Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

X-ray tomography and coordinate metrology benchmark for internal cavities in additively manufactured alloys

Josselin Paccoud
École de Technologie Supérieure
Hristo Valtchanov
hristo.valtchanov@polymtl.ca
Polytechnique Montréal
Eric Fournier
Comet Technologies Canada Inc.
Vladimir Brailovski
École de Technologie Supérieure
François Guibault
Polytechnique Montréal

Article

Keywords:

Posted Date: November 4th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11393438/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

X-ray tomography and coordinate metrology benchmark for internal cavities in additively manufactured alloys

Josselin Paccoud[1], Hristo Valtchanov[2,*], Eric Fournier[3], Vladimir Brailovski[1], François Guibault[2]

[1]École de technologie supérieure (ETS), Montréal, QC, Canada

[2]Department of Computer Engineering and Software Engineering, Polytechnique Montréal, Montréal, QC, Canada

[3]Comet Technologies Canada Inc., Montréal, QC, Canada

*Correspondence: Hristo.Valtchanov@polymtl.ca

Abstract

X-ray computed tomography (CT) is widely used for dimensional inspection of inaccessible internal geometry, but multi-modal public benchmark datasets with reference metrology remain scarce, particularly for internal cavities. This paper releases a CT–coordinate measuring machine (CMM) benchmark for wall-thickness measurements using CT surface extraction of an artefact with a complex internal cavity that is disassembled for tactile probing. Four cases span Ti-6Al-4V and Inconel 625 in rough and chemically milled states. Each case includes nominal computer-aided design (CAD) geometry aligned to the CT reconstruction, CT projections, reconstructed volumes, CT surface segmentations, CMM point clouds with predefined thickness-pair vectors, and derived surface- and thickness-deviation labels. The paper defines the benchmark tasks and reporting convention and reports baseline results for reconstruction-based and CAD-informed routes. The benchmark permits validation of CT-based dimensional metrology in noisy and clean environments, including CAD-constrained surface reconstruction approaches that work entirely in projection space, such as elastic registration.

1 Introduction

Dimensional validation of internal cavities remains difficult because the features of interest are often inaccessible to tactile and optical metrology[1,2]. CT (also denoted XCT) can inspect internal and external geometry in one acquisition and derive myriad quality assurance metrics, but the dimensional result depends on acquisition, reconstruction, surface determination, alignment, and the reference convention[1,3,4]. The benchmark presented here links

these steps to an external reference on a curved internal cavity across two materials and two surface states.

1.1 Related work

For CT metrology, dimensional error is shaped not only by voxel size but also by beam hardening, cone-beam artefacts, extinction, radiographic image quality, and the surface-determination route used to convert the reconstruction into a measurement-ready surface[1,4–6]. These effects become more severe in dense alloys, narrow internal cavities, and regions with poor local contrast. Recent studies make this dependence more explicit. Lifton et al. showed that XCT-based internal roughness measurements vary with the thickness of surrounding material, even when the nominal surface is unchanged[7]. Chen et al. related measurable surface scale to radiographic image quality, while also showing that the final surface error still depends on surface determination[8]. Klos et al. argued that CT enhancement workflows should be validated against task-specific metrics such as porosity or roughness preservation rather than visual quality alone[9]. CT dimensional validation is also a surface-determination problem, and while global or manually tuned thresholding can be fast, it is sensitive to artefacts and density variation. Local-gradient methods can better follow boundaries in heterogeneous data, and learning-based segmentation can improve robustness in difficult regions when training data and parameterization are appropriate[1,5,10].

For high-attenuation parts, or for regions where local image quality is substantially worse than elsewhere in the same scan, the surface-determination route can be as important as the acquisition settings themselves. The reference side of the comparison is also condition-dependent. On rough additively manufactured surfaces, CMM probing mechanically filters the surface by contacting peaks and bypassing valleys, so CT–CMM differences depend on stylus geometry, surface morphology, and feature accessibility[3,11–16]. A useful benchmark must therefore document the CT protocol, probing strategy, and alignment convention, and it must include cases in which surface state or attenuation materially change the comparison. Existing artefacts and datasets address parts of this problem[3,6,17–21], but no public aligned release exists for the same curved internal cavity across multiple materials and surface states that aligns nominal CAD, raw XCT projections, CMM reference data and comparison labels. Such a dataset would for example permit the benchmarking of a CAD-constrained surface reconstruction purely from XCT projections to standard CT metrology pipelines.

1.2 Dataset challenges and study rationale

For internal-cavity validation, the benchmark must preserve the link between nominal CAD, XCT data, CT-derived surfaces or volumes, and the CMM reference so that reconstruction, surface determination, and registration routes can be tested under the same conditions. Related aligned additive-manufacturing datasets already show the value of registration-ready multi-source releases[22–25] covering process monitoring and post-build inspection and alignment, rather than the curved-cavity CT–CMM comparison across materials and surface states considered here. Surface state must also be part of the benchmark because roughness affects both CT boundary localization, projection and reconstruction quality and the tactile reference measurements. To our knowledge, no public dataset combines CT and CMM

measurements before and after surface modification on the same part family.

1.3 Contribution and scope

This paper releases a CT–CMM benchmark built around a two-part cavity artefact manufactured in Ti-6Al-4V and Inconel 625, each in rough and chemically milled states. Each primary case contains aligned nominal CAD, XCT projections, reconstructed volumes, CT surface segmentations, CMM point clouds, thickness-pair vectors, and derived deviation labels. The paper defines the benchmark tasks, reporting convention, and baseline outputs. Elastic registration (ER) is included only as a CAD-informed benchmark execution; its detailed analysis is left to the companion paper[26]. The experimental system was first documented in an internal project report by Josselin Paccoud; the present manuscript consolidates the benchmark release.

2 Methods

2.1 Artefact development and benchmark cases

The benchmark is centered on a two-part artefact composed of a thicker fixed half (Fix) on the left and a mobile thinner half (Mob) on the right. In the assembled state it contains an internal "hourglass" cavity region with varying wall thickness and surface roughness; after disassembly the same region becomes accessible to tactile probing. This permits CT measurement in the assembled state and CMM reference measurement after disassembly on the same nominal geometry and digital realignment. The nominal geometry and the measurement region are shown in Figure 1.

The geometry was designed to represent cavity-rich inspection problems while remaining manufacturable by laser powder bed fusion (LPBF). Overhangs were limited to reduce support-removal damage, and the reference interfaces were machined to improve assembly repeatability and datum quality. For Ti-6Al-4V, the released reconstruction-based surfaces use fall-off snap factor (FSF) settings of 0.2, 0.4, and 0.6. For Inconel 625, the release includes FSF 0.6 and a deep-learning (DL) route to counter extinction artefacts. Thickness-pair counts are identical across all four cases: 215 on Fix and 184 on Mob. An additional dataset, a Chinook rotor hub blade, is included here to provide another reference point. We do not perform CT/CMM comparison on this part because its surface roughness was qualitatively high and the part could not be cut to access its internal geometry. The main CMM/CT benchmark is therefore confined to the titanium and Inconel artefacts. The cone-beam CT (CBCT) data for the Chinook rotor hub, a titanium 3D-printed blade-like part referenced in[26], are included with independently acquired linear diode array (LDA) XCT reconstruction data used here for comparison. No CMM measurement, chemical milling, or surface roughness characterization was performed on the Chinook dataset. The cases shown in this study and their corresponding reference data are given in Table 1.

Table 2 summarizes the benchmark tasks and primary reporting quantities.

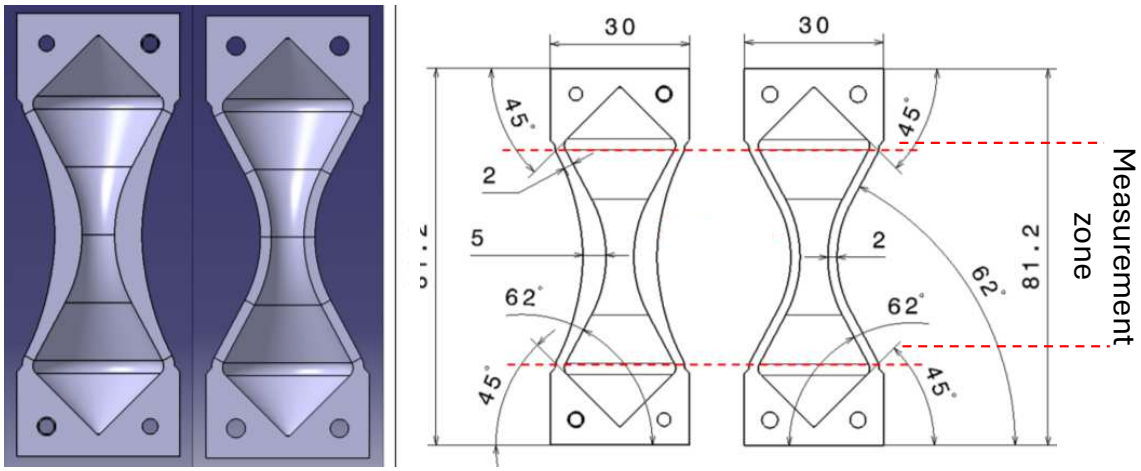


Figure 1: Fix (Left) and Mob (right) artefact halves, nominal dimensions, and the measurement region used for wall-thickness validation. Dimensions are given in mm.

Table 1: Benchmark cases and benchmark data objects.

Case	Material	State	Reconstruction-based CT surfaces	Reference data
Ti64-Rough	Ti-6Al-4V	Rough	FSF 0.2/0.4/0.6	CMM cloud, surface roughness, labels
Ti64-Polished	Ti-6Al-4V	Polished	FSF 0.2/0.4/0.6	CMM cloud, surface roughness, 3D scans, labels
IN625-Rough	Inconel 625	Rough	FSF 0.6, DL	CMM cloud, surface roughness, labels
IN625-Polished	Inconel 625	Polished	FSF 0.6, DL	CMM cloud, surface roughness, 3D scans, labels
Chinook	Ti-6Al-4V	Rough	FSF 0.6	LDA CT Data

Table 2: Benchmark tasks and primary reporting quantities.

Task	Inputs	Reference object	Primary metrics
Surface comparison	CT-derived mesh, documented alignment	CMM point cloud	signed surface deviations, deviation maps
Thickness comparison	CT-derived mesh, thickness-pair vectors	CMM thickness pairs	mean and std of Δt, P95 of $\lvert\Delta t\rvert$
Robustness study	same task repeated across 4 cases	all cases	change with material and surface state

2.2 Benchmark definition and reporting convention

In this work, *surface reconstruction* denotes the full CT-to-surface workflow, *surface determination* the local-gradient Steinbeiss-style step that places the boundary within the grey-level transition, and *CT surface segmentation* the process of extracting the surface mesh from the CT reconstruction volume. The benchmark is defined by fixed released objects and a fixed comparison convention. Each primary case contains nominal CAD, raw XCT projections, reconstructed volumes, CT surface segmentations, CMM point clouds, and thickness-pair vectors. CMM data are given as raw point clouds at the measured points; paired-point differences are used to compute the material thickness and measurement directions on the segmentation surface (see Fig. 5). Thickness comparison in the hourglass region is the primary endpoint; surface-deviation maps are supporting diagnostics, and robustness across material and surface state is secondary. Rough and polished cases are separate cases because chemical milling changes the geometry. Here, "polished" denotes the chemically milled state and is retained in the case identifiers and figure labels. Direct comparison with the released baseline requires preserving the released thickness-pair indexing and the documented alignment methodology. Recomputed alignment should be reported together with the case, reconstruction route, key parameters, and outputs.

2.3 Data generation

2.3.1 Artefact manufacturing, machining, and surface-state preparation

Two artefacts of the same nominal geometry were manufactured by laser powder bed fusion (LPBF), one in Ti-6Al-4V and one in Inconel 625. The Inconel artefact was produced on a Trumpf TruPrint 1000. The Ti-6Al-4V artefact was produced on an EOSINT M 280, stress relieved at 800 °C for 2 h, and both artefacts were removed from the build plates by electrical discharge machining.

Post-build machining improved contact-interface flatness. For Ti-6Al-4V, Mob surface flatness changed from 0.078 mm to 0.014 mm and Fix from 0.063 mm to 0.036 mm. For Inconel 625, Mob surface flatness changed from 0.047 mm to 0.021 mm and Fix from 0.065 mm to 0.023 mm. Shoulder screws were then used for repeatable assembly.

The parts underwent chemical milling at Tech Met Inc. in Montréal, Quebec, to reduce surface roughness and the effect of mechanical filtering during CMM measurement, with reported material removal of 0.23–0.25 mm. The mid-plane interfaces and holes were masked

to preserve assembly quality. The initial and post-processed artefact halves are shown in Fig.2.

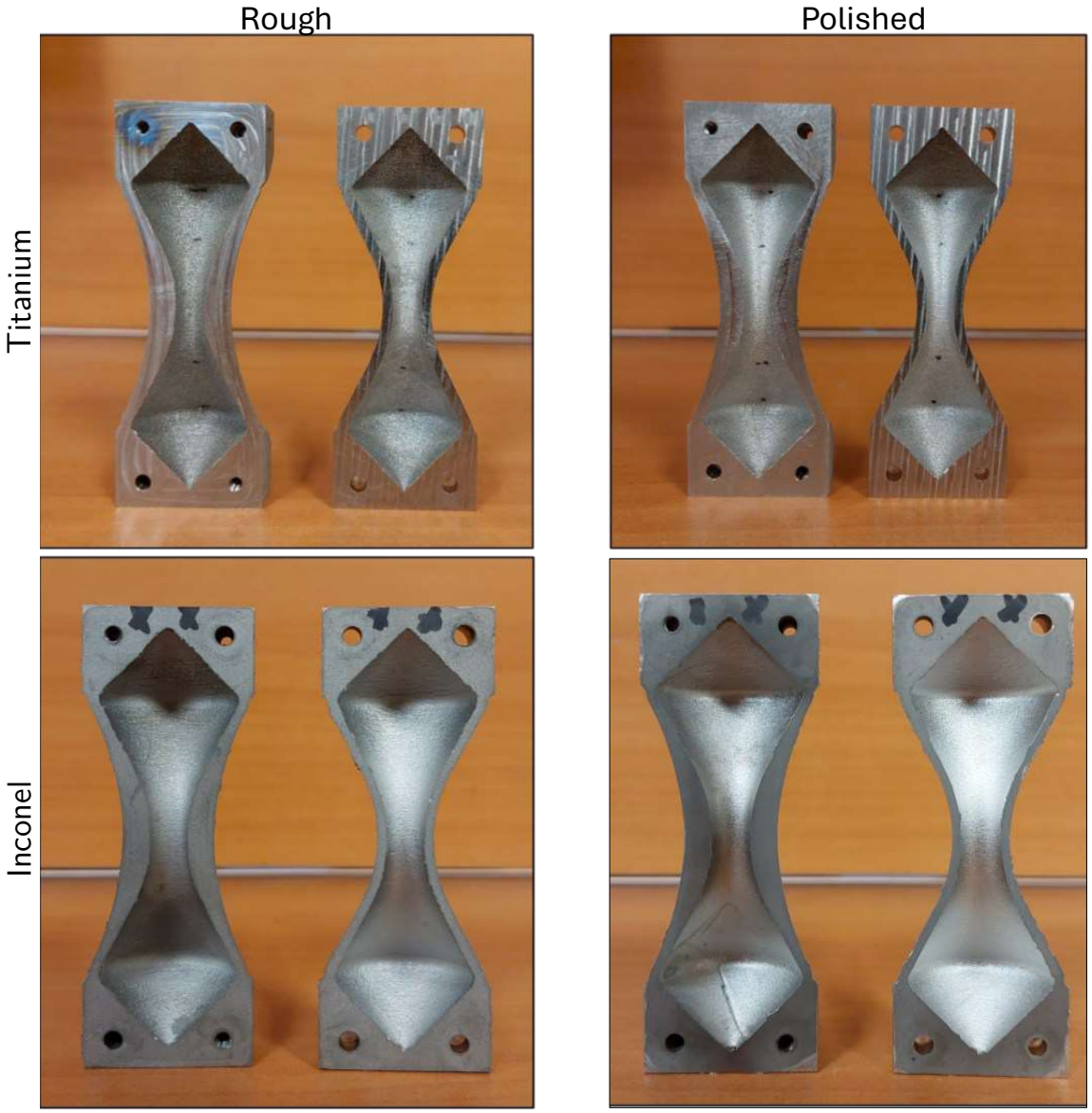


Figure 2: Ti-6Al-4V and Inconel 625 artefacts in their rough and polished states.

2.3.2 Roughness characterization

Surface state affects both the CMM reference, through mechanical filtering, and the CT result, at the surface-extraction stage. To quantify this effect on CT–CMM deviations, the roughness parameters R_a, R_z, and R_p were measured on the Mob half of both prototypes before and after chemical milling. R_a, R_z, and R_p correspond to the arithmetic mean roughness, maximum profile height, and maximum profile peak height, respectively[27]. Measurements were acquired with a Mitutoyo SJ-410 roughness tester using a $5\,\mu$m radius stylus. Four

zones were measured (Int High, Int Low, Ext High, Ext Low). These zones were selected so that the angle with the horizontal plane remained constant within each zone, giving a more homogeneous surface condition over each trace. For each zone, five separate traces were acquired, each with a sampling length of 8 mm and a total evaluation length of 40 mm (five sampling lengths per trace). Table 4 reports the arithmetic mean of the five trace measurements for each zone. This procedure corresponds to ISO 4288:1996 for $R_a > 10\,\mu$m[28]. The same settings were retained for zones with $R_a < 10\,\mu$m so that all measurements remained directly comparable. The measurement arrangement and sampled zones are shown in Figure 3.

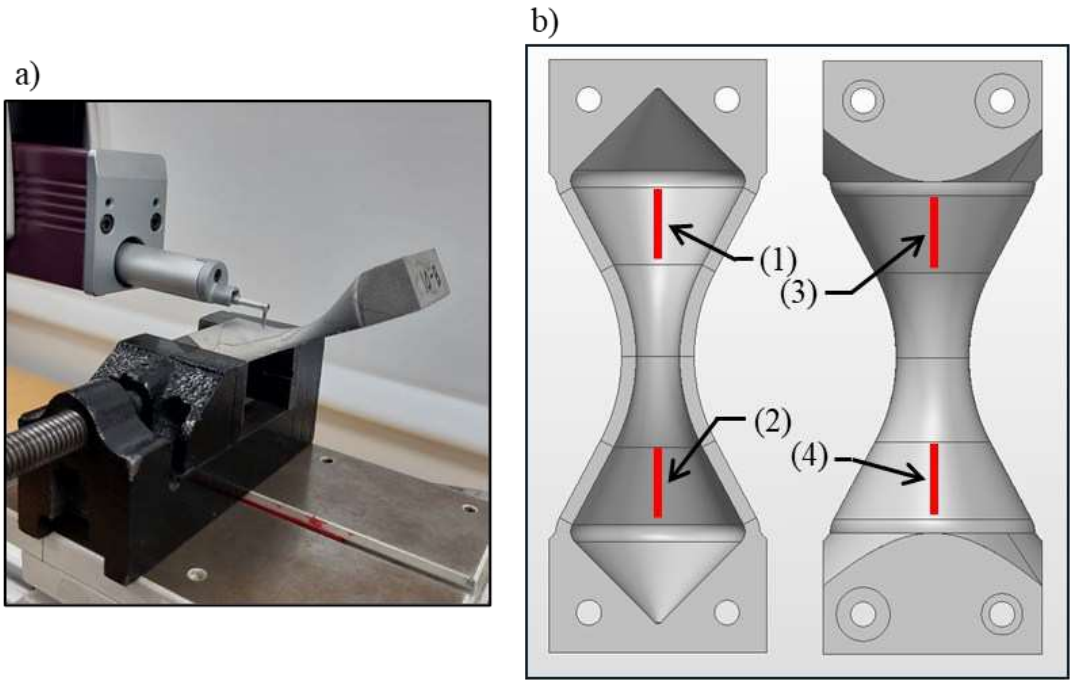


Figure 3: a) Attachment of parts for surface roughness measurement, b) Roughness measurement areas, and: (1) Int High, (2) Int Low, (3) Ext High, (4) Ext Low.

2.3.3 CMM reference measurement and thickness-pair definition

Reference measurements were acquired on a Hexagon GLOBAL S Green 7.10.7 CMM equipped with an HH-A-T7.5 articulated head and an HP-S-X1H probe. The probing configuration used a 95 mm stylus with a 2 mm spherical tip diameter. PolyWorks Metrology Suite 2024 IR9 was used for execution and post-processing of the CMM data as well as alignment of the CMM point clouds with the CT-derived surfaces.

The benchmark is defined around wall-thickness evaluation in the internal hourglass region. Thickness pairs were constructed by sampling the outer surface on a grid and projecting inward along the local surface normal. This produced 215 pairs on Fix and 184 on Mob, or 399 pairs per case and approximately 1,600 pairs across the four primary cases. Parts were

secured in a vice for access, while the machined mid-plane and shoulder-screw bores were used as datum features for the initial CAD-to-CMM setup.

2.3.4 CT acquisition

All CBCT projection data were acquired on a Comet Yxlon FF35 system. Each of the four primary artefact cases used the final tilted-support configuration: the assembled artefact was scanned with 4,500 projections over a full 360° rotation, with three integrated images per projection and a copper filter to mitigate beam hardening. A 0.22 mm plastic foil was inserted between Fix and Mob to improve visibility of the interface during surface determination. The acquisition settings for the primary artefacts and the auxiliary Chinook CBCT scan are summarized in Table 3.

Table 3: CT acquisition settings used for the released benchmark cases.

Parameter	Art. Titanium	Art. Inconel	Chinook CBCT
Voxel size (μm)	33.3	33.3	37.5
Inclination angle	20°	20°	20°
Number of projections	4,500	4,500	2,500
Magnification	4.5	4.5	4
Filter	Cu (1 mm)	Cu (0.4 mm)	Cu (1 mm)
Detector-source distance (mm)	900	900	900
Object-source distance (mm)	200	200	225
Voltage (kV)	180	190	195
Current (μA)	135	120	120
Power (W)	24.3	22.8	23.4
Frame rate (Hz)	3.5	3.0	2
Exposure time (ms)	285.714	333.333	433.4
Integrated images / projection	3	3	3
Detector sensitivity	50	25	50
Nominal scan duration (min)	64.2	75.0	62.5

The Chinook linear diode array (LDA) tomographic data were acquired independently at Pratt and Whitney (USA) on a proprietary custom dual-energy LDA rig and are not described by these CBCT settings. Only the supplied reconstruction metadata are reported here: 383 images of 480×630 pixels, with a recorded in-plane pixel spacing of 0.10 mm and an axial increment of 0.25 mm.

The tilted support was adopted after an upright configuration produced strong cone-beam artefacts on the mounting blocks and poor visibility of the Fix/Mob interface. The final support configuration and the probing setup are shown together in Figure 4.

2.3.5 Reconstruction and benchmark CT surface segmentations

The released primary-case CT volumes were reconstructed from 4,500 projections using CERA with a simple single-material beam-hardening correction, assuming Ti for the titanium cases and Ni for the Inconel cases, and without scatter correction. The auxiliary

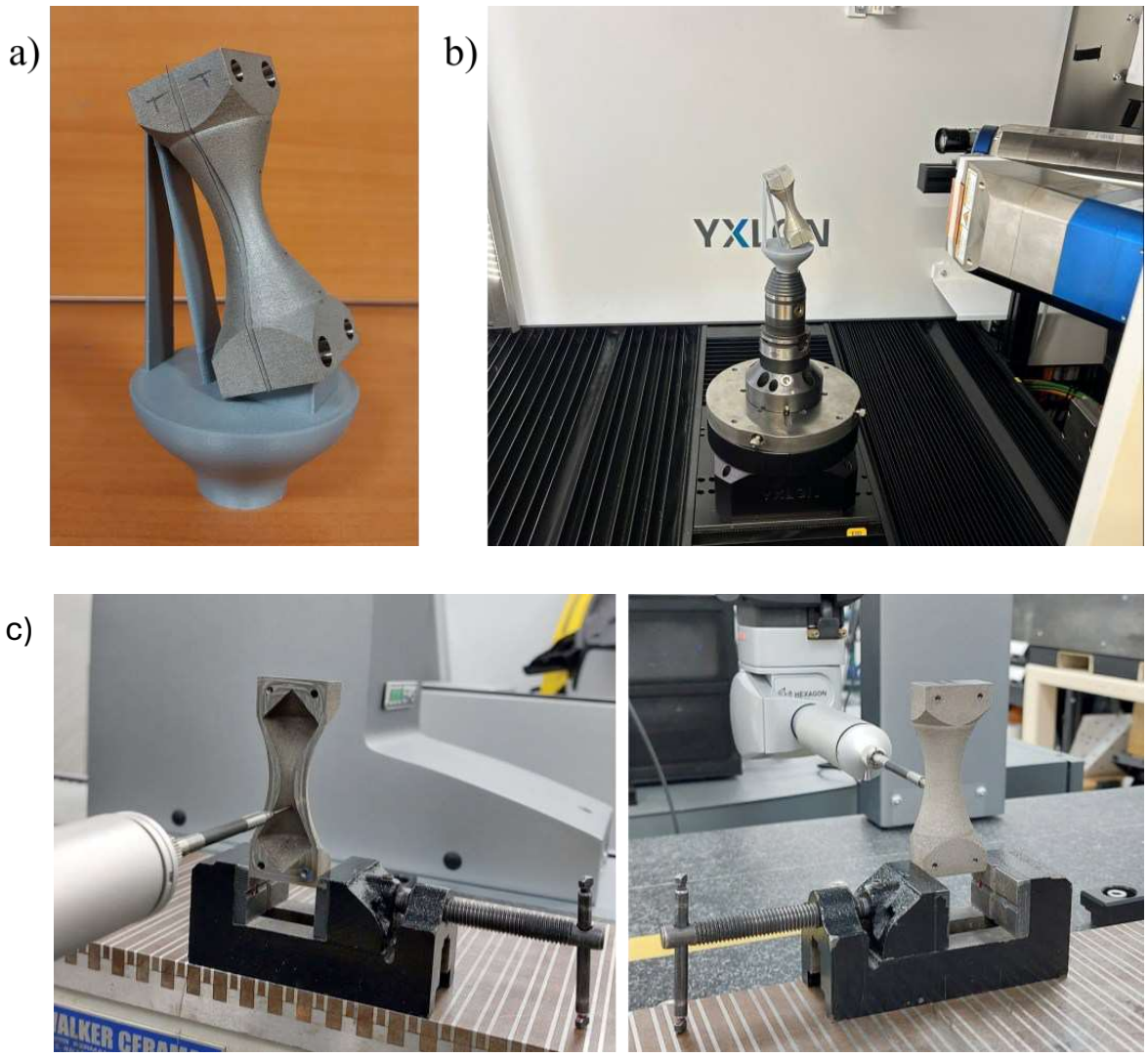


Figure 4: Fixture-related setup images: (a) tilted poly(lactic acid) support, (b) CT positioning with the 20° support and plastic foil, and (c) probing setup on the CMM.

Chinook CBCT volume was reconstructed from 2,500 projections, and its released surface used FSF 0.6. Benchmark CT surface segmentations were generated in Dragonfly 3D World 2025.1. The reconstruction-based CT-segmentation routes reported here are a local-gradient surface-determination workflow derived from the Steinbeiss approach[29] and, for Inconel, a deep-learning (DL) route operating in the probability space of a U-Net-like model[10]. The DL model with proprietary weights is released as part of the Dragonfly World 2025.1 software package. In the local-gradient route, the boundary is placed within the reconstructed grey-level transition; the fall-off snap factor (FSF) shifts that position and therefore the effective surface. The DL route was used in the Inconel cases where dense-material image degradation due to the extinction artefact made the local-gradient extraction less robust for extracting the real geometry.

The data-generation steps above define the released physical measurements and CT-derived objects. The benchmark itself is operationalized through a fixed alignment and comparison convention.

2.3.6 3D scans of the polished Inconel and titanium artefacts

To provide an additional validation point, and because the CMM studies were metrologically focused such that the CMM points were specifically aligned to capture thickness, complete surface point clouds of the artefacts were not acquired with the CMM; instead, probing focused on a zone of interest indicated in Fig. 5. More generally, the most problematic CT artefacts were found to occur outside the CMM-validated measurement zone, missing an opportunity to validate against artefacts that are difficult to reconstruct. Consequently, we also provide 3D scans of the polished and disassembled titanium and Inconel artefacts. These scans were acquired using a Revopoint Mini 3D scanner, for which the manufacturer reports single-frame precision up to 0.02 mm[30], and the resulting point clouds were aligned to the CMM dataset by via (point-plane) iterative closest point search aligning the point cloud to the segmentation meshes.

2.4 Benchmark protocol

2.4.1 Alignment and comparison convention

The geometric comparison convention used to generate the benchmark outputs is illustrated in Figure 5. For CT-to-CMM comparison, CMM point clouds were aligned to each CT-derived mesh in PolyWorks using a best-fit iterative closest point (ICP) registration in the measurement region, minimizing the distance between the measured CMM points and the corresponding CT-derived surface. This alignment was chosen to reduce the localized bias observed when using standard CAD-referenced alignment due to extensive deformation during the LPBF process, which exhibited large localized distortions relative to the scale of the measurement deviations.

After alignment, surface deviations were evaluated along the released vectors and exported from PolyWorks. MATLAB R2024b was used to calculate the reported descriptive summary statistics. Thickness deviation for pair i was computed as

$$\Delta t_i = t_i^{CT} - t_i^{CMM}. \tag{1}$$

Positive Δt_i indicates that the CT-derived thickness exceeds the CMM thickness.

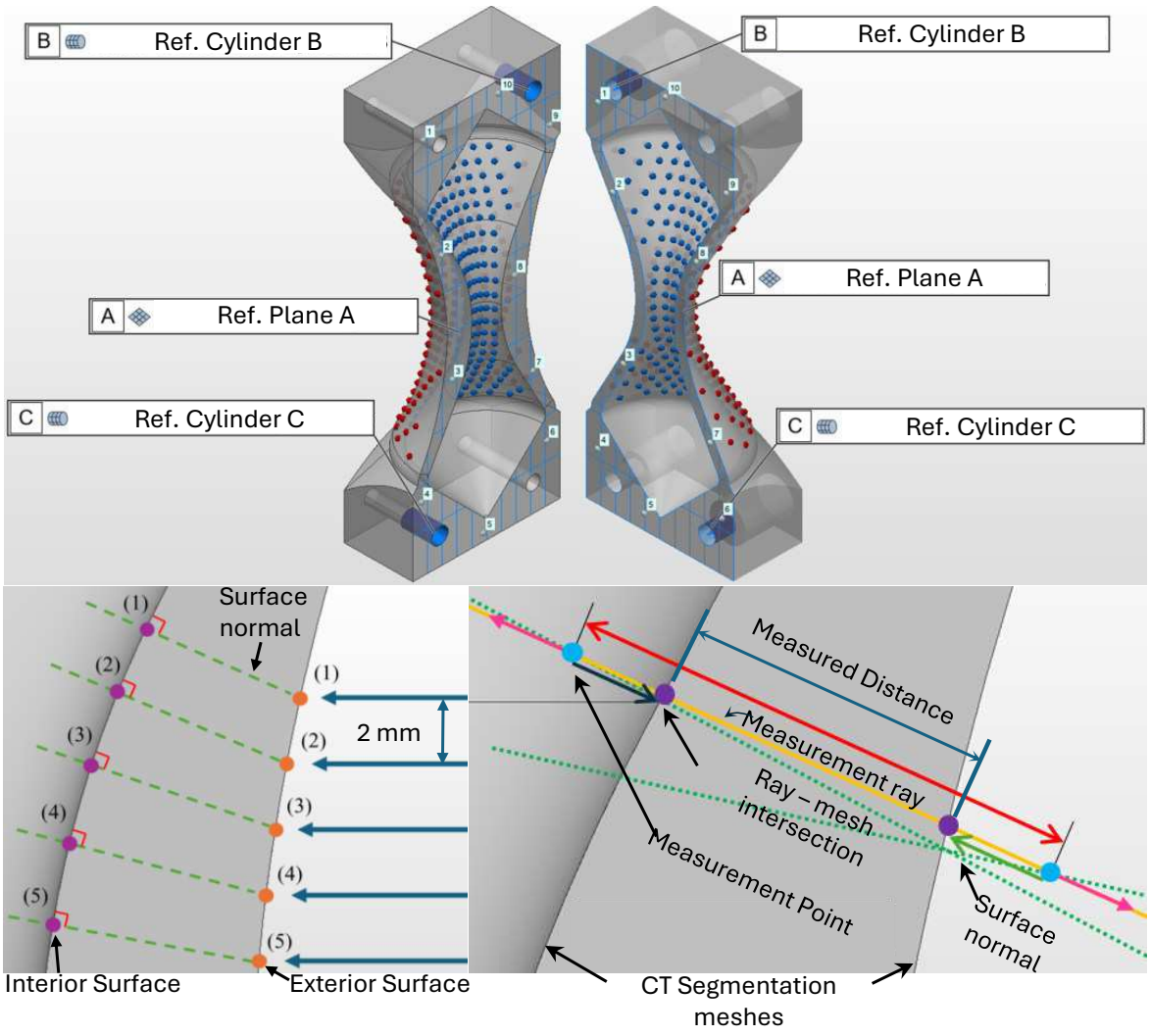


Figure 5: Benchmark comparison schematic showing rigid alignment of CMM points to CT-derived surfaces and computation of surface and thickness deviations.

3 Results

This section reports roughness, reconstruction and surface-determination challenges, surface roughness measurements, and CT–CMM comparison results. Reconstruction-based CT surface segmentations provide the released baselines; ER surfaces appear only as a CAD-informed comparison under the same convention.

Table 4: Measured roughness values R_a, R_z, and R_p on the Mob part before and after chemical milling. Units are μm.

Material	State	Zone	R_a	R_z	R_p
Ti-6Al-4V	Rough	Int High	19.392	163.629	87.212
Ti-6Al-4V	Rough	Int Low	14.553	136.050	72.982
Ti-6Al-4V	Rough	Ext High	24.980	227.796	126.797
Ti-6Al-4V	Rough	Ext Low	14.773	141.102	76.734
Inconel 625	Rough	Int High	8.145	80.824	50.633
Inconel 625	Rough	Int Low	11.131	94.501	52.337
Inconel 625	Rough	Ext High	12.512	139.920	96.111
Inconel 625	Rough	Ext Low	13.160	124.301	90.414
Ti-6Al-4V	Polished	Int High	9.276	97.246	56.527
Ti-6Al-4V	Polished	Int Low	7.691	66.356	25.613
Ti-6Al-4V	Polished	Ext High	10.732	77.764	45.997
Ti-6Al-4V	Polished	Ext Low	6.294	51.029	21.675
Inconel 625	Polished	Int High	5.524	45.039	18.926
Inconel 625	Polished	Int Low	5.567	51.752	18.673
Inconel 625	Polished	Ext High	4.625	41.039	19.977
Inconel 625	Polished	Ext Low	4.355	42.493	14.715

3.1 Surface-state characterization

Table 4 reports the measured roughness values R_a, R_z, and R_p on the Mob part before and after chemical milling. Chemical milling reduces all three parameters across all four zones for both materials.

3.2 Qualitative benchmark execution and reconstruction challenges

Figure 6 shows representative reconstructions for the rough titanium and Inconel artefacts, whereas Figure 7 shows the Chinook rotor hub dataset, discussed in[26]. The titanium case exhibits clearer metallic-air boundaries, and low noise in the cone region. The rough Inconel case shows substantially higher noise, especially in the cone region of the assembly block, where extinction artefacts and low signal to noise ratio lower the transmitted signal and make the surface harder to localize, with stronger cone-beam CT artefacts around the dense base. The CMM-measured thickness pair lie outside of the cone, however, so these local failures are not directly sampled by the thickness statistics, but they do substantially affect local surface reconstruction and alignment. Cone beam artefacts are also prevalent in the Chinook CBCT data, particularly at the upper tip, whereas the boundaries are far clearer in the LDA dataset, despite its lower resolution providing overall better accuracy.

Figure 8 shows representative surface outputs for the titanium case. The three FSF settings are visually close, but changing FSF shifts the recovered boundary by about one pixel. The ER surface is visually close to the FSF surfaces in this example. The Fix/Mob interface remains sensitive because the foil separator and cavity boundary must both be resolved.

For Inconel, extinction is the dominant reconstruction challenge. Figure 9 shows how the cone region in the high assembly block becomes difficult to recover with local-gradient surface determination when the detector receives insufficient transmitted intensity. A DL segmentation route was therefore used for the Inconel case. Outside the cone, where attenuation is less severe, the surfaces are much closer. Elastic registration does not explicitly

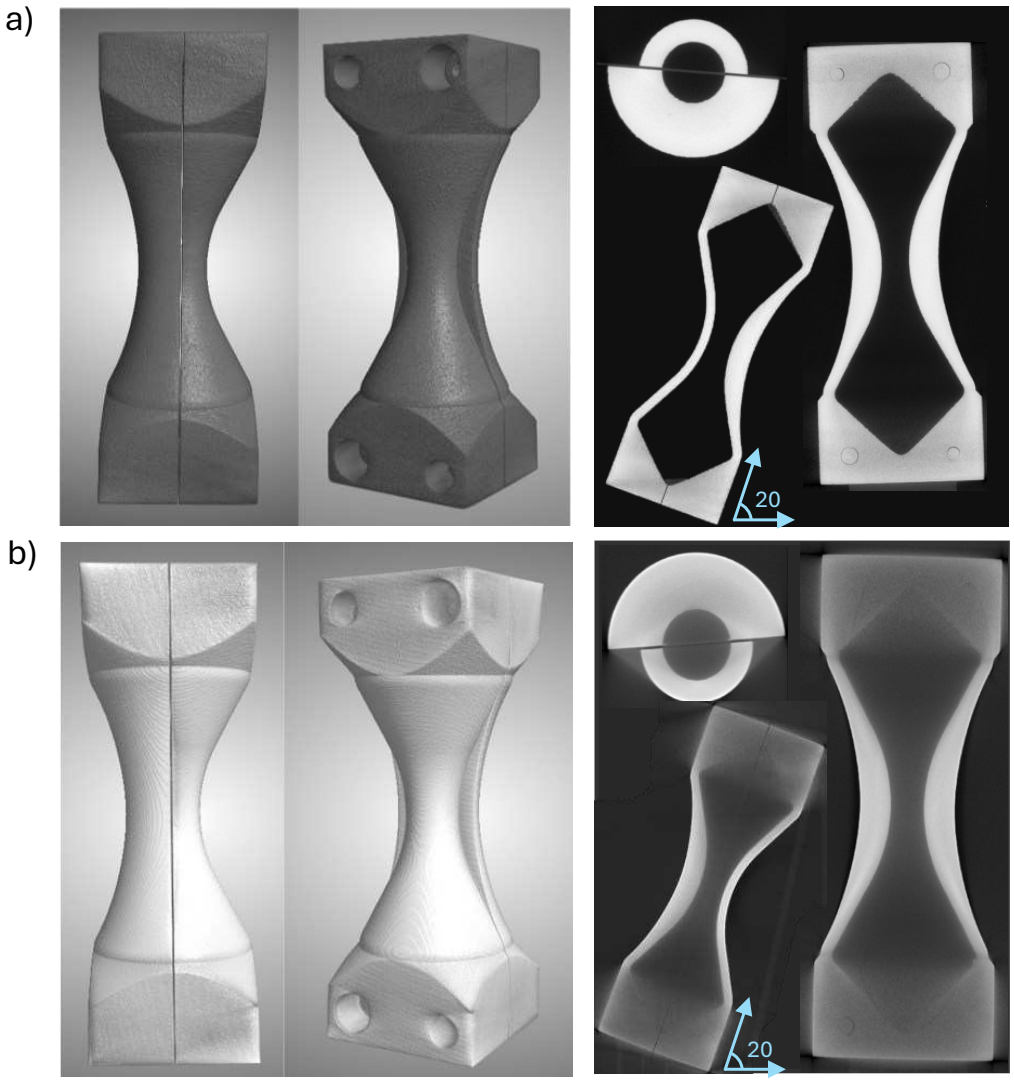


Figure 6: Representative CT reconstructions for the rough artefacts: (a) Ti-6Al-4V and (b) Inconel 625. Three-dimensional reconstructed views are shown on the left and two-dimensional reconstructed slices on the right, including the tilted acquisition orientation.

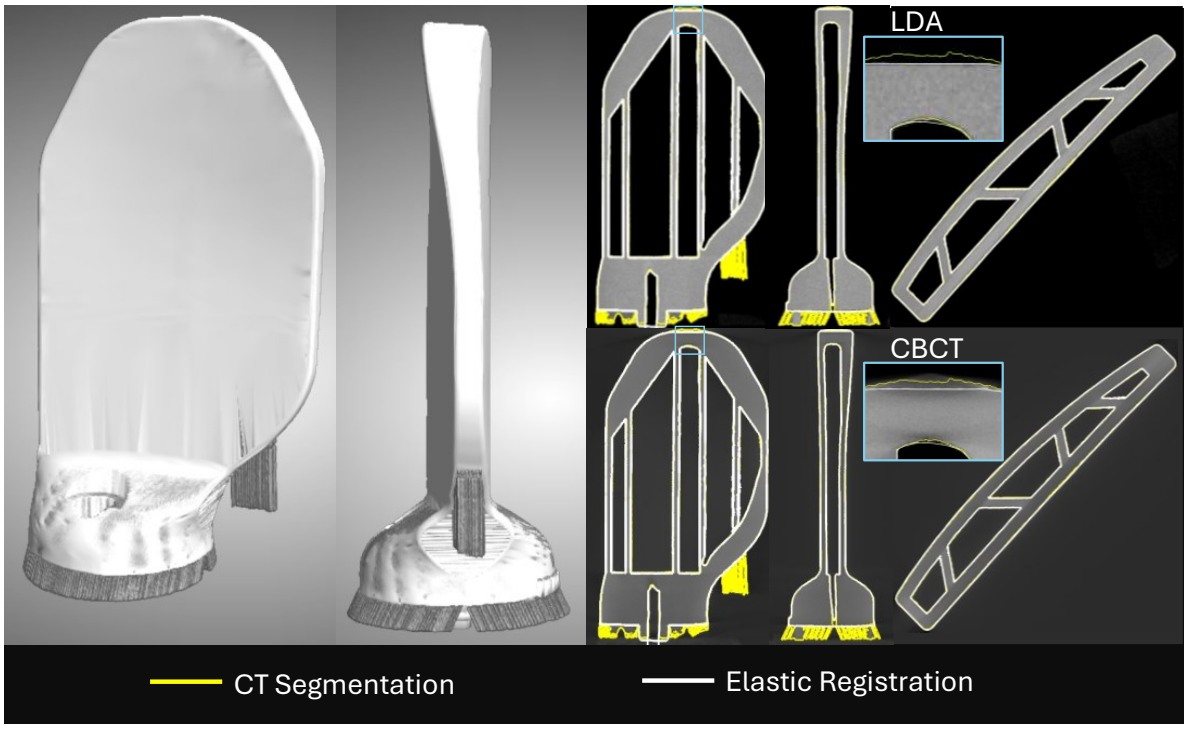


Figure 7: Overview and visual representation of the aligned Chinook Rotor Hub CBCT and LDA reconstructions. CT segmentation is done using Dragonfly FSF 0.6 on the CBCT reconstruction, whereas elastic registration was executed using the CBCT raw projections.

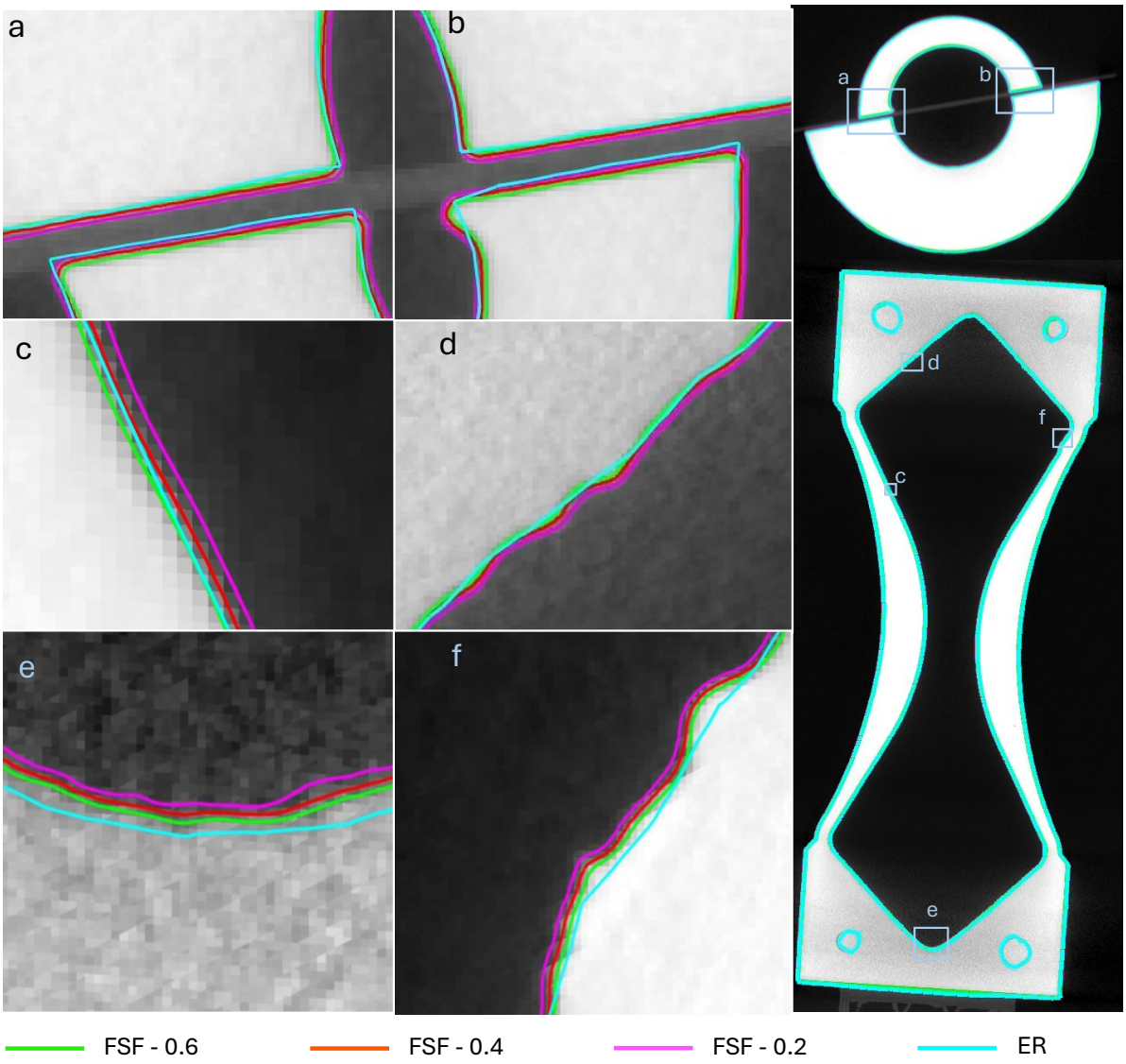


Figure 8: Representative surface extractions for the titanium artefact. Panels (a–f) enlarge the matching labelled regions in the overview slices at right: (a,b) the Fix/Mob interface, (c) the inner wall, (d,e) the upper and lower cone regions, and (f) the outer wall. Colours identify FSF 0.6 (green), FSF 0.4 (orange), FSF 0.2 (magenta), and ER (cyan). The FSF-based surfaces are very close overall.

Table 5: Baseline thickness-deviation statistics for the reconstruction-based CT surface segmentations and ER comparison surfaces. 399 CMM measurement point pairs were used per artefact case.

State	Config	Mean (mm)	Std (mm)	P95 (mm)
Rough	Ti-0.2	-0.007	0.028	0.062
Rough	Ti-0.4	-0.068	0.030	0.123
Rough	Ti-0.6	-0.159	0.034	0.223
Rough	In-0.6	-0.002	0.025	0.046
Rough	In-DL	0.036	0.032	0.085
Polished	Ti-0.2	0.074	0.026	0.114
Polished	Ti-0.4	0.014	0.026	0.054
Polished	Ti-0.6	-0.033	0.026	0.074
Polished	Ti-ER	-0.012	0.029	0.034
Polished	In-0.6	-0.027	0.024	0.064
Polished	In-DL	0.039	0.026	0.083
Polished	In-ER	0.027	0.026	0.065

correct cone-beam artefacts and relies on edge extraction of the raw CBCT projections and prior-known CAD geometry to better estimate surface location. In this example, its cone outline is visually closer to the nominal shape than the local-gradient result, but as other cases remains affected by the diminished signal and noise. This region lies outside the CMM-validated zone, but can be compared with the 3D scan data.

3.3 Quantitative baseline CT–CMM comparison

Figure 10 summarizes surface deviation maps for the four benchmark cases using the reconstruction-based baseline surface segmentations selected for the smallest absolute mean thickness deviation in each case, while Figure 11 compares thickness-deviation histograms for all applicable FSF and DL routes, with ER included for the polished cases. In well-resolved regions, surface deviations are usually within about $\pm 25\,\mu$m; local extraction failures can exceed $150\,\mu$m. Rough configurations tend to show negative mean thickness deviations consistent with mechanical filtering. After chemical milling, several distributions move toward zero and the gap between the different surfaces declines markedly.

From the deviation maps in Figure 10 it is also seen that generally speaking deviations are greatest near the neck, where thickness is greatest and we have concave curvature, which is particularly evident in the Inconel case. Likewise, on the thinner Mob case, deviations are overall greater than in the Fix case illustrating spatial variation with wall thickness and local geometry. Thickness and curvature co-vary here, so these maps do not isolate a thickness effect.

Table 5 summarizes the benchmark behavior. Mean and standard deviation refer to signed thickness deviation, Δt; P95 is the 95th percentile of absolute thickness deviation, $|\Delta t|$. Rough-state mean offsets are substantial, reaching $-0.159\,$mm for Ti-0.6. In the rough

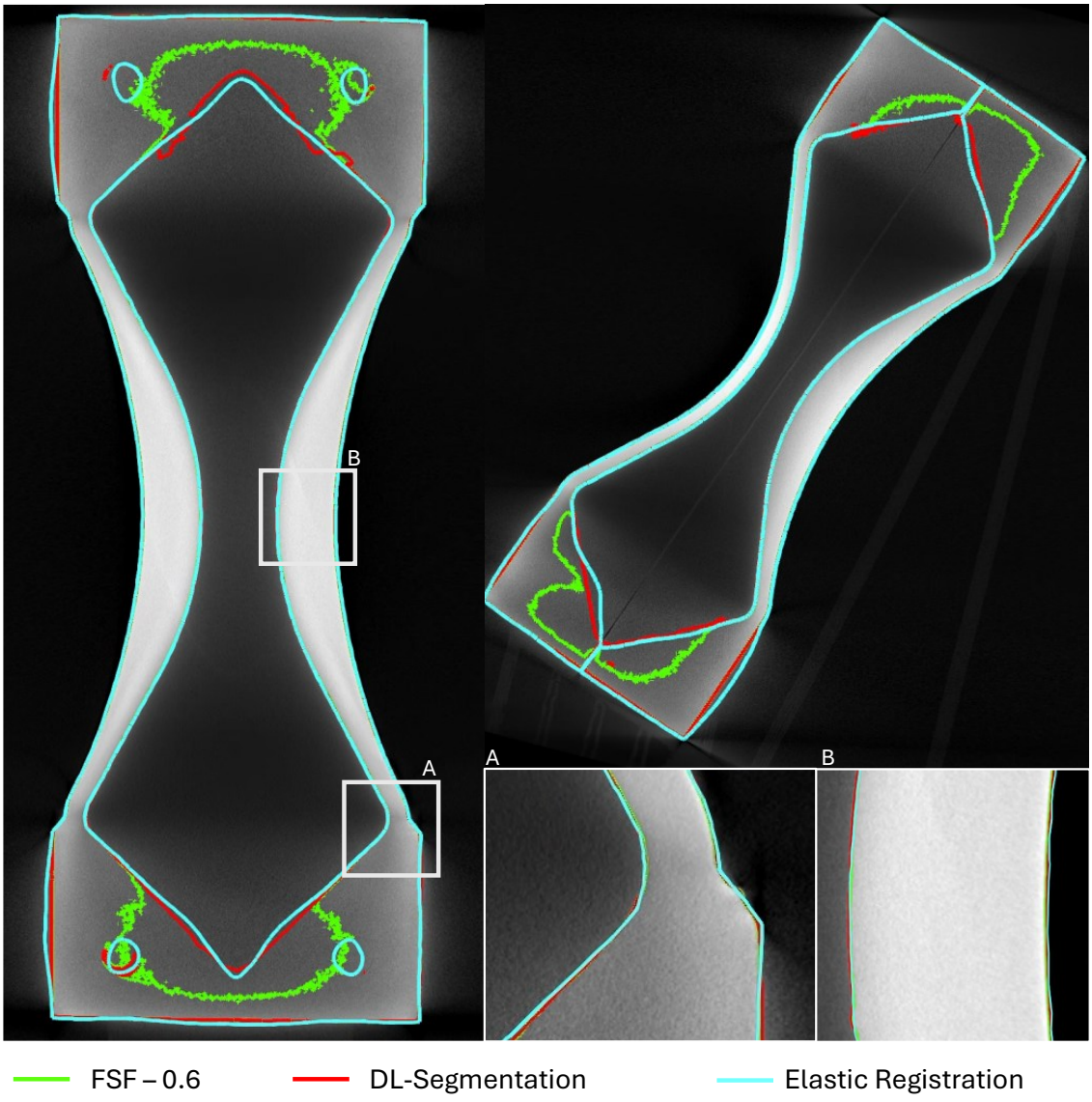


Figure 9: Extinction-driven extraction challenge in the Inconel artefact. Green, red, and cyan contours show FSF 0.6, DL segmentation, and ER, respectively. The lower-right insets A and B enlarge the corresponding boxed regions of the left-hand slice. The cone region is difficult for local-gradient surface determination, so a DL route was also used.

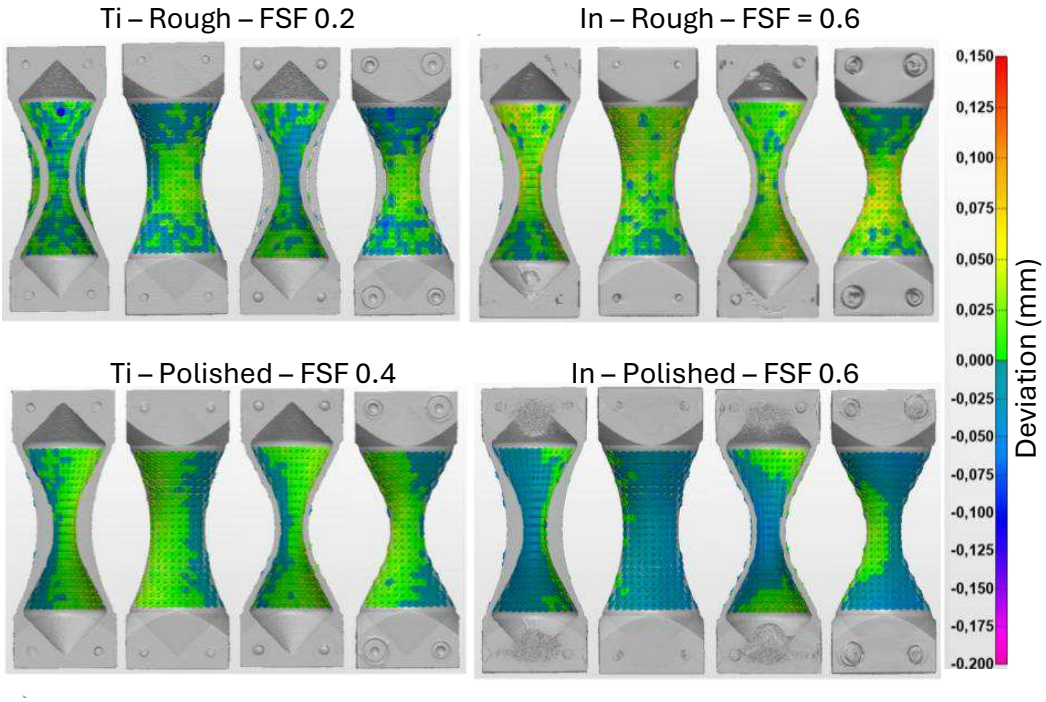


Figure 10: Surface deviation maps for the benchmark cases, grouped by material and surface state, using the reconstruction-based surface segmentation selected for the smallest absolute mean thickness deviation in each case. From left to right in each case, deviations between CMM and surface Segmentations for the Fix inside, Fix outside, Mob inside, Mob outside are shown for each case.

condition, FSF values of 0.2 and 0.6 produce the smallest absolute mean thickness deviations for titanium and Inconel, respectively. After chemical milling, FSF values of 0.4 and 0.6 yield the smallest absolute mean thickness deviations for titanium and Inconel, respectively. For titanium, linear interpolation between the polished-case deviation means at FSF 0.4 and 0.6 estimates a zero-mean setting of $0.4+0.2(0.014)/(0.014+0.033) \simeq 0.460$. Applying the same FSF to the rough-case yields an estimated mean offset of $-0.095\,\mathrm{mm}$, which is consistent with the measured surface roughness. These are interpolated estimates, not measurements from an additional reconstructed surface. The latter magnitude is of the same order as the zone-averaged maximum profile peak height R_p of 0.091 mm for rough titanium, not its mean R_a. Across the reported cases, no single tested FSF setting yields zero-mean CT–CMM agreement.

3.4 3D scans of polished artefacts

Figure 12 shows deviation maps between the 3D scans of the polished titanium and Inconel artefacts and the titanium FSF 0.4 and Inconel DL segmentation surfaces, respectively.

Mean unsigned deviations between 3D scans and segmentations are 0.1182 mm and 0.1462 mm for the titanium Fix and Mob parts, respectively, and 0.10422 mm and 0.117 mm for

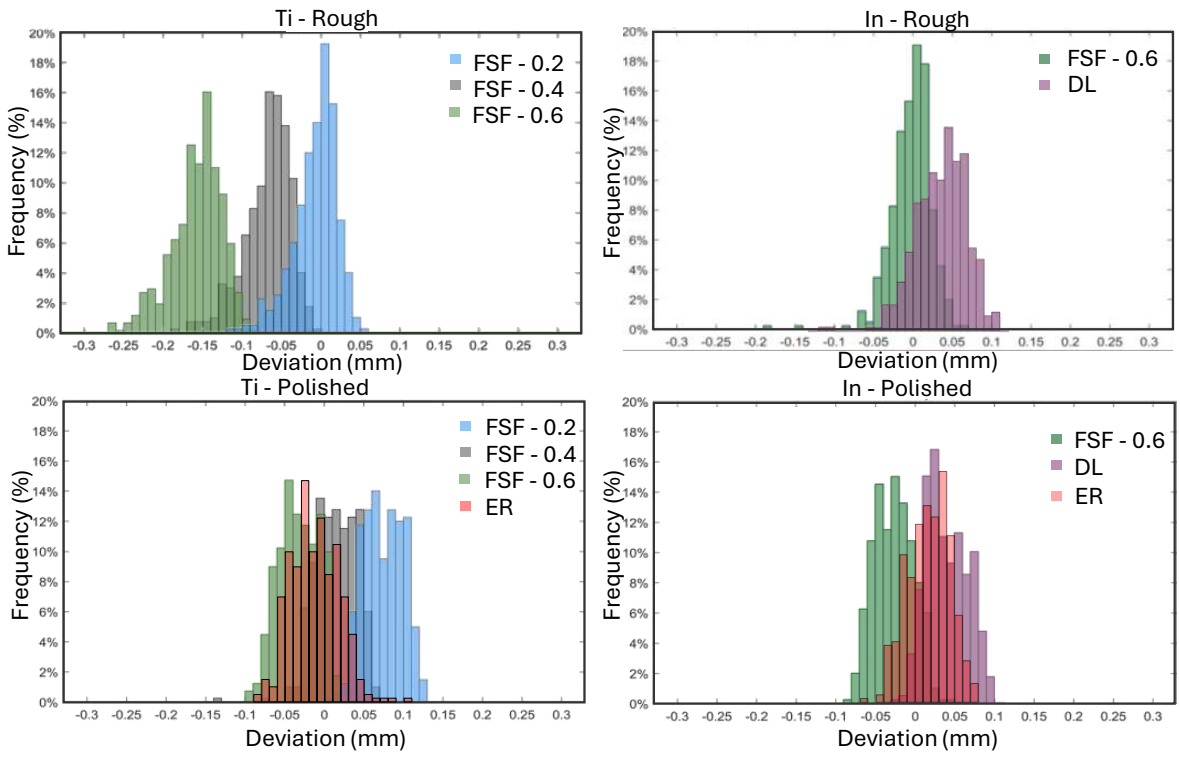


Figure 11: Thickness-deviation histograms between measurements using XCT-based 3D surface reconstruction and CMM across rough and polished cases. All applicable FSF and DL routes are shown, with ER comparison surfaces included for the polished cases.

the corresponding Inconel parts. These differences include scan precision, segmentation, and ICP registration accuracy and do not represent 3D scanner precision. ICP registration inaccuracy can be caused by inclusion of all 3D scan points including clear outliers, as well as segmentation failures failing to capture surfaces and thus biasing the registration. Indeed, the Inconnel Fix part in particular has minor registration error as the blue/red shift caused by misalignment of the scan and segmentation in the cone. Thus these numbers represent a rough accuracy estimate of the registered 3D scans, and use of these data should take into account these effects.

4 Discussion

The XCT-based evaluation of thickness, using several segmentation approaches and compared against CMM measurements, highlights several interesting findings. By design the study compares material density and surface roughness using the same nominal geometry. Chemical milling changes the realized geometry, and scan settings differ between materials, so these effects are not isolated.

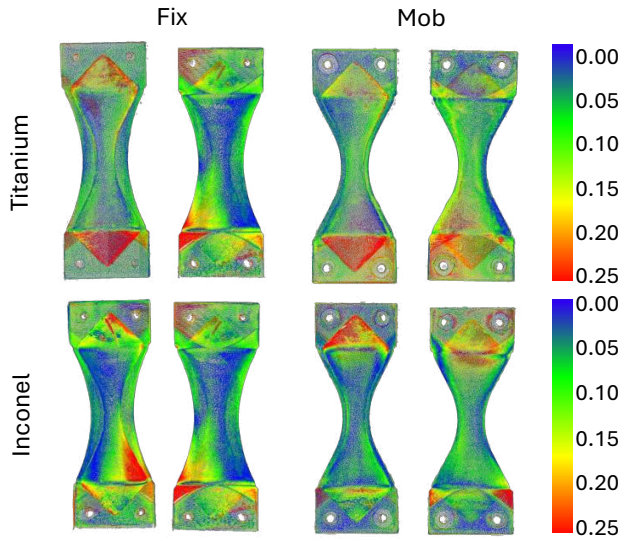


Figure 12: Unsigned point-to-plane deviations between 3D scanned parts and CBCT 3D reconstruction segmentation surfaces for titanium and Inconel polished cases. Segmentation surfaces used were the titanium FSF 0.4 and Inconel DL surfaces, respectively. Deviations are in millimetres.

4.1 Effect of material and CT artefacts

Figure 9 shows that the Inconel CT reconstruction suffers severe local extinction-driven failure in the cone region, where the reconstructed surface does not recover the cone geometry. In the illustrated case, deep learning and elastic registration recover a recognisable cone outline; the DL route partially captures the cone but misses the tip while the ER result shows local boundary departures reflecting its dependence on noisy projection edge maps. Since the cone region lies outside the CMM measurement zone, the accuracy of either method in this area is not quantified directly in the benchmark deviations. More broadly, because cone-beam artefacts are most pronounced near the base in all cases, their direct influence on the paired thickness comparison is expected to be limited, although indirect effects through alignment remain possible. The paired measurements therefore characterize a region less affected by the most severe visible CT artefacts, rather than validating the full surface. Although CMM measurements are not executed on the entire surface of the titanium and Inconel artefacts, 3D scans (mean unsigned CT surface-to-scan deviations of 0.104,22–0.146,2 mm) of the polished titanium and Inconel artefacts are provided and included in the dataset for reference.

4.2 Effect of surface roughness on benchmark results

Surface roughness affects both CT boundary localization and the practical CMM reference. Several rough configurations exhibit larger mean offsets, whereas chemical milling shifts some

distributions closer to zero in Fig. 11. This observation is consistent with the expectation that mechanical filtering over-predicts thickness because the probe tip rides over roughness peaks. The FSF setting interpolated to zero mean in the polished titanium case gives an estimated rough-case mean thickness deviation of approximately -0.095 mm, of the same order as the zone-averaged maximum profile peak height R_p of 0.091 mm in Table 4 - a comparison of magnitudes rather than an equivalence between R_p and thickness bias. Cone-beam artefacts, extinction artefacts, and scan noise also appear qualitatively stronger in the rough cases than in the polished cases, though as we do not quantify these effects, relating them directly to surface roughness is left to future work. For titanium, the standard deviation averaged over the three FSF settings in Table 5 decreases by about 15% as mean R_a decreases by about 50%, consistent with reduced deviation scatter after chemical milling.

4.3 Surface-determination FSF parameter sensitivity

The FSF setting giving the smallest absolute mean CT–CMM deviation varies among the reported material and surface-state cases. In the smooth cases, the FSF = 0.6 surfaces produce offsets of $-0.033\,\mathrm{mm}$ for titanium and $-0.027\,\mathrm{mm}$ for Inconel, respectively (Table 5), showing a modest difference between these two cases. Similarly, between the rough and chemically milled Inconel cases, the distribution of the FSF = 0.6 surface deviations shifts left in Fig. 11. This trend is opposite to the expected effect of mechanical filtering and shows that the effect of changing surface state cannot be explained by mechanical filtering alone. The FSF parameter corresponds to the locally normalized intensity transition used to place the surface, which can be affected by CT artefacts and noise. The observed CT–CMM agreement is therefore case-specific; these comparisons do not isolate material or roughness effects from differences in geometry and acquisition settings. No single tested FSF setting yields zero-mean CT–CMM deviation across these cases. Benchmark submissions should therefore report the surface-reconstruction route, parameter settings, any smoothing or post-processing, and whether the documented alignment was reused or recomputed. 3D surface reconstruction by elastic registration, which matches edges directly in the projection space, and segmentation by deep learning, which probabilistically infers object presence, do not place the boundary using an FSF parameter applied directly to the reconstructed grey-level transition. For the reported chemically milled cases, these approaches also yield mean CT–CMM deviations within 0.04 mm, although these mean offsets alone do not establish accuracy and their standard deviations are comparable to or higher than those of the FSF routes.

4.4 Limitations and recommended use

The CMM data are a high-quality practical reference rather than absolute geometric truth. Mechanical stylus filtering, datum choice, and alignment of the CMM data to the CT all contribute to the CT–CMM difference. The raw unaligned CMM point clouds are provided to permit different alignment schemes. Residual alignment uncertainty remains, and this study does not establish a full measurement uncertainty budget or certified metrological traceability. Chemical milling also changes the geometry of the polished cases, so rough and polished configurations are not direct before and after correspondences. Furthermore, though smoothness is high in the polished cases and roughness is quantified, mechanical filtering is

always present, and the zone-averaged R_p values in Table 4 provide a characteristic surface-peak scale, rather than a limit on CMM precision (never-the-less measuring betwen 0.54 and 1.12 voxels for the chemically milled Inconel and titanium cases, respectively). Users should report whether they reused the raw or aligned point clouds, whether they regenerated surfaces from the raw projections or used a different segmentation technique, and which thickness statistics are primary used. Direct comparison with the baseline tables requires the released thickness-pair indexing and application of the same alignment methodology.

5 Dataset contents and availability

The benchmark comprises the four primary cases listed in Table 1, each linking nominal CAD, raw CT projections, reconstructed CT volumes, CT surface segmentations, CMM point clouds, thickness-pair definitions, and derived deviation labels. Table 6 summarizes the object classes described by the dataset record.

Table 6: Benchmark objects for the four primary benchmark cases.

Object class	Contents
Nominal CAD	Fix/Mob nominal geometry
CT projections	projection stack and scan metadata
CT volumes	reconstructed CERA volumes
CT surfaces	Ti FSF 0.2/0.4/0.6; Inconel FSF 0.6 and DL segmentation surfaces for the chemically milled primary cases
CMM data	point clouds and thickness-pair vectors
Benchmark labels	surface and thickness deviations

For reproducible benchmark use, the reporting convention is stated explicitly. Geometric quantities and deviation statistics are reported in millimetres; voxel-scale comparisons for the primary cases use the nominal voxel size of $33.3\,\mu$m. Benchmark reports should state which released objects were used, whether alignment was reused or recomputed, which surface-reconstruction route and parameters were applied, and which summary statistics and deviation maps were compared.

6 Conclusion

This paper publishes a CT–CMM benchmark dataset for dimensional validation of a part with a complex internal cavity in LPBF Ti-6Al-4V and Inconel 625 artefacts. The benchmark couples assembled-state CT with disassembled-state CMM probing through a documented acquisition, alignment, and labelling protocol across two materials and two surface states. Thickness comparison is the primary quantitative endpoint, with surface-deviation maps as supporting diagnostics. Baseline results show case-specific differences in CT–CMM agreement across materials, surface states, and surface-reconstruction routes. Their standard deviations are approximately on the voxel scale, whereas mean offsets and local deviations can be larger. No single tested FSF setting yields zero-mean CT–CMM agreement across

the reported roughness and material states; this does not isolate roughness or material effects from the accompanying geometry and acquisition differences. The dataset therefore provides a practical basis for evaluating volume-reconstruction, surface-reconstruction, and CAD-informed comparison methods on complex internal cavities.

Data Availability

The benchmark dataset has a persistent Borealis record at `https://doi.org/10.5683/SP3/5XLJVS`[31]. Version 1.0 was published on 9 September 2026, and the record lists a CC0 1.0 licence. The Borealis record provides metadata and links to the data files hosted in TomoBank[32], with the catalogue and access links at `https://tomobank.readthedocs.io/en/latest/source/data/docs.data.metrology.html`. The TomoBank accessions map to the cases as follows: tomo_00110, rough Ti-6Al-4V; tomo_00111, chemically milled Ti-6Al-4V; tomo_00112, rough Inconel 625; tomo_00113, chemically milled Inconel 625; and tomo_00114, the auxiliary Chinook CBCT/LDA case. Each accession links to its corresponding TomoBank Globus folder. The four primary cases share CAD geometry, CT projections and reconstructions, surface segmentations, CMM measurements and thickness-pair definitions, roughness measurements, and derived deviations; 3D scans are included for the chemically milled artefacts. The auxiliary Chinook case contains CT data without a CMM thickness reference.

Code Availability

No previously unreported custom code or algorithm was used to generate or process the released reference data. CT reconstructions were generated in CERA, CT surface segmentations in Dragonfly 3D World 2025.1, and CMM post-processing and CMM-to-CT alignment in PolyWorks Metrology Suite 2024 IR9. MATLAB R2024b was used only to apply Eq. (1) and calculate the reported descriptive summary statistics. The separately published CADER method[26] was used only to generate the ER comparison surfaces. CERA, PolyWorks, MATLAB, and Dragonfly are third-party software and are not distributed with this work. Dragonfly is available from the developer at `https://dragonfly.comet.tech/`; eligible non-commercial academic users may apply for a free academic licence.

References

[1] J. P. Kruth, M. Bartscher, S. Carmignato, R. Schmitt, L. De Chiffre, and A. Weckenmann. Computed tomography for dimensional metrology, 2011. URL `https://doi.org/10.1016/j.cirp.2011.05.006`. CIRP Annals , 60 (2), 821–842.

[2] T. Hovell, J. Petzing, W. Guo, C. Gill, L. Justham, N. Lohse, and P. Kinnell. Measurement techniques for three-dimensional metrology of high aspect ratio internal features–a review, 2023. URL `https://doi.org/10.3390/metrology3020009`. Metrology , 3 (2), 139–168.

[3] H. Villarraga-Gomez, C. Lee, and S. T. Smith. Dimensional metrology with x-ray ct: A comparison with cmm measurements on internal features and compliant structures, 2018. URL https://doi.org/10.1016/j.precisioneng.2017.08.021. Precision Engineering , 51 , 291–307.

[4] H. Villarraga-Gómez, E. L. Herazo, and S. T. Smith. X-ray computed tomography: from medical imaging to dimensional metrology, 2019. URL https://doi.org/10.1016/j.precisioneng.2019.06.007. Precision Engineering , 60 , 544–569.

[5] Jennifer M. Sietins, William H. Green, and Justin S. Jones. Materials evaluation using x-ray computed tomography. In *Comprehensive Structural Integrity*, pages 159–199. Elsevier, 2023. doi: 10.1016/B978-0-12-822944-6.00038-4. URL https://doi.org/10.1016/B978-0-12-822944-6.00038-4.

[6] W. Sun, C. Giusca, S. Lou, X. Yang, X. Chen, T. Fry, X. Jiang, A. Wilson, S. Brown, and H. Boulter. Establishment of x-ray computed tomography traceability for additively manufactured surface texture evaluation, 2022. URL https://doi.org/10.1016/j.addma.2021.102558. Additive Manufacturing , 50 , 102558.

[7] J. J. Lifton, Y. Liu, Z. J. Tan, B. Mutiargo, X. Q. Goh, and A. A. Malcolm. Internal surface roughness measurement of metal additively manufactured samples via x-ray ct: the influence of surrounding material thickness, 2021. URL https://doi.org/10.1088/2051-672X/ac0e7c. Surface Topography: Metrology and Properties , 9 (3), 035008.

[8] X. Chen, S. Lou, W. Sun, P. Scott, and X. Jiang. Investigation of the correlation between radiographic image quality and surface measurement quality of xct using frequency response analysis, 2025. URL https://doi.org/10.1016/j.cirp.2025.04.031. CIRP Annals , 74 (1), 713–717.

[9] A. Klos, L. Salvo, and P. Lhuissier. Super-resolution x-ray tomography using deep learning applied to the 3d quantification of defects in lattice structures, 2025. URL https://doi.org/10.1038/s41598-025-20372-4. Scientific Reports, 15, 36664.

[10] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation., 2015. URL https://doi.org/10.48550/arXiv.1505.04597.

[11] A. Weckenmann, T. Estler, G. Peggs, and D. McMurtry. Probing systems in dimensional metrology, 2004. URL https://doi.org/10.1016/S0007-8506(07)60034-1. CIRP Annals , 53 (2), 657–684.

[12] S. Lou, S. B. Brown, W. Sun, W. Zeng, X. Jiang, and P. J. Scott. An investigation of the mechanical filtering effect of tactile cmm in the measurement of additively manufactured parts, 2019. URL https://doi.org/10.1016/j.measurement.2019.04.066. Measurement , 144 , 173–182.

[13] V. Aloisi and S. Carmignato. Influence of surface roughness on x-ray computed tomography dimensional measurements of additive manufactured parts, 2016. URL

https://doi.org/10.1016/j.csndt.2016.05.005. Case Studies in Nondestructive Testing and Evaluation , 6 , 104–110.

[14] S. Carmignato, V. Aloisi, F. Medeossi, F. Zanini, and E. Savio. Influence of surface roughness on computed tomography dimensional measurements, 2017. URL https://doi.org/10.1016/j.cirp.2017.04.067. CIRP Annals , 66 (1), 499–502.

[15] W. Liu, X. Chen, W. Zeng, W. Sun, D. Gorman, A. Wilson, Q. Qi, P. Scott, X. Jiang, and S. Lou. Comparison of x-ray computed tomography and coordinate-measuring system dimensional measurement for additive manufacturing parts using physical and simulation methods, 2024. URL https://doi.org/10.1016/j.measurement.2024.114414. Measurement , 229 , 114414.

[16] W. Liu, X. Chen, W. Zeng, D. Tang, P. J. Scott, X. J. Jiang, and S. Lou. An advanced method to assess mechanical filtering effect of tactile coordinate-measuring system in dimensional measurement of additive manufacturing parts, 2025. URL https://doi.org/10.1016/j.procir.2025.01.013. Procedia CIRP , 132 , 74–79.

[17] S. Moylan, A. Cooke, K. Jurrens, J. Slotwinski, and M. A. Donmez. A review of test artifacts for additive manufacturing process measurement and control, 2012. URL https://doi.org/10.6028/NIST.IR.7858. NIST Interagency/Internal Report (NISTIR) 7858.

[18] V. M. Rivas Santos, A. Thompson, D. Sims-Waterhouse, I. Maskery, P. Woolliams, and R. Leach. Design and characterisation of an additive manufacturing benchmarking artefact following a design-for-metrology approach, 2020. URL https://doi.org/10.1016/j.addma.2019.100964. Additive Manufacturing , 32 , 100964.

[19] D. A. Broadbent, D. M. Gorman, W. Zeng, and S. Lou. Modular system of additive manufacturing benchmarking artefacts for xct inspection using a design-for-metrology approach, 2026. URL https://doi.org/10.1016/j.precisioneng.2025.10.012. Precision Engineering , 97 , 706–721.

[20] F. Zanini, M. Sorgato, E. Savio, and S. Carmignato. Dimensional verification of metal additively manufactured lattice structures by x-ray computed tomography: Use of a newly developed calibrated artefact to achieve metrological traceability, 2021. URL https://doi.org/10.1016/j.addma.2021.102229. Additive Manufacturing , 47 , 102229.

[21] A. M. Müller and T. Hausotte. Improving template-based ct data evaluation by integrating cmm reference data into a cad model-based high fidelity triangle mesh, 2022. URL https://doi.org/10.1080/10589759.2022.2091135. Nondestructive Testing and Evaluation , 37 (5), 692–706.

[22] B. Lane and H. Yeung. Process monitoring dataset from the additive manufacturing metrology testbed (ammt): Three-dimensional scan strategies, 2019. URL https://doi.org/10.6028/jres.124.033. Journal of Research of the National Institute of Standards and Technology , 124.

[23] B. Lane and H. Yeung. Process monitoring dataset from the additive manufacturing metrology testbed (ammt): Overhang part x4, 2020. URL https://doi.org/10.6028/jres.125.027. Journal of Research of the National Institute of Standards and Technology , 125.

[24] M. Praniewicz, B. Lane, F. Kim, and C. Saldana. X-ray computed tomography data of additive manufacturing metrology testbed (ammt) parts: Overhang part x4, 2020. URL https://doi.org/10.6028/jres.125.031. Journal of Research of the National Institute of Standards and Technology , 125.

[25] S. C. Feng, Y. Lu, and A. T. Jones. Measured data alignments for monitoring metal additive manufacturing processes using laser powder bed fusion methods, 2020. URL https://doi.org/10.1115/DETC2020-22478. Proceedings of the ASME 2020 IDETC/CIE Conference, Volume 9: 40th Computers and Information in Engineering Conference (CIE), V009T09A022.

[26] H. Valtchanov, W. Liu, S. Singh, J. Paccoud, A. Al Sheghri, J. Dompierre, C. Desrosiers, N. Piché, V. Brailovski, and F. Guibault. Cader: Cad-constrained elastic registration for voxel-accurate sparse-view 3d surface reconstruction in minutes, 2026. URL https://doi.org/10.1016/j.ndteint.2026.103791. NDT & E International, 163, 103791.

[27] International Organization for Standardization. ISO 4287:1997, Geometrical Product Specifications (GPS) — Surface texture: Profile method — Terms, definitions and surface texture parameters, 1997. URL https://www.iso.org/standard/10132.html.

[28] International Organization for Standardization. ISO 4288:1996, Geometrical Product Specifications (GPS) — Surface texture: Profile method — Rules and procedures for the assessment of surface texture, 1996. URL https://www.iso.org/standard/2096.html.

[29] H. P. Steinbeiß. *Dimensionelles messen mit mikro-computertomographie*. PhD thesis, Technische Universität München, Munich, Germany, 2005.

[30] Revopoint 3D Technologies Inc. *MINI 3D Scanner Quick Start Guide*, 2023. URL https://download.revopoint3d.com/support/download/mini/mini-quickstartguide-en-v3.0-2023.7.27.pdf.

[31] Josselin Paccoud, Hristo Valtchanov, Eric Fournier, Vladimir Brailovski, and François Guibault. A ct–cmm benchmark dataset for dimensional validation of internal cavities in lpbf titanium and inconel artefacts. Borealis, version 1.0, 2026. URL https://doi.org/10.5683/SP3/5XLJVS. Data files hosted in TomoBank, records tomo_00110–tomo_00114.

[32] Francesco De Carlo, Doğa Gürsoy, Daniel J. Ching, K. Joost Batenburg, Wolfgang Ludwig, Lucia Mancini, Federica Marone, Rajmund Mokso, Daniël M. Pelt, Jan Sijbers, and Mark Rivers. Tomobank: a tomographic data repository for computational x-ray science. *Measurement Science and Technology*, 29(3):034004, 2018. doi: 10.1088/1361-6501/aa9c19. URL https://doi.org/10.1088/1361-6501/aa9c19.

Author Contributions

Josselin Paccoud: Data curation, Methodology, Formal Analysis, Investigation, Writing – original draft. Hristo Valtchanov: Conceptualization, Methodology, Formal Analysis, Supervision, Writing – original draft. Eric Fournier: Methodology, Investigation (execution of CT surface segmentations). Vladimir Brailovski: Project administration, Resources, Supervision, Funding acquisition, Writing – review & editing. François Guibault: Conceptualization, Project administration, Resources, Supervision, Funding acquisition, Writing – review & editing.

Competing Interests

Eric Fournier is affiliated with Comet Technologies Canada Inc., which provided technical resources and segmentation expertise. Pratt & Whitney Canada provided project funding and the Chinook LDA scans. The authors declare no other competing interests.

Acknowledgements

The authors gratefully acknowledge Pratt & Whitney Canada for insight into metrology best practices and LDA scans of the Chinook blade, and Comet Technologies Montreal (Dragonfly) for technical resources and segmentation expertise.

Funding

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Alliance program (grant ALLRP 577578-22), PRIMA Québec (project R23-13-010), and Pratt & Whitney Canada.