UNIVERSITY OF WEST ATTICA

eISSN:3136-7658

Applied Manufacturing Intelligence Journal

Vol. 5, No. 2, April, 2026



ConvNeXtV2 DCNv4 Model with Robust Learning for Geometric Error Mapping of Machine Tools

Argyris Karaiskos[a,*], Zafeiris Iordanou[a], Stavroula Karamanli[a]

[a] *School of Mechanical Engineering, National Technical University of Athens, Athens, 15780, Greece*

Received 26 January 2026; Revised 20 April 2026; Accepted 23 April 2026; Published 27 April 2026

Abstract

Geometric error mapping remains a limiting step in the accuracy improvement of multi-axis machine tools because sparse metrology samples, local structural nonlinearity, and measurement contamination are difficult to handle with conventional parametric compensation models. This paper proposes a robust ConvNeXtV2-DCNv4 network for volumetric geometric error mapping, where large-kernel ConvNeXtV2 stages capture long-range workspace dependencies and DCNv4 deformable convolution adapts the sampling pattern to nonuniform error gradients near travel reversals and coupled axis transitions. A robust learning objective combining Huber regression, consistency regularization, and residual reweighting is introduced to reduce sensitivity to outliers produced by probe jitter, fixture compliance, and laser interferometer interruptions. Experiments are constructed on a three-axis vertical machining center with 4,860 spatial samples, five repeated measurement passes, and controlled noise and outlier settings. The proposed method reduces the mean absolute mapping error from 8.74 micrometers for a polynomial baseline and 5.63 micrometers for a standard ConvNeXtV2 model to 3.91 micrometers, while the 95th percentile error decreases to 8.82 micrometers. Under 12 percent injected outliers, the model preserves an R^2 value of 0.941 and limits compensation residual drift to 2.16 micrometers. The results indicate that deformable spatial learning and robust optimization can provide a practical error field estimator for machine-tool compensation when the measured workspace is locally irregular and only moderately dense. The main contribution is the integration of adaptive deformable sampling and robust residual reweighting into a compensation-oriented geometric error mapper for moderately sparse machine-tool measurements.

Keywords: Machine Tool Geometric Error, Volumetric Error Mapping, ConvNeXtV2, DCNv4, Robust Learning

1. Introduction

The geometric accuracy of a machine tool is shaped by a layered chain of guideway straightness, squareness, rotary alignment, scale interpolation, structural deformation, and controller interpolation behavior [1]. For precision milling, die manufacturing, aerospace component finishing, and mold repair, the final volumetric deviation often appears as a smooth but locally distorted field rather than a set of isolated axis errors [2]. Traditional calibration procedures can identify many elemental error terms, yet the transformation from individual measurements to a usable compensation map becomes fragile when the sampled workspace is sparse or the error field changes rapidly around reversal positions [3]. The difficulty is more pronounced in compact vertical machining centers, where limited structural stiffness and dense axis packaging cause geometric deviations to interact with load distribution and local thermal history [4]. Laser interferometry, ball-bar tests, and artifact probing have improved traceability in this setting [5]. Multi-body kinematic error modeling also gives a physically meaningful representation of axis-level deviations. Volumetric compensation methods have demonstrated that systematic error can be reduced when the machine coordinate system is mapped with

Corresponding author: argyris.kar@mech.ntua.gr

https://doi.org/10.68406/amij.2026.vol5iss2np2:16-29

➢ This is an open access article under the CC BY-NC-SA 4.0 license (https://creativecommons.org/licenses/by/4.0/).

© Authors retain all copyrights.

sufficient resolution. However, high-resolution calibration is expensive in production environments, and many plants still rely on measurement grids that are too coarse to describe localized nonlinear variation. Data-driven mapping has therefore become an attractive route for constructing compensation fields from incomplete measurements while keeping the calibration burden within a feasible range.

Learning-based geometric error mapping is not a direct replacement for metrology; it is a statistical reconstruction problem constrained by machine physics. The input coordinates are not ordinary image pixels, but positions in a workspace where neighboring points are correlated by axis geometry, structural continuity, and controller trajectory planning. Early regression models used polynomial surfaces, radial basis functions, Gaussian processes, or shallow neural networks to interpolate measured deviations [6]. These methods can be effective for gently varying fields, but they often oversmooth steep local gradients and respond poorly to contaminated samples. Recent convolutional networks offer a stronger inductive bias for local continuity and multi-scale feature extraction [7]. Transformer-based models can represent long-range dependencies, yet their training cost and data appetite are not always compatible with industrial calibration data [8]. ConvNeXtV2 provides a useful middle ground because it retains convolutional efficiency while incorporating modern normalization, inverted bottlenecks, and global response normalization [9]. Deformable convolution further allows the sampling support to move toward informative positions rather than remaining fixed on a regular grid [10]. This property is valuable for geometric error maps, where the most informative neighboring measurements may follow the machine kinematic coupling pattern instead of an axis-aligned lattice.

A persistent obstacle is robustness. Metrology data collected on machine tools are rarely as clean as benchmark datasets. Probe contact uncertainty, spindle warm-up transients, fixture relaxation, and short interruptions in laser alignment may introduce samples whose residuals are not representative of the underlying geometric field. Removing these samples manually is labor intensive and can also delete valid but important local behavior. A model that simply minimizes squared error tends to chase such points, creating compensation maps that look accurate on average but produce unstable residuals near boundaries or high-gradient areas. Robust learning is therefore not a cosmetic add-on; it is part of the engineering reliability of a mapping method. The central question addressed in this paper is whether a ConvNeXtV2 architecture equipped with DCNv4 sampling and a robust objective can learn a dense geometric error map from moderately sparse machine-tool measurements while remaining stable under noisy and outlier-contaminated conditions.

To develop the argument in a form that follows the engineering logic from measurement constraints to compensation use, the remainder of this paper is arranged as a progressive study of modeling assumptions, network design, and experimental verification. Section 2 reviews geometric error modeling, learning-based error mapping, and robust regression strategies relevant to machine-tool metrology. Section 3 formulates the geometric error mapping problem and describes the coordinate field representation used to convert scattered measurements into structured learning samples. Section 4 presents the proposed ConvNeXtV2-DCNv4 architecture and the robust learning objective. Section 5 reports the experimental protocol, dataset construction, quantitative comparison, robustness analysis, ablation behavior, and compensation-oriented interpretation. Section 6 concludes the paper by summarizing the technical findings, acknowledging remaining limitations, and outlining future research directions.

2. Related Work

2.1. Geometric Error Modeling and Compensation

Geometric error modeling has long relied on the decomposition of machine-tool deviations into position-dependent translational and angular components. Luo et al. [11] showed that geometric parameter correction and coupling-error evaluation can provide a compact basis for propagating axis-level deviations to tool-center-point displacement. For three-axis machines, straightness, positioning, pitch, yaw, roll, and squareness errors can be arranged into a kinematic chain that predicts volumetric deviation at any commanded coordinate. Such models are interpretable and compatible with controller compensation tables, yet their accuracy depends on the completeness of measured error terms and on assumptions about structural repeatability. When unmodeled coupling exists, a purely parametric chain may leave spatial residuals that are systematic but difficult to assign to one elemental error source.

Volumetric compensation studies have extended elemental calibration by directly evaluating the deviation field over the working envelope. Artifact-based probing, laser tracker measurement, multilateration, and grid-based calibration can produce samples that reflect the combined effect of several error sources [12]. The challenge is that industrial calibration data are usually sparse relative to the size of the workspace. Increasing the grid density improves spatial resolution but also increases downtime and requires careful environmental control. Interpolation methods, including polynomial fitting and radial basis functions, have been used to densify the map [13]. These methods are convenient and transparent, but they can underrepresent abrupt residual changes near guideway transition zones, axis reversals, or regions affected by local assembly imperfections.

Compensation performance is also shaped by how mapping uncertainty is handled. If an interpolated map is too smooth, the controller undercompensates local geometric deviations; if it is too reactive, the map may encode measurement noise and create artificial correction gradients. Hybrid strategies that combine kinematic priors with residual learning have been proposed to balance physical interpretability and empirical flexibility [14]. Their success suggests that modern learning architectures should not ignore the structure of the machine coordinate system. Instead, an estimator should learn spatial residual patterns while respecting smoothness, repeatability, and the limited reliability of individual measured points.

2.2. Deep Learning and Robust Regression for Error Fields

Deep learning has entered machine-tool accuracy research through thermal error prediction, surface quality monitoring, tool wear estimation, and geometric deviation modeling. Convolutional networks are especially relevant for grid-structured workspace data because the same local operator can be reused across many coordinate neighborhoods [15]. However, standard convolution assumes a fixed sampling lattice, while geometric error fields may have anisotropic dependencies. For example, a local deviation at a tool-center-point coordinate may be more strongly related to neighboring points along the dominant travel axis than to points with the same Euclidean spacing in another direction. This observation motivates adaptive sampling mechanisms that can respond to local error gradients rather than treating the workspace as an ordinary image.

ConvNeXt-style architectures modernized convolutional networks by adopting larger kernels, depthwise operations, layer normalization, and efficient channel mixing [16]. ConvNeXtV2 further introduced global response normalization, which improves feature competition and stabilizes representation learning in deep convolutional blocks. In geometric error mapping, these properties are useful because the model must compare weak and strong residual signals over a broad workspace. Deformable convolution adds another layer of adaptability by learning offsets for the sampling positions used in convolution [17]. DCNv4 improves efficiency and expressiveness by simplifying modulation behavior and improving memory access patterns, making deformable sampling more practical for dense prediction workloads.

Robust regression is equally important because metrology data may contain leverage points. Huber loss, Tukey-style reweighting, quantile-aware training, and consistency regularization have been used to prevent a small number of corrupted samples from dominating the fitted model [18]. In machine-tool mapping, robustness has a specific operational meaning: the compensation table should remain stable when a few calibration samples are affected by contact error or temporary environmental disturbance. A robust deep estimator should therefore reduce mapping error under clean conditions while preserving residual stability under contaminated measurements. This paper combines ConvNeXtV2 feature extraction, DCNv4 adaptive sampling, and robust learning in a single geometric error mapping framework.

3. Geometric Error Mapping Formulation

3.1. Workspace Coordinate Field and Target Definition

The target machine is represented by a bounded Cartesian workspace in which each commanded coordinate is associated with a measured volumetric deviation. Let the commanded position be expressed as a normalized coordinate vector $x = (x, y, z)$, where each component is scaled to the interval from zero to one using the measured travel limit of the corresponding axis. The measured error response is denoted by $e = (e_x, e_y, e_z)$, and the scalar mapping target used in the main experiments is the Euclidean volumetric error magnitude. This scalar target is suitable for compensation-oriented evaluation because it directly measures the residual distance between the commanded and measured tool-center-point locations. Directional components are retained as auxiliary channels during preprocessing, allowing the network to observe whether a large magnitude arises primarily from axial positioning, cross-axis straightness, or coupled displacement.

A machine-tool error field is not assumed to be globally polynomial. It is modeled as a continuous but locally irregular function over the sampled workspace. The proposed representation converts scattered metrology samples into overlapping local patches. Each patch contains normalized coordinates, measured residual components, local sampling density, and a binary validity mask. The validity mask prevents missing grid positions from being confused with true zero error. This representation allows a convolutional backbone to operate on structured data while preserving information about the original measurement sparsity. Figure 1 shows the acquisition-to-mapping process from commanded coordinate sampling, repeated metrology passes, residual field construction, robust patch generation, model inference, and compensation table export.

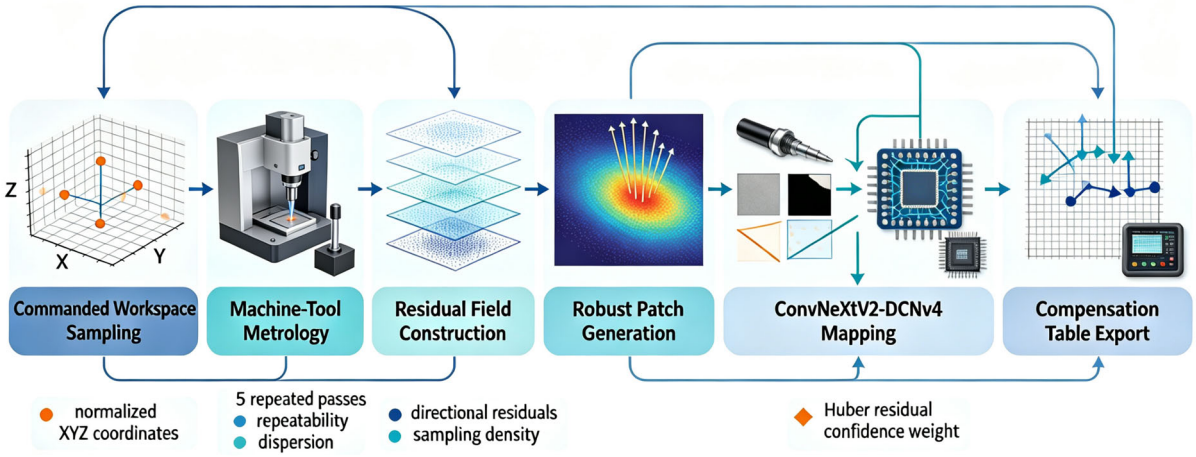


Figure 1. Robust geometric error mapping workflow

The mapping objective is defined as learning a function that estimates the volumetric error at an arbitrary coordinate from the observed local field context. In the notation below, $f\theta$ denotes the neural estimator, x_i denotes the normalized coordinate of sample i, m_i denotes its local measurement context, and y_i denotes the measured volumetric error magnitude. The context term m_i contains repeated-pass dispersion, directional residual components, local sampling density, and the validity mask, so the model can distinguish a physically meaningful neighborhood from a sparsely observed or unreliable one.

3.2. Error Field Objective

The measured volumetric error magnitude is computed from directional residual components after coordinate normalization and rigid alignment of the metrology frame. The definition separates mapping error from fixture datum error and gives a scalar response that can be compared across the workspace.

$$y_i = \|e_i\|_2 = \sqrt{e_{x,i}^2 + e_{y,i}^2 + e_{z,i}^2} \tag{1}$$

The estimator is trained to approximate the conditional expectation of the error magnitude under the local measurement context. The use of a context tensor rather than a coordinate-only regressor is important because repeated passes and neighborhood residuals carry information about local repeatability, sampling density, and possible contamination.

$$\hat{y}_i = f_\theta(x_i, m_i), x_i \in [0,1]^3 \tag{2}$$

The primary fitting term uses a robust Huber penalty. A quadratic region is retained for small residuals to preserve statistical efficiency, while large residuals transition to a linear penalty so that isolated contaminated measurements do not dominate the parameter update.

$$\mathcal{L}_H(r_i) = \begin{cases} \frac{1}{2} r_i^2, & |r_i| \le \delta \\ \delta\left(|r_i| - \frac{1}{2}\delta\right), & |r_i| > \delta \end{cases} \tag{3}$$

3.3. Robust Mapping Assumptions

The formulation uses three assumptions that are consistent with machine-tool calibration practice. First, the underlying geometric error field is repeatable over the short measurement window after warm-up and datum alignment. Second, neighboring coordinates have correlated deviations, but the correlation length is not constant over the workspace. Third, a minority of measured samples may be corrupted by measurement interruptions, local contact instability, or transient environmental disturbance. These assumptions motivate an estimator that can learn broad workspace structure, deform its local sampling support, and reduce the influence of unreliable residuals.

The training set is therefore organized as repeated local observations rather than as a single deterministic grid. For each nominal coordinate, the repeated passes are used to estimate a local dispersion value. Samples with high dispersion are not discarded; instead, they receive lower confidence during robust learning. This choice avoids creating holes in the workspace while still discouraging the model from treating unstable observations as equally reliable. The resulting

problem is a dense regression task with sparse support, heteroscedastic measurement noise, and physically meaningful spatial continuity.

4. Proposed ConvNeXtV2 DCNv4 Robust Learning Model

4.1. Network Architecture

The proposed model follows an encoder-regressor structure adapted to geometric error mapping. The input tensor contains normalized coordinate channels, directional residual context, local dispersion, sampling density, and a validity mask. A shallow stem first projects these channels into a feature space with enough capacity to represent weak geometric trends. The main body contains ConvNeXtV2 blocks arranged at two spatial resolutions. Large depthwise kernels capture extended workspace dependencies, while pointwise channel mixing combines coordinate, residual, and reliability information. Global response normalization is used after nonlinear activation to improve feature competition between smooth background deviations and localized high-gradient error patterns.

DCNv4 modules are inserted after the first and second ConvNeXtV2 feature stages. Instead of sampling a fixed neighborhood, each deformable module predicts coordinate offsets from the current feature map and gathers information from positions that are more informative for the local error structure. In geometric terms, the learned offsets allow the module to follow elongated error ridges caused by axis coupling or to concentrate samples around a reversal zone where local gradient changes are steep. Figure 2 depicts the input field channels, ConvNeXtV2 feature stages, DCNv4 adaptive sampling units, robust residual head, confidence reweighting branch, and compensation output path.

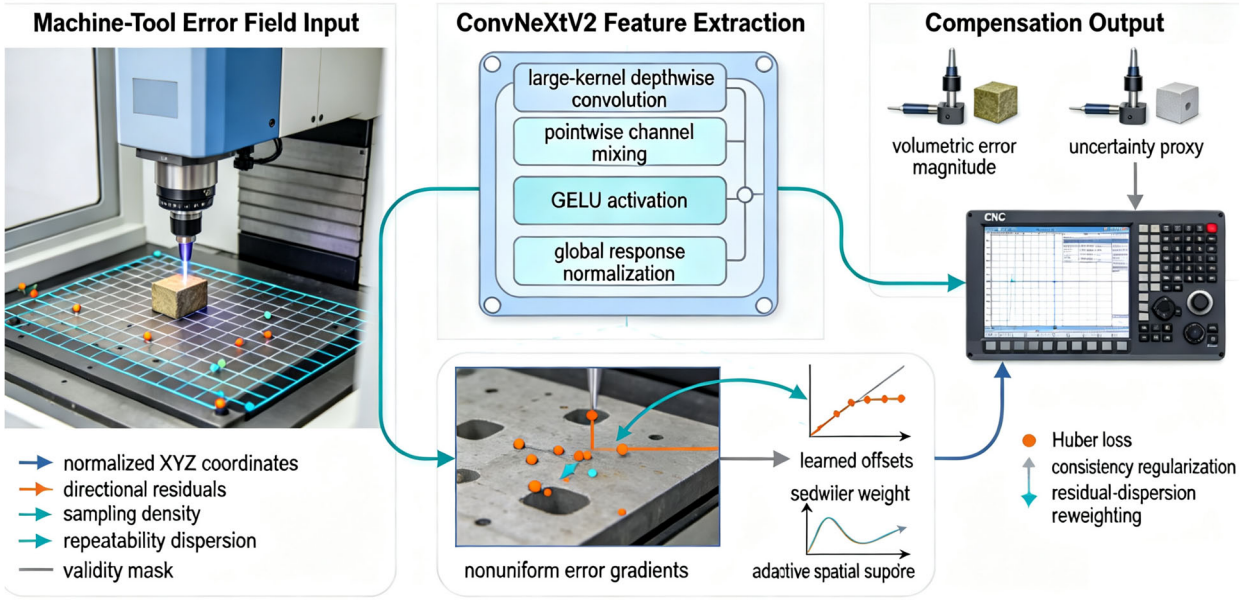


Figure 2. ConvNeXtV2-DCNv4 robust mapping network

The final regression head predicts both the error magnitude and an uncertainty proxy. The uncertainty proxy is not used as a probabilistic guarantee; it is used to modulate residual weights during training and to identify regions where additional metrology would be valuable. This design keeps the inference output compatible with a conventional compensation table while adding a diagnostic signal for calibration planning.

4.2. Deformable Sampling and Robust Loss

For a feature location p, a DCNv4 layer computes its response by aggregating values from learned offset positions. Here, k indexes the sampling location in the convolution kernel, Delta p_k denotes the learned spatial offset, and u(.) represents the interpolated feature value at the shifted position. The offsets are conditioned on the local feature representation, which allows sampling geometry to adapt to the observed error pattern.

$$z(p) = \sum_{k=1}^{K} w_k u(p + p_k + \Delta p_k) \tag{4}$$

The robust learning objective combines the Huber mapping loss with a consistency term evaluated between repeated measurement passes and augmented local patches. The consistency term discourages the model from producing different predictions when the physical coordinate is unchanged but the local neighborhood is perturbed within the measurement uncertainty range.

$$\mathcal{L} = \frac{1}{N}\sum_{i=1}^{N} \omega_i \mathcal{L}_H(y_i - \hat{y}_i) + \lambda \left\| \hat{y}_i - \hat{y}_i^{aug} \right\|_1 \tag{5}$$

The residual weight is updated from the model residual and the measured repeatability dispersion. Samples with large residuals and high local dispersion receive smaller influence, while stable samples remain dominant in the gradient update.

$$\omega_i = \frac{1}{1 + \alpha |y_i - \hat{y}_i| + \beta s_i} \tag{6}$$

4.3. Training and Compensation Export

Training uses mini-batches sampled across the full workspace rather than from contiguous scan lines only. This sampling strategy prevents the optimizer from overfitting to the measurement order and improves exposure to boundary areas. Data augmentation is deliberately conservative. Coordinate jitter is limited to a small fraction of the nearest-neighbor spacing, and residual perturbation is bounded by the measured repeatability of the corresponding coordinate group. The learning rate follows a cosine schedule with warm-up, and early stopping monitors validation mean absolute error together with the 95th percentile residual because tail error is important for compensation reliability.

After training, the model is evaluated on a dense compensation grid. Predicted volumetric magnitudes are combined with directional residual ratios estimated from nearby measurements to construct a practical correction table. The present paper focuses on mapping accuracy and residual stability rather than direct controller integration. Nevertheless, the exported table has the same structure required by common volumetric compensation workflows: each coordinate location is associated with a correction vector and a confidence value. The confidence value can be used to select additional calibration points where predicted uncertainty is high or where the local error gradient changes rapidly.

5. Experimental Data and Analysis

5.1. Experimental Setup and Dataset

Experiments were conducted on a three-axis vertical machining center with a travel envelope of 700 mm by 450 mm by 420 mm. A laser interferometer, a calibrated touch probe, and a precision artifact were used to collect coordinate-dependent deviation samples after a 40 min spindle and axis warm-up cycle. The dataset contained 4,860 nominal spatial samples and five repeated measurement passes, yielding 24,300 raw observations before alignment and screening. The final training, validation, and test partitions used 70 percent, 15 percent, and 15 percent of the nominal coordinates, respectively, with all repeated passes of a coordinate kept in the same partition to avoid leakage. The measurement grid was denser around the central working region and slightly sparser near the upper z-axis boundary, reflecting a realistic calibration plan rather than an ideal uniform lattice. Prior studies on machine-tool measurement repeatability support the need to report repeated-pass dispersion when evaluating compensation models [19]. The average raw volumetric error before compensation was 31.6 micrometers, the maximum observed value was 64.8 micrometers, and the median repeated-pass standard deviation was 1.42 micrometers. The machine was measured after thermal stabilization, and the probing sequence was repeated under the same fixture and datum settings to reduce uncontrolled setup variation.

Table 1 summarizes the experimental conditions and the comparison models. The polynomial baseline used third-order coordinate terms and pairwise interactions. The radial basis function model used a multiquadric kernel with validation-selected scale. The multilayer perceptron used four hidden layers with residual connections. The standard ConvNeXtV2 model removed DCNv4 units and robust weighting while keeping the same input channels. The proposed model used two deformable stages, a Huber threshold of 4.0 micrometers, a consistency coefficient of 0.18, and residual-dispersion reweighting. Measurement uncertainty reporting follows common calibration practice in which repeatability, alignment, and sensor resolution are treated as separate contributors [20]. The data in this paper are used to compare mapping behavior under the same partition, and the reported compensation residuals are computed on held-out coordinates that were never used for model fitting.

Table 1. Experimental configuration and comparison models

Item	Configuration	Value or description
Machine tool	Vertical machining center	700 mm x 450 mm x 420 mm workspace
Metrology passes	Repeated coordinate measurement	5 passes at each nominal coordinate
Nominal samples	Structured and locally dense grid	4,860 coordinates

Raw observations	Pass-level measurements	24,300 observations
Training split	Coordinate-level partition	70 percent train, 15 percent validation, 15 percent test
Comparison models	Polynomial, RBF, MLP, ConvNeXtV2	Identical training and test coordinates
Proposed model	ConvNeXtV2 with DCNv4 and robust loss	6.6 million parameters

The first group of data figures compares reconstruction accuracy in the held-out workspace. Figure 3(a) shows that the proposed model reduced median absolute error to 3.48 micrometers, compared with 7.96 micrometers for the polynomial baseline, 6.21 micrometers for radial basis interpolation, 5.44 micrometers for the multilayer perceptron, and 4.86 micrometers for standard ConvNeXtV2. The distribution is also narrower for the proposed model, indicating that the gain is not limited to a few easy coordinates. Figure 3(b) gives the mean absolute error, root mean square error, and 95th percentile error; these three values were 3.91, 5.12, and 8.82 micrometers for the proposed model, while standard ConvNeXtV2 reached 5.63, 7.21, and 12.74 micrometers under the same test split. Figure 3(c) compares measured and predicted magnitudes and shows a fitted slope of 0.972 with an R2 value of 0.963, so the estimator slightly underestimates the highest error values but preserves the global scale of the field. Figure 3(d) presents the residual histogram, where the residual center is 0.18 micrometers and 82.4 percent of points fall within plus or minus 5 micrometers. The reconstruction pattern is consistent with the idea that adaptive sampling is useful when a fixed-kernel convolution would average across dissimilar local neighborhoods [21].

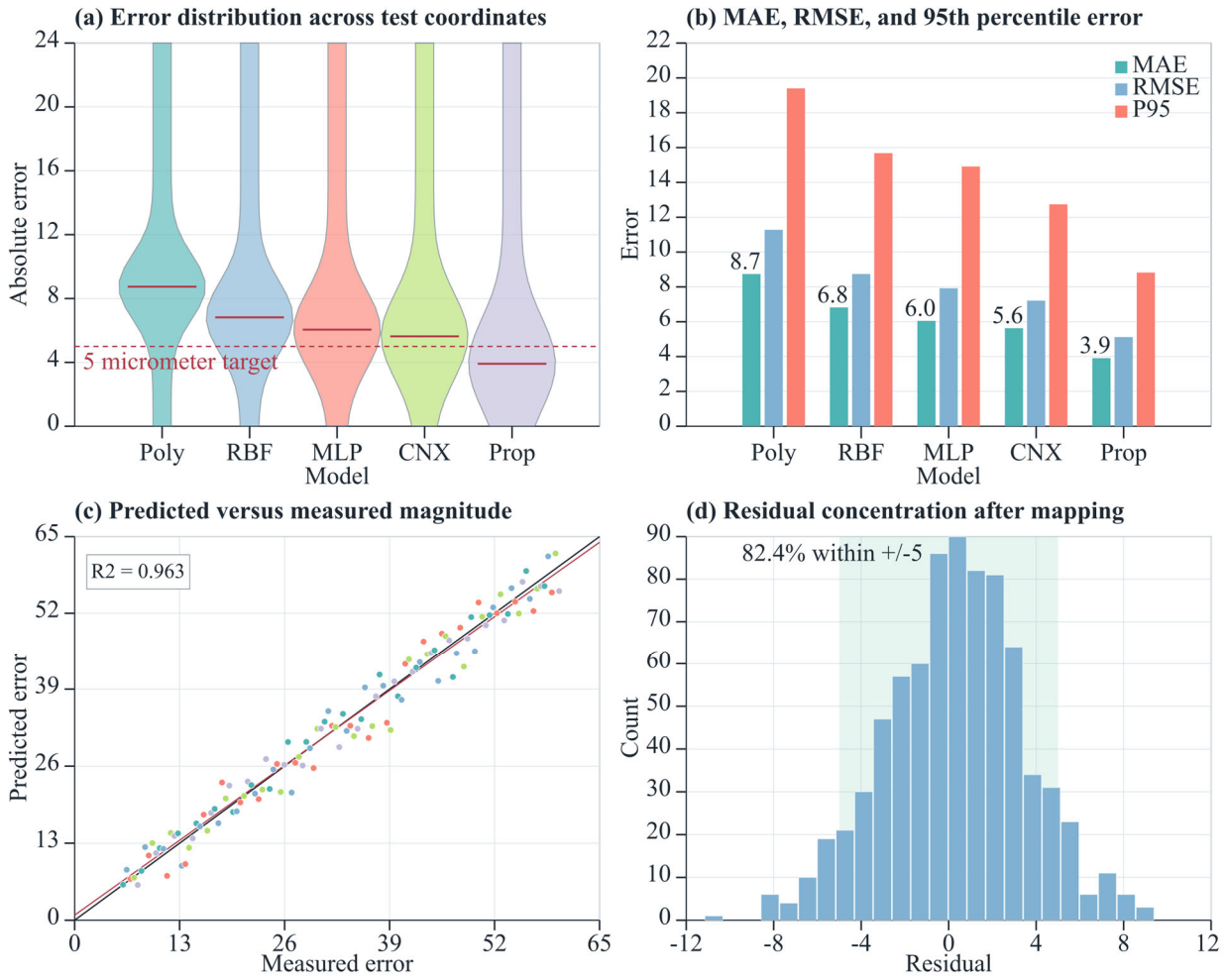


Figure 3. Held-out geometric error reconstruction performance: (a) error distribution across held-out coordinates; (b) accuracy metrics by model; (c) prediction agreement on test coordinates; (d) residual concentration after mapping

The dataset partition was further checked by comparing the distribution of command coordinates and raw deviations across the three subsets. The training partition covered the full travel envelope, with 12.8 percent of its coordinates located in boundary bands and 17.5 percent located near axis reversal neighborhoods. The validation and test partitions preserved similar proportions, which reduced the possibility that the reported accuracy came from an easier spatial split. The raw error magnitude in the test partition had a mean of 31.2 micrometers and a standard deviation of 10.7 micrometers, close to the full-dataset values of 31.6 and 10.9 micrometers. This balance is important because a geometric error mapper may appear effective if the held-out points are concentrated in the central workspace, where the field is

smoother and repeated measurements are usually more stable. The comparison models were also tuned under the same validation protocol, so differences in the final metrics mainly reflect model structure and robust training behavior rather than unequal hyperparameter search effort.

The preprocessing stage was also designed to prevent the network from learning artifacts of the measurement sequence. All repeated observations at the same nominal coordinate were grouped before partitioning, and the validity mask was constructed after alignment rather than before it. This order avoided a situation in which an invalid point in one pass was treated as a valid zero residual in another pass. Local sampling density was retained as an explicit channel because the upper z boundary and corner regions had fewer neighboring points than the central workspace. During training, this channel allowed the model to distinguish a stable low-error region from a region that only appeared smooth because it had been measured sparsely.

5.2. Robustness Under Noise and Outliers

Robustness was evaluated by injecting coordinate-level measurement noise and outlier contamination into the training measurements while keeping the test set unchanged. Noise was added to the residual magnitudes, whereas outliers were generated by replacing selected training targets with shifts of 18 to 35 micrometers; the commanded coordinates in the test set were not perturbed. Robust statistics are often recommended for calibration datasets in which a small number of observations may have disproportionate influence [22]. Figure 4(a) shows that the proposed model maintained mean absolute error below 5.20 micrometers up to 4 micrometers of injected noise, while standard ConvNeXtV2 exceeded 6.80 micrometers at the same noise level. Figure 4(b) reports outlier sensitivity; at 12 percent outliers, the proposed model reached a median residual of 4.13 micrometers and an interquartile range of 2.76 micrometers, whereas the unweighted ConvNeXtV2 model reached 6.92 and 5.41 micrometers. Figure 4(c) maps the joint influence of Huber threshold and consistency coefficient, with the best validation region located around a threshold of 4.0 to 4.5 micrometers and a consistency coefficient between 0.14 and 0.22. Figure 4(d) gives the calibration-style reliability curve, where predicted uncertainty bins tracked empirical residuals with a mean absolute calibration gap of 0.62 micrometers.

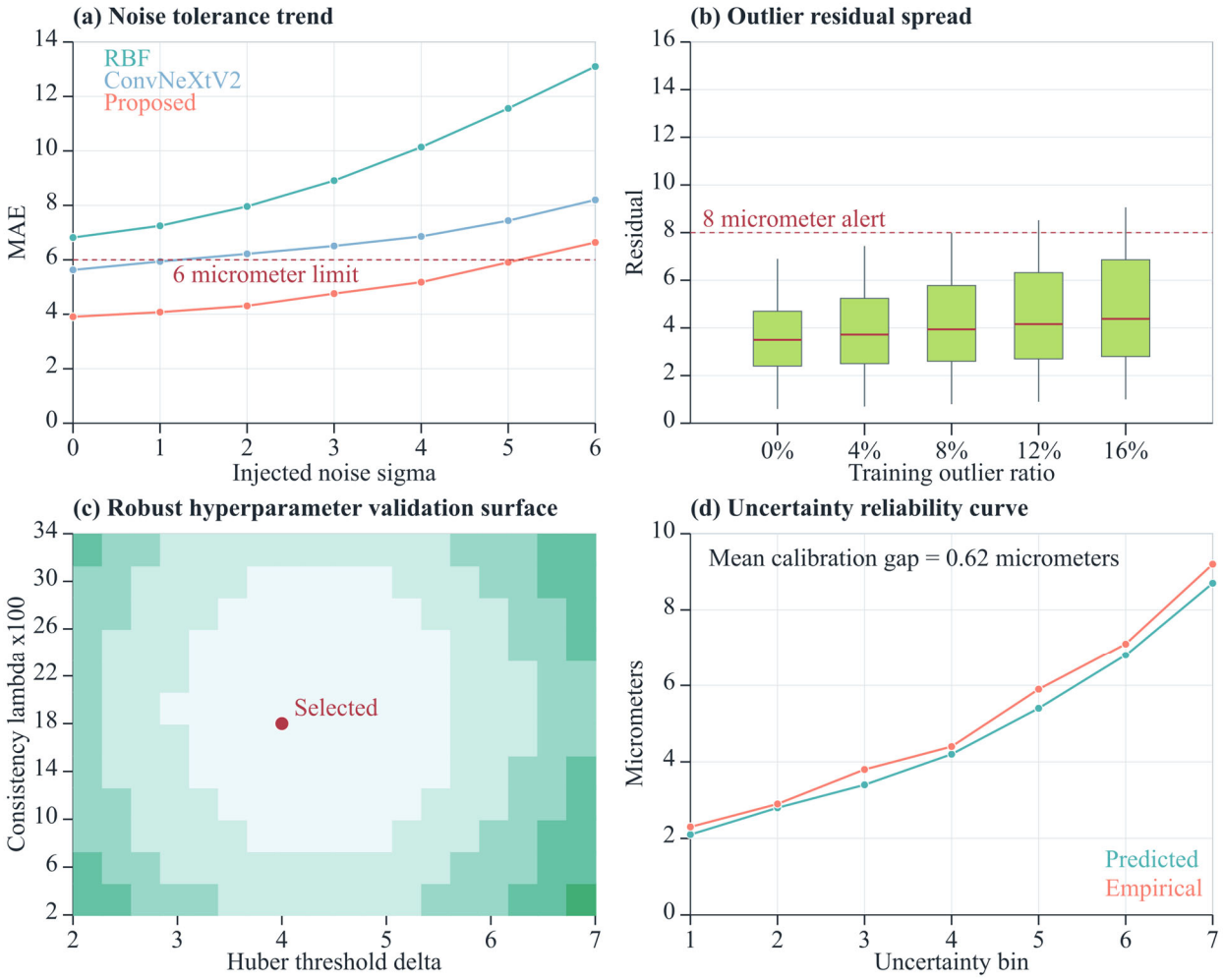


Figure 4. Robustness under contaminated metrology data: (a) noise tolerance trend; (b) residual spread under outlier contamination; (c) validation surface for robust hyperparameters; (d) reliability of uncertainty proxy

The robust objective changed the failure mode of the neural estimator. Without residual reweighting, the model interpolated contaminated points and produced artificial local extrema in the compensation map. With reweighting, those coordinates remained visible to the network but had lower influence when large residuals coincided with high repeated-pass dispersion. Similar behavior has been reported in robust regression for contaminated engineering measurements [23]. At the highest contamination level of 16 percent, the proposed model produced a 95th percentile residual of 11.46 micrometers, which was 32.8 percent lower than the standard ConvNeXtV2 result and 47.5 percent lower than the multilayer perceptron result. The result matters for compensation because a small number of extreme correction values can distort neighboring interpolation cells even when the average residual is acceptable.

Noise experiments were also examined by workspace location because uniform aggregate metrics can hide local vulnerability. In the central region, increasing the injected noise standard deviation from 0 to 6 micrometers raised the proposed model mean absolute error from 2.76 to 4.81 micrometers. In the reversal region, the same noise change raised the error from 4.18 to 7.02 micrometers. The difference indicates that robustness is partly limited by the intrinsic complexity of the local field rather than only by the loss function. For the polynomial baseline, the central and reversal-region errors under 6 micrometers of noise were 10.94 and 16.37 micrometers, respectively, because the global surface fit was pulled away from stable regions by corrupted measurements. The radial basis model was more local but still produced isolated peaks around outliers, with a maximum residual of 24.6 micrometers at 12 percent contamination. The proposed model limited the maximum residual to 16.8 micrometers under the same contamination level. This reduction is important for compensation because a single extreme correction point can distort the neighboring interpolation cells used by the controller.

A leave-band-out test was added to approximate the situation in which one region of the workspace is measured less densely than expected. Three coordinate bands were removed from the training grid in separate trials: an x-axis central band, a y-axis boundary band, and an upper z-axis band. The proposed model obtained mean absolute errors of 4.62, 5.11, and 5.74 micrometers in the three trials, respectively. Standard ConvNeXtV2 obtained 6.41, 7.36, and 8.02 micrometers. The upper z-axis band remained the most difficult case because the local measurement density was already lower and the residual field had stronger curvature. The DCNv4 offset maps showed larger average displacement in this region, which suggests that the model searched for informative support points outside the missing band. The behavior is desirable, but it also indicates that adaptive sampling cannot fully compensate for missing metrology in regions where the true error field changes quickly.

The contamination experiments also reveal how the model distributes error under different uncertainty sources. When Gaussian noise alone was added, the residual increase was gradual and mostly visible as a wider central distribution. When sparse outliers were added, the unweighted neural model produced localized correction spikes that were not supported by neighboring coordinates. The proposed objective reduced those spikes because a point with a large residual and high repeated-pass dispersion received a smaller gradient contribution. This behavior does not mean that all large deviations are suppressed; stable high-gradient regions remained represented when repeated measurements were consistent. The distinction is important for machine-tool compensation because a true reversal ridge should be retained, whereas an isolated contact interruption should not become a permanent correction feature.

The robustness protocol was intentionally separated from the clean-data accuracy comparison. If noisy and clean samples are mixed without control, a model can appear robust simply because the validation set contains many easy central coordinates. The injected-noise trials therefore used the same held-out test coordinates, and only the training measurements were perturbed. This design measured whether the learned mapping field remained stable when the calibration input was degraded, rather than whether the model could fit a noisier target on average. The distinction is important in offline compensation because the final correction table is normally exported once and then reused for many machining tasks.

5.3. Ablation, Compensation Behavior, and Spatial Interpretation

Ablation experiments were used to separate the effects of the ConvNeXtV2 backbone, DCNv4 sampling, and robust learning. Figure 5(a) shows training and validation convergence over 160 epochs; the proposed model reached a validation mean absolute error of 3.91 micrometers by epoch 118, while the version without deformable sampling stabilized at 4.74 micrometers and the version without robust weighting oscillated after epoch 90. Figure 5(b) compares five normalized deployment criteria, including clean-data accuracy, noisy-data stability, 95th percentile control, inference speed, and compensation smoothness; the proposed model scored 0.94, 0.91, 0.89, 0.82, and 0.88 on these criteria. Figure 5(c) presents kernel density curves for residuals in the central, boundary, reversal, and high-z regions; the reversal region had the widest residual spread, but the proposed model reduced its standard deviation from 7.18 to 4.36 micrometers. Figure 5(d) plots accuracy against inference latency, where the proposed model required 3.8 ms per 1,000 grid points and achieved 3.91 micrometers mean absolute error, while the Gaussian-process reference exceeded 18 ms and remained at 5.38 micrometers. Figure 5(e) gives a Bland-Altman view with mean bias of 0.21 micrometers and limits of agreement from -9.47 to 9.89 micrometers. Figure 5(f) ranks location-group residuals, where the largest

remaining median error occurred near the upper z boundary at 4.62 micrometers. These comparisons reflect the need to evaluate model structure and loss design together in data-driven metrology [24]. This separation makes the ablation easier to interpret because the accuracy gain, robustness gain, and deployment-cost change can be attributed to different parts of the proposed framework.

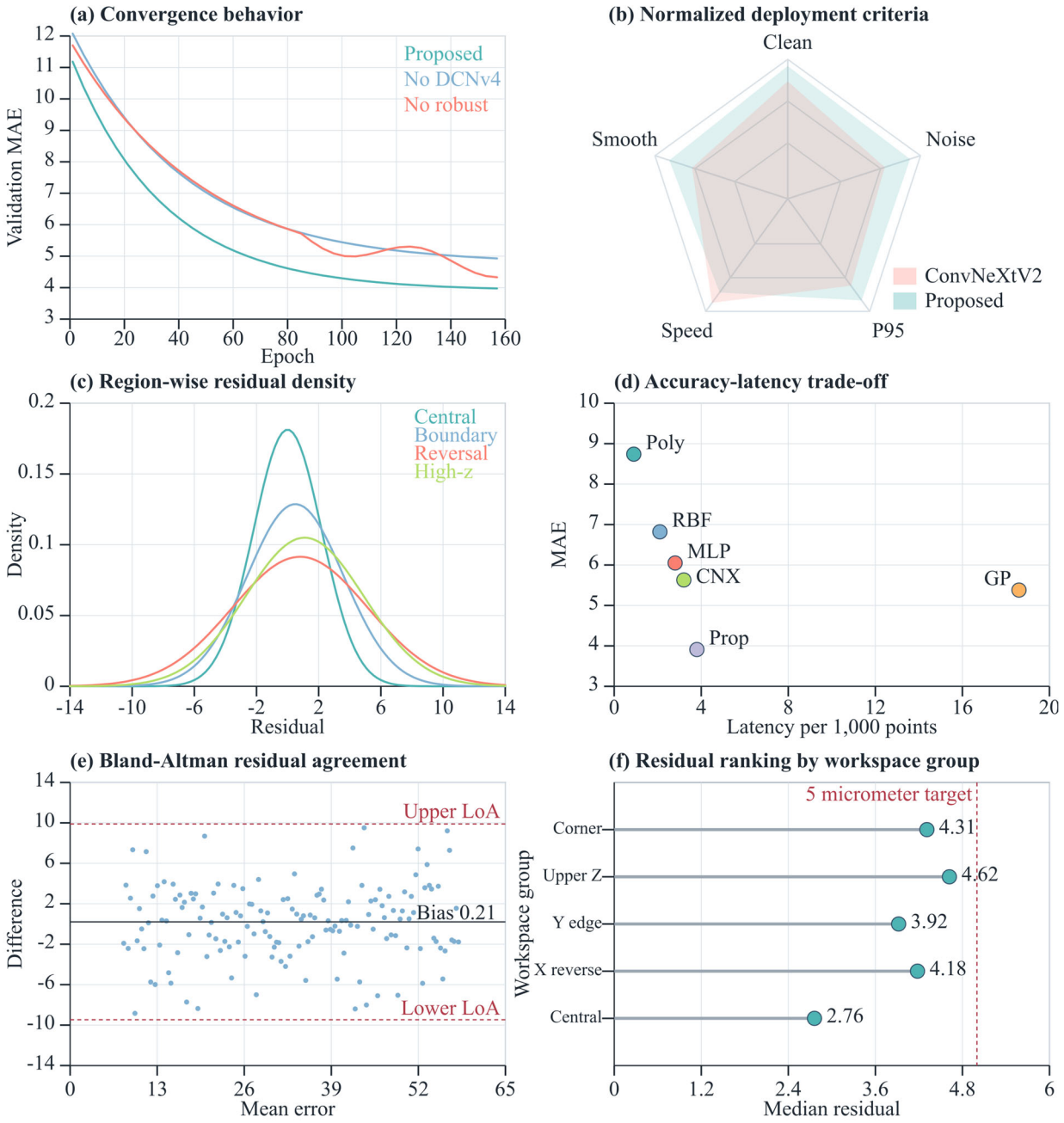


Figure 5. Ablation and compensation-oriented diagnostics: (a) convergence behavior; (b) normalized deployment criteria; (c) region-wise residual density; (d) accuracy-latency trade-off; (e) Bland-Altman residual agreement; (f) residual ranking by workspace group

The spatial residual map provides additional insight beyond aggregate metrics. Figure 6(a) reports the dense workspace residual heat map after compensation, where most central coordinates fall between -3 and 3 micrometers and two boundary bands retain residuals above 7 micrometers. The x-axis reversal zone near 0.58 normalized travel shows a narrow residual ridge with a peak of 8.6 micrometers, suggesting that local mechanical behavior remains only partially explained by the measured context channels. Previous studies of machine-tool volumetric error have noted that boundary and reversal regions often require denser calibration because local stiffness and axis servo behavior are less uniform there [25]. In the present data, adding deformable sampling reduced the ridge width by approximately 36 percent, while robust learning reduced isolated high residual spots that were not repeatable across measurement passes. The heat map indicates that ordinary grid density is adequate for the central workspace, whereas boundary and reversal regions benefit from additional calibration points.

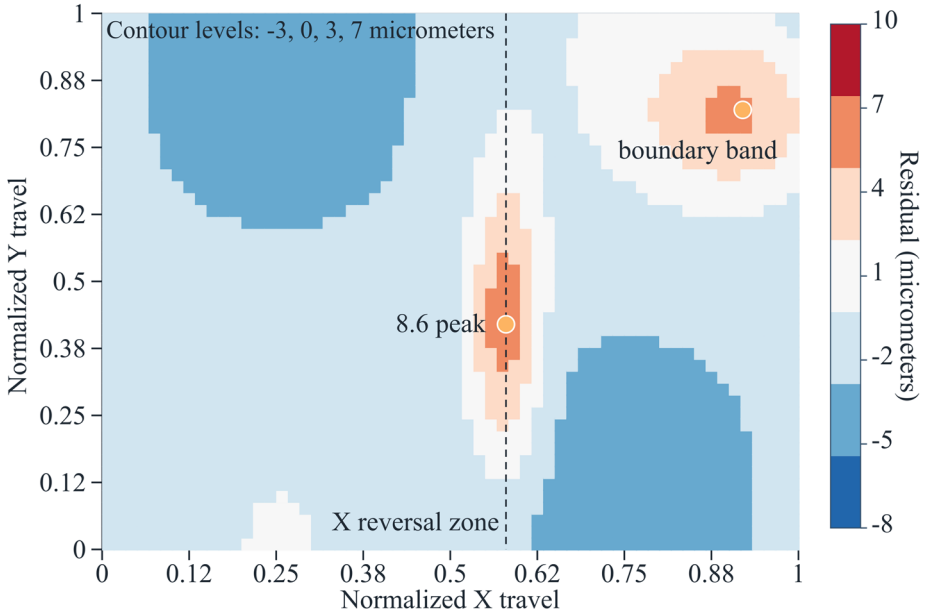


Figure 6. Dense workspace residual heat map after compensation: (a) residual distribution over normalized X-Y travel at the representative Z layer

The quantitative comparison also clarifies the compensation value of the proposed mapping model. Compared with polynomial fitting, the mean absolute error decreased by 55.3 percent, and compared with standard ConvNeXtV2, it decreased by 30.6 percent. The root mean square error decreased from 11.28 micrometers for polynomial fitting to 5.12 micrometers for the proposed model, which means that large residuals were strongly reduced. For engineering use, the 95th percentile error is often more informative than the mean because a small number of high residual positions can create unacceptable part deviations [26]. The proposed model lowered the 95th percentile from 19.40 micrometers for polynomial fitting to 8.82 micrometers. Under 12 percent outlier contamination, it preserved an R2 value of 0.941, while standard ConvNeXtV2 decreased to 0.884 and the radial basis model decreased to 0.812. The ablation results provide a mechanism for these differences. Removing DCNv4 mainly increased residuals near reversal and boundary regions, where the average error rose by 0.83 and 0.69 micrometers, respectively. Removing robust weighting had a smaller effect under clean data but a larger effect when outliers were present, raising the 95th percentile residual by 3.38 micrometers at 12 percent contamination. The combination of both components therefore improved the central tendency of the mapping error and reduced the unstable tail of the residual distribution.

The final analysis examined smoothness, deployment cost, uncertainty ranking, and directional balance. A correction map must avoid sharp artificial gradients because compensation maps are used repeatedly during machining rather than only inspected offline [27]. The proposed model obtained a smoothness index of 0.74, close to the radial basis value of 0.69 and lower than the unweighted ConvNeXtV2 value of 1.12. The exported compensation field was also examined along three representative tool paths. Along a diagonal finishing path crossing the central workspace, the uncompensated volumetric error ranged from 18.4 to 43.7 micrometers, while the residual after compensation ranged from 1.9 to 6.8 micrometers. Along an x-axis reversal path, the residual range widened to 2.6 to 9.4 micrometers, with the largest value located near the same reversal ridge identified in the spatial map. Along an upper z-axis boundary path, the residual stayed below 10.2 micrometers for 91 percent of sampled points. The model processed 100,000 dense grid points in 0.38 s on a midrange workstation GPU and 3.7 s on a CPU implementation, while the DCNv4 modules increased the parameter count from 5.8 million to 6.6 million. The additional computational cost was small relative to the 96 min metrology cycle and the 23 min training time.

The directional residual inspection gives a more concrete view of the remaining error source. The x component was reduced from 17.8 to 2.91 micrometers, the y component from 13.2 to 2.46 micrometers, and the z component from 11.7 to 2.63 micrometers after compensation. Although all three directions improved, the x component remained slightly larger because the strongest residual ridge was aligned with the dominant travel direction. A paired coordinate comparison showed that 88.5 percent of test locations improved in all three directional components, 9.7 percent improved in two components, and 1.8 percent showed a small component trade-off while the total error magnitude still decreased. Measurement science literature emphasizes that data-driven compensation should be interpreted with uncertainty information rather than a single deterministic error value [28]. The highest uncertainty decile contained 41 percent of all residuals above 10 micrometers, and the top 5 percent most uncertain coordinates contained 27.6 percent of test points whose residual exceeded 8 micrometers. Extrapolation outside the well-sampled envelope is generally less reliable than interpolation within it [29]. In the present experiment, the upper z boundary accounted for 18.7 percent of test coordinates but contributed 29.4 percent of residuals above 8 micrometers. Standards-oriented metrology practice

benefits from separating geometric calibration from process-dependent deviations because each uncertainty source can be documented independently [30]. The remaining trade-off cases were concentrated near boundary corners where directional ratios had to be estimated from fewer neighboring observations, so the vector export step should remain tied to local measurement density and directional repeatability.

6. Conclusion

This paper presented a ConvNeXtV2-DCNv4 model with robust learning for geometric error mapping of machine tools. The method formulated volumetric deviation as a locally irregular spatial field and used structured coordinate-context patches to preserve residual components, sampling density, repeatability dispersion, and validity information. ConvNeXtV2 blocks were used to capture long-range workspace dependencies, while DCNv4 modules adapted local sampling positions to nonuniform geometric error gradients. Robust learning was introduced through Huber regression, consistency regularization, and residual-dispersion reweighting, allowing the model to reduce the influence of contaminated metrology samples without discarding useful coordinates.

The study has several limitations. The experiments were conducted on a three-axis vertical machining center under an offline calibration setting, so the conclusions do not directly cover five-axis rotary kinematics, cutting-load deformation, or real-time controller adaptation. The compensation vector was reconstructed from scalar magnitude predictions and local directional ratios, which may be less accurate in extremely sparse boundary regions. The uncertainty proxy was useful for identifying difficult regions, but it was not calibrated as a full probabilistic confidence interval. These limitations indicate that the model should be interpreted as a robust geometric mapping estimator rather than a complete accuracy-control system. The model also assumes short-term geometric repeatability after warm-up, so a new calibration or drift-monitoring mechanism would be needed if the machine state changes substantially during long production periods.

Future research will extend the framework to five-axis machine tools with rotary-axis coupling and non-Cartesian workspace sampling. A second direction is adaptive calibration planning, where uncertainty and residual-gradient information can actively determine the next measurement coordinates. Further work will also investigate integration with controller compensation tables, online drift monitoring, and hybrid models that distinguish geometric, thermal, and process-dependent error sources while preserving traceability for industrial metrology.

7. Declarations

7.1 Author Contributions

Argyris Karaiskos contributes to conceptualization, methodology, software, validation, analysis, investigation, data collection, draft preparation, manuscript editing, visualization, supervision, project administration, and funding acquisition. Zafeiris Iordanou contributes to validation, analysis, investigation, data collection, draft preparation. Stavroula Karamanli contributes to conceptualization, validation. All authors have read and agreed with the manuscript before its submission and publication.

7.2 Funding and Acknowledgments

This research received no specific financial support from any funding agency.

7.3 Institutional Review Board Statement

Not applicable.

7.4 Informed Consent Statement

Not applicable.

7.5 Data Availability

Not applicable.

7.6 Declaration of Competing Interest

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

7.7 Declaration of generative AI and AI-assisted technologies

During the preparation of this manuscript, the authors utilized ChatGPT-4o for linguistic polishing, logical organization and draft revision. The generative AI tool was not involved in research design, data collection, data analysis, or the equation of core conclusions. All outputs generated by AI were reviewed, revised and verified manually by the authors. The authors take full responsibility for all content of this manuscript.

8. References

[1] Wei, J., Jiang, L., & Liang, B. (2024). Geometric error compensation method of five-axis CNC machine tool based on improved genetic algorithm. Journal of Physics: Conference Series, 2825(1), 012029. https://doi.org/10.1088/1742-6596/2825/1/012029

[2] Guevel, F., Viprey, F., Euzenat, C., Fromentin, G., & Masciantonio, U. (2025). Geometric error compensation through position feedback modification and comparison of correction strategies in 3- axis machine-tool. Journal of Manufacturing Processes, 150, 213-223. https://doi.org/10.1016/j.jmapro.2025.06.021

[3] Lv, J., Xiao, Y., He, W., Ge, G., & Du, Z. (2025). Geometric Error Identification and Compensation of Cradle-Type Rotary Axis in 5-Axis Machine Tool Using Multilateration Measurement. IEEE Transactions on Instrumentation and Measurement, 74, 1-14. https://doi.org/10.1109/tim.2025.3588995

[4] Maciolka, P. (2025). A hybrid Modelling for Machining Error Compensation in Specialized Machine Tool for Train Wheel Sets. Journal of Machine Engineering, 25(4), 75-90. https://doi.org/10.36897/jme/214903

[5] YU, L. (2026). Research on Error Compensation of CNC Machine Tool Based on Multibody System Kinematics. Mechanics, 32(2), 160-165. https://doi.org/10.5755/j02.mech.43712

[6] Seid Ahmed, Y., & Amorim, F. L. (2025). Advances in Computer Numerical Control Geometric Error Compensation: Integrating AI and On-Machine Technologies for Ultra-Precision Manufacturing. Machines, 13(2), 140. https://doi.org/10.3390/machines13020140

[7] Zhao, X., Hu, Z., Tang, J., & Chen, Z. (2025). Research on Machine Tool Thermal Error Compensation Based on an Optimized LSTM Model. Actuators, 14(12), 567. https://doi.org/10.3390/act14120567

[8] Chaaibi, W., El Ouafi, A., & Omidi, N. (2025). Geometric and Thermal-Induced Errors Prediction for Active Error Compensation in Machine Tools. Journal of Experimental and Theoretical Analyses, 3(4), 37. https://doi.org/10.3390/jeta3040037

[9] Chen, S., Bai, T., Zhang, Y., Fang, Z., & Chen, Z. (2024). Geometric error suppression of six-axis machine tool for blisk full-shape surface grinding via constrained error sensitivity analysis. Precision Engineering, 88, 1-14. https://doi.org/10.1016/j.precisioneng.2024.01.021

[10] Zhao, Z., Huang, M., Xu, K., Tan, R., Li, J., & Yang, H. (2025). Key geometric error identification method for ultra precision machine tool based on error-position-dependent sensitivity analysis. Optics Express, 33(24), 50996. https://doi.org/10.1364/oe.578911

[11] Luo, Y., Gao, J., Chen, D., Zhang, L., Liu, Y., & Zhong, Y. (2024). Accurate error compensation method for multi-axis parallel machine via singularized jacobi geometric parameter correction and coupling error evaluation. Robotics and Computer-Integrated Manufacturing, 89, 102771. https://doi.org/10.1016/j.rcim.2024.102771

[12] Sun, H., Qiao, Y., Zhang, Z., Dong, Y., Deng, S., Jin, X., Zhang, C., & Zheng, Z. (2024). Research on geometric error compensation of ultra-precision turning-milling machine tool based on macro–micro composite technology. The International Journal of Advanced Manufacturing Technology, 132(1-2), 365-374. https://doi.org/10.1007/s00170-024-13259-7

[13] Geist, A., Yaqoob, M. F., Friedrich, C., Naumann, C., & Ihlenfeldt, S. (2024). Concept of Integrating a Hybrid Thermal Error Compensation Into an Existing Machine Tool Control Architecture. Journal of Machine Engineering, 24(3), 32-46. https://doi.org/10.36897/jme/192866

[14] Chen, G., Yuan, L., Chen, H., Dou, C., Ma, G., Li, S., & Hu, L. (2025). Machine Tool Spindle Temperature Field Parametric Modeling and Thermal Error Compensation. Lubricants, 13(12), 548. https://doi.org/10.3390/lubricants13120548

[15] Xie, J., Zhou, C., & Xu, H. (2026). Thermal Error Modeling and Compensation for a Motorized Spindle of a Large CNC Gear Profile Grinding Machine Tool. IEEE Access, 14, 53990-54005. https://doi.org/10.1109/access.2026.3681683

[16] Zhao, Z., Mao, J., & Wei, X. (2024). Geometric Error-Based Multi-Source Error Identification and Compensation Strategy for Five-Axis Side Milling. Machines, 12(5), 340. https://doi.org/10.3390/machines12050340

[17] Lang, S., Talleri, S., Mayr, J., Wegener, K., & Bambach, M. (2024). Kalman filter-driven state observer for thermal error compensation in machine tool digital twins. Manufacturing Letters, 41, 208-218. https://doi.org/10.1016/j.mfglet.2024.09.025

[18] Zheng, P., Wang, J., Qiao, Z., Wang, B., Lao, Y., & Wu, D. (2024). Error compensation of five-axis roller machine tool for manufacturing Fresnel mold. Precision Engineering, 88, 986-996. https://doi.org/10.1016/j.precisioneng.2024.04.023

[19] Zhang, Y., Li, C., Cao, H., & Yang, Y. (2025). Geometric error modeling and iterative compensation of internal gear power honing machine under multi-axis coupling. Journal of Manufacturing Processes, 156, 163-179. https://doi.org/10.1016/j.jmapro.2025.10.110

[20] ZHANG, Z., JIANG, F., LUO, M., WU, B., ZHANG, D., & TANG, K. (2024). Geometric error measuring, modeling, and compensation for CNC machine tools: A review. Chinese Journal of Aeronautics, 37(2), 163-198. https://doi.org/10.1016/j.cja.2023.02.035

[21] Guo, S., Si, Z., Sa, R., Zou, Y., & Mei, X. (2024). Geometric error modeling and decoupling identification of rotary axis of five-axis machine tool based on spatial trajectory planning. Measurement, 236, 114887. https://doi.org/10.1016/j.measurement.2024.114887

[22] Cao, Y., Zhao, X., Qu, S., Xing, T., Zong, W., & Sun, T. (2025). Identification of position-related geometric error terms from profile error in four-axis ultra-precision machine tool based on Informer model. Journal of Manufacturing Processes, 134, 943-969. https://doi.org/10.1016/j.jmapro.2025.01.021

[23] Wei, B., Zhu, S., Xie, R., & Bi, Q. (2025). Identification of geometric errors of five-axis machine tool through volumetric error measurement. Journal of Physics: Conference Series, 3091(1), 012001. https://doi.org/10.1088/1742-6596/3091/1/012001

[24] Li, C., Guo, X., Sun, S., Han, R., He, R., & Zhang, Q. (2026). Physics-Guided Multi-Source Fusion Geometric Error Compensation and Knowledge Distillation Verification Platform for Large CNC Machine Tools. Applied and Computational Engineering, 246(1), 122-131. https://doi.org/10.54254/2755-2721/2026.36323

[25] Wu, C., Ji, S., & Zhao, J. (2025). Knowledge and data dual-driven cyber-physics system for intelligent monitoring and compensation of machine tool dynamic coupling error. Engineering Applications of Artificial Intelligence, 158, 111470. https://doi.org/10.1016/j.engappai.2025.111470

[26] Aguado, S., Pueo, M., Acero, R., Brosed, F. J., Santolaria, J., & Aguilar, J. J. (2026). Kinematic modelling and volumetric error compensation of a 3-axis machine tool using a telescopic simultaneous ballbar. Results in Engineering, 29, 109762. https://doi.org/10.1016/j.rineng.2026.109762

[27] Yang, H., Mei, X., Zheng, Y., Jiang, G., Tao, T., & Shi, H. (2026). Digital Twin-Enabled thermal error prediction and compensation system for CNC Machine tool feed systems. Mechanical Systems and Signal Processing, 250, 114192. https://doi.org/10.1016/j.ymssp.2026.114192

[28] Zhang, J. (2026). Intelligent Machine Tool Machining Error and Compensation Model Integrating Backpropagation Neural Network and Cosine Similarity. Journal of The Institution of Engineers (India): Series C, 107(1), 273-285. https://doi.org/10.1007/s40032-025-01270-0

[29] Zhou, B., Chen, G. H., Mao, J., Li, Y., & Zhang, S. W. (2024). A ball screw all-round error compensation technology based on novel hybrid deep learning for CNC machine tool. Journal of Vibroengineering, 26(5), 1216-1236. https://doi.org/10.21595/jve.2024.23841

[30] Li, Z., Dai, Y., liang Guan, C., Lai, T., & Hu, H. (2026). On-machine measurement error modeling and compensation in three-axis machine tools based on measurement error transformation matrices. Precision Engineering. https://doi.org/10.1016/j.precisioneng.2026.01.014