyst would need to investigate in priority order before reaching the true fracture initiator: for $\bar{R} = 0.132$ and a mean defect count of ≈14, this corresponds to examining approximately 2–3 defects on average.

Overall, the results demonstrate that geometry-based models, particularly the Tammas-Williams model, achieve ranking accuracy comparable to or exceeding that of the best simulation-based approaches, despite requiring orders of magnitude less computational effort. The additional complexity of individual FE sub-model simulations does not translate into a measurably higher ranking accuracy for the present dataset.

3.4.1. Computational cost and input complexity

Table 9 summarises the input data required for each model category. The size model requires the least amount of parameters, followed by the Murakami model. The other models have similar requirements, with some simulation-based metrics needing further parameters for correlations and the definition of characteristical lengths. The Tammas-Williams model has the highest parameter requirements, due to the combination of geometry-based analysis and a-priori FE simulations.

The geometry-based models operate directly on the X-ray CT reconstruction and the defect-free stress field; their evaluation required less than 1 min per specimen on a single CPU core, though the exact time scales with the number of detected defects per specimen (1 to 31 in the present dataset). The simulation-based models additionally require volume meshing and individual FE solutions for each defect. With the LE material model, the mean computation time was approximately 10 min per specimen (including meshing, solving and post-processing for all defects) with 4 CPU cores on 16 GB memory; the EP model increased this to approximately 30 min per specimen due to the iterative nonlinear solution. All simulation-based ranking metrics (k_g, $\bar{W}$, χ^*, P_{SWT}, ε_{pl}, σ_{TCD}) are extracted from the same FE results, so the marginal cost of evaluating additional metrics is negligible once the simulation is completed. The simulation-based models require approximately one to two orders of magnitude more computation time than the geometry-based approaches, yet this additional effort does not translate into a measurably higher ranking accuracy for the present dataset.

3.4.2. Root cause analysis

A detailed analysis of the discriminative power of all metrics (Area Under the ROC Curve (AUC), partial correlations controlling for $\sqrt{\text{Area}}$, and within-specimen Coefficient of Variation (CoV) comparison) is provided in Supplementary Material S4. The key finding is that all FE-based metrics retain negligible independent information after controlling for defect size ($|\rho_{\text{partial}}| \leq 0.039$), confirming that $\sqrt{\text{Area}}$ is the dominant criticality driver. This conclusion is not sensitive to the specific literature values used for the material length scales L, R_c, and a^*. Since all three parameters ($L = 8.72\,\mu\text{m}$, $R_c = 6.62\,\mu\text{m}$, $a^* = 0.07\,\text{mm}$) are smaller than or comparable to the lower end of the observed defect size range ($\sqrt{\text{Area}}$: 15 µm to 200 µm), the averaging domains of the respective metrics are dominated by the defect geometry itself rather than by the precise parameter values. Furthermore, the near-zero partial correlations of the TCD, SED and support-gradient metrics ($|\rho_{\text{partial}}| \leq 0.031$, 0.005, and 0.023, respectively; Supplementary Material S4) confirm that none of these models provides independent ranking information beyond defect size, irrespective of the choice of L, R_c or a^* within physically plausible ranges. The EP material model reduces the within-specimen metric spread by 32 %, explaining its consistently lower ranking performance. Combined ranking metrics achieve a maximum hit rate of 10/16, suggesting a performance ceiling governed by stochastic microstructural effects.

3.5. Part B: Fatigue-life prediction

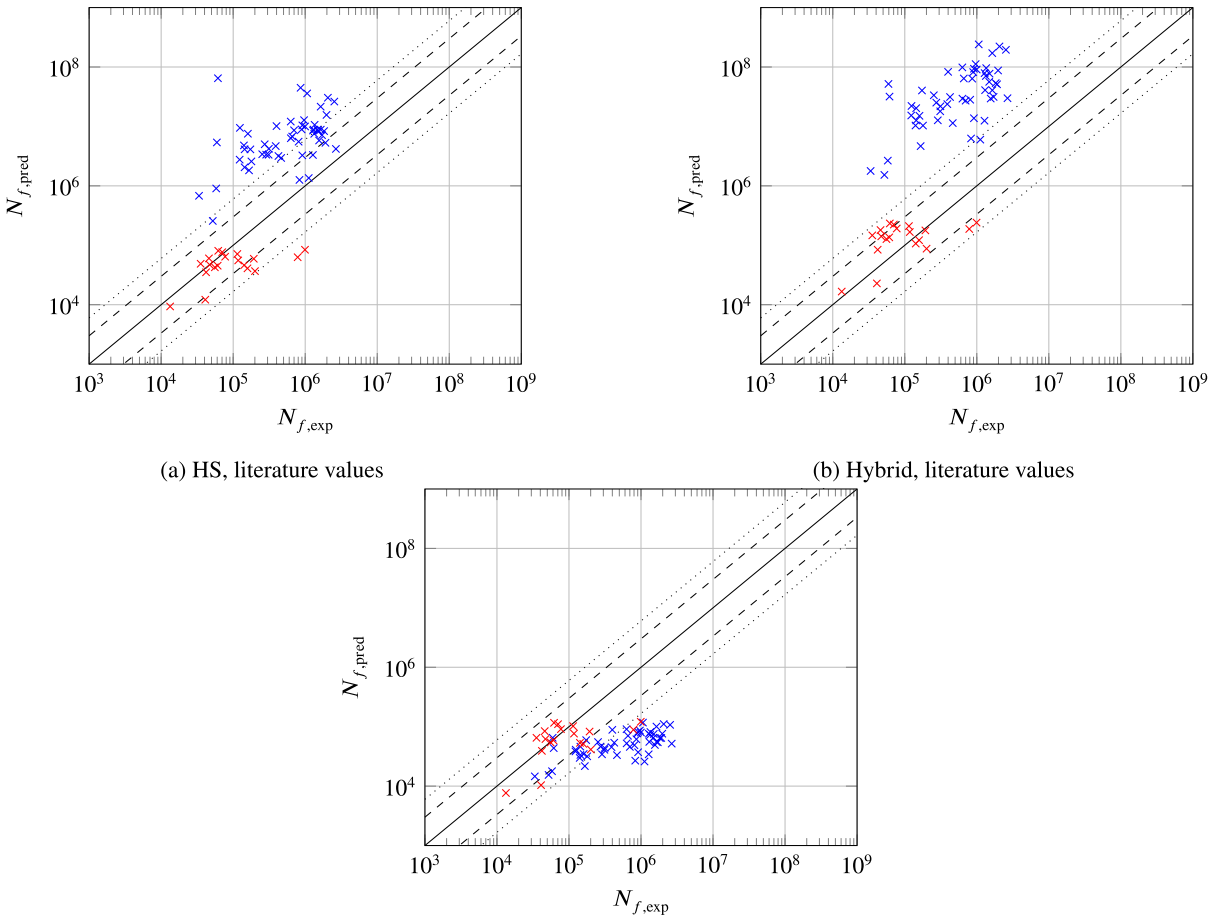
The presentation follows a logical progression: Section 3.5.1 first applies all models with literature parameters to establish the baseline performance (Phase 1); Section 3.5.2 then fits only the internal crack growth parameters to the present PBF-LB/M dataset (Phase 2); Section 3.5.3 quantifies the sensitivity of the fitted parameters to key modelling assumptions; and Section 3.5.4 provides a direct comparison of the threshold formulations.

3.5.1. Phase 1: Direct comparison without parameter fitting

The specimen-level input data ($\sqrt{\text{Area}}$, l_n, $N_{f,\text{exp}}$) used for all life predictions are listed in Supplementary Material S5. Fig. 9 shows the predicted versus experimentally observed cycles to failure for all specimens, using the crack growth parameters from Table 4 without any fitting to the present dataset. For all models, the initial crack size is taken as $a_0 = \sqrt{\text{Area}}$ from the dominant defect identified on the fracture surface via optical microscopy at 500× magnification, and the geometry factor is $Y = 0.65$ for surface defects and $Y = 0.5$ for internal defects following Murakami [13].

Table 10 summarises the prediction accuracy of the three models. The metrics include the root-mean-square error of the logarithmic life ratio ($\text{RMSE}_{\log}$), the percentage of predictions falling within factor-3 and factor-6 scatter bands, and the mean logarithmic error (bias $\bar{\eta}$).

The unfitted results reveal a clear performance split between surface and internal defect subsets (Table 10 and Fig. 9). For surface-initiated



(a) HS, literature values

(b) Hybrid, literature values

(c) ME, literature values

Fig. 9. Predicted versus experimental fatigue-life for all three crack growth models with literature parameters (no parameter fitting, Phase 1). Results shown for all 73 specimens; red markers: surface-initiated failure (n = 19); blue markers: internally initiated failure (n = 54). Solid line: 1:1 correspondence; dashed lines: factor-3 scatter band; dotted lines: factor-6 scatter band.

Table 10
Phase 1 prediction accuracy metrics (literature parameters, no parameter fitting to present dataset). Results shown separately for surface-initiated (n = 19) and internally initiated (n = 54) failures.

Environment	Model	$\bar{\eta}$	$\mathrm{RMSE}_{\log}$	Within	
				±3× (%)	±6× (%)
Surface	HS	−0.300	0.475	63.158	89.474
	ME	−0.199	0.428	73.684	89.474
	Hybrid	0.135	0.406	68.421	100
Internal	HS	1.050	1.153	9.259	25.926
	ME	−1.012	1.086	7.407	29.630
	Hybrid	1.771	1.820	0	1.852

failures, the HS model achieves an $\mathrm{RMSE}_{\log}$ of 0.475 with 63.2 % of predictions within the factor-3 scatter band and 89.5 % within factor 6. The small negative bias ($\bar{\eta}$ = −0.300) indicates a slight systematic underprediction of fatigue-life, i.e. the model is marginally conservative for surface defects. The ME model achieves an $\mathrm{RMSE}_{\log}$ of 0.428 for surface defects, with a similarly conservative bias ($\bar{\eta}$ = −0.199) and the highest fraction within the factor-3 band (73.7 %). The hybrid model exhibits the lowest surface $\mathrm{RMSE}_{\log}$ = 0.406 with an unconservative overprediction for ($\bar{\eta}$ = +0.135) and 68.4 % within factor 3, but achieves 100 % within factor 6.

For internally initiated failures, all three models perform substantially worse. The HS model strongly overpredicts fatigue-life ($\bar{\eta}$ = +1.050, $\mathrm{RMSE}_{\log}$ = 1.153), with only 9.3 % of predictions within factor 3 and 25.9 % within factor 6. This indicates that the internal (vacuum) parameters from Han et al. [41], derived from wrought TiAl6V4, predict excessively slow crack growth for the present PBF-LB/M material. The ME model underpredicts internal fatigue lives ($\bar{\eta}$ = −1.012, $\mathrm{RMSE}_{\log}$ = 1.086), with 7.4 % within factor 3 and 29.6 % within factor 6. Since the ME model does not distinguish between air and self-contained environments, this underprediction reflects its inability to capture the reduced crack growth rate in the internal self-contained environment. The hybrid model shows the largest discrepancy for internal defects ($\bar{\eta}$ = +1.771, $\mathrm{RMSE}_{\log}$ = 1.820), with 1.9 % of predictions falling within the factor-6 band. This extreme overprediction arises because the combination of the Murakami-derived evolving threshold (which is higher than the fixed HS threshold for larger defects) with the unfitted Han et al. vacuum parameters produces an excessively low predicted crack growth rate.

These results confirm two key observations: (i) the surface-in-air parameters from Jones et al. [40] transfer reasonably well to the present PBF-LB/M TiAl6V4 without refitting, with the hybrid model achieving the lowest surface $\mathrm{RMSE}_{\log}$ = 0.406, followed by ME (0.428) and HS (0.475); and (ii) the internal crack growth parameters require material-specific parameter fitting, as all three models show $\mathrm{RMSE}_{\log}$ > 1.08 for internal defects. This motivates the parameter fitting of the internal parameters in Phase 2 (Section 3.5.2).

3.5.2. Phase 2: Parameter fitting of internal crack growth parameters

Given the discrepancy observed for internal defects in Phase 1, the internal crack growth parameters are fitted to the present dataset while

Table 11
Fitted internal crack growth parameters compared to literature values by Han et al. [41].

Parameter	Model	Literature value	Fitted value
D_{int}	HS	2.897×10^{-13} MPa$\sqrt{\text{m}}$	2.471×10^{-13} MPa$\sqrt{\text{m}}$
p_{int}	HS	4.18	4.564
$\Delta K_{\text{thr,int}}$	HS	2.0 MPa$\sqrt{\text{m}}$	0.5 MPa$\sqrt{\text{m}}$
D_{int}	Hybrid	2.897×10^{-13} MPa$\sqrt{\text{m}}$	3.388×10^{-11} MPa$\sqrt{\text{m}}$
p_{int}	Hybrid	4.18	3.12

Table 12
Phase 2 prediction accuracy metrics after parameter fitting of internal crack-growth parameters to the present PBF-LB/M TiAl6V4 dataset. ME model shown unchanged for reference.

Model	$\bar{\eta}$	$\text{RMSE}_{\log}$	Within ±3× (%)	Within ±6× (%)
HS (fitted internal)	−0.080	0.425	75.34	94.52
Hybrid (fitted internal)	0.033	0.386	76.71	95.89
ME (unchanged)	−0.800	0.959	24.66	45.21

keeping the surface-in-air parameters from Jones et al. [40] unchanged. This approach is motivated by the observation that internal crack growth occurs in a self-contained environment whose characteristics may differ between wrought and PBF-LB/M material due to differences in microstructure, residual porosity, and local chemistry at the crack tip.

For the HS model, three parameters are fitted: D_{int}, p_{int}, and $\Delta K_{\text{thr,int}}$, with $A_{\text{HS}} = 90\,\text{MPa}\sqrt{\text{m}}$ held fixed (its insensitivity in HCF is confirmed in Section 3.5.3 and Supplementary Material S3). For the hybrid model, only D_{int} and p_{int} are fitted, since ΔK_{thr} is determined by the Murakami formulation at each crack size. The parameter fitting minimises the mean squared logarithmic life error over the internal-defect subset:

$$\text{obj} = \frac{1}{n_{\text{int}}} \sum_{i=1}^{n_{\text{int}}} \left[\log_{10} \left(\frac{N_{f,\text{pred},i}}{N_{f,\text{exp},i}} \right) \right]^2 . \tag{34}$$

Table 11 compares the fitted internal parameters with the literature values. The fitted HS coefficient $D_{\text{int}} = 2.471 \times 10^{-13}\,\text{MPa}\sqrt{\text{m}}$ is slightly smaller than the Han et al. value of $2.897 \times 10^{-13}\,\text{MPa}\sqrt{\text{m}}$, indicating comparable internal crack growth rates. The dominant change lies in the substantially reduced threshold $\Delta K_{\text{thr,int}}$, which allows crack propagation at lower driving forces and thus accelerates the predicted life consumption in the near-threshold regime. The exponent p_{int} increases slightly from 4.18 to 4.564, steepening the near-threshold sensitivity. Most notably, the fitted threshold $\Delta K_{\text{thr,int}} = 0.5\,\text{MPa}\sqrt{\text{m}}$ is substantially lower than the literature value of $2.0\,\text{MPa}\sqrt{\text{m}}$. This suggests that internal cracks in PBF-LB/M material begin propagating at lower driving forces than in wrought material, consistent with the irregular, tortuous morphology of process-induced defects that may locally elevate the effective stress intensity factor.

For the hybrid model, the fitted D_{int} is approximately two orders of magnitude larger than in the HS model ($3.388 \times 10^{-11}\,\text{MPa}\sqrt{\text{m}}$ vs. $2.471 \times 10^{-13}\,\text{MPa}\sqrt{\text{m}}$), compensating for the higher evolving threshold imposed by the Murakami formulation. The lower exponent ($p = 3.12$ vs. 4.564) reflects the different functional form of the driving force near the threshold.

After parameter fitting, Table 12 and Fig. 10 present the improved prediction accuracy for the full dataset. Both parameter fitted models achieve $\text{RMSE}_{\log} < 0.45$ over all specimens, with more than 90 % of predictions falling within the factor-6 scatter band.

The parameter fitting reduces the $\text{RMSE}_{\log}$ for internally initiated specimens from 1.153 (Phase 1) to 0.406 for the HS model and from 1.820 to 0.379 for the hybrid model, while leaving surface-initiated predictions unchanged ($\text{RMSE}_{\log} = 0.475$ for HS, $\text{RMSE}_{\log} = 0.406$ for hybrid, since the surface parameters are identical). The overall improvement is driven entirely by the internal subset, confirming that the surface-in-air parameters from Jones et al. [40] require no modification for the present material.

Table 13
Leave-one-out cross-validation results for the parameter fitted crack growth models (internal defect subset, $n = 54$).

	$\text{RMSE}_{\log}$			CoV (%)		
Model	In-sample	LOOCV	Δ	D	p	ΔK_{thr}
HS	0.406	0.425	0.019	13.9	2.4	1.0
Hybrid	0.379	0.391	0.011	2.7	1.2	–

The ME model remains unchanged in Phase 2, as it does not use separate environmental parameters. Its overall $\text{RMSE}_{\log}$ of 0.959 reflects the systematic underprediction of internal lives ($\bar{\eta} = -0.800$) combined with reasonable surface performance ($\bar{\eta} = -0.199$). This confirms that a single-environment model cannot simultaneously capture both failure populations.

The hybrid model achieves the lowest overall $\text{RMSE}_{\log}$ (0.386) with one fewer free parameter than the HS model (ΔK_{thr} is not fitted), demonstrating that the Murakami-derived evolving threshold provides a physically motivated and effective replacement for the empirical threshold. Its slight positive overall bias ($\bar{\eta} = +0.033$) indicates near-unbiased predictions, compared to the slightly conservative HS model ($\bar{\eta} = -0.080$).

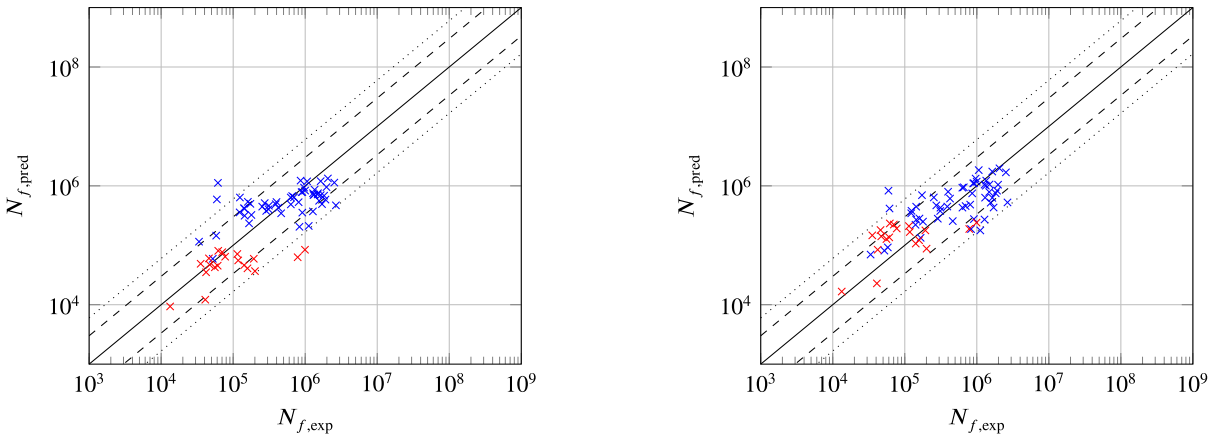
It is important to emphasise that the fitted internal parameters (D_{int}, p_{int}, $\Delta K_{\text{thr,int}}$) should be interpreted as *effective parameters* of the proposed modelling framework, derived from total fatigue-life observations and post-mortem fracture-origin measurements. They were not obtained from direct observation of internal crack-growth kinetics and may implicitly absorb contributions from short-crack behaviour, defect morphology variability, uncertainty in $\sqrt{\text{Area}}$ and l_n, the gradual transition from internal to surface crack growth, and simplifications in the adopted SIF formulation (Eq. (26)). This framing does not weaken the practical utility of the approach, but the parameters should not be transferred to other material conditions without independent validation.

The hybrid and HS models are internally validated for the present PBF-LB/M TiAl6V4 dataset at $R = -1$. Broader generalisation to other build parameters, specimen geometries, stress ratios, material conditions or defect populations would require independent validation on a separate dataset not used in the fitting.

Taken together, the results of Sections 3.1–3.5.2 represent a self-contained experimental chain for defect-tolerant fatigue assessment of PBF-LB/M TiAl6V4: pre-test X-ray CT inspection, fatigue testing to failure, post-mortem fractographic identification of fracture origins, CT/fractography matching, and fitting of effective internal crack-growth parameters from total fatigue-life data. The systematic combination of all five steps on a single, well-defined specimen population constitutes a substantive experimental contribution independent of the model benchmark results. Even where individual steps have been reported in previous studies, their consistent execution at the population level, across 84 specimens, two CT resolutions, and two specimen geometries, provides a dataset whose value extends beyond the specific model comparisons presented here.

Cross-validation. To assess the generalisability of the fitted internal parameters, leave-one-out cross-validation (LOOCV) was performed on the $n_{\text{int}} = 54$ internally initiated specimens. In each fold, one specimen is held out, the internal parameters are refitted on the remaining $n_{\text{int}} - 1$ specimens, and the held-out specimen's life is predicted with the fold-specific parameters. Table 13 summarises the results.

The cross-validated $\text{RMSE}_{\log}$ exceeds the in-sample value by $\Delta = +0.011$ for the hybrid model and $\Delta = +0.019$ for the HS model, confirming that overfitting is negligible for both formulations despite the small fitting dataset. The parameter stability across folds confirms that both parameter fits are robust: the HS model achieves CoV values of 13.9 % for D, 2.4 % for p, and 1.0 % for ΔK_{thr}, while the hybrid model achieves CoV < 3.0 % for both D and p. Both models thus yield uniquely



(a) HS, fitted internal

(b) Hybrid, fitted internal

Fig. 10. Predicted versus experimental fatigue-life after fitting of internal crack-growth parameters (Phase 2). Surface parameters unchanged from Jones et al. [40]. Red markers: surface-initiated failure ($n = 19$); blue markers: internally initiated failure ($n = 54$). Solid line: 1:1 correspondence; dashed lines: factor-3 scatter band; dotted lines: factor-6 scatter band.

identified parameter sets on the present dataset (two vs. three free parameters), which constitutes a practical advantage when the fitting dataset is small. The negligible LOOCV degradation ($\Delta < 0.02$ for both models) confirms that neither formulation overfits to the fitting data.

3.5.3. Sensitivity studies

The robustness of the predictions was assessed through systematic sensitivity studies on the key model parameters: the ligament threshold l_n, the cyclic fracture toughness A_{HS}, the failure criterion, and the fractographic measurement uncertainty. The principal findings are:

- The ligament threshold $l_{n,\mathrm{thr}}$ governs the surface/internal classification. For $l_{n,\mathrm{thr}} \leq 20\,\mu\mathrm{m}$, prediction accuracy is insensitive to the exact value ($\mathrm{RMSE}_{\log}$ varies by approximately 0.04); beyond 50 µm, accuracy degrades significantly as internal defects are misclassified.
- The cyclic fracture toughness A_{HS} is confirmed to be insensitive in HCF in terms of *prediction quality*: across $A_{\mathrm{HS}} \in 50\,\mathrm{MPa}$ to $170\,\mathrm{MPa}\sqrt{\mathrm{m}}$, the $\mathrm{RMSE}_{\log}$ varies by less than 0.003. Within the physically relevant range ($A_{\mathrm{HS}} \geq 60\,\mathrm{MPa}\sqrt{\mathrm{m}}$), the fitted parameters are remarkably stable (D: < 20 %, p: < 5 %, $\Delta K_{\mathrm{thr}} \approx 0.5\,\mathrm{MPa}\sqrt{\mathrm{m}}$). Below $A_{\mathrm{HS}} = 60\,\mathrm{MPa}\sqrt{\mathrm{m}}$, the optimiser converges to a qualitatively different solution. This confirms that A_{HS} neither affects the life prediction nor the fitted parameters within the Jones et al. [40] range.
- The fractographic measurement uncertainty propagates into a mean life uncertainty of approximately 1.5× for the parameter fitted models (excluding classification-boundary specimens), which is well within the factor-3 scatter band. The defect size ($\sqrt{\mathrm{Area}}$) is the dominant contributor (∼1.5×), while the ligament distance l_n has negligible influence (∼1.0×).
- The failure criterion has negligible influence: $\mathrm{RMSE}_{\log}$ varies by less than 0.022 for the parameter fitted models across all tested criteria.

Full details are provided in Supplementary Material S3.

3.5.4. Comparison of threshold formulations

A key distinction between the HS and hybrid models lies in the treatment of the crack growth threshold. The HS model uses a fixed ΔK_{thr} (fitted or taken from literature), while the hybrid model derives a crack-size-dependent threshold $\Delta K_{\mathrm{thr}}(a)$ from the Murakami fatigue limit $\sigma_w(a)$ at the current crack size (Eq. (30)):

$$\Delta K_{\mathrm{thr}}(a) = Y \cdot f(R) \cdot \sigma_w(a) \cdot \sqrt{\pi a}\,. \tag{35}$$

Fig. 11 compares the Murakami-derived threshold at the initial defect size $a_0 = \sqrt{\mathrm{Area}}$ with the fixed HS thresholds (literature value: $2.0\,\mathrm{MPa}\sqrt{\mathrm{m}}$; fitted value: $0.5\,\mathrm{MPa}\sqrt{\mathrm{m}}$). The Murakami-derived values range from 1.68 MPa to $3.95\,\mathrm{MPa}\sqrt{\mathrm{m}}$. For surface defects, the range is 2.29 MPa to $3.58\,\mathrm{MPa}\sqrt{\mathrm{m}}$, exceeding the fixed HS threshold of $2.0\,\mathrm{MPa}\sqrt{\mathrm{m}}$. All values exceed the fitted HS threshold of $0.5\,\mathrm{MPa}\sqrt{\mathrm{m}}$. The threshold increases monotonically with defect size through the $a^{1/3}$ dependence inherent in the $\sqrt{\mathrm{Area}}$ parameter model ($\Delta K_{\mathrm{thr}} \propto \sigma_w \cdot \sqrt{a} \propto a^{-1/6} \cdot a^{1/2} = a^{1/3}$). This observation helps explain the different behaviour of the unfitted and fitted hybrid models.

First, in the *unfitted* hybrid model, the Murakami-derived threshold is consistently higher than the fixed threshold used in the HS model. As a result, the effective driving force ($\Delta K - \Delta K_{\mathrm{thr}}$) is reduced. This directly lowers the predicted crack growth rate. Combined with the already slow vacuum parameters from Han et al. [41], this leads to a strong overprediction of fatigue-life in Phase 1 ($\mathrm{RMSE}_{\log} = 1.820$ for internal defects).

Second, after parameter fitting, this effect is largely compensated by a larger fitted coefficient D_{int} (Table 11). The parameter fitted hybrid model offsets the higher evolving threshold by increasing the overall crack growth rate. The net effect is that the *fitted* hybrid model achieves slightly better prediction accuracy ($\mathrm{RMSE}_{\log} = 0.386$) than the fitted HS model ($\mathrm{RMSE}_{\log} = 0.425$), despite the higher threshold.

A key difference remains, however, in how both models represent near-threshold behaviour. The HS model relies on a constant fitted threshold, whereas the hybrid model uses an evolving threshold $\Delta K_{\mathrm{thr}}(a)$ that increases with crack size. This gives the hybrid formulation a more physically motivated near-threshold response, which may explain its slightly better overall fit.

For surface defects (red markers in Fig. 11), the Murakami-derived thresholds range from 2.29 MPa to $3.58\,\mathrm{MPa}\sqrt{\mathrm{m}}$, lying 15 % to 79 % above the fixed value of $2.0\,\mathrm{MPa}\sqrt{\mathrm{m}}$. For internal defects (blue markers), the range extends to $3.95\,\mathrm{MPa}\sqrt{\mathrm{m}}$ due to the larger typical defect sizes. The internal defects also exhibit higher threshold values because the subsurface location factor ($C = 1.56$ vs. $C = 1.43$ for surface) increases the Murakami fatigue limit and, consequently, the derived threshold.

The fitted HS threshold of $0.5\,\mathrm{MPa}\sqrt{\mathrm{m}}$ lies below all Murakami-derived values. The parameter fitted HS model allows crack growth at lower driving forces than the hybrid model's evolving threshold. This difference is compensated by the lower HS coefficient ($D_{\mathrm{int,HS}} = 2.471 \times 10^{-13}\,\mathrm{MPa}\sqrt{\mathrm{m}}$ vs. $D_{\mathrm{int,hybrid}} = 3.388 \times 10^{-11}\,\mathrm{MPa}\sqrt{\mathrm{m}}$), such that both parameter fitted models produce similar integrated lives.

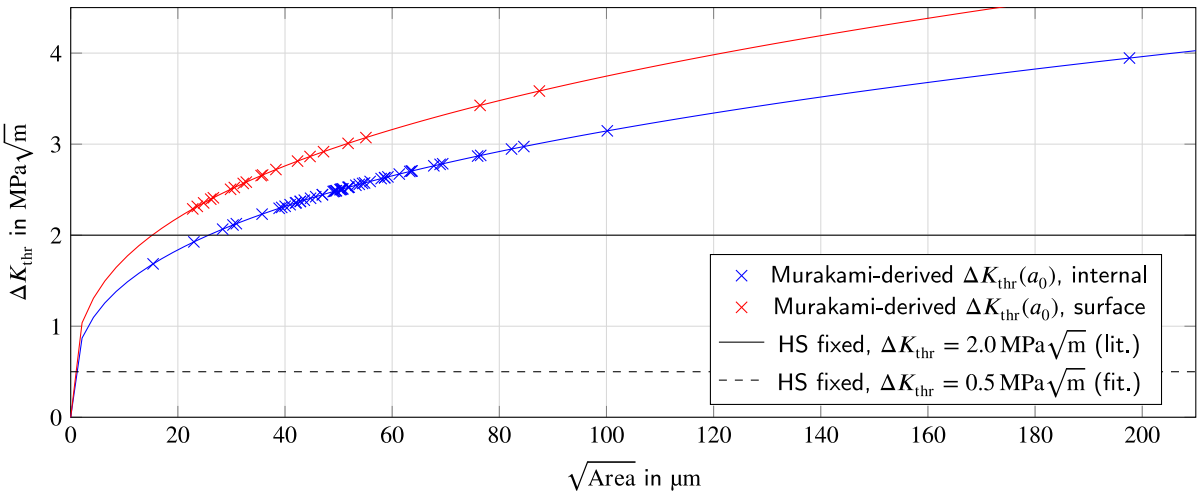


Fig. 11. Comparison of initial threshold stress intensity factors ΔK_{thr} as a function of initial defect size $\sqrt{Area}$. Markers: Murakami-derived evolving threshold (Eq. (30)) evaluated at the initial defect size for each specimen. Horizontal lines: fixed thresholds used in the HS model.

Table 14
Summary comparison of fatigue-life prediction models. "Fitted parameters" refers to parameters fitted to the present dataset; all other parameters are taken from literature or are universal constants.

Model	Fitted parameters	Total parameters	Requires HV	Environment distinction	$RMSE_{log}$
HS (dual env.)	3 (D, p, ΔK_{thr})	8[a]	No	Yes	0.425
ME (single env.)	0	3[b]	Yes	No	0.959
Hybrid	2 (D, p)	6[c]	Yes	Yes	0.386

[a] D, p, ΔK_{thr}, A_{HS} per environment × 2 environments = 8; surface parameters from Jones et al. [40] (not fitted), A_{HS} fixed.
[b] $C^* = 10^{-4}$, $m^* = 2$, $n^* = 1$; claimed material-independent [42].
[c] D, p per environment + A_{HS} (fixed) + $\sigma_w(a, HV)$ from Eq. (29); ΔK_{thr} replaced by Murakami formulation.

3.5.5. Summary of model comparison

Table 14 provides a consolidated comparison of all three crack growth models, including the number of parameters requiring fitting to the present dataset, the input requirements beyond the defect geometry, and the prediction performance after fitting of internal parameters.

In summary, the parameter fitted hybrid model achieves the lowest overall $RMSE_{log}$ (0.386) with only two free internal parameters, confirming that the Murakami-derived evolving threshold provides a physically motivated and effective replacement for the empirical ΔK_{thr}. The HS model offers a marginally conservative alternative ($RMSE_{log}$ = 0.425, $\bar{\eta} = -0.080$) at the cost of one additional fitting parameter. The ME model, while requiring no parameter fitting, cannot capture the reduced internal crack growth rate and is therefore limited to a conservative first-order estimate for surface-initiated failures (surface $RMSE_{log}$ = 0.428). These findings, together with the ranking results from Part A, are synthesised into practical recommendations in Section 4.

The choice of model depends on the available input data and the acceptable prediction uncertainty. If crack growth data (or fitting specimens) are available, the **hybrid model** is recommended: it requires only HV, the defect geometry ($\sqrt{Area}$, l_n), and two internal fitting parameters. If no fitting data are available but HV is known, the **ME model** provides a conservative estimate for surface-initiated failure without any curve fitting. If a conservative prediction is desired with parameter fitting, the **HS model** provides a slight systematic underprediction that constitutes an implicit safety margin.

The results of Part B complement the ranking benchmark of Part A: the geometry-based defect criticality models (Section 3.4) provide a computationally inexpensive identification of the most critical defect, and the hybrid crack growth model then translates that defect's characteristics into a fatigue-life estimate with an accuracy of ≈0.4 decades ($RMSE_{log}$) and 96 % of predictions within a factor of 6.

4. Conclusion

This work presents a systematic benchmark of nine defect criticality models and three crack growth models for defect-tolerant fatigue assessment of PBF-LB/M TiAl6V4, evaluated on 84 fatigue specimens (two series) inspected by X-ray CT prior to testing. Within the scope of the present dataset, the following main conclusions are drawn:

1. **Geometry-based defect criticality models perform comparably to simulation-based approaches in the present matched dataset.** For the 21 specimens successfully matched between CT and fractography, the geometry-based models, in particular the Tammas-Williams model, achieve ranking accuracy comparable to or better than the best simulation-based approaches. The geometry-based models require less than 1 min per specimen, compared to 10 min to 30 min for the simulation-based workflow (LE/EP), yet the additional computational effort does not translate into measurably higher ranking accuracy in the present dataset.
2. **Defect size is the dominant criticality descriptor in the present benchmark.** Partial correlation analysis demonstrates that all FE-based metrics retain negligible independent discriminative power ($|\rho_{partial}| \leq 0.039$) after controlling for $\sqrt{Area}$. This indicates that, for the investigated specimens at the given CT resolution, defect size is the primary driver of ranking performance, while more elaborate local-field metrics add only limited additional information. The ranking accuracy ceiling of approximately 60 % is governed by the dominance of defect size rather than by the fidelity of the stress analysis.
3. **EP material models reduce ranking performance.** Stress redistribution through local yielding homogenises the stress–strain response across defects (mean CoV reduction of 32 %), diminishing the discriminative power of simulation-based metrics. LE

models consistently outperform EP variants for defect ranking purposes.

4. **Dual-environment crack growth modelling is essential for life prediction accuracy.** The distinction between surface crack growth in air and internal crack growth in a self-contained environment improves the logarithmic prediction accuracy by more than a factor of two ($RMSE_{log} < 0.45$ for the dual-environment models vs. 0.959 for the single-environment ME model). Surface-in-air parameters from the literature [40] transfer directly to the present PBF-LB/M material without refitting, whereas internal parameters require material-specific fitting. The fitted internal parameters indicate faster crack growth and a lower threshold compared to wrought TiAl6V4, consistent with the irregular morphology of process-induced defects in the present build. These fitted values should be understood as effective parameters of the proposed framework (Section 2.8.5) rather than as independently measured intrinsic crack-growth constants, as they are inferred from total fatigue-life data and may absorb effects beyond the modelled crack-growth mechanism.
5. **The proposed hybrid model achieves the best prediction accuracy with the fewest fitted parameters.** By replacing the fixed threshold ΔK_{thr} with the Murakami-derived, crack-size-dependent formulation, one free parameter is eliminated without loss of accuracy ($RMSE_{log} = 0.386$, 96 % within factor 6). Where no fitting data are available, the parameter-free ME model provides a conservative baseline for surface-initiated failures (surface $RMSE_{log} = 0.428$).
6. **CT resolution governs the practical applicability of the ranking approach.** The matching rate between CT-detected and fractographically identified defects increases from 12 % (25.6 µm resolution) to 37 % (14 µm resolution), highlighting the need for high-resolution CT inspection in defect-tolerant assessment frameworks. This dependence of the matching rate on CT resolution is not only a practical limitation but also a substantive experimental finding: it demonstrates the practical difficulty of closing the complete chain from non-destructive inspection to fracture-origin validation, and quantifies the CT resolution required for reliable CT-to-fractography matching in defect-tolerant fatigue assessment.

Combining both parts, the results suggest the following workflow for defect-tolerant fatigue assessment of PBF-LB/M components: (i) inspect by X-ray CT at a resolution sufficient to resolve the expected critical defect size, (ii) rank defects using the Tammas-Williams model or, equivalently, the Murakami $\sqrt{\text{Area}}$ model with interaction corrections, and (iii) predict the fatigue-life of the critical defect using the hybrid crack growth model with dual-environment parameters fitted to a small number of representative specimens.

Limitations of this study include the restriction to uniaxial loading ($R = -1$), a single material system (TiAl6V4), and a defect population dominated by near-equiaxed gas pores with moderate lack-of-fusion contributions. For components with highly irregular LOF-dominated defect populations, multiaxial stress states, or significantly higher defect densities, the relative advantage of simulation-based approaches may increase. Future work should extend the benchmark to multiaxial loading conditions, alternative AM materials, and larger matched datasets to further validate the observed ranking equivalence and the hybrid crack growth formulation.

CRediT authorship contribution statement

Sebastian Mansky: Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Dirk Herzog:** Writing – review & editing, Supervision. **Ingomar Kelbassa:** Writing – review & editing, Supervision.

Funding

The work described is financed with funds from the German Federal Ministry of Education and Research (BMBF) under Reference No: 01LY2113C

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.

Acknowledgements

The authors would like to thank Mr. Adithya Kalliath from Imprintec GmbH and Dr. Frank Herold from VisiConsult X-ray Systems & Solutions GmbH for the excellent scientific collaboration in the research project Enabl4D, as well as Dr. Geert Schellenberg from MPA Stuttgart for the fatigue testing. The authors would also like to thank Mr. Sari Elhalees and Mr. Milan Dreyer for the microstructural preparation.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.tafmec.2026.105875.

Data availability

Data will be made available on request.

References

[1] D. Herzog, V. Seyda, E. Wycisk, C. Emmelmann, Additive manufacturing of metals, Acta Mater. 117 (2016) 371–392, http://dx.doi.org/10.1016/j.actamat.2016.07.019.
[2] T. DebRoy, H.L. Wei, J.S. Zuback, T. Mukherjee, J.W. Elmer, J.O. Milewski, A.M. Beese, A. Wilson-Heid, A. De, W. Zhang, Additive manufacturing of metallic components – Process, structure and properties, Prog. Mater. Sci. 92 (2018) 112–224, http://dx.doi.org/10.1016/j.pmatsci.2017.10.001.
[3] N. Sanaei, A. Fatemi, Defects in additive manufactured metals and their effect on fatigue performance: A state-of-the-art review, Prog. Mater. Sci. 117 (2021) 100724, http://dx.doi.org/10.1016/j.pmatsci.2020.100724.
[4] A. Yadollahi, N. Shamsaei, Additive manufacturing of fatigue resistant materials: Challenges and opportunities, Int. J. Fatigue 98 (2017) 14–31, http://dx.doi.org/10.1016/j.ijfatigue.2017.01.001.
[5] U. Zerbst, G. Bruno, J.-Y. Buffiere, T. Wegener, T. Niendorf, T. Wu, X. Zhang, N. Kashaev, G. Meneghetti, N. Hrabe, M. Madia, T. Werner, K. Hilgenberg, M. Koukolíková, R. Procházka, J. Džugan, B. Möller, S. Beretta, A. Evans, R. Wagener, K. Schnabel, Damage tolerant design of additively manufactured metallic components subjected to cyclic loading: State of the art and challenges, Prog. Mater. Sci. 121 (2021) http://dx.doi.org/10.1016/j.pmatsci.2021.100786.
[6] M. Gorelik, Additive manufacturing in the context of structural integrity, Int. J. Fatigue 94 (2017) 168–177, http://dx.doi.org/10.1016/j.ijfatigue.2016.07.005.
[7] H. Masuo, Y. Tanaka, S. Morokoshi, H. Yagura, T. Uchida, Y. Yamamoto, Y. Murakami, Effects of defects, surface roughness and HIP on fatigue strength of Ti-6Al-4V manufactured by additive manufacturing, Procedia Struct. Integr. 7 (2017) 19–26, http://dx.doi.org/10.1016/j.prostr.2017.11.055.
[8] N. Sanaei, A. Fatemi, Analysis of the effect of surface roughness on fatigue performance of powder bed fusion additive manufactured metals, Theor. Appl. Fract. Mech. 108 (2020) 102638, http://dx.doi.org/10.1016/j.tafmec.2020.102638.
[9] J.L. Bartlett, X. Li, An overview of residual stresses in metal powder bed fusion, Addit. Manuf. 27 (2019) 131–149, http://dx.doi.org/10.1016/j.addma.2019.02.020.
[10] S. Leuders, M. Thöne, A. Riemer, T. Niendorf, T. Tröster, H.A. Richard, H.J. Maier, On the mechanical behaviour of titanium alloy TiAl6V4 manufactured by selective laser melting: Fatigue resistance and crack growth performance, Int. J. Fatigue 48 (2013) 300–307, http://dx.doi.org/10.1016/j.ijfatigue.2012.11.011.
[11] A. Du Plessis, I. Yadroitsava, I. Yadroitsev, Effects of defects on mechanical properties in metal additive manufacturing: A review focusing on X-ray tomography insights, Mater. Des. 187 (2020) 108385, http://dx.doi.org/10.1016/j.matdes.2019.108385.

[12] L. Schild, L. Weiser, K. Höger, G. Lanza, Analyzing the error of computed tomography-based pore detection by using microscope images of matched cross-sections, Precis. Eng. 81 (2023) 192–206, http://dx.doi.org/10.1016/j.precisioneng.2023.01.013.
[13] Y. Murakami, Metal Fatigue: Effects of Small Defects and Nonmetallic Inclusions, second ed., Academic Press, Oxford, 2019.
[14] S. Romano, A. Brückner-Foit, A. Brandão, J. Gumpinger, T. Ghidini, S. Beretta, Fatigue properties of AlSi10Mg obtained by additive manufacturing: Defect-based modelling and prediction of fatigue strength, Eng. Fract. Mech. 187 (2018) 165–189, http://dx.doi.org/10.1016/j.engfracmech.2017.11.002.
[15] S. Romano, A. Brandão, J. Gumpinger, M. Gschweitl, S. Beretta, Qualification of AM parts: Extreme value statistics applied to tomographic measurements, Mater. Des. 131 (2017) 32–48, http://dx.doi.org/10.1016/j.matdes.2017.05.091.
[16] S. Beretta, S. Romano, A comparison of fatigue strength sensitivity to defects for materials manufactured by AM or traditional processes, Int. J. Fatigue 94 (2017) 178–191, http://dx.doi.org/10.1016/j.ijfatigue.2016.06.020.
[17] S. Tammas-Williams, P.J. Withers, I. Todd, P.B. Prangnell, The influence of porosity on fatigue crack initiation in additively manufactured titanium components, Sci. Rep. 7 (1) (2017) 7308, http://dx.doi.org/10.1038/s41598-017-06504-5.
[18] N. Sanaei, A. Fatemi, Defect-based fatigue life prediction of L-PBF additive manufactured metals, Eng. Fract. Mech. 244 (2021) 107541, http://dx.doi.org/10.1016/j.engfracmech.2021.107541.
[19] B. Torries, A. Imandoust, S. Beretta, S. Shao, N. Shamsaei, Overview on microstructure- and defect-sensitive fatigue modeling of additively manufactured materials, JOM 70 (9) (2018) 1853–1862, http://dx.doi.org/10.1007/s11837-018-2987-9.
[20] M. Awd, L. Saeed, F. Walther, A review on the enhancement of failure mechanisms modeling in additively manufactured structures by machine learning, Eng. Fail. Anal. 151 (2023) 107403, http://dx.doi.org/10.1016/j.engfailanal.2023.107403.
[21] P. Zhang, D. Z.Zhang, B. Zhong, Constitutive and damage modelling of selective laser melted Ti-6Al-4V lattice structure subjected to low cycle fatigue, Int. J. Fatigue 159 (2022) 106800, http://dx.doi.org/10.1016/j.ijfatigue.2022.106800.
[22] M. Cao, Y. Liu, F.P. Dunne, A crystal plasticity approach to understand fatigue response with respect to pores in additive manufactured aluminium alloys, Int. J. Fatigue 161 (2022) 106917, http://dx.doi.org/10.1016/j.ijfatigue.2022.106917.
[23] X. Cai, K. Tang, P. Ferro, F. Berto, Coordinated effect of microstructure and defect on fatigue accumulation in dual-phase Ti-6Al-4V: Quantitative characterization, Int. J. Fatigue 167 (2023) 107305, http://dx.doi.org/10.1016/j.ijfatigue.2022.107305.
[24] E. Mikkola, G. Marquis, J. Solin, Mesoscale modelling of crack nucleation from defects in steel, Int. J. Fatigue 41 (2012) 64–71, http://dx.doi.org/10.1016/j.ijfatigue.2011.12.022.
[25] P. Li, P.D. Lee, D.M. Maijer, T.C. Lindley, Quantification of the interaction within defect populations on fatigue behavior in an aluminum alloy, Acta Mater. 57 (12) (2009) 3539–3548, http://dx.doi.org/10.1016/j.actamat.2009.04.008.
[26] Y.X. Gao, J.Z. Yi, P.D. Lee, T.C. Lindley, The effect of porosity on the fatigue life of cast aluminium-silicon alloys, Fatigue Fract. Eng. Mater. Struct. 27 (7) (2004) 559–570, http://dx.doi.org/10.1111/j.1460-2695.2004.00780.x.
[27] L. Afroz, S.B. Inverarity, M. Qian, M. Easton, R. Das, Analysing the effect of defects on stress concentration and fatigue life of L-PBF AlSi10Mg alloy using finite element modelling, Prog. Addit. Manuf. (2023) http://dx.doi.org/10.1007/s40964-023-00457-0.
[28] Z. Li, A. Gut, I. Burda, S. Michel, D. Gwerder, P. Schütz, D. Romancuk, C. Affolter, Investigation of individual lack-of-fusion defects in the fatigue performance of laser-powder bed fusion Ti6Al4V alloys: A finite element analysis, Theor. Appl. Fract. Mech. 134 (2024) 104735, http://dx.doi.org/10.1016/j.tafmec.2024.104735.
[29] P. Lazzarin, F. Berto, Some expressions for the strain energy in a finite volume surrounding the root of blunt V-notches, Int. J. Fract. 135 (1–4) (2005) 161–185, http://dx.doi.org/10.1007/s10704-005-3943-6.
[30] F. Berto, P. Lazzarin, A review of the volume-based strain energy density approach applied to V-notches and welded structures, Theor. Appl. Fract. Mech. 52 (3) (2009) 183–194, http://dx.doi.org/10.1016/j.tafmec.2009.10.001.
[31] E. Siebel, M. Stieler, Ungleichförmige spannungsverteilung bei schwingender beanspruchung, VDI-Z 97 (5) (1955) 121–126.
[32] K.N. Smith, P. Watson, T.H. Topper, Stress-strain function for the fatigue of metals, J. Mater. 5 (1970) 767–778.
[33] D. Taylor, The theory of critical distances, Eng. Fract. Mech. 75 (7) (2008) 1696–1705, http://dx.doi.org/10.1016/j.engfracmech.2007.04.007.
[34] L. Susmel, The theory of critical distances: a review of its applications in fatigue, Eng. Fract. Mech. 75 (7) (2008) 1706–1724, http://dx.doi.org/10.1016/j.engfracmech.2006.12.004.
[35] S. Cicero, D. Taylor, L. Susmel, Theoretical and practical implications derived from the formulation of the theory of critical distances, Fatigue & Fract. Eng. Mater. Struct. (2026) http://dx.doi.org/10.1111/ffe.70282.
[36] L. Barricelli, L. Patriarca, A. Du Plessis, S. Beretta, A comparison of fatigue analysis methods for L-PBF net-shape surfaces in Ti6Al4V parts, Theor. Appl. Fract. Mech. 128 (2023) 104143, http://dx.doi.org/10.1016/j.tafmec.2023.104143.
[37] B. Gillham, A. Yankin, F. McNamara, C. Tomonto, D. Taylor, R. Lupoi, Application of the theory of critical distances to predict the effect of induced and process inherent defects for SLM Ti-6Al-4V in high cycle fatigue, CIRP Ann 70 (1) (2021) 171–174, http://dx.doi.org/10.1016/j.cirp.2021.03.004.
[38] E. Salvati, A. Tognan, L. Laurenti, M. Pelegatti, F. de Bona, A defect-based physics-informed machine learning framework for fatigue finite life prediction in additive manufacturing, Mater. Des. 222 (2022) 111089, http://dx.doi.org/10.1016/j.matdes.2022.111089.
[39] S. Mansky, M. Becker, D. Herzog, I. Kelbassa, Enhancing computed tomography-based pore mesh models through matching with microscope cross-section images, J. Nondestruct. Eval. 44 (3) (2025) 1–11, http://dx.doi.org/10.1007/s10921-025-01240-7, URL: https://link.springer.com/article/10.1007/s10921-025-01240-7.
[40] R. Jones, J.G. Michopoulos, A.P. Iliopoulos, R.K. Singh Raman, N. Phan, T. Nguyen, Representing crack growth in additively manufactured Ti-6Al-4V, Int. J. Fatigue 116 (2018) 610–622, http://dx.doi.org/10.1016/j.ijfatigue.2018.07.019.
[41] S. Han, T.D. Dinh, I. de Baere, M. Boone, I. Josipovic, W. van Paepegem, Study of the effect of defects on fatigue life prediction of additive manufactured Ti-6Al-4V by combined use of micro-computed tomography and fracture-mechanics-based simulation, Int. J. Fatigue 180 (2024) 108110, http://dx.doi.org/10.1016/j.ijfatigue.2023.108110.
[42] Y. Murakami, M. Endo, Prediction model of S-N curve without fatigue test or with a minimum number of fatigue tests, Eng. Fail. Anal. 154 (2023) 107647, http://dx.doi.org/10.1016/j.engfailanal.2023.107647.
[43] M. Endo, Y. Murakami, Prediction model of S-N curve under mean stress without fatigue test or with a minimum number of fatigue tests, Theor. Appl. Fract. Mech. 138 (2025) 104918, http://dx.doi.org/10.1016/j.tafmec.2025.104918.
[44] DIN Media GmbH, DIN EN 6072:2011-06, luft- und raumfahrt - metallische werkstoffe - prüfverfahren - ermüdungstest mit konstanter amplitude; deutsche und englische fassung EN 6072:2010, 2011, http://dx.doi.org/10.31030/1757524.
[45] S. Romano, P.D. Nezhadfar, N. Shamsaei, M. Seifi, S. Beretta, High cycle fatigue behavior and life prediction for additively manufactured 17-4 PH stainless steel: Effect of sub-surface porosity and surface roughness, Theor. Appl. Fract. Mech. 106 (2020) 102477, http://dx.doi.org/10.1016/j.tafmec.2020.102477.
[46] G. Kasperovich, J. Hausmann, Improvement of fatigue resistance and ductility of TiAl6V4 processed by selective laser melting, J. Mater. Process. Technol. 220 (2015) 202–214, http://dx.doi.org/10.1016/j.jmatprotec.2015.01.025.
[47] N. Derimow, J.T. Benzing, H. Joress, A. McDannald, P. Lu, F.W. DelRio, N. Moser, M.J. Connolly, A.I. Saville, O.L. Kafka, C. Beamer, R. Fishel, S. Sarker, C. Hadley, N. Hrabe, Microstructure and mechanical properties of laser powder bed fusion Ti-6Al-4V after HIP treatments with varied temperatures and cooling rates, Mater. Des. 247 (2024) 113388, http://dx.doi.org/10.1016/j.matdes.2024.113388.
[48] B. Vrancken, L. Thijs, J.-P. Kruth, J. van Humbeeck, Heat treatment of Ti6Al4V produced by selective laser melting: Microstructure and mechanical properties, J. Alloys Compd. 541 (2012) 177–185, http://dx.doi.org/10.1016/j.jallcom.2012.07.022.
[49] DIN Media GmbH, DIN EN ISO 17295, additive manufacturing - general principles - part positioning, coordinates and orientation (ISO 17295:2023), 2023, http://dx.doi.org/10.31030/3412793.
[50] DIN Media GmbH, DIN 50100:2022-12, schwingfestigkeitsversuch - durchführung und auswertung von zyklischen versuchen mit konstanter lastamplitude für metallische werkstoffproben und bauteile, 2022, http://dx.doi.org/10.31030/3337109.
[51] C.A. Schneider, W.S. Rasband, K.W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis, Nature Methods 9 (7) (2012) 671–675, http://dx.doi.org/10.1038/nmeth.2089.
[52] M. Deshpande, G. Karypis, Item-based top- N recommendation algorithms, ACM Trans. Inf. Syst. 22 (1) (2004) 143–177, http://dx.doi.org/10.1145/963770.963776.
[53] J.L. Herlocker, J.A. Konstan, L.G. Terveen, J.T. Riedl, Evaluating collaborative filtering recommender systems, ACM Trans. Inf. Syst. 22 (1) (2004) 5–53, http://dx.doi.org/10.1145/963770.963772.
[54] H. Moon, P.J. Phillips, Computational and performance aspects of PCA-based face-recognition algorithms, Perception 30 (3) (2001) 303–321, http://dx.doi.org/10.1068/p2896.
[55] E.M. Voorhees, et al., The TREC-8 question answering track report, in: Trec, vol. 99, 1999, pp. 77–82.
[56] H. Müller, W. Müller, D.M. Squire, S. Marchand-Maillet, T. Pun, Performance evaluation in content-based image retrieval: overview and proposals, Pattern Recognit. Lett. 22 (5) (2001) 593–601, http://dx.doi.org/10.1016/S0167-8655(00)00118-5.
[57] D. Rigon, G. Meneghetti, An engineering estimation of fatigue thresholds from a microstructural size and Vickers hardness: application to wrought and additively manufactured metals, Int. J. Fatigue 139 (2020) 105796, http://dx.doi.org/10.1016/j.ijfatigue.2020.105796.
[58] D. Greitemeier, F. Palm, F. Syassen, T. Melz, Fatigue performance of additive manufactured TiAl6V4 using electron and laser beam melting, Int. J. Fatigue 94 (2017) 211–217, http://dx.doi.org/10.1016/j.ijfatigue.2016.05.001.
[59] D. Rigon, G. Meneghetti, Engineering estimation of the fatigue limit of wrought and defective additively manufactured metals for different load ratios, Int. J. Fatigue 154 (2022) 106530, http://dx.doi.org/10.1016/j.ijfatigue.2021.106530.

[60] V. Cain, L. Thijs, J. van Humbeeck, B. van Hooreweder, R. Knutsen, Crack propagation and fracture toughness of Ti6Al4V alloy produced by selective laser melting, Addit. Manuf. 5 (2015) 68–76, http://dx.doi.org/10.1016/j.addma.2014.12.006.

[61] M.H. El Haddad, K.N. Smith, T.H. Topper, Fatigue crack propagation of short cracks, J. Eng. Mater. Technol. 101 (1) (1979) 42–46, http://dx.doi.org/10.1115/1.3443647.

[62] S.-P. Zhu, W.-L. Ye, J.A. Correia, A.M. Jesus, Q. Wang, Stress gradient effect in metal fatigue: Review and solutions, Theor. Appl. Fract. Mech. 121 (2022) 103513, http://dx.doi.org/10.1016/j.tafmec.2022.103513.

[63] J. Mei, S. Xing, A. Vasu, J. Chung, R. Desai, P. Dong, The fatigue limit prediction of notched components – A critical review and modified stress gradient based approach, Int. J. Fatigue 135 (2020) 105531, http://dx.doi.org/10.1016/j.ijfatigue.2020.105531.

[64] W.-L. Ye, S.-P. Zhu, X.-P. Niu, J.-C. He, Q. Wang, Fatigue life prediction of notched components under size effect using stress gradient-based approach, Int. J. Fract. 234 (1–2) (2022) 249–261, http://dx.doi.org/10.1007/s10704-021-00580-5.

[65] H. Oguma, T. Nakamura, Fatigue crack propagation properties of Ti–6Al–4V in vacuum environments, Int. J. Fatigue 50 (2013) 89–93, http://dx.doi.org/10.1016/j.ijfatigue.2012.02.012.

[66] F. Yoshinaka, T. Nakamura, A. Takeuchi, M. Uesugi, K. Uesugi, Initiation and growth behaviour of small internal fatigue cracks in Ti–6Al–4V via synchrotron radiation microcomputed tomography, Fatigue Fract. Eng. Mater. Struct. 42 (9) (2019) 2093–2105, http://dx.doi.org/10.1111/ffe.13085.

[67] A. Junet, A. Messager, A. Weck, Y. Nadot, X. Boulnat, J.-Y. Buffiere, Internal fatigue crack propagation in a Ti-6Al-4V alloy: An in situ study, Int. J. Fatigue 168 (2023) 107450, http://dx.doi.org/10.1016/j.ijfatigue.2022.107450.

[68] P. Paris, F. Erdogan, A critical analysis of crack propagation laws, J. Basic Eng. 85 (4) (1963) 528–533, http://dx.doi.org/10.1115/1.3656900.

[69] E08 Committee, Test method for measurement of fatigue crack growth rates, ASTM E647-24, 2024, http://dx.doi.org/10.1520/E0647-24.

[70] J.A. Nelder, R. Mead, A simplex method for function minimization, Comput. J. 7 (4) (1965) 308–313, http://dx.doi.org/10.1093/comjnl/7.4.308.

[71] M.H. Wright, Direct search methods: Once scorned, now respectable, Pitman Res. Not. Math. Ser. (1996) 191–208.