

SYSTEM FOR ROTARY-AXIS DIAGNOSTICS IN FIVE-AXIS CNC MACHINING CENTERS

MARCIN BARSZCZ¹, JERZY JÓZWIK², PAWEŁ TOMIŁO³, DARIA
SAŁAMACHA², KRZYSZTOF DZIEDZIC¹, TOMASZ WIŚNIEWSKI⁶

^{1,5}Department of Computer Science, Lublin University of
Technology, Lublin, Poland

^{2,4}Department of Production Engineering, Lublin University of
Technology, Lublin, Poland

³Department of Marketing, Lublin University of Technology,
Lublin, Poland

⁶Charger, LLC, Świdnik, Poland

DOI: 10.17973/MMSJ.2026_10_2026066

m.barszcz@pollub.pl

Five-axis CNC machine tools are indispensable for manufacturing parts with complex geometry, yet the conversion of raw kinematic measurements into reliable maintenance knowledge is still often fragmented and manual. This article presents a Python-based monitoring environment that reads R-Test export files, invokes one of three selectable machine-learning models – Multilayer Perceptron (MLP), Kolmogorov-Arnold Network (KAN), or Multi-Output Gaussian Process (MOGP) – and automatically produces bilingual diagnostic PDF reports. A heuristic rule layer interprets displacement trajectories and assigns probable fault categories based on signal signatures, without relying on labelled fault datasets. Validation on 1568 R-Test records from two five-axis milling machines shows that MOGP remains the most accurate inference backend. The main contribution of the paper is the integration of data ingestion, inference, heuristic interpretation, visual analytics, and report generation into a deployable condition-monitoring platform rather than a standalone model-benchmarking study.

KEYWORDS

CNC machine tool, rotary axis, R-Test, machine learning, diagnostic system

1 INTRODUCTION

Rotary-axis positioning deviations are among the major contributors to accuracy loss in five-axis CNC machining centres. Mechanical backlash, thermally driven drift, geometric misalignment, and runout propagate through the machine kinematic chain and appear as dimensional deviations in critical workpiece features – such as sockets, bores, and shank interfaces – where tight tolerances determine assembly quality, service life, and operational reliability [Cheluszka 2021]. The R-Test method records tri-axial relative motion of a reference sphere during a complete worktable rotation and is now regarded as a high-information-density calibration approach capable of separating position-independent and position-dependent geometric errors within a single measurement campaign [Hsieh 2024].

The literature on CNC rotary-axis diagnostics may be grouped into three broad areas: measurement, modelling, and compensation [Zhang 2024]. From the measurement standpoint, laser-assisted R-Test procedures enable angular calibration directly linked to controller-level compensation [Hsieh 2024], while transposed-matrix identification approaches improve the separation of linear and rotary error components [Tang 2025] and constraint-aware formulations increase the numerical stability of transformation matrices under non-ideal

installation conditions [Chen 2025]. Simultaneous identification strategies for dual rotary axes reduce calibration effort by combining multiple error sources within one experimental campaign [Yao 2023]. From the modelling perspective, Gaussian-process regression delivers not only point predictions but also calibrated uncertainty estimates [Beckers 2021], and review papers confirm that GP-based models are advantageous when datasets are moderate in size and output variables are correlated [Mu 2025]. Related work has demonstrated uncertainty-gated compensation schemes [Kaftan 2025] as well as real-time surrogate architectures that preserve high predictive fidelity under CNC loop timing constraints [Huang 2025]. The Kolmogorov-Arnold Network (KAN) [Liu 2024] has emerged as a flexible alternative to classical feed-forward neural networks by employing learnable basis functions in place of fixed activations.

Despite these advances, a clear research gap remains: comparatively little attention has been paid to industrial-grade condition-monitoring software that unifies data ingestion, machine-learning inference, heuristic fault interpretation, visual analytics, multilingual reporting, and traceable session management in one deployable application. Most published studies concentrate on calibration accuracy, standalone model benchmarking, or compensation performance, without addressing the software-integration layer that converts ML inference into a practical maintenance workflow. Earlier publications from the same research group established that MOGP provides the most reliable predictions on R-Test datasets [Józwik 2025, Barszcz 2025]; the present article moves beyond model comparison to fill this integration gap.

The specific objectives of this study are therefore: (i) to design and implement a three-module Python-based diagnostic platform covering file ingestion, ML inference, heuristic fault classification, visualisation, and bilingual PDF reporting; (ii) to validate the platform on 1568 R-Test records collected from two industrial five-axis milling machines; and (iii) to demonstrate that MOGP, as the default inference backend, provides the highest and most uniform predictive fidelity across all monitored axes.

2 SYSTEM REQUIREMENTS AND DESIGN GOALS

The system was designed to import one or multiple R-Test files, validate numerical records, execute a user-selected machine-learning model, reconstruct displacement trajectories, assign probable fault categories, visualise the results, export a bilingual PDF report, and preserve session history. Non-functional requirements included runtime language switching, extensibility of the fault taxonomy, deployment on standard Windows and Linux systems, and operation without a dedicated metrology runtime beyond the R-Test device.

3 OBJECT-ORIENTED ARCHITECTURE – THREE-MODULE PIPELINE

The software architecture is organised as a three-stage pipeline. The first module performs data ingestion and validation. The second executes the selected ML model and reconstructs displacement trajectories. The third provides heuristic interpretation, visualisation, and report generation. This separation reduces coupling between pre-processing, inference, and user-facing diagnostics.

4 DATA INGESTION AND FEATURE ENGINEERING

The ingestion layer accepts plain-text R-Test export files and begins parsing after the MEASUREMENT_DATA delimiter. For each valid row, the parser extracts the numerical fields required

for subsequent analysis, while validation routines verify angular consistency and physically plausible displacement ranges; suspicious records are excluded from inference. When several files are loaded simultaneously, the merge stage preserves the process-parameter tuple used during model training. Finally, the input features are linearly scaled to the interval $[-1, 1]$ using extrema derived from the training dataset. Fig. 1 presents example displacement trajectories for the case $R = 150$ mm and $v_f = 2000$ mm/min. X' and Y' exhibit a quasi-sinusoidal form, whereas Z' shows lower amplitude and comparatively denser oscillatory content.

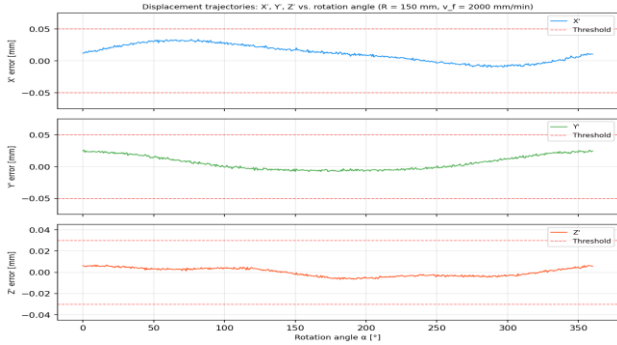


Figure 1. Representative X' , Y' , and Z' trajectories versus rotation angle α at $R = 150$ mm and $v_f = 2000$ mm/min; dashed lines denote tolerance bands

5 MACHINE-LEARNING INFERENCE ENGINE

5.1 Per-Axis Predictive Fidelity

Tab. 1 summarises the R^2 values for each output axis on the held-out test subset. MOGP outperforms the remaining models consistently across X' , Y' , and Z' . By contrast, the MLP model deteriorates most strongly on Z' , which is the least variable and therefore relatively more sensitive to noise. KAN occupies an intermediate position, but its axis-to-axis spread remains larger than that of MOGP, suggesting stronger dependence on the exact spline-knot configuration. The same per-axis comparison is presented graphically in Fig. 2, which highlights the consistency advantage of MOGP across all three outputs.

Model	$R^2 - X'$ axis	$R^2 - Y'$ axis	$R^2 - Z'$ axis	$R^2 - \text{Avg.}$	R^2 range (max-min)
MLP	0.756	0.786	0.740	0.761	0.046 (widest)
KAN	0.983	0.967	0.972	0.974	0.016
MOGP	0.993*	0.991*	0.989*	0.991*	0.004 (narrowest)

Table 1. Per-axis R^2 values for the evaluated ML models on the test subset ($n = 314$)

The * symbol denotes the best-performing model in a given axis. Of particular importance is the intra-model R^2 spread: 0.004 for MOGP, compared with 0.016 for KAN and 0.046 for MLP. Such uniformity indicates that the cross-axis covariance structure captured by MOGP [Bonilla 2007] yields balanced predictions across all outputs, which is practically important because the worktable should be considered conforming only when every monitored axis remains inside tolerance.

Fig. 3 complements the R^2 analysis by comparing the mean percentage error (MPE) and mean squared error (MSE) obtained for the evaluated models.

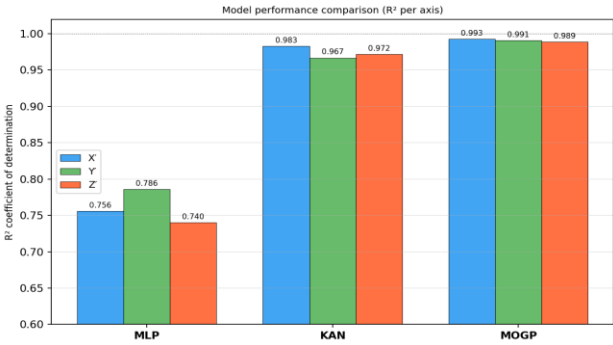


Figure 2. R^2 values for X' , Y' , and Z' obtained with MLP, KAN, and MOGP

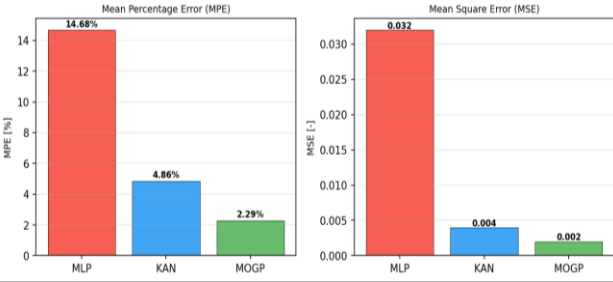


Figure 3. Comparison of MPE and MSE for the evaluated models

5.2 Training Efficiency and Inference Latency

Tab. 2 extends the comparison beyond predictive fidelity and contrasts the three models in terms of convergence behaviour, per-sample inference latency, and the availability of uncertainty estimates. Together, these factors determine which backend is best suited for deployment in a diagnostic application rather than in a purely academic benchmark. The MLP and MOGP backends employ the ADAM optimiser [Kingma 2014], while KAN uses L-BFGS; all backends were run for 100 epochs. Inference latency was measured in Python 3.9 on a standard workstation-class CPU and should therefore be interpreted as an application-level timing indicator rather than as a hardware-independent benchmark.

Property	MLP	KAN	MOGP
Training algorithm	ADAM (100 epochs)	L-BFGS (100 epochs)	ADAM + MLL (100 epochs)
Loss function	MSE	MSE	Marginal Log-Likelihood
Convergence epoch (approx.)	~60 (plateau from epoch 65)	~80 (slow early descent)	~70 (smooth monotone)
Serialised weight file size	~12 kB	~38 kB	~5 kB (kernel params only)
Inference latency (per sample, μs)	< 0.5	~12 (spline eval)	~45 (matrix inversion)
Uncertainty output	No	No	Yes (posterior std. dev.)
Default in application	No	No	Yes

Table 2. Deployment-oriented comparison of the evaluated ML backends

Although MOGP is computationally heavier at inference time ($\sim 45 \mu\text{s}$ per sample versus $< 0.5 \mu\text{s}$ for MLP), this overhead is entirely acceptable in the offline diagnostic scenario considered

here, where a typical analysis covers only hundreds or at most a few thousand samples. More importantly, MOGP uniquely provides posterior stand-ard deviations. These can be converted into confidence intervals and used to tighten or relax acceptance thresholds in angular sectors where predictive uncertainty is elevated, thereby supporting adaptive decision rules at the level of maintenance planning.

6 HEURISTIC FAULT CLASSIFIER

Once ML inference is complete, a rule-based interpretation stage assigns probable fault categories without relying on labelled examples. Because the checks are inde-pendent, more than one label may be assigned. Three signal-level descriptors are evaluated: (i) Backlash: the largest absolute jump around direction-reversal indices is compared with 50% of the axis-specific tolerance ϵ . (ii) Thermal drift: the Pearson correlation between the absolute predicted displacement and the sample index, used here as a time surrogate, is computed; values above 0.7 are treated as indicative of monotonic thermal growth. (iii) Geometric error: spectral energy in the AC part of the FFT exceeding 30% of the dominant spectral peak is treated as indicative of periodic runout. Because these checks are independent, the classifier naturally supports com-pound-fault assignment.

The heuristic layer should be interpreted as a decision-support mechanism based on signal-shape descriptors rather than as a fully validated fault-identification model trained on labelled maintenance cases. The adopted thresholds were selected to pro-vide practically useful discrimination in the present application scenario, but they should be recalibrated when the system is transferred to different machine classes, measurement setups, or maintenance policies.

Fig. 4 illustrates the signature patterns associated with the three principal diag-nostic categories: a step-like amplitude change at reversal points for backlash, a steadily growing envelope for thermal drift, and periodic oscillation at twice the fun-damental frequency for geometric runout.

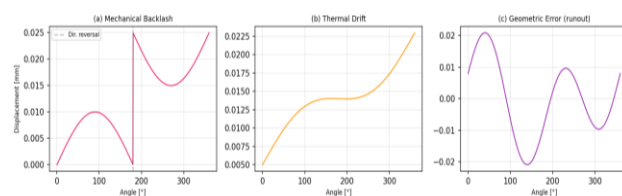


Figure 4. Typical displacement signatures used by the heuristic classifier: (a) mechanical backlash, visible as a sharp jump near direction reversal; (b) thermal drift, represented by a monotonically increasing envelope; (c) geometric error, expressed as periodic oscillation consistent with elliptic runout

7 VISUALISATION SUBSYSTEM AND ERROR MAP

The visualisation subsystem embeds a 2×2 Matplotlib canvas through FigureCanvasTkAgg. The upper-left, upper-right, and lower-left panels display the X' , Y' , and Z' displacement trajectories together with the corresponding tolerance lines, while the lower-right panel contains a two-dimensional contour map of the resultant displacement. To create this map, the system computes $R_d = \sqrt{X'^2 + Y'^2 + Z'^2}$ for each loaded sample, projects the values into the (X', Y') plane, interpolates them over a regular grid using `scipy.interpolate.griddata` with cubic interpolation, and finally renders the surface as a filled contour using the viridis colour scale. This representation exposes angular sectors with elevated positional error and therefore supports targeted maintenance action.

All plot titles, axis labels, and legend entries are obtained from the bilingual translation dictionary and refreshed in place whenever the language toggle is activated. Fig. 5 presents an example contour map produced from MOGP-based predictions.

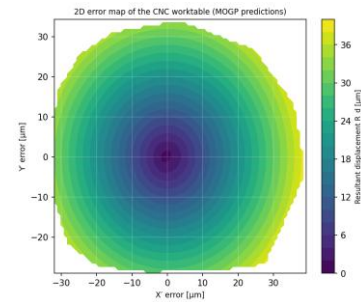


Figure 5. Contour map of resultant worktable displacement R_d [μm] obtained from MOGP predictions

8 REPORT GENERATION AND AUDIT TRAIL

The report-generation component assembles a bilingual PDF document containing the active model, threshold settings, axis-level summary statistics, interpreted fault categories, and the generated figures. To support traceability, each analysis session is stored with its timestamp, file name, selected model, and predicted fault labels.

9 EXPERIMENTAL VALIDATION

9.1 Measurement Campaign

Validation measurements were acquired on two five-axis milling machines at Lublin University of Technology: monoBLOCK 65 and Lasertec 65, both controlled by the Sinumerik 840D family. The R-Test tri-axial capacitive measuring system was employed. The campaign covered complete 360° worktable rotations sampled every 30° for measurement radii $R \in \{75, 150, 225, 300\}$ mm, feed rates $v_f \in \{500, 2000, 3500, 5000\}$ mm/min, and both rotation directions. In total, 1568 measurement records were collected and later divided according to an 80/20 training/test split.

The resulting split yielded 1254 training records and 314 test records. Feature scaling parameters were estimated on the training subset and then applied unchanged to the test subset. The split was performed at the record level and was intended to pro-vide a representative distribution of radii, feed rates, and rotation directions across both subsets.

All experiments were carried out using the same preprocessing and evaluation pipeline for the three compared backends, so that the reported differences reflect model behaviour rather than differences in data handling.

9.2 Statistical Properties of the Error Dataset

Tab. 3 reports the distribution characteristics of the positioning errors after stratification by measurement radius. Presenting the statistics in this way, rather than only in aggregate form, makes it possible to see how radius influences the spread of each axis and which operating conditions reveal the largest geometric deviations of the worktable.

R [mm]	$\sigma(X')$	$ \max (X')$	$\sigma(Y')$	$ \max (Y')$	$\sigma(Z')$	$ \max (Z')$	Dominant axis
75	0.018	0.042	0.018	0.031	0.003	0.007	$X' \approx Y'$
150	0.020	0.058	0.018	0.047	0.004	0.011	X'
225	0.018	0.044	0.015	0.045	0.007	0.020	Y'
300	0.019	0.047	0.020	0.047	0.011	0.022	Z' growing

Table 3. Standard deviation and absolute maximum deviation of X' , Y' , and Z' by measurement radius R

The clearest pattern in Tab. 3 is the growth of Z' variability from a standard deviation of 0.003 mm at $R = 75$ mm to 0.011 mm at $R = 300$ mm, which corresponds to a 3.7-fold increase. By comparison, X' and Y' remain relatively stable over the same range. This suggests that the Z' axis is more strongly amplified at larger radii and that its contribution to the resultant displacement R_d becomes operationally relevant rather than negligible noise when compensation policies are being defined.

9.3 Software Acceptance Testing

Tab. 4 presents the software acceptance-testing matrix. Rather than serving as a simple pass/fail checklist, the table links each function under test to a concrete stimulus and an explicit quantitative acceptance threshold, which makes later regression testing reproducible after code modifications.

Test ID	Function under test	Test stimulus	Acceptance threshold	Result
T-01	File parser: field count	Single R-Test file	Exactly 12 numeric fields per valid row	PASS
T-02	Multi-file merge	4 files (4 radii)	DataFrame row count = sum of individual counts	PASS
T-03	Range validation	File with $\alpha = 400^\circ$ (invalid)	Record excluded; warning displayed	PASS
T-04	Feature normalisation	Raw input dataset	All normalised values $\in [-1.0, 1.0]$	PASS
T-05	MOGP R^2 fidelity	Test split (n = 314)	Avg. $R^2 \geq 0.985$	PASS (0.991)
T-06	KAN R^2 fidelity	Test split (n = 314)	Avg. $R^2 \geq 0.960$	PASS (0.974)
T-07	MLP R^2 fidelity	Test split (n = 314)	Avg. $R^2 \geq 0.700$	PASS (0.761)
T-08	Backlash classifier	Synthetic signal: $\Delta err = 0.6 \cdot \epsilon$ at reversal	BACKLASH label assigned	PASS
T-09	Thermal drift classifier	Synthetic signal: $r = 0.85$	THERMAL_DRIFT label assigned	PASS
T-10	Geometric error classifier	Synthetic $2\times$ periodic signal	GEOMETRIC_ERROR label assigned	PASS
T-11	Bilingual toggle completeness	toggle_language() once	All 28 string tokens updated; no restart	PASS
T-12	PDF report integrity	Complete analysis session	Timestamp, model ID, stats, faults present	PASS
T-13	Session history ordering	Three sequential analyses	Entries in reverse-chronological order	PASS
T-14	MOGP uncertainty output	Single-sample inference	Posterior std. dev. returned per axis	PASS

Table 4. Software acceptance-testing matrix with explicit numerical criteria and associated test stimuli

All 14 acceptance tests were completed successfully. For the heuristic fault classifier, acceptance tests T-08 through T-10 were performed on synthetic signals designed to reproduce canonical signature patterns for backlash, thermal drift, and periodic geometric error. These tests verify internal decision consistency of the rule layer, but they do not replace validation against independently labelled maintenance cases. The thresholds selected for T-05 through T-07 were chosen to preserve a practically meaningful minimum performance level while allowing moderate variability across future measurement campaigns.

10 DISCUSSION

The per-axis R^2 results in Tab. 1 provide a more informative perspective than global averages alone. In particular, the very small intra-model spread observed for MOGP indicates that modelling cross-axis covariance improves predictive stability across all three channels rather than only boosting a single

output. Tab. 2 shows that this advantage is obtained at the cost of higher inference latency; however, the absolute value remains negligible in the offline diagnostic scenario considered here, while the availability of calibrated uncertainty estimates adds practical value for engineering interpretation. At present, the main limitations of the platform are the dependence on manually exported R-Test files and the fact that the heuristic fault-classification layer has not yet been validated against a large labelled set of maintenance cases. Future work should therefore address tighter integration with measurement hardware, for example through socket-based or OPC-UA interfaces, and should examine classifier robustness across broader machine populations, operating regimes, and independently confirmed fault scenarios.

11 CONCLUSIONS

This paper presented a bilingual condition-monitoring platform for CNC rotary-axis diagnostics, integrating ML inference,

heuristic fault classification, visualisation, and PDF reporting into a single deployable application. Among the evaluated backends, MOGP proved most accurate (avg. $R^2 = 0.991$) and is recommended as the default; MLP suits latency-critical scenarios, KAN offers a middle ground. The rule-based classifier detected backlash, thermal drift, and geometric runout without la-belled data, and all 14 acceptance tests passed on 1568 records from two industrial machines. Future work should address direct OPC-UA acquisition and validation across a broader machine population.

ACKNOWLEDGMENTS

The research was funded under the project Lublin University of Technology – Excellent Science: 'Investing in Potential' (grant no. 8/IP/2025/F).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

REFERENCES

- [Barszcz 2025] Barszcz, M., et al. An Integrated System for CNC Machine Tool Rotary-Axis Error Diagnostics with ML Inference. In: Proc. IEEE MetroAeroSpace, 2025 (in press).
- [Beckers 2021] Beckers, T. An introduction to Gaussian process models. arXiv:2102.05497, 2021. <https://doi.org/10.48550/arXiv.2102.05497>.
- [Bonilla 2007] Bonilla, E.V., et al. Multi-task Gaussian process prediction. NeurIPS, 200, Vol. 20. <https://api.semanticscholar.org/CorpusID:10790217>.
- [Cheluszka 2021] Cheluszka, P., et al. Conical picks of mining machines with increased utility properties. Acta Montanistica Slovaca, 2021, Vol.26, No.2, pp 195-204. <https://doi.org/10.46544/AMS.v26i2.02>.
- [Chen 2025] Chen, H., et al. Geometric errors identification of machine tool rotary axes considering R-test theoretical model constraints. Discover Applied Sciences, 2025, Vol.7, p 323. <https://doi.org/10.1007/s42452-025-06807-7>
- [Hsieh 2024] Hsieh, T.H., et al. Geometric error compensation method using the laser R-test. Int. J. Adv. Manuf. Technol., 2024, Vol.31, pp 4035-4053. <https://doi.org/10.1007/s00170-024-13099-5>
- [Huang 2025] Huang, Y., et al. Real-time surrogate compensation architecture for machine-tool thermal error. Discover Applied Sciences, 2025, Vol.7, No.8. <https://doi.org/10.1007/s42452-024-06389-w>.
- [Józwik 2025] Józwik, J., et al. Identification of rotary axis positioning errors in the machining of mining equipment components using ML techniques. Acta Montanistica Slovaca, 2025, Vol.30, No.3, pp 759-775. <https://doi.org/10.46544/AMS.v30i3.16>.
- [Kaftan 2025] Kaftan, P., et al. Reducing thermal errors with confidence: uncertainty-based compensation. CIRP J. Manuf. Sci. Technol., 2025, Vol.54, pp 400-409. <https://doi.org/10.3929/ethz-b-000747253>.
- [Kingma 2014] Kingma, D.P. and Ba, J. Adam: A method for stochastic optimization. ICLR, 2014. arXiv:1412.6980. <https://doi.org/10.48550/arXiv.1412.6980>.
- [Liu 2024] Liu, Z., et al. KAN: Kolmogorov-Arnold Networks. arXiv:2404.19756, 2024. <https://doi.org/10.48550/arXiv.2404.19756>
- [Mu 2025] Mu, S., et al. A review of ML-based thermal error modeling methods for CNC machine tools. Machines, 2025, Vol.13, No.2, p 153. <https://doi.org/10.3390/machines13020153>.
- [Tang 2025] Tang, Y., et al. Measurement and identification of geometric errors in five-axis machine tools based on transposed matrix. Int. J. Adv. Manuf. Technol., 2025. <https://doi.org/10.1007/s00170-025-16692-4>.
- [Yao 2023] Yao, S., et al. Simultaneous identification for geometric error of dual rotary axes in five-axis machine tools. Measurement, 2023, Vol.220, p 113368. <https://doi.org/10.1016/j.measurement.2023.113368>.
- [Zhang 2024] Zhang, Z., et al. Geometric error measuring, modeling, and compensation for CNC machine tools: A review. Chinese J. Aeronautics, 2024, Vol.37, No.2, pp 163-198. <https://doi.org/10.1016/j.cja.2023.02.035>.

CONTACTS:

Marcin Barszcz, Ph.D.

Department of Computer Science, Lublin University of Technology
Nadbystrzycka 36A, 20-618 Lublin, Poland
m.barszcz@pollub.pl | www.pollub.pl

Jerzy Józwik, Prof.

Department of Production Engineering, Lublin University of Technology
Nadbystrzycka 36A, 20-618 Lublin, Poland
j.jozwik@pollub.pl | www.pollub.pl

Paweł Tomiło, Ph.D.

Department of Marketing, Lublin University of Technology
Nadbystrzycka 36A, 20-618 Lublin, Poland
p.tomilo@pollub.pl | www.pollub.pl

Daria Sałamacha, M.Sc.

Department of Production Engineering, Lublin University of Technology
Nadbystrzycka 36A, 20-618 Lublin, Poland
s97261@pollub.pl | www.pollub.pl

Krzysztof Dziedzic, Ph.D.

Department of Computer Science, Lublin University of Technology
Nadbystrzycka 36A, 20-618 Lublin, Poland
k.dziedzic@pollub.pl | www.pollub.pl

Tomasz Wiśniewski, M.Sc.

Charger, LLC
Świdnik, Poland
t.wisniewski@rapsody.com.pl

LICENSE CREATIVE COMMONS:

The article is published under the terms and conditions of the Creative Commons Attribution 4.0 International License (CC BY 4.0).