

Uncovering hidden bias in neutron diffraction residual strain measurements

Cole Franz,^{a*} Michael Prime,^b Jeffrey Bunn,^c Andrew Payzant^c and Katharine Page^a^aTickle College of Engineering, University of Tennessee at Knoxville, Knoxville, Tennessee 37920, USA, ^bLos Alamos National Laboratory, Los Alamos, New Mexico, USA, and ^cOak Ridge National Laboratory, Oak Ridge, Tennessee, USA.

*Correspondence e-mail: cfranz@vols.utk.edu

Received 25 March 2026

Accepted 17 April 2026

Edited by J. Keckes, Montanuniversität Leoben, Austria

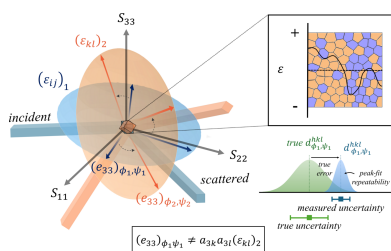
Keywords: residual strain; residual stress; neutron diffraction; full tensors; solid-state additive manufacturing.**Supporting information:** this article has supporting information at journals.iucr.org/j

When calculating residual strain via neutron or X-ray diffraction, uncertainties propagated from the peak fit are often inadequate to describe the true scatter of measurements about a singular strain state, such as one that should describe a macroscopic continuum. Because diffraction is inherently a selective process, orientation-dependent scatter arises from the sub-sampling of strong microstructure and strain gradients. This paper investigates the appropriateness of propagated uncertainties with reference to their original intention, *i.e.* noise about a mean value. Thirty-six unique orientations of strain measurements are taken at multiple locations within an additive friction-stir deposition component with fine-scale gradients ($\sim 200\ \mu\text{m}$) of plastic strain, texture and residual elastic strain. Multiple strain and stress calculation pathways are compared: direct substitution of three measurements into Hooke's law, direct inversion of any six unique orientations into the strain state tensor and thirty-six measurement least-squares estimation. For the last two cases, the appropriateness of the uncertainty interval is statistically evaluated on the basis of a physical constraint: common agreement under the strain transformation law. For this sample, the direct inversion of six measurements retains a conservative estimate of the uncertainty. However, propagated uncertainties in the least-squares solution greatly underestimate the true experimental scatter. A simple pathway to estimate appropriate uncertainty intervals is suggested. These results demonstrate that the interpretation of uncertainty in residual strain is strongly dependent on intrinsic sample-dependent effects, and that oversampling orientations and statistical analysis can give more accurate results with realistic uncertainties.

1. Introduction

Residual stress calculations made with neutron diffraction are central to qualifying advanced components, including additively manufactured (AM) materials. For these calculations, uncertainties are typically reported as one-standard-deviation intervals derived from peak fitting. When propagated through the strain or stress tensor, these uncertainties are expected to describe the variability of the solution. For example, two independently computed solutions should agree within their one-standard-deviation intervals approximately 68% of the time. In practice, however, agreement is often significantly poorer, particularly in AM components that exhibit steep gradients in elastic strain, stored plastic strain and crystallographic texture.

This discrepancy reflects a broader issue common to many inverse problems: propagated measurement uncertainties do not necessarily capture the true variability of the solution. In residual stress analysis, inverse methods are used at multiple stages, including diffraction peak fitting and tensor reconstruction. It is well established that inverse problems may contain inherent and often unknowable bias, particularly when



regularization or simplifying assumptions are required. For example, incremental relaxation methods introduce bias through smoothing, leading to uncertainty estimates that underestimate the true error (Beghini & Grossi, 2024; Beghini *et al.*, 2023). In such cases, the inadequacy of uncertainties obtained by simply propagating measurement uncertainties through the analysis has long been recognized (Prime & Hill, 2006). Diffraction methods also rely on inverse procedures, most notably the fitting of diffraction peaks to assumed peak shapes. In neutron diffraction, it is conventional to propagate uncertainties based on the counting statistics of a peak fit. However, an experimental study comparing measurements from multiple neutron instruments on the same welded specimen found that these propagated uncertainties were ‘significantly lower than the actual uncertainties’ (Akrivos *et al.*, 2020).

In neutron diffraction, uncertainty underestimation arises from both measurement-related and analysis-related effects. During measurement, systematic deviations may arise from peak-fitting assumptions, instrument alignment or, more fundamentally, from the interaction between the neutron beam and populations of aligned crystallites. In samples with strong texture or attenuation gradients, the ensemble of diffracting crystallites may not represent the geometric center of the measurement volume, leading to anomalous peak shifts. These effects can increase point-to-point scatter well beyond the nominal peak-fit uncertainty (Holden *et al.*, 2015; Hutchings *et al.*, 2005). While experimental strategies such as scanning detectors can mitigate some of these effects, a more fundamental issue remains. Specifically, diffraction measurements do not always sample a true average of the underlying strain continuum. Because neutron diffraction probes relatively large volumes (of the order of cubic millimetres), measurements in materials with fine-scale gradients often

represent only a partial sampling of the strain field. In such cases, the fitted peak position may deviate from the true volumetric average, not because of measurement noise but because different orientations sample different subsets of the microstructure. This unrepresentative volume averaging is exacerbated by texture and attenuation gradients and reflects the absence of a representative volume element (RVE) within the measurement volume (Şeren *et al.*, 2023). As a result, the variability between measurements may be dominated by physical heterogeneity rather than random noise. This concept is captured in Fig. 1. Thus, the residual strains should be sampled in a manner that acknowledges the variability and self-equilibration of the strain continuum, which occurs at increasingly smaller length scales (Withers & Bhadeshia, 2001*a,b*).

Additional complications arise during the strain and stress tensor calculation. Since the tensor is reconstructed from a limited number of directional measurements, the problem is inherently inverse and may be poorly conditioned. Small deviations in the measured strains can therefore produce large variations in the calculated tensor components. A common approach to mitigating this instability is to combine many measurements in a least-squares framework, which reduces variance by averaging. This was initially done by Winholtz & Cohen (1988) and followed up by Bunn *et al.* (2014) to understand the affects of multi-axial loading. While effective at reducing scatter, this approach implicitly assumes that all measurements represent a single underlying strain state and does not account for systematic or individual variability among them. As a result, the propagated uncertainties from least-squares estimation may underestimate the true variability of the measurements.

Bias cannot be meaningfully evaluated unless the measurements are checked against a physical constraint. Mathematically, biases introduced by inverse solutions are generally unknowable unless additional physical constraints on the true solution can be applied (Beghini *et al.*, 2023). In neutron diffraction measurements, however, such a constraint exists. Since stress and strain are tensors, that constraint is the coordinate transformation law, which all measured lattice strains must satisfy if they represent components of a single strain tensor. In other words, the measured strains are not independent but must satisfy the strain transformation equation, often illustrated through Mohr’s circle. A previous study on an aluminium forging demonstrated that neutron peak-fit uncertainties had to be more than doubled for the measured strains to agree with the transformation equation within uncertainty (Prime *et al.*, 2014). That study involved only limited redundant measurement, leaving open questions about how broadly this issue applies.

In this work, we investigate these effects using a representative component produced by solid-state AM, which exhibits steep gradients in texture, plastic strain and elastic strain ($\leq 200 \mu\text{m}$) (Metz *et al.*, 2023, 2024). Three common stress calculation pathways are examined: (i) direct substitution of three orthogonal strain measurements into Hooke’s law, (ii) inversion of six unique orientations using the strain

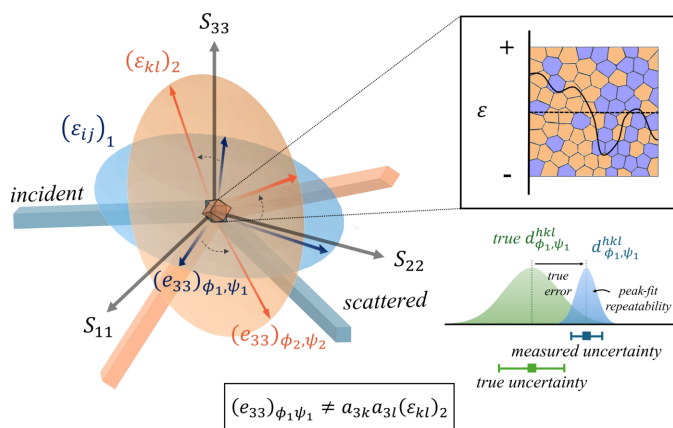


Figure 1 Schematic showing a case where two complimentary strain measurements produce unrepresentative volume averages, due to sampling of unique strain states within volumes of strong texture and strain gradients. The incident and scattered beams, and each orientation of measured strain and its sampled strain state (shown as strain ellipsoids), are distinguished by blue and orange colors, respectively. A true volumetric average and its associated uncertainty are more closely obtained by ensuring compatibility of measurements under the strain transformation law.

transformation equations and (iii) least-squares estimation of the strain tensor using 36 measurements. In this study, the peak-fit uncertainties are treated as accurate representations of measurement precision, but not necessarily of the variability of the underlying strain field. The results show that these one-standard-deviation peak-fit uncertainties behave as intended when propagated through the six-measurement direct inversion, capturing the empirical spread of solutions. In contrast, when many measurements are combined in a least-squares estimator, the propagated uncertainties converge to values that are no longer representative of the observed variability. This occurs because the assumption of a single strain state is violated, even though the transformation law itself remains valid. These findings demonstrate that, in materials with fine-scale heterogeneity, residual strains obtained by diffraction and represented as continuum strains must be interpreted with care and given sufficiently conservative uncertainties. The estimation of sufficient uncertainty depends strongly on both the measurement strategy and the analysis method.

2. Methods

2.1. Deposition process

Extruded AA6061-T6 bar stock (12.7 × 12.7 mm) was deposited onto a rolled AA6061-T6 baseplate (152.4 × 152.4 × 9.5 mm) using an L3 discrete feed additive friction-stir deposition (AFSD) machine (MELD Mfg, Christiansburg, Virginia, USA). In total, 11 unidirectional layers were deposited. Additional process specifications are given in the supporting information. A diagram of the deposition procedure and AFSD coordinate system is given in Fig. 2.

2.2. Neutron diffraction

Neutron diffraction was performed on the High Intensity Diffractometer for Residual Stress Analysis (HIDRA) at the High-Flux Isotope Reactor, located at Oak Ridge National Laboratory (Bunn *et al.*, 2023). The incident and scattering slits were set up to isolate a 3 × 3 × 3 mm volume in the AFSD sample. The scattering wavelength ($\lambda = 1.54 \text{ \AA}$) was selected to probe the face-centered cubic Al (311) interplanar spacing with a near-isotropic gauge volume ($2\theta \simeq 78^\circ$). Seven locations were sampled, approximately 2.54 mm apart (*i.e.* the AFSD deposition layer height), starting from the top surface

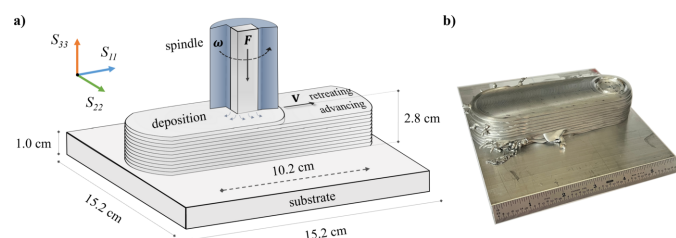


Figure 2

(a) Process diagram and coordinate system for AFSD. (b) Photograph of the AFSD component.

Table 1

Sampled orientations of lattice parameters.

	ϕ ($^\circ$)	ψ ($^\circ$)
$d^{(\phi, \psi)}$	0	90, 105, 120, 135, 150, 180
	15	90, 105, 120, 135, 150
	45	90, 105, 120, 135, 150
	60	90, 105, 120, 135, 150
	75	90, 105, 120, 135, 150
	90	90, 105, 120, 135, 150
$d_0^{(\phi, \psi)}$	0	90
	45	90
	69	90
	90	90

($S_{33} = 0$) and moving down two-thirds of the component height ($S_{33} = -2.80, -5.21, -7.62, -10.16, -12.70, -15.24$ and -17.78 mm). At each location, 36 orientations of strain measurements were taken over one-quarter of the total orientation space.

The peaks were fitted with pseudo-Voigt models using the *pyRS* analysis package (Fancher *et al.*, 2021). The diffraction and sample coordinate systems and angular conventions are shown in Fig. 3, alongside a photograph of the experimental setup. The in-plane and out-of-plane angle pairs (ϕ, ψ) for each sampled orientation are given in Table 1. The signal was counted for 10 min to achieve an average fitted peak position uncertainty equivalent to ± 50 – $100 \mu\epsilon$.

It was assumed that a strain-free reference was achieved by mechanical relief. For this purpose, a 5 mm diameter pillar was sectioned from an identical build, covering the same volumes as those sampled in the un-sectioned component (see supporting information, Fig. S1). To investigate sample texture, pole figures were collected from the 220, 311, 222 and 331 reflections using the strain-free reference pillar mounted on a miniature eulerian cradle. Texture was only collected at the bottom-most and top-most volumes ($S_{33} = -2.80$ and -17.78 mm). Orientation distribution functions (ODFs) were reconstructed from the measured pole figures using established conventions in the *MTEX* toolbox for MATLAB (Bachmann *et al.*, 2010; Hielscher & Schaeben, 2008).

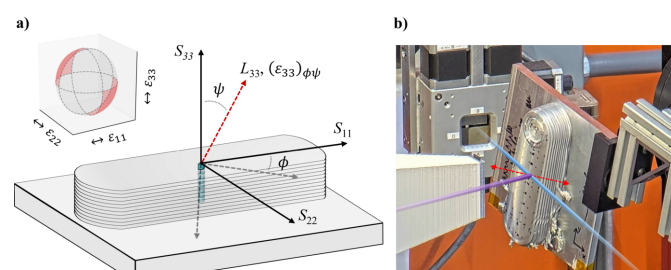


Figure 3

(a) Conventions used for the sample and diffraction coordinate systems. The inset in the top left-hand corner shows the region of strain space sampled ($0^\circ < \phi < 90^\circ, 90^\circ < \psi < 180^\circ$). Each approximately isotropic gauge volume is highlighted in blue. (b) Photograph of the AFSD sample within the neutron diffraction setup, which consists of multiple rotation stages mounted in tandem to rotate the sample about the center of each gauge volume. The transmitted beam, diffracted beam and diffraction (strain) vector are shown in blue, purple and red, respectively.

2.3. Strain and stress tensor calculations

In the following subsections, three different methods of calculation are compared. Each method is named according to the number of independent strain measurements used (*e.g.* three, six or 36 strains). The conventions for stress, standard deviation and uncertainty are written as σ_{ij} , s and μ , respectively.

The measured strains are first calculated by considering displacement of the (311) interplanar spacing with respect to its reference value,

$$(\varepsilon_{33})_{\phi\psi} = \frac{d^{\phi\psi} - d_0}{d_0}. \quad (1)$$

The total variation in d_0 for the 311 reflection was equivalent to $\pm 87 \mu\text{e}$, which is quite small. Thus, an average d_0 was used for all strain calculations in this paper.

2.3.1. Method 1 – three measurements

The most common method of residual stress mapping involves substitution of three mutually orthogonal strains into Hooke's law for isotropic elasticity,

$$\sigma_{ii} = \frac{E}{1-\nu} \left[\varepsilon_{ii} + \frac{\nu}{1-2\nu} (\varepsilon_{11} + \varepsilon_{22} + \varepsilon_{33}) \right]. \quad (2)$$

These strains are often measured along the component axes, but the equation is valid for any orthogonal set of directions. To calculate stress, strains measured along orientations (ϕ , ψ) parallel to the three sample axes, S_{11} (0° , 90°), S_{22} (90° , 90°) and S_{33} (0° , 180°), are substituted into equation (2). In this dataset, there is only a single instance of this solution for each depth. Elastic constants $E = 70 \text{ GPa}$ and $\nu = 0.35$ were used, based on the elastic behavior predicted from single-crystal properties and the sample ODFs (Hielscher & Schaeben, 2008; Bachmann *et al.*, 2010) (supporting information, Fig. S2).

2.3.2. Method 2 – six measurements

The previous stress calculation is perfectly valid for isotropic elasticity even if there are shear strains, although if the principal components of strain/stress are misaligned from the three chosen orientations, equation (2) will not capture the potentially significant influence of the shear stresses on mechanical behavior. To obtain all six independent components of the strain tensor directly, any six strains measured in the diffraction coordinate system $(\varepsilon_{33})_{\phi\psi}$ can be used to calculate the sample strain state in the sample coordinate system. These will be referred to as the 'sample strains'. First, the sample strains are expressed in terms of the measured strains using the fundamental law of strain transformation,

$$\begin{aligned} (\varepsilon_{33})_{\phi\psi} = & \varepsilon_{11} \cos^2 \phi \sin^2 \psi + \varepsilon_{22} \sin^2 \phi \sin^2 \psi + \varepsilon_{33} \cos^2 \psi \\ & + \varepsilon_{12} \sin 2\phi \sin^2 \psi + \varepsilon_{13} \cos \phi \sin 2\psi \\ & + \varepsilon_{23} \sin \phi \sin 2\psi. \end{aligned} \quad (3)$$

Computationally, this system of equations is solved using linear algebra. In this case, equation (3) in matrix notation is written as

$$\mathbf{e}_i = \mathbf{f}_j(\phi_i, \psi_i) \boldsymbol{\varepsilon}_j, \quad (4)$$

where \mathbf{e}_i is a 6×1 vector of measured strains $(\varepsilon_{33})_{\phi\psi}$, $\mathbf{f}_j(\phi_i, \psi_i)$ is a 6×6 matrix of transformation coefficients (*i.e.* trigonometric functions of ϕ and ψ) and $\boldsymbol{\varepsilon}_j$ is 6×1 vector containing the sample strains (*i.e.* the components of the strain tensor). The sample strains are found by matrix inversion,

$$\mathbf{f}_j(\phi_i, \psi_i)^{-1} \mathbf{e}_i = \boldsymbol{\varepsilon}_j. \quad (5)$$

This direct solution requires that the six chosen measurement orientations be linearly independent; otherwise, the system is ill-posed due to correlation among measurements. In practice, the numerical conditioning of the transformation matrix $\mathbf{f}_j(\phi_i, \psi_i)^{-1}$ determines how strongly the measurement uncertainties are amplified in the recovered tensor components. Well spaced orientations yield a well conditioned system, whereas nearly redundant orientations can lead to unreasonable solutions and uncertainties.

To demonstrate the effects of numerical conditioning, all possible strain tensor components were calculated using unique combinations of six strain measurements drawn from the available set of 36,

$$\frac{n!}{k!(n-k)!}, \quad (6)$$

where $n = 36$ is the total number of measured strains at a given location and $k = 6$ is the number of strains required for a direct tensor inversion. This yields 1947792 unique combinations per location, of which approximately 0.8% are non-invertible owing to linear dependence within $\mathbf{f}_j(\phi_i, \psi_i)^{-1}$. For each valid combination, the full sample strain tensor was recovered by solving the linear system in equation (5). Since the number of feasible solutions is exceedingly large, visualization was performed by randomly subsampling 50000 solutions per component and estimating their continuous distributions using an adaptive weighted kernel density estimation (KDE),

$$\hat{f}(x) = \frac{1}{\sum_{i=1}^n w_i h_i'} \sum_{i=1}^n w_i K\left(\frac{x-x_i}{h_i}\right). \quad (7)$$

Here, K is a Gaussian kernel, $w_i = 1/\mu_i^2$ are inverse-variance weights based on measurement uncertainty and h_i' are local adaptive bandwidths containing the measurement uncertainty,

$$h_i' = \sqrt{h_i^2 + \sigma_i^2}, \quad (8)$$

where h_i is the initial density estimate using Abramson's rule.

Hooke's law for isotropic elasticity was applied to calculate stress,

$$\sigma_{ij} = \frac{E}{1+\nu} \varepsilon_{ij} + \frac{\nu E}{(1+\nu)(1-2\nu)} \delta_{ij} \varepsilon_{kk}, \quad (9)$$

where δ_{ij} is the Kronecker delta. The resulting stress tensor components were likewise visualized using adaptive KDEs, with preprocessing steps (outlier rejection, normalization and adaptive bandwidth selection) identical to those used for the strain distributions. Outlier removal for all analyses is based on the interquartile method, which removes points beyond 1.5 times the interquartile range from the first and third quartiles.

2.3.3. Method 3 – thirty-six measurements

To incorporate all 36 strain measurements simultaneously, a weighted least-squares formulation was adopted following Winholtz & Cohen (1988). The closed-form solution is expressed as

$$\boldsymbol{\varepsilon}_j = [\mathbf{f}_j(\phi_i, \psi_i)^T \mathbf{W} \mathbf{f}_j(\phi_i, \psi_i)]^{-1} \mathbf{f}_j(\phi_i, \psi_i)^T \mathbf{W} \mathbf{e}_i, \quad (10)$$

where \mathbf{W} is the diagonal weight matrix, $\mathbf{W} = \text{diag}(1/\mu_{\mathbf{e}_i}^2)$, \mathbf{e}_i is a 36×1 vector of measured strains $(\varepsilon_{33})_{\phi\psi}$, $\mathbf{f}(\phi_i, \psi_i)$ is a 36×6 matrix of transformation coefficients (*i.e.* trigonometric functions of ϕ and ψ) and $\boldsymbol{\varepsilon}_j$ is 6×1 vector containing the sample strains (*i.e.* the components of the strain tensor). This approach assigns greater influence to measurements with smaller uncertainty. For reference, the detectability limit of the HIDRA beamline is equivalent to ± 100 to $150 \mu\text{e}$.

2.4. Uncertainty propagation

Uncertainties were evaluated following ISO/IEC Guide 98-3:2008 (GUM) (Joint Committee for Guides in Metrology, 2020). The uncertainties in each measured strain were obtained by the law of propagation of uncertainty,

$$u^2(\varepsilon_{hkl,i}) = \left(\frac{\partial \varepsilon}{\partial d_i}\right)^2 u^2(d_i) + \left(\frac{\partial \varepsilon}{\partial d_{0,i}}\right)^2 u^2(d_{0,i}) + 2 \frac{\partial \varepsilon}{\partial d_i} \frac{\partial \varepsilon}{\partial d_{0,i}} \text{cov}(d_i, d_{0,i}). \quad (11)$$

Under the working assumption that d_i and $d_{0,i}$ are independent,

$$u(\varepsilon_{hkl,i}) = \sqrt{\frac{u^2(d_i)}{d_{0,i}^2} + \frac{d_i^2}{d_{0,i}^4} u^2(d_{0,i})} \quad (12)$$

and the uncertainty of the estimated tensor components follows from the covariance of $\boldsymbol{\varepsilon}_j$. With uncorrelated measurements,

$$\text{cov}(\boldsymbol{\varepsilon}) = [\mathbf{f}_j(\phi_i, \psi_i)^T \mathbf{W} \mathbf{f}_j(\phi_i, \psi_i)]^{-1}. \quad (13)$$

Thus, the component standard uncertainties are the square roots of the diagonal elements,

$$u(\varepsilon_{ij}) = \sqrt{[\text{cov}(\boldsymbol{\varepsilon})]_{kk}}. \quad (14)$$

3. Results

3.1. Correlation of crystallite orientation and signal

In Fig. 4(a), the measured signals for the 311 reflection are reported alongside predictions of the strength of the (311) crystallographic orientation. Each figure is based on interpolation of the data within the breadth of the sampled orientation space (ϕ, ψ) onto an equal-area projection. Variations in strength of preferred orientation, reported as multiples of random distribution (MRD), are weak for the (311) plane, which is why it is typically recommended for strain measurement (Hutchings *et al.*, 2005). Peak area (*i.e.* signal) is well correlated with crystallite orientation and each

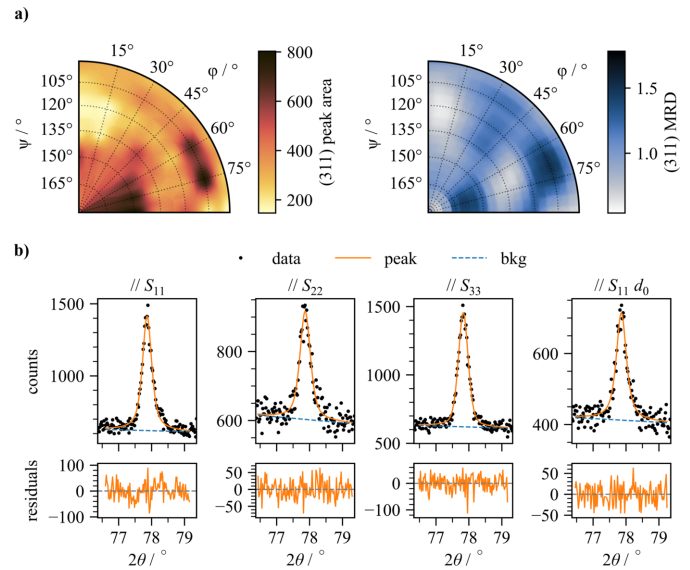


Figure 4

(a) The measured signal (left), computed as peak area, correlates well with the predicted crystallographic orientation of the 311 reflection (right), based on the ODF reconstruction (data are shown for the top volume, $S_{33} = -2.8 \text{ mm}$). (b) Raw data and pseudo-Voigt peak fits from the top volume, shown along three orthogonal directions, including data for d_0 .

of these metrics is inversely correlated with the fitted peak uncertainty. Representative peak fits along the three sample axes are shown in Fig. 4(b). The 311 peak height-to-background ratio varies modestly across the measured orientation space, with values spanning roughly 0.44–2.94 (median $\simeq 1.0$). A conservative estimate was made for the number of grains within each volume. If the grain size is assumed to be $20 \mu\text{m}$ [based on a similar alloy and deposition process (Metz *et al.*, 2024)] within the $3 \times 3 \times 3 \text{ mm}$ measurement volumes, the total number of grains within each volume is of the order of 3.4×10^6 . Due to fine-scale gradients in texture, only some fraction of these grains participate in diffraction (Metz *et al.*, 2024).

3.2. Variation of the reference interplanar spacing

Fig. 5 shows all measured (311) interplanar spacings. The total variation in interplanar spacing is approximately an order of magnitude larger than that of the reference values ($10^{-3} \mu\text{e}$ versus $10^{-4} \mu\text{e}$ equivalent). Accordingly, an average reference value of 1.21656 \AA was adopted for all strain calculations. At the extreme, this assumption could incur an error of up to $164 \mu\text{e}$, corresponding to 10 MPa in annealed AA6061.

Using this fixed reference, most of the calculated lattice strains are tensile, consistent with prior neutron diffraction measurements in similar regions of friction-stir deposited AA6061 (Zhu *et al.*, 2023). In that study, the total variation in d_0 was equivalent to $700 \mu\text{e}$; however, that component was approximately twice the size of the present build and experienced roughly six times more depositions. Another investigation of friction-stir deposited AA7075 reported a total

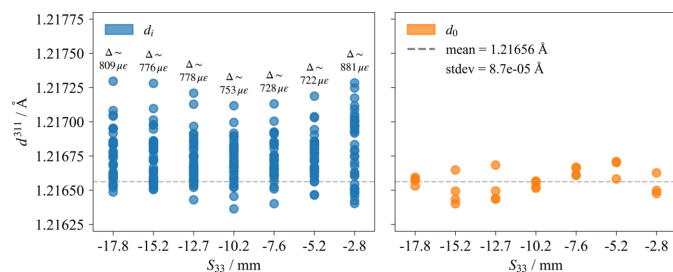


Figure 5 Scatterplots of (left) all measured sample lattice parameters and (right) reference lattice parameters. The total variation in sample lattice parameters, denoted above each location in microstrain equivalent, is an order of magnitude greater than the variation in the reference parameters. The average reference lattice parameter used in all calculations is shown with a dashed line.

variation in d_0 equivalent to 300 $\mu\epsilon$ for a build of comparable geometry and number of depositions to this study (Metz *et al.*, 2025). The total possible variation in the (311) interplanar spacing, based on matrix solute concentrations (whether from solute supersaturation or precipitation), may vary up to 1400 $\mu\epsilon$ equivalent in AA6061 (see supporting information, Fig. S3). However, that is not the case here.

3.3. Verification of physical bias

When calculating the strain tensor via the strain transformation law, it is assumed that all measurements represent the same strain state. When that is not true, bias is incurred within each measurement depending on its deviation from the true volumetric average of residual strain within the gauge volumes. The existence of physical biases within the strain measurements, primarily due to centroid misalignment and unrepresentative volume averaging, was investigated. From the available data, the relative influence of each type of bias is indistinguishable. However, the combined influence is evident by examining the behavior of a $\sin^2\psi$ plot. Such a plot examines the relationship between diffraction geometry ($\sin^2\psi$) and the measured strains ($\epsilon_{33})_{\phi\psi}$. For a single representative stress state, the $\sin^2\psi$ plots exhibit either linear (biaxial stress state) or symmetric curvilinear behavior (triaxial stress state). The latter is called psi-splitting (Noyan & Cohen, 1987). Oscillatory behavior implies that strong strain gradients are sampled, violating the RVE assumption and potentially compromising the accuracy of the data when checked against the strain transformation law. In addition to $\sin^2\psi$ plots, the influence of strongly heterogeneous intergranular strains was investigated. It was assumed that variations in fitted integral breadth (peak area divided by height) were positively correlated with variations in intergranular strains.

The results of the $\sin^2\psi$ and integral breadth tests are shown in Fig. 6. Each test revealed strongly oscillatory behavior, indicating strong orientation dependence of the measured strain state. This verifies that some combined influence of centroid misalignment and unrepresentative volume averaging produces strong biases in the measured strains. These

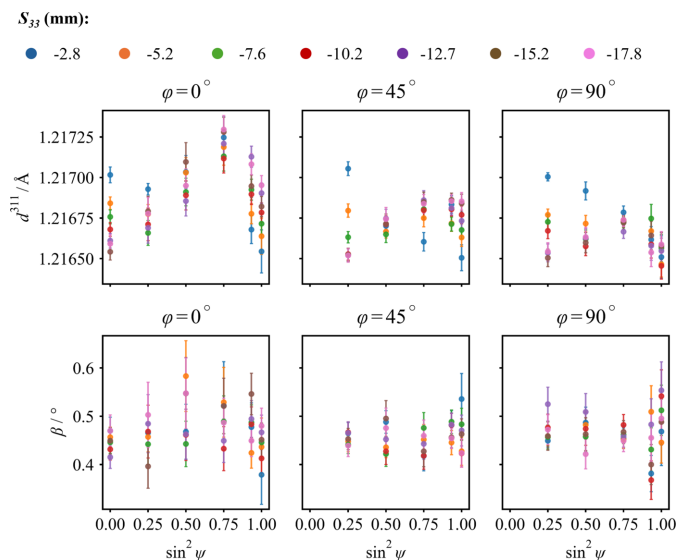


Figure 6 The top row shows three $\sin^2\psi$ tests with oscillatory behavior, suggesting that the measurements do not consistently sample a single representative strain state. The bottom row shows a strong influence of the measured orientation on the integral breadth, which is assumed to be correlated with intergranular strain variations. Both tests suggest that significant biases exist within the strain measurements, which, if left unaccounted for, will strongly influence the calculated stresses and lead to underestimated uncertainties.

variations in integral breadth and the variations in crystallite orientation previously described both influence peak shape and together explain the variability in fitted peak uncertainties (± 50 to 100 $\mu\epsilon$).

3.4. Strain and stress tensor calculations

3.4.1. Method 1 – three measurements

Fig. 7 shows the results from substituting three orthogonally measured strains directly into Hooke’s law. The average uncertainty in the strain is $\pm 100 \mu\epsilon$, approximately equal to the reported detectability of the instrument (Bunn *et al.*, 2023). The average uncertainty in the stress is 20 MPa.

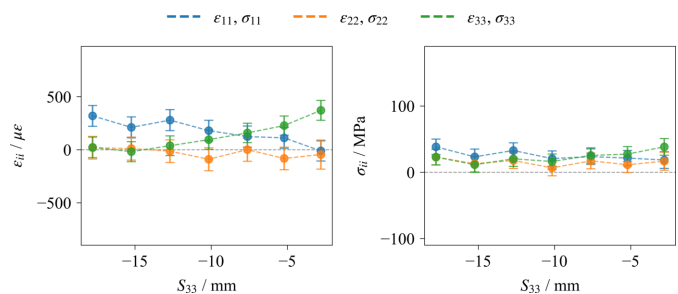


Figure 7 Strain and stress tensor results from the three-component analysis plotted against depth. The calculated strains and stresses are predominantly tensile.

3.4.2. Method 2 – six measurements

Fig. 8 shows continuous distributions of all calculable sample strain/stress tensor components, based on combinations of six orientations from 36 available measurements. The strains/stresses can be either tensile or compressive depending on the choice of orientations, spanning a range greater than five times the 62 MPa yield strength of annealed or friction-stir deposited AA6061 (Anderson *et al.*, 2019; Rutherford *et al.*, 2020). Averaging the standard deviation of each distribution across the depths shows that the shear-strain components (ϵ_{12} , ϵ_{13} , ϵ_{23}) exhibit markedly larger dispersion than the normal-strain components (ϵ_{11} , ϵ_{22} , ϵ_{33}): 293 $\mu\epsilon$ versus 148 $\mu\epsilon$ (1.98 \times). In contrast, the shear-stress components (σ_{12} , σ_{13} , σ_{23}) are much less dispersed than the normal-stress components (σ_{11} , σ_{22} , σ_{33}): 26 MPa versus 103 MPa (0.26 \times).

Two effects may explain these trends. Firstly, when solving the six-component inversion, the measurement operator that maps rotated lattice-spacing observations to sample-frame strains [*i.e.* $\mathbf{f}_j(\phi_i, \psi_i)^{-1}$] has uneven leverage across components. Many six-angle subsets provide strong constraints on the normal terms but weak constraints on the shear terms. In the strain transformation law, nearly coplanar or symmetrically clustered orientations make the corresponding columns of the design matrix $\mathbf{f}_j(\phi_i, \psi_i)$ partially collinear, amplifying noise and yielding broader KDEs for ϵ_{shear} . Second, the isotropic stress equation $\boldsymbol{\sigma} = \lambda \text{tr}(\boldsymbol{\epsilon})\mathbf{I} + 2\mu\boldsymbol{\epsilon}$ (with λ and μ being Lamé's first and second parameters) mixes the volumetric strain into all three normal stresses, so uncertainty

in $\text{tr}(\boldsymbol{\epsilon})$ is broadcast to (σ_{11} , σ_{22} , σ_{33}) with weight λ , whereas shear stresses depend only on the shear strains via the factor 2μ . Because λ is of the same order as 2μ for aluminium, variance in $\text{tr}(\boldsymbol{\epsilon})$ (and in how ϵ_{11} , ϵ_{22} , ϵ_{33} partition that trace) inflates σ_{normal} relative to σ_{shear} , even though ϵ_{shear} itself is more weakly constrained. Together, these conditioning and propagation effects yield the observed pattern: a large spread in ϵ_{shear} , a modest spread in σ_{shear} (set mainly by 2μ) and the largest spread in σ_{norm} due to the added volumetric coupling.

These complex uncertainty interactions make this data set an ideal test case. If the underlying peak-fit uncertainties are accurate and representative, uncertainty propagation through the equations will correctly capture the uncertainty in calculated quantities such as stress. The average solution for each of the distributions is plotted in Fig. 9. The uncertainties are calculated as the average uncertainties of the ten most common solutions and are approximately an order of magnitude larger than those obtained from the three-measurement method, primarily due to noise amplification arising from the geometric conditioning of the design matrix for non-ideal sets of orientations.

In most six-measurement combinations, the selected orientations do not align with the sample axes; thus, the transformation required to recover the strains along these axes amplifies the uncertainties. To be clear, the uncertainties shown in Fig. 9 represent the uncertainty of a single typical six-measurement solution constructed from an arbitrary set of orientations. They do not represent a fundamental lower bound on the uncertainty of the strain tensor, nor do they imply that six measurements inherently produce larger uncertainty than three measurements. Rather, they reflect the fact that most randomly selected orientation sets are not optimally conditioned for resolving the tensor components along the sample axes. This distinction is important. If the three orthogonal (sample-axis aligned) strain measurements

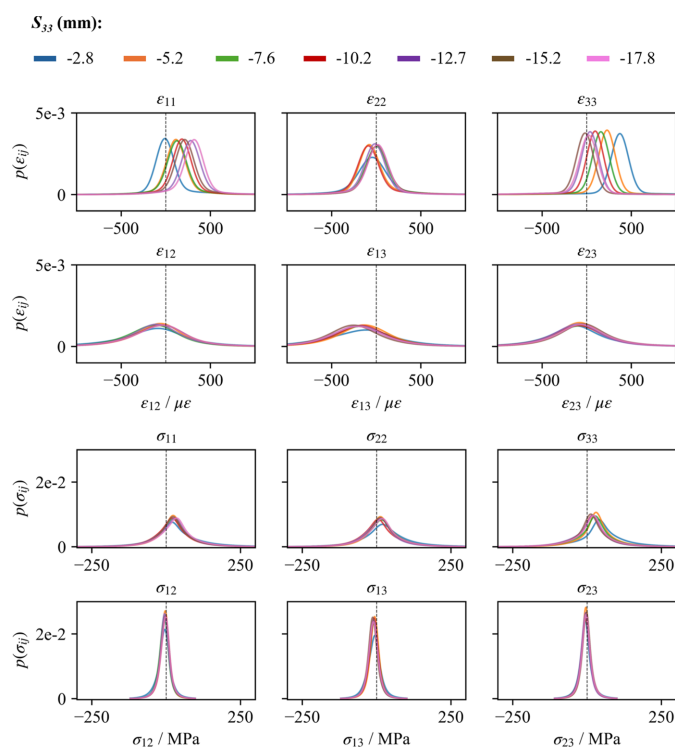


Figure 8
Continuous distributions (KDEs) of all calculable six-component solutions for strain (top two rows) and stress (bottom two rows) are extremely broad, indicating that taking a random combination of six orientations might not even yield the correct sign for strain or stress.

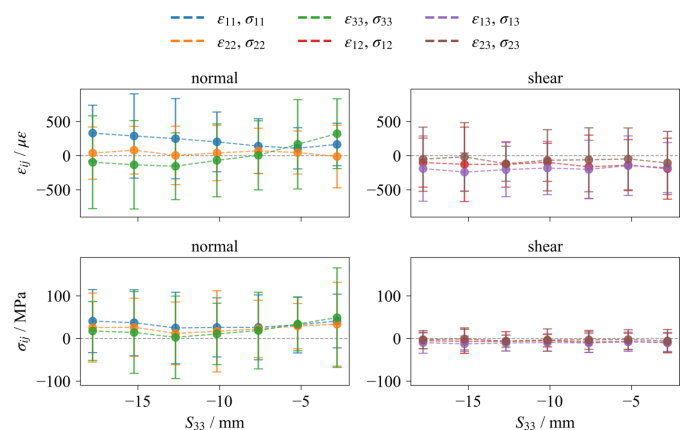


Figure 9
Results of the average solution for the strain and stress tensors (averages of the KDEs) given a choice of six orientations of strain measurements. The uncertainties are calculated as the average uncertainties of the ten most common solutions. Uncertainties for the six-measurement method are an order of magnitude larger than those for the three-measurement method, due to conditioning of the inversion process when obtaining sample strains.

are included and supplemented with three additional orientations, the uncertainty in the normal strain components remains the same as for the three-measurement method. In this case, the additional measurements do not degrade the uncertainty because the primary components are already directly measured and the system is well conditioned. However, such a set of ideal orientations represents only a small subset of all possible six-measurement combinations. For most orientation sets, the strain tensor must be reconstructed through a transformation from non-axis-aligned measurements, and the associated conditioning of the design matrix leads to amplification of uncertainty in the recovered components. Thus, the larger uncertainties observed in Fig. 9 are primarily a consequence of measurement geometry rather than an inherent limitation of the six-measurement approach. The effects of measurement selection, geometric conditioning and the appropriateness of the resulting uncertainty magnitudes are further investigated in the discussion below.

3.4.3. Method 3 – thirty-six measurements

The least-squares procedure was used to incorporate all 36 strain measurements into a single solution [equation (10)], the results of which are shown in Fig. 10. As might be expected, the stress and strain values are very similar to those plotted in Fig. 9, which are an average of the most common solutions for the six-measurement analysis. The propagated uncertainties are lower than those of the three- and six-measurement analyses by factors of 2× and 5×, respectively. Also, the point-to-point scatter in the curves is drastically reduced compared with the three-measurement analysis. This is the intended effect of incorporating many imprecise measurements, which decreases solution variance yet potentially increases bias, because the underlying variability in the strain continuum is not represented within the uncertainty intervals. The $\sin^2\psi$ test clearly showed that the sampled strain states vary more strongly than these solutions suggest. This motivates the need to inflate the measurement uncertainties.

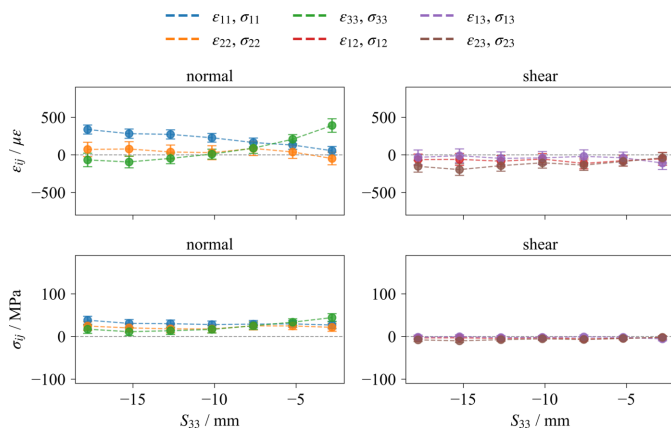


Figure 10 Results of the least-squares analysis that includes all 36 measured strains. These profiles are smoother and with smaller apparent uncertainties than the preceding analyses. This probably indicates that strain/stress now more closely resemble continuum quantities at the macro-scale.

4. Discussion

Given the presence of bias in the measured strains, this discussion examines how uncertainty propagates to stress for both the six-measurement direct solution and the 36-measurement least-squares solution. The key distinction between these approaches lies in how their propagated uncertainties should be interpreted. It will be shown that the six-measurement method propagates uncertainties that retain the meaning of a one-standard-deviation measure of random noise. In contrast, the least-squares method produces a lower-variance estimate of the strain and stress state, but the propagated uncertainties no longer reflect the variability present in the original measurements. As a result, the interpretation of uncertainty differs between the two methods, particularly with respect to their relationship to the unknown ground truth for this dataset.

4.1. Uncertainty analysis of direct solutions

First, the six-measurement procedure was investigated to assess whether propagated one-standard-deviation uncertainties reflect the empirical spread of solutions. For this purpose, normalized deviations are formed for each tensor component ij and solution i ,

$$D_{ij} = \frac{\varepsilon_{ij}^{(i)} - \bar{\varepsilon}_{ij}}{\mu_{\varepsilon_{ij}^{(i)}}}, \quad (15)$$

where $\varepsilon_{ij}^{(i)}$ is the individual strain value, $\bar{\varepsilon}_{ij}$ is the mean of all strains and $\mu_{\varepsilon_{ij}^{(i)}}$ is the uncertainty for that value. The same format is used for the stress deviation. This metric quantifies how far each solution lies from the mean relative to its own uncertainty. Values with a magnitude less than one indicate agreement within one standard deviation, regardless of the absolute difference. This deviation metric is specifically designed to test whether the propagated uncertainties capture the observed spread of solutions. It does not penalize solutions with large uncertainties, provided those uncertainties are sufficient to explain the deviation from the mean. Importantly, this formulation does not assume that the mean represents the true volumetric strain state. Rather, the analysis is performed in a relative sense, which is appropriate because the strain transformation law constrains only relative values.

The means and standard deviations of the normalized deviations are shown in Fig. 11. If a standard deviation is greater than 1, the values are dispersed more broadly than predicted by the individual uncertainties, indicating that those uncertainties are too small on average. The mean is harder to interpret. Since ground truth is not known for the strain state, a non-zero mean does not necessarily indicate a bias in the solutions. It only indicates that the deviations are systematically shifted to one side of the mean. Nonetheless, some interesting patterns appear in the means and are thus discussed.

Apart from the shallowest location, the uncertainty estimate is mildly conservative. Both the strain and stress components have standard deviations below 1, indicating

propagated uncertainties that slightly exceed the spread of values. This result gives confidence in the peak-fit uncertainties, a result not necessarily expected. At the shallowest depth, the standard deviation is predominantly greater than zero, with typical values around 1.2, which is investigated further shortly.

The normal strain components show consistently non-zero means – negative for ε_{11} and ε_{22} and positive for ε_{33} . These non-zero means are suppressed when mapped to stress. This sign structure may be compatible with a small systematic effect in the normal strains (*e.g.* the common d_0 or angular coverage) to which shear is insensitive and which is diluted by the isotropic Hooke relation through the $\lambda \text{tr}(\varepsilon)$ term. Overall, the stress deviations have means closer to zero than the strain deviations, suggesting that interpretation and reporting at the stress level are more robust to systematic effects from the measured strains, and carrying a shared d_0 may suppress some systematic effects.

To diagnose the physical and experimental drivers behind departures from the ideal normal distribution $D \simeq \mathcal{N}(0, 1)$, particularly at the shallowest depth, each solved σ_{ij} component is split into inliers $|D| \leq 1$, where the deviation is close to 1, and outliers $|D| > 1$, where the deviation is greater than 1. The following metrics are compared: (i) peak area (counting statistics), (ii) integral breadth (IB, fit stability/line broadening), (iii) a texture proxy [MRD mapped to (ϕ, ψ) , only for S_{33} : -2.8 and -17.8 mm] and (iv) a geometry metric that summarizes angular coverage of the six directions used in each inversion [a proxy for conditioning of the design matrix $\mathbf{f}_j(\phi_i, \psi_i)$]. Cellular maps, similar to Fig. 11, are given in Figs. S4 and S5 for each statistic across inliers and outliers.

It is assumed that maximizing the angular coverage of measurements leads to better conditioning of the design

matrix. Thus, a geometry metric of primary interest is the minimum spanning tree (MST) angle sum for the six unit directions $\{\hat{n}_k\}_{k=1}^6$,

$$S_{\text{MST}} = \sum_{(i,j) \in \text{MST}} \arccos(|\hat{n}_i \cdot \hat{n}_j|), \quad (16)$$

i.e. the total great-circle separation of the five edges that connect all six directions with minimum total length. This favors globally spread sets (better conditioning of the design matrix) while down-weighting clustered angles.

To test whether outliers differ systematically from inliers, we use an exact two-sided sign test on paired component–depth differences. For a given metric M and for each component–depth cell (c, d) (six components per depth), the paired difference is defined as

$$\Delta_{c,d} = \overline{M}_{\text{outlier}}(c, d) - \overline{M}_{\text{inlier}}(c, d). \quad (17)$$

Let k be the number of positive $\Delta_{c,d}$ among n non-tied pairs and assume the null of no systematic difference,

$$H_0 : \mathbb{P}(\Delta > 0) = \mathbb{P}(\Delta < 0) = \frac{1}{2}, \quad (18)$$

under which $k \simeq \text{Binomial}(n, 0.5)$. Here, H_0 is the null hypothesis of no systematic difference between outliers and inliers. The two-sided p value is

$$p = 2 \min\{\mathbb{P}(K \leq k), \mathbb{P}(K \geq k)\} \\ = 2 \min\left\{\sum_{r=0}^k \binom{n}{r} 2^{-n}, \sum_{r=k}^n \binom{n}{r} 2^{-n}\right\}. \quad (19)$$

This non-parametric test makes no normality assumption. It requires only approximate independence of the pairs across component–depth cells and symmetry of signs under H_0 . With small n (here $n = 6$ within a depth and $n = 12$ when pooling two depths, such as the top-most and bottom-most), p values are discrete (*e.g.* the smallest two-sided p at $n = 6$ is 0.03125).

Table 2 pools both depths (12 pairs). Outliers have a slightly larger S_{MST} on average that is statistically significant (1.2%, $p = 0.0386$). Peak area, IB, d -spacing uncertainty and MRD all show very small average differences yet are not statistically significant. Table 3 resolves the comparison by depth: at the shallow location, outliers are weakly associated with lower area (-2.26% , $p = 0.031$), higher IB (0.57%, $p = 0.031$) and higher d -spacing uncertainty (2.58%, $p = 0.031$). The MRD differences are uniformly small. Overall, there are some statistically significant differences between inliers and outliers and between the top-most and bottom-most sampling locations, yet the actual magnitude of those differences is quite small ($< 3\%$). The small differences between the top-most and bottom-most locations are probably due to differences in thermal cycling. That is, the top-most layer has not undergone the additional recovery and recrystallization of the prior layers and thus retains a greater degree of accumulated plastic strain (*i.e.* integral breadth) and deformation texture strength (*i.e.* MRD).

Alongside the deviation analysis, these findings suggest that variability in the calculated strain/stress tensors is not due to

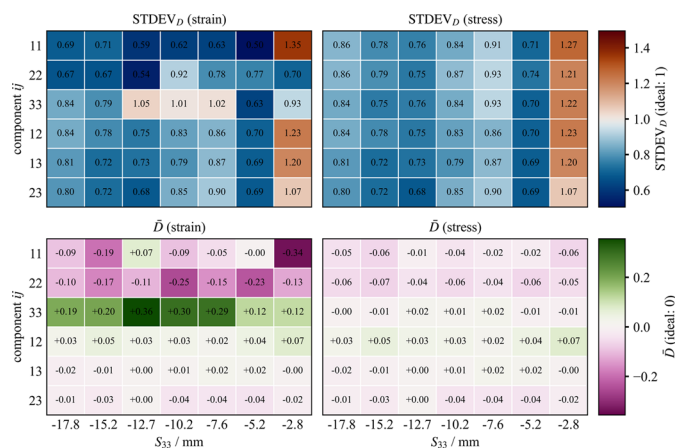


Figure 11 Statistical measures of systematic effects and the appropriateness of the propagated uncertainties, based on the mean and standard deviation of the normalized deviations to the six-measurement solutions [equation (15)]. Other than for the shallowest depth, the generally low values for the standard deviations indicate that the propagated peak-fit uncertainties give somewhat conservative estimates of the actual variance. Because the ground truths for the measured strains and stresses are unknown, it is difficult to assign significance to the deviation means significantly greater than zero.

Table 2
Cross-depth summary (both depths pooled; $n = 12$ pairs).

Metric	Mean Δ (outlier – inlier)	Relative Δ (%)	Two-sided p	Comments
Peak area (a.u.)	-3.48	-1.0%	0.774	Small difference
Integral breadth (IB)	0.0007	0.16%	1.000	No consistent difference
d -spacing uncertainty ($\mu\epsilon$)	0.50	1.0%	1.000	No consistent difference
Texture (MRD)	0.002	0.17%	0.774	Small difference
S_{MST} ($^\circ$)	1.58	1.2%	0.0386	Small cross-depth effect with consistent direction

Table 3
Depth-resolved summary (component-paired within depth; $n = 6$ pairs per depth).

Metric	$S_{33} = -2.8$ mm Δ (relative %)	p	$S_{33} = -17.8$ mm Δ (relative %)	p
Peak area (a.u.)	-2.26%	0.031	0.53%	0.219
Integral breadth (IB)	0.57%	0.031	-0.26%	0.031
d -spacing uncertainty ($\mu\epsilon$)	2.58%	0.031	-0.65%	0.031
Texture (MRD)	-0.33%	0.219	0.73%	0.031
S_{MST} ($^\circ$)	1.43 $^\circ$	0.219	1.73 $^\circ$	0.219

Notes: Δ denotes outlier minus inlier means within each component–depth cell before pairing; relative changes are expressed as percentages of the inlier mean. Exact ties (none here) would be dropped from n for that comparison.

any single variable. Rather, the variability is due to the sampling of discretely different strain states, characterized by averages of select grain populations. This demonstrates that in samples with complex and fine-scale variations in microstructure and strain the calculated solutions and uncertainties do not reflect the true volumetric average, even with infinitely accurate measurements. For this experimental setup, any one solution may be either tensile or compressive depending on the choice of measurements used (at least for the six-measurement method).

Building on the deviation analysis above, we next quantify how the choice of six orientations governs uncertainty amplification in the direct inversion of the design matrix $\mathbf{f}_j(\phi_i, \psi_i)$ (which we now write as \mathbf{F} for brevity). Uncertainties in the measured strains are propagated to the sample strains through the information matrix,

$$\text{cov}(\epsilon_j) = (\mathbf{F}^T \mathbf{W} \mathbf{F})^{-1}, \tag{20}$$

where $\mathbf{W} = \text{diag}[1/\text{var}(e_i)]$. The diagonal entries $[(\mathbf{F}^T \mathbf{W} \mathbf{F})^{-1}]_{jj}$ are per-component variance multipliers. Larger values indi-

cate that the six chosen orientations tend to amplify noise into component j more strongly (poorer geometric leverage on that component), while smaller values indicate better leverage.

To visualize how each orientation participates in conditioning of the information matrix $[(\mathbf{F}^T \mathbf{W} \mathbf{F})^{-1}]_{jj}$, contour plots were constructed as follows. For each orientation (ϕ, ψ) , we collect all solutions that contain that measurement and compute for each component j the median of $[(\mathbf{F}^T \mathbf{W} \mathbf{F})^{-1}]_{jj}$. Lower values mark orientations that tend to participate in well constrained inversions, while higher values mark orientations that tend to participate in poorly constrained inversions. The results of this analysis for each location yielded similar results. Thus, only the results of the shallowest location are visualized in Fig. 12.

This analysis shows how the observability of a particular strain/stress component along a particular measurement orientation is the dominant constraint on that component’s solution. No particular variability in measurement uncertainty has any notable influence on noise amplification when inverting the information matrix. Rather, the choice of orientations strongly constrains the solution. The relative consistency of $[(\mathbf{F}^T \mathbf{W} \mathbf{F})^{-1}]_{jj}$ for the shear components is a primary reason for the large variance of the solutions. It appears that optimal geometric conditioning of orientations is achieved by (i) covering the three sample axes S_{11} , S_{22} and S_{33} , and (ii) choosing three interior angles with adequate spacing and a lack of redundant information due to symmetry. These findings reinforce intuitions regarding data collection.

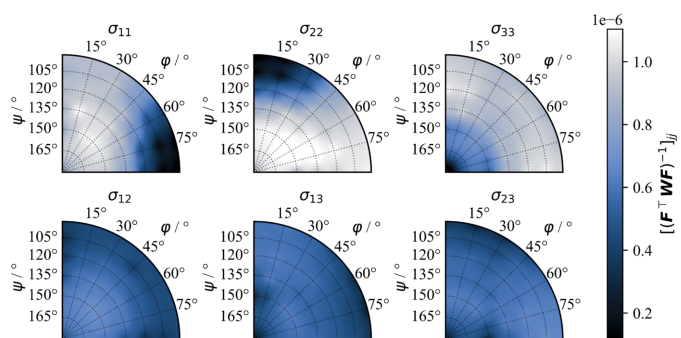


Figure 12
Per component conditioning metric $[(\mathbf{F}^T \mathbf{W} \mathbf{F})^{-1}]_{jj}$, colored according to its median value across all solutions for a given orientation. These plots demonstrate that optimizing the observability of each component is a necessary conditioning metric, for which there are no preferred orientations for the shear components.

4.2. Uncertainty analysis of least-squares method

Incorporating 36 measurements into a least-squares procedure reduces the variance of solutions. One important caveat to acknowledge is that the propagated uncertainties of the least-squares estimator only describe the confidence of the fit. This is opposed to the confidence of the solution with respect to the ground truth, which cannot be obtained by a single least-squares estimation. As more measurements are

included, the propagated uncertainty will converge to a small value relative to the individual measurement uncertainties. This phenomenon of uncertainty convergence is demonstrated in Fig. 13 and Table S4. For example, the spread of nominal solutions and the average uncertainty for ε_{11} fall below $\pm 100 \mu\text{e}$ once about 27 measurements are used. After 36 measurements are used, the uncertainty falls below $\pm 50 \mu\text{e}$.

As additional measurements are included in the least-squares fit, the propagated uncertainty on the fitted tensor decreases, as expected for an averaging estimator. However, the resulting uncertainty is only accurate if the per-measurement uncertainties fully describe the variability of the measurements about a single strain tensor. We test that assumption by examining the standardized residuals. These residuals are defined as

$$R_w = (e_{i,\text{obs}} - e_{i,\text{pred}}) \sqrt{w_i} = \frac{e_{i,\text{obs}} - e_{i,\text{pred}}}{\mu(e_i)}, \quad (21)$$

where $e_{i,\text{obs}}$ is the measured strain in the diffraction coordinate system, $e_{i,\text{pred}}$ is the corresponding strain predicted by inserting the least-squares sample tensor back into the strain transformation law at the measured (ϕ, ψ) , and $\mu(e_i)$ is the one-standard-deviation uncertainty of the strains based on the peak-fit uncertainty. If the 36 measurements are consistent with a single strain tensor and the uncertainties are complete, then the residuals should have zero mean and unit variance. In that case, the goodness-of-fit statistic

$$\chi^2 = \sum_{i=1}^N R_{w,i}^2, \quad (22)$$

$$\chi^2_v = \frac{\chi^2}{N - p} \quad (23)$$

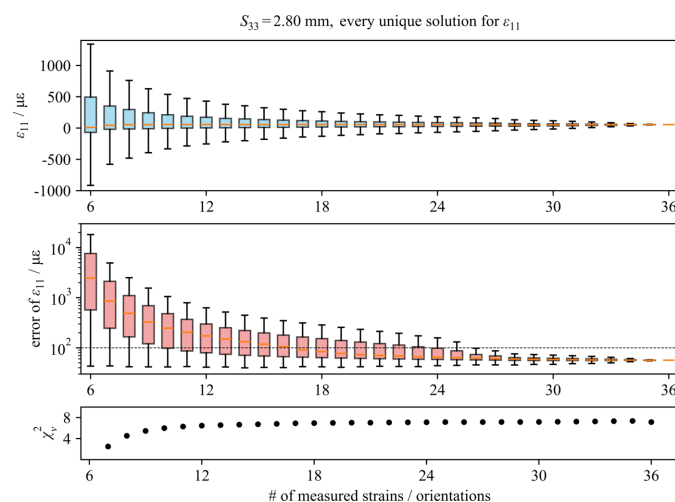


Figure 13 Box plots of all calculable solutions and uncertainties for ε_{11} , computed from the least-squares method using orientations within the top volume. The median reduced χ^2_v is shown on the bottom-most plot, demonstrating that, as more measurements are included, the assumption of a single sampled strain state breaks down. The box plots are constructed from the medians (orange) and the 25th and 75th percentiles of data.

should be consistent with a χ^2 distribution with $N - p = 30$ degrees of freedom.

Fig. 14 shows the distributions of standardized residuals for each sampling location. Clearly, the spread of residuals is much broader than what would be predicted if the measurements were representative of the same strain state. Formalized χ^2 and t tests were conducted at each location to assess the significance of high variance and systematic effects. The reduced χ^2 values range from 2.34 to 7.15, far above the expected value of 1. See Table S4 for statistics at each depth. The t tests do not reject a zero mean of residuals, indicating that there are probably no significant systematic effects relative to the best-fit tensor. Overall, the residuals are centered near zero, but their spread is substantially broader than expected from the measurement uncertainties. Thus, while the least-squares tensor sits near the center of the data, the peak-fit uncertainties are too small to explain the observed residual scatter about a single tensor. In standard statistical interpretation, such elevated values of reduced χ^2 would suggest either that the uncertainties are incomplete or that the underlying model does not describe the data.

We contend that the measurement uncertainties should be inflated until the reduced χ^2 equals 1. This is justified on the basis of three assumptions, which are not always valid: (i) that the measurement errors (deviations from truth) behave randomly; (ii) that the model is linear; and (iii) that the model is correct and represents the underlying data (Andrae *et al.*, 2010). Often, these assumptions are not valid, as has been noted in the field of astronomy, where uncertainty estimation is paramount (Andrae, 2010; Andrae *et al.*, 2010). In addition, it is common for the model to be incorrect. However, our use case satisfies all three assumptions. First, the deviation analysis in Tables 2 and 3 shows no significant correlation between measurement variables and a solution's deviation from the average solution. Second, the strain-transformation equation is indeed a linear transformation of strains and trigonometric

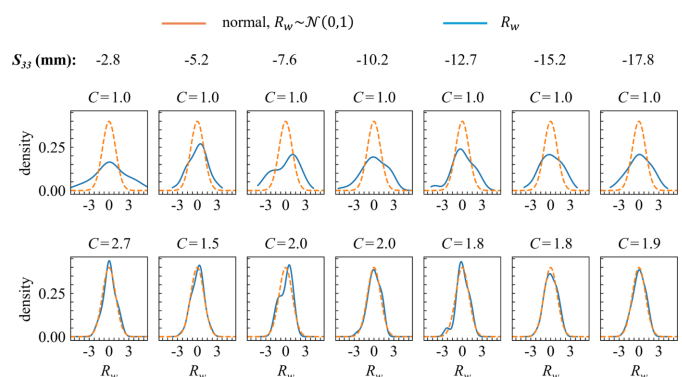


Figure 14 Distributions of residuals expressed as continuous functions (KDEs). If the measurements represent only random noise about a singular strain state, the weighted residuals should model a normal distribution, $R_w \approx \mathcal{N}(0, 1)$. The top row shows that the measured uncertainties yield weighted residuals inconsistent with random noise, whereas the bottom row shows that inflating the uncertainties produces residuals consistent with random noise. This suggests that inflating the measured uncertainties is required to represent the experimental scatter.

coefficients. When the first two assumptions are satisfied, the χ^2 statistics often reveal the inadequacy of simplified physical models via under- or over-fitting of the data (*i.e.* $\chi_v^2 \gg 1$ or $\chi^2 < 1$, respectively). In our case, the strain transformation equations are a law, not a model, and they must be satisfied by any true macroscopic strain state. While each measurement does sample different populations of crystallites, giving rise to the observed deviations, it is justified to inflate uncertainties via the reduced χ^2 test because the results will be interpreted as macroscopic continuum strains (for structural integrity applications). Simply put, the uncertainties should be large enough to capture variations relative to the average strain state. In that context, measurement uncertainties may be inflated by a common factor until the standardized residuals pass the reduced χ^2 test. This means the inflated uncertainties propagated through the least-squares fit are more justified and conservative.

This interpretation is also supported by three prior observations. First, oscillatory behavior of the $\sin^2\psi$ plots indicates heterogeneity in the sampled strain states. Second, the non-zero and solution-dependent deviation means $\mu_{D_{ij}^e}$ in the six-measurement analysis suggest an orientation-dependent scatter. Third, the suppression of non-zero means from the strain deviations $\mu_{D_{ij}^e}$ to zero means for the stress deviations $\mu_{D_{ij}^s}$ implies a common systematic effect (*e.g.* d_0).

Accordingly, a more complete uncertainty description recognizes that each solution to the strain tensor is offset from the true solution by a common bias b_{ij} , an orientation-dependent bias $\delta_{ij}^{(k)}$ and random noise,

$$\varepsilon_{ij}^{(k)} = \varepsilon_{ij}^{\text{true}} + b_{ij} + \delta_{ij}^{(k)} + \text{noise}^{(k)}. \quad (24)$$

If these additional contributions are ignored, the uncertainty propagated into the least-squares estimator no longer retains its interpretation as a one-standard-deviation measure of variability (*i.e.* noise). Thus, the repeatability of the experiment is not ensured.

The common bias b_{ij} is incalculable when the true strain is not known. However, the relative magnitude of the orientation-dependent biases may be estimated by introducing an additional scaling factor C to the measurement uncertainties,

$$\mu_{e_i, \text{eff}} = C\mu_{e_i}, \quad (25)$$

where C is chosen such that the reduced χ^2 satisfies $\chi_v^2 = 1$ (see Table S4). Depending on location, the required scale factor ranges from $C = 1.53$ to $C = 2.67$. The spread indicates a location- (microstructure-)dependent effect. More specifically, the intensity of small-scale microstructural gradients is greatest for the top-most volume $S_{33} = -2.8$ mm. This top-most volume has undergone less thermal cycling and recrystallization than the locations below, a difference noted previously for the results of Table 3. This scaling does not alter the nominal least-squares solution. The distributions of standardized residuals, before and after uncertainty inflation, are shown alongside each other in Fig. 14. The results of the least-squares estimation after uncertainty inflation are given in Fig. 15. The uncertainty intervals are similar to the three-

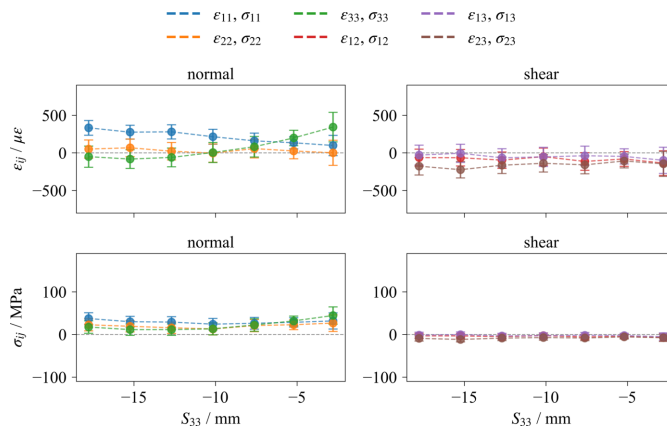


Figure 15 Results of the strain/stress solutions and their errors after inflating measurement uncertainties to account for hidden biases. The uncertainties are two to three times larger than those of the un-inflated least-squares estimator, yet they are similar to the three-measurement solutions. This implies that the three-measurement uncertainties are too small since the measurement biases could not be evaluated. In addition, the uncertainty appears to scale with the degree of orientation-dependent scatter. This is evident at the top-most location ($z = -2.8$ mm), where the deviations and residuals exhibited a greater spread in both the six- and 36-measurement analyses.

measurement results, which suggests that the uncertainties for the latter – which is the common method – are underestimated.

4.3. Suggestions for ideal count times

A key implication of this work is that an ideal diffraction measurement is defined not solely by precision but by whether its reported uncertainty meaningfully represents the variability of the underlying strain state. In conventional practice, a minimal set of orientations are measured and count times are increased to reduce peak-fit uncertainty. As shown here, this approach can produce uncertainties that are artificially small and unrepresentative of the true experimental scatter when measurements sample different strain states.

Instead, a more appropriate strategy is to measure redundant orientations and evaluate consistency through the strain transformation law. In this framework, least-squares estimation reduces the potential systematic effects of orientation-dependent variability, while statistical tests such as the reduced χ^2 provide a means to assess whether the propagated uncertainties remain valid. When they do not, the uncertainties can be adjusted to reflect the effective variability of the measurements. This approach reframes uncertainty as a quantity that must be validated against physical consistency, rather than minimized through counting statistics alone.

However, such strategies introduce practical constraints, as measurement time is limited and cannot be arbitrarily increased across many orientations. This motivates the need for optimized measurement strategies that balance orientation coverage, count time and the reliability of the resulting uncertainty estimates. Thus, the following strategy is proposed.

For the sake of argument, it is assumed that peak-fit uncertainty is proportional to count time based on counting statistics of random noise,

$$\mu \propto \frac{1}{\sqrt{N}}. \quad (26)$$

Here, N is the total number of counts, which increases linearly with time. Therefore, uncertainty should decrease with time according to

$$\mu = \mu_0 \sqrt{\frac{t_0}{t}}, \quad (27)$$

where t_0 is the original count time with associated uncertainty μ_0 , and t is the new count time. If uncertainties μ should decrease by a certain factor f from their original value μ_0 , then the new time is given by

$$t_{\text{new}} = \frac{t_0}{f^2}, \quad (28)$$

based on substituting the relation $\mu(t_{\text{new}}) = f\mu_0$ into equation (27). Since the uncertainties need to be inflated by at least a factor of 1.5, the time for each measurement could be reduced from 10 to 4.4 min (assuming the same measurements are taken).

It is possible to define an optimization scheme where one measures a large number of strains (*e.g.* 36) at various angles (ϕ_i, ψ_i) within a representative microstructure of the component. According to these measurements, an objective is set based on the desired count time or uncertainty in the best-fit strain tensor. For example, the desired maximum uncertainty in the best-fit strain tensor can be set to some reasonable fraction of the material's yield strength (*e.g.* for annealed AA6061, approximately $100 \mu\epsilon$). A global minimization strategy is performed for a given objective function, such as, 'Given a target maximum uncertainty in the best-fit strain tensor (*e.g.* $\max \mu_{\epsilon_{ij}} \leq 100 \mu\epsilon$), what is the minimum number of orientations and the minimum count time per orientation needed?' On the basis of this objective, an algorithm will set a stop limit during an iterative process whereby the algorithm will loop through the appropriate number of measured orientations (ϕ_i, ψ_i) and associated count times to achieve this metric. Such a process would make full-tensor calculations more confident and economical.

5. Conclusions

This work demonstrates that the interpretation of propagated uncertainty in diffraction-based residual strain analysis depends critically on both the calculation pathway and the underlying assumption of a single representative strain state. In the additively manufactured component studied here, fine-scale gradients in texture, plastic strain and elastic strain lead to orientation-dependent sampling of distinct strain states. As a result, one-standard-deviation peak-fit uncertainties do not consistently represent random noise about a common mean once propagated to the solution.

For the six-measurement direct inversion, propagated uncertainties remain meaningful and generally conservative with respect to the empirical spread of solutions. In contrast, the 36-measurement least-squares approach reduces variance through averaging but produces uncertainty intervals that underestimate the true variability of the measurements. In this case, the reported uncertainties reflect overconfidence in the fit rather than capturing the physical scatter of the data.

These results show that uncertainty underestimation in diffraction is not solely due to instrumental or fitting limitations but is strongly influenced by intrinsic sample-dependent effects. Consequently, agreement with the strain transformation law provides a powerful physical check on the validity of propagated uncertainties. An important implication is that oversampling does not inherently improve the reliability of uncertainty estimates. While least-squares methods stabilize the solution, they may obscure physically meaningful variability when the assumption of a singular strain state is compromised. For structural integrity applications, this distinction is critical: a low-variance estimate is not necessarily a conservative one.

Finally, this work provides a practical pathway for improving uncertainty estimates through oversampling the number of measured orientations and statistical consistency checks. By evaluating residuals and enforcing agreement under the strain transformation law, measurement uncertainties can be adjusted to better reflect the effective variability of the experiment. Reliable residual stress assessment in heterogeneous materials therefore requires not only precise measurements but also explicit validation of the assumptions underlying uncertainty propagation.

We suggest that more accurate strain estimates with more realistic uncertainties may be achieved by oversampling orientations with reduced count time per measurement, rather than by acquiring a minimal set of orientations with longer count times.

Future work should repeat these analyses on an ideal reference standard, such as a ring and core, across multiple beamlines in a round-robin style experiment. This would improve calibration and reporting of fundamental uncertainties for neutron diffraction beamlines designed for strain determination.

6. Related literature

For further literature related to the supporting information, see Shin *et al.* (2017).

Acknowledgements

This work makes use of perceptually uniform colormaps developed by Crameri *et al.* (2020). This research used resources at the High Flux Isotope Reactor, a DOE Office of Science User Facility operated by Oak Ridge National Laboratory. The beam time was allocated to the HIDRA instrument on proposal No. IPTS-33702.1. The authors are thankful for Rob Patterson's help creating the AFSD samples.

The authors also thank Dr Benjamin Wing for assistance at the High Flux Isotope Reactor. Author contributions are as follows: conceptualization, CF, MP; formal analysis, CF; investigation, CF, MP, JB, AP, KP; writing (original draft), CF; writing (review and editing), MP, JB, AP, KP; funding acquisition, KP.

Funding information

This material is based upon work supported by DEVCOM Army Research Laboratory (Cooperative Agreement W911NF21200 to KP).

References

- Akrivos, V., Wimpory, R. C., Hofmann, M., Stewart, B., Muransky, O., Smith, M. C. & Bouchard, J. (2020). *J. Appl. Cryst.* **53**, 1181–1194.
- Anderson, K., Weritz, J. & Kaufman, J. G. (2019). *6061 and Alclad 6061 Datasheet*. Materials Park: ASM International.
- Andrae, R. (2010). *arXiv*, 1009.2755.
- Andrae, R., Schulze-Hartung, T. & Melchior, P. (2010). *arXiv*, 1012.3754.
- Bachmann, F., Hielscher, R. & Schaeben, H. (2010). *Solid State Phenom.* **160**, 63–68.
- Beghini, M. & Grossi, T. (2024). *Exp. Mech.* **64**, 851–874.
- Beghini, M., Grossi, T., Prime, M. B. & Santus, C. (2023). *Exp. Mech.* **63**, 495–516.
- Bunn, J. R., Fancher, C. M., Payzant, E. A., Cornwell, P. A., Bailey, W. B. & Gregory, R. (2023). *Rev. Sci. Instrum.* **94**, 035101.
- Bunn, J. R., Penumadu, D., Lou, X. & Hubbard, C. R. (2014). *Metall. Mater. Trans. A* **45**, 3806–3813.
- Crameri, F., Shephard, G. E. & Heron, P. J. (2020). *Nat. Commun.* **11**, 5444.
- Fancher, C. M., Bunn, J. R., Bilheux, J., Zhou, W., Whitfield, R. E., Borreguero, J. & Peterson, P. F. (2021). *J. Appl. Cryst.* **54**, 1886–1893.
- Hielscher, R. & Schaeben, H. (2008). *J. Appl. Cryst.* **41**, 1024–1037.
- Holden, T. M., Traore, Y., James, J., Kelleher, J. & Bouchard, P. J. (2015). *J. Appl. Cryst.* **48**, 582–584.
- Hutchings, M. T., Withers, P., Holden, T. M. & Lorentzen, T. (2005). *Introduction to the Characterization of Residual Stress by Neutron Diffraction*, pp. 149–202. Taylor & Francis.
- Joint Committee for Guides in Metrology (2020). *JCGM GUM-6:2020, Guide to the Expression of Uncertainty in Measurement – Part 6: Developing and Using Measurement Models*. BIPM, Paris, France. https://www.bipm.org/documents/20126/2071204/JCGM_GUM_6_2020.pdf.
- Metz, P. C., Arwood, Z., Franz, C., Heikkinen, E., Chawla, V., Babu, S. S., Penumadu, D. & Page, K. (2023). *Materialia* **30**, 101799.
- Metz, P. C., Franz, C., Kincaid, J., Schmitz, T., Lass, E. A., Babu, S. S. & Page, K. (2024). *Additive Manufact.* **81**, 103989.
- Metz, P. C., Miller, L., Kincaid, J., Charles, E., Wood, A. T., Sims, Z. C., Drewry, S., Houston, A., Bunn, J. R., Compton, B., Schmitz, T., Lass, E. A., Penumadu, D. & Page, K. (2025). *Mater. Charact.* **229**, 115600.
- Noyan, I. C. & Cohen, J. B. (1987). *Residual Stress: Measurement by Diffraction and Interpretation*. Springer-Verlag New York.
- Prime, M. B., DeWald, A. T., Hill, M. R., Clausen, B. & Tran, M. (2014). *Eng. Fracture Mech.* **116**, 158–171.
- Prime, M. B. & Hill, M. R. (2006). *J. Eng. Mater. Technol.* **128**, 175–185.
- Rutherford, B. A., Avery, D. Z., Phillips, B. J., Rao, H. M., Doherty, K. J., Allison, P. G., Brewer, L. N. & Jordon, J. B. (2020). *Metals* **10**, 947.
- Şeren, M. H., Pagan, D. C. & Noyan, I. C. (2023). *J. Appl. Cryst.* **56**, 1144–1167.
- Shin, D., Roy, S., Watkins, T. R. & Shyam, A. (2017). *Comput. Mater. Sci.* **138**, 149–159.
- Winholtz, R. A. & Cohen, J. B. (1988). *Aust. J. Phys.* **41**, 189–200.
- Withers, P. J. & Bhadeshia, H. K. D. H. (2001a). *Mater. Sci. Technol.* **17**, 355–365.
- Withers, P. J. & Bhadeshia, H. K. D. H. (2001b). *Mater. Sci. Technol.* **17**, 366–375.
- Zhu, N., Avery, D. Z., Chen, Y., An, K., Jordon, J. B., Allison, P. G. & Brewer, L. N. (2023). *J. Mater. Eng. Perform.* **32**, 5535–5544.