https://doi.org/10.31449/inf.v50i2.10897 Informatica **50** (2026) 215–220 **215**

# Forecasting Rectangular Pocket Deviations in Birch Plywood for Milling Process Optimization

Primož Kocuvan[1], Rok Struna[2], Vinko Longar[2]
[1]Department of Intelligent Systems, Jožef Stefan Institute, Jamova cesta 39, 1000 Ljubljana, Slovenia
[2]Rudolfovo – znanstveno in tehnološko središče Novo mesto, Podbreznik 15, 8000 Novo mesto, Slovenia
E-mail: primoz.kocuvan@ijs.si, rok.struna@rudolfovo.eu, vinko.longar@rudolfovo.eu

**Keywords:** milling process, regression, rectangular pocket deviation, optimization, augmentation

**Received:** July 4, 2026

*We explored the field of predicting and minimizing dimensional deviations when milling rectangular pockets in birch plywood, with the aim of improving the accuracy of manufacturing processes. We examined various machining conditions altering feed rate, depth of cut, and spindle speed and acquired high-resolution 3D scans to quantify pocket-shape deviations against the prescribed tolerances. To enhance the model's robustness, the empirical dataset was augmented via synthetic data-generation techniques. We then compared several approaches using cross-validation. MLP proved the most accurate. These findings demonstrate the utility of readily available machine-learning algorithms for precise deviation prediction and lay the groundwork for future integration of automated hyperparameter tuning and feature-based process optimization.*

*Povzetek***:** *Raziskava je pokazala, da je mogoče z metodami strojnega učenja, zlasti z modelom MLP, učinkovito napovedovati in zmanjševati dimenzijska odstopanja pri rezkanju vezanega lesa.*

## 1 Introduction

Dimensional accuracy is a central quality requirement in CNC machining of wood-based panels, particularly in furniture, interior products, and modular assemblies where milled pockets, slots, and mating features must fit reliably without excessive post-processing. In plywood and other engineered wood materials, dimensional deviations may arise from the combined effects of cutting parameters, tool deflection, local material heterogeneity, layer orientation, moisture-related variability, and machine dynamics. These deviations are especially relevant in pocket milling, where small errors in edge position or pocket geometry can affect assembly precision, visual quality, and repeatability in batch production. Recent studies on CNC machining of engineered wood panels have shown that parameters such as depth of cut and feed rate can significantly influence dimensional accuracy and surface quality, confirming the need for systematic parameter selection and process control. Although the effects of machining parameters on surface roughness and dimensional accuracy have been studied for several wood-based materials, much of the available work relies on conventional experimental comparison or statistical analysis. Data-driven prediction of local dimensional deviations in plywood pocket milling remains less explored, particularly when high-resolution 3D scanning is used to quantify the machined geometry and when limited experimental datasets are expanded using data-augmentation techniques. This creates a practical research opportunity: if dimensional deviations can be predicted from machining parameters before or during production, CNC settings can be adjusted more systematically to improve accuracy, reduce trial-and-error tuning, and support automated process optimization.

In the DIGITOP project, this problem was investigated through the fabrication of a custom-engineered plywood substrate intended for the assembly of a headphone stand prototype. The substrate consists of a single panel of commercial-grade birch plywood, 10 mm in thickness, with overall dimensions of 100 × 76 cm. Birch was selected on account of its high specific strength and stiffness relative to softwood alternatives, as well as its dimensional stability under controlled indoor conditions; however, its susceptibility to moisture and ultraviolet degradation limits its suitability for long-term exterior applications [1]. Alternative hardwood or engineered wood composites may be substituted without fundamentally altering the downstream processing methodology.

Prior to milling, the full-size panel was subdivided into smaller tiles by $CO_2$ laser cutting. The laser parameters employed were a traverse speed of 3 $mm \cdot s^{-1}$, 95% of a nominal 130 W output power, and an air-assist pressure of 1.5 bar to optimize kerf quality and minimize thermal char. These settings were determined experimentally to achieve clean, burr-free edges while maintaining acceptable processing throughput. In the subsequent milling process, rectangular pockets were produced in the plywood to form the functional geometry of the headphone stand. A 6 mm spiral cutter made of HWM material, i.e., solid tungsten carbide, was used. This cutter was selected for its high wear resistance and stiffness, which minimize tool deflection and contribute to tighter pocket-edge tolerances; preliminary trials showed that softer cutter materials increased radial run-out by up to 15 µm under identical feed conditions.

This paper makes the following contributions:

• A systematic dataset of 17 real and 483 SMOGN-augmented observations linking feed rate, depth of cut, and spindle speed to rectangular-pocket deviation in birch plywood.

• A comparative evaluation of four off-the-shelf regression algorithms using 5-fold cross-validation, identifying MLP as the top-performing model, with $R^2$ = 0.9316.

Compared with existing ML-based machining prediction studies, which most commonly address surface roughness, tool wear, chatter, energy consumption, or dimensional accuracy in metal machining, this work focuses specifically on local dimensional and surface-topography deviations in CNC pocket milling of birch plywood. The novelty of the study lies in combining high-resolution 3D-scanner-derived deviation metrics with a small-data SMOGN augmentation workflow to model a wood-specific machining problem that has received limited attention in the existing ML machining literature.

The study contributes to the state of the art by combining 3D-scanner-based evaluation of plywood surface topography with standard machine learning regression techniques to predict the local-scale variability of dimensional accuracy, expressed through its standard deviation. To the best of the authors' knowledge, the specific combination of the applied 3D scanning setup, plywood material, machined geometry, and data-augmentation pipeline has not been previously reported in this context. The developed regression model provides a basis for automated process optimization by supporting the adjustment of feed rate, depth of cut, and spindle speed to achieve improved surface quality.

# 2 Related work

Birch lumber possesses superior mechanical performance compared to softwood species. With the cross-laminated veneers configuration, birch plywood is considered promising for structural use. Due to its architecture and attendant enhancement of stiffness, strength, and dimensional stability, birch plywood represents a highly promising candidate for load-bearing structural applications [2],[3]. The similar process of milling a rectangular pocket is boring. In the course of bore machining under finishing conditions, the parameters that determine the quality of the surface layer represented by the roughness and quality grade are particularly important. The roughness of the treated surface is mainly affected by the vibrations of the manufacturing system [4]. The boring operation may be performed on a variety of machine-tool platforms, including general-purpose or universal equipment such as lathe (turning) centers and milling machines.

Sutçu and Karagöz evaluated how spindle speed, feed rate, stepover and depth of cut influence the surface topography of pocket-milled MDF using a CNC router. Employing roughness average (Ra), mean peak-to-valley height (Rz) and root-mean-square height (Rq) as quality metrics, they found that raising the spindle speed improves the finish (lower Ra, Rz, Rq), whereas increasing the feed rate, stepover or depth of cut leads to a rougher surface [5].

In a recent investigation of CNC-router machining, a two-factor factorial design was employed to assess how spindle speed, feed rate, cutting depth and wood species affect surface-finish quality. The authors report that the wood's intrinsic structure exerts a significant influence on the machined surface, with higher spindle speeds yielding smoother finishes, whereas increased feed rates produce noticeably rougher textures [6].

İşleyen and Karamanoğlu (2019) conducted a factorial study varying spindle radial speed (18 000 $min^{-1}$ vs. 24 000 $min^{-1}$), feed rate (2 500–10 000 mm/min), tool diameter (4 vs. 6 mm), and depth of cut (4 vs. 6 mm) to evaluate their effects on Ra and Rz surface-roughness parameters of 18 mm MDF using a stylus profilometer (Handysurf E-35) with ten measurements per sample [7].

José and Mathew (2024) propose a comprehensive IIoT-based platform that captures and unifies real-time data from traditional control systems (PLCs, DCSs), mobile devices, and ERP systems, leveraging edge processing alongside cloud-based machine learning and AI to compute and visualize key performance indicators (KPIs). The article has similar objective to ours, because they both apply data-driven, predictive-analytics strategies to improve manufacturing accuracy and efficiency [8].

Feng et al. (2024) and similarly Hwang (2026) reviewed machine-learning applications in wood materials, including property prediction, defect detection, health monitoring, and material-design optimization. They showed that ML can reduce experimental effort by modeling complex links between wood structure, processing, and performance. However, they also identified limited data availability, data quality, and model interpretability as key challenges. Since dimensional accuracy prediction in wood machining was not specifically addressed, this remains a relevant research gap for CNC wood manufacturing [9], [10].

Purwanto et al. (2026) investigated CNC nesting of High Moisture Resistance (HMR) panels, focusing on the effects of depth of cut and feed rate on dimensional accuracy and surface roughness. They found that depth of cut significantly affected dimensional accuracy, while feed rate mainly influenced surface roughness. The best machining condition was obtained at a 2 mm depth of cut and 33 mm/s feed rate. However, the study used experimental and ANOVA-based analysis rather than ML-based prediction, leaving room for data-driven models to predict dimensional deviations in CNC wood machining [11].

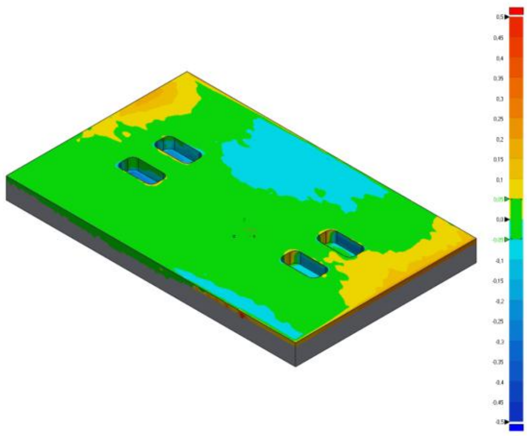
# 3 Experiment

We implemented a controlled laboratory study to predict rectangular pocket deviation as a function of key process variables. Our goal was to identify the combination of settings that minimizes deviation quantified here by the forecasted standard deviation of rectangular pocket diameter, which serves as a proxy for machining error. To that end, we supplied our predictive models with three principal factors known to affect surface roughness and structural integrity within a milled rectangular pocket:

- Feed (move) speed (mm/min). Governs the axial advancement rate of the cutting tool relative to the workpiece, directly influencing heat generation and chip formation.
- Milling depth of spiral cutter (mm). Defines the radial engagement of the cutter per pass; deeper cuts typically increase cutting forces and deflection.
- Spindle radial speed (% of maximum, where 100 % corresponds to 24 000 $min^{-1}$). Controls the circumferential velocity of the tool, affecting surface finish and tool–workpiece interaction dynamics. Here we can also set the radial speed to over 100 %. At some measurements we set it to 120 %.

A 6 mm-diameter spiral cutter (Figure 1 [12]) was used in conjunction with an electrical spindle (Hiteco; Figure 1 [13]). Birch specimens were digitized with a Raptor 3DX Standard (5 MP) structured-light scanner. Scans were executed in automatic mode on the patented 2 + 1-axis robotic cradle (360° rotation, 90° arm swing, ±45° sensor tilt). Using the FOV 330 optics (270 × 200 × 200 mm, point spacing 0.16 mm), a single specimen was captured from ten poses. Raw point clouds were auto-registered in 3DX Scan and exported as STL meshes [14]. The Raptor 3DX system provides a fully automated solution for 3D metrology encompassing quality control and reverse-engineering workflows. Figure 2 presents a representative 3D scan of a birch plywood specimen. Surface deviations within a ±0.1 mm tolerance band are rendered in green; yellow indicates slight positive deviations, red denotes larger positive deviations beyond tolerance, and blue corresponds to negative deviations. The adjacent legend quantifies the deviation thresholds associated with each color.


Figure 1: Spiral cutter 6mm CMT 192.860.11 (right) and



Electrical spindle Hiteco S-13/18 24 ER25 DX BT (left)

Figure 2: Headphone stand made out of Birch plywood - 3D surface (colormap)

Seventeen birch-plywood specimens were prepared, each characterized by a distinct combination of feed (move) speed, milling depth and tool spindle speed. The environment variables are not applicable. Because this sample size was insufficient for robust regression modeling, we augmented the dataset by generating synthetic observations using the SMOGN (Synthetic Minority Over-Sampling Technique for Regression with Gaussian Noise) library in Python [15]. Table 1 summarizes the SMOGN parameter settings. These values were selected to minimize estimation error while promoting a greater dispersion of the synthesized

samples. These SMOGN settings were selected based on a preliminary sensitivity analysis, which showed that a relevance cutoff of 0.90 and noise level of 0.30 best balanced synthetic-data diversity against fidelity to real measurements.
The quality of the synthetically generated records was assessed against the critical process parameter spindle radial speed. Experimental machining trials showed that cutter fracture occurs whenever spindle radial speed falls below 34 %. Accordingly, the data-augmentation pipeline was constrained so that no synthetic sample violated this lower bound, and all subsequent evaluations were performed exclusively within the safe operating range ($\geq$ 34 %). This approach ensures that the augmented dataset reflects only feasible machining conditions and does not introduce unrealistic failure cases that would never be allowed in practice.

The first parameter means each seed point can be interpolated with more distant neighbors, broadening the convex hull in which new points lie hence more dispersed synthetic samples. The second parameter specifies the magnitude of Gaussian jitter added to interpolated points (in our case 30% of the feature range). The third parameter sets the relevance cutoff $\varphi(y) \geq 0.90$ for defining “rare” (to be oversampled) vs. “normal” (to be under-sampled) observations. The fourth chooses between two sampling schemes “balance” or “extreme”.

Table 1: Most important parameters for SMOGN library configuration

| Name of the parameter | Configured value |
|---|---|
| Number of nearest neighbors to use when generating synthetic points via interpolation (default is 5). | 15 |
| Proportion of Gaussian noise (relative to each feature’s standard deviation) | 0.30 |
| Relevance cutoff (0–1) defining which target values are considered “rare” and thus eligible for oversampling. | 0.90 |
| Strategy for selecting which rare cases to oversample here `"extreme"` focuses on the most extreme targets. | “extreme” |

In total, 483 synthetic data points (other 17 a real data points) were produced, a quantity chosen empirically.

# 4 Evaluation and results

We evaluated four regression algorithms multilayer perceptron regression, linear regression, Bayesian ridge regression, and support vector regression using the scikit-learn Python library (latest stable release version 1.9.0 [16]) with different hyperparameter settings. To quantify generalization performance, we performed 5-fold cross-validation by partitioning the dataset into five equal folds and iteratively training each model on four folds and validating on the remaining fold. For each fold we computed the coefficient of determination ($R^2$), defined as (Equation 1) and the mean squared error (MSE) (Equation 2), and then averaged these metrics across folds; the aggregated results are presented in Table 2. An $R^2$ value closer to one indicates better predictive accuracy, whereas a negative $R^2$ implies that a model performs worse than a baseline predictor that always outputs the mean of the target variable. According to the results in Table 2, the MLP regressor model attained the highest mean coefficient of determination ($R^2$) and the lowest mean squared error (MSE). For hyperparameter optimization we applied GridSearchCV using the default parameter grid provided. The tuned models exhibited a mean increase in the coefficient of determination ($R^2$) of just 0.003, an improvement that is practically negligible. Table 3 presents an additional sanity-check experiment conducted exclusively on the 17 real experimental specimens without cross-validation. In this sanity-check experiment, the models were trained on the SMOGN-augmented dataset and evaluated on the 17 original measured specimens, which were not used during training. The purpose of this analysis was to assess whether the predictive behavior observed with the SMOGN-augmented dataset is also reflected when only original measured samples are considered. The strongly negative $R^2$ obtained by the SVR model suggests that its kernel-based representation did not match the structure of this small and augmented dataset, leading to predictions that were less accurate than simply using the mean response value.

Table 2: ML models and results on mixed synthetic and real data

| Model | Mean MSE | Mean $R^2$ |
|---|---|---|
| MLP Regressor | 0.000193 | 0.931626 |
| Bayesian Ridge | 0.000332 | 0.881769 |
| Linear Regression | 0.000332 | 0.871411 |
| SVR | 0.003192 | -0.135114 |

Table 3: ML models and results on real data as a test set (17 specimens)

| Model | Mean MSE | Mean $R^2$ |
|---|---|---|
| MLP Regressor | 0.000072 | 0.929800 |
| Bayesian Ridge | 0.000278 | 0.767827 |
| Linear Regression | 0.000277 | 0.767573 |
| SVR | 0.004445 | -2.712226 |

(1) $R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y}_i)^2}$

(2) $MSE = \frac{1}{n} \sum (y_i - \hat{y}_i)^2,$

where $\hat{y}, \bar{y}$ are predicted values and sample mean of the observation, respectively.

Because the process of 3D scanning the birch plywood specimens is slow (each scan takes about 20 min), we scanned only 17 specimens and augmented additionally 483 specimens.

# 5 Discussion

Our results demonstrate that classical supervised ML techniques can predict the dimensional error incurred during rectangular-pocket milling with low numerical error, while also indicating that there remains scope for further improvement. Specifically, the evaluated models produced mean squared error (MSE) values of very low magnitude, on the order of 0.0001 in the squared-millimeter domain, reflecting the intrinsically small variability of the milling error within the investigated experimental space. This low MSE corresponds to minor deviations in the predicted dimensional error and confirms that both the predictive discrepancies and the actual process deviations were relatively small. However, these performance values should be interpreted with caution. Real and synthetic samples were pooled and randomly shuffled before cross-validation. The folds were generated from the combined dataset, so the assignment of samples to folds was not based on whether samples were real or synthetic. This procedure avoided a systematic separation of real and synthetic samples that could bias or inflate the performance metrics.

Nevertheless, the comparison between model performance on synthetic and real examples suggests that the reported metrics may still be somewhat optimistic. The models generally performed better on the SMOGN-augmented synthetic samples than on the original experimental samples, indicating that the synthetic data may represent a smoother or more regular approximation of the process than the real measurements. Although SMOGN is useful for addressing data imbalance in small experimental datasets, the generated samples are primarily obtained by interpolation within the observed feature space and therefore may not fully capture the complex non-linear physical behavior of the milling process. This is a known limitation in small-data experimental settings, where the available measurements cover only a narrow range of process parameters and where synthetic augmentation cannot replace additional physical experiments.

Despite these limitations, the results provide evidence that ML-based regression models can support the prediction of dimensional deviations in rectangular-pocket milling. The study is limited by the narrow parameter ranges and by the omission of factors such as tool wear, plywood moisture content, and ambient conditions. Future research will extend the dataset to include these variables, explore physics-informed feature engineering, and integrate real-time deviation forecasting into CNC-controller feedback loops for adaptive milling. In a production setting, whether in furniture, cabinetry, or prototype fabrication, integrating such regression models could allow operators to adjust feed rate, depth of cut, and spindle speed to remain within tight tolerance bands, thereby reducing scrap and rework. The approach may also support predictive maintenance by identifying parameter regimes that accelerate tool wear and could be embedded into real-time control loops for adaptive compensation of geometric errors. Beyond cost and time savings, these capabilities may enhance quality consistency across batches of birch plywood components and contribute to the development of more automated, high-precision woodworking workflows. From a practical industrial perspective, these findings indicate that ML-based prediction can support more informed selection of CNC parameters before machining, reducing reliance on trial-and-error setup procedures.

# Acknowledgements

The research was supported by DIGITOP project which is funded by Ministry of Higher Education, Science and Innovation of Slovenia, Slovenian Research and Innovation Agency, and EU NextGenerationEU under Grant TN-06-0106. We thank prof. dr. Matjaž Gams for proof-reading the article and mentorship support within DIGITOP project.

# References

[1] Wang, Yue & Wang, Tianxiang & Crocetti, Roberto & Wålinder, Magnus. (2023). Effect of moisture on the edgewise flexural properties of acetylated and unmodified birch plywood: a comparison of strength, stiffness and brittleness properties. European Journal of Wood and Wood Products. 82. 10.1007/s00107-023-02014-6.

[2] Tianxiang Wang, Yue Wang, Roberto Crocetti, Magnus Wålinder, In-plane mechanical properties

of birch plywood, Construction and Building Materials, Volume 340, 2022, 127852, ISSN 0950-0618, https://doi.org/10.1016/j.conbuildmat.2022.127852.

[3] Yue Wang, Tianxiang Wang, Roberto Crocetti, Michael Schweigler, Magnus Wålinder, Embedment behavior of dowel-type fasteners in birch plywood: Influence of load-to-face grain angle, test set-up, fastener diameter, and acetylation, Construction and Building Materials, Volume 384, 2023, 131440, ISSN 0950-0618, https://doi.org/10.1016/j.conbuildmat.2023.131440.

[4] Ignatov, A.V., Tagil'tsev, S.V. & Namazova, A.I. Bore making using vibration-resistant boring arbors assembled with the use of adhesive materials. *Polym. Sci. Ser. D* **10**, 106–110 (2017). https://doi.org/10.1134/S1995421217020071

[5] Sütcü, A.; Karagöz, Ü. Effect of Machining Parameters on Surface Quality after Face Milling of MDF. *Wood Res.* 2012, 57, 231–240.

[6] Koç, K.H.; Erdinler, E.S.; Hazır, E.; Öztürk, E. Effect of CNC Application Parameters on Wooden Surface Quality. In *Proceedings of the 58th International Convention of Society of Wood Science and Technology*, Grand Teton National Park, Jackson, WY, USA, 7–12 June 2015; pp. 1–8.

[7] İşleyen, Ü.K.; Karamanoğlu, M. The Influence of Machining Parameters on Surface Roughness of MDF in Milling Operation. BioResources 2019, 14, 3266–3277.

[8] Jeeva Jose, Vijo Mathew (2024). Internet of Things – A Model for Data Analytics of KPI Platform in Continuous Process Industry. *Informatica*, 48(1), 119–130. https://doi.org/10.31449/inf.v48i1.3826

[9] Feng, Y., Mekhilef, S., Hui, D., Chow, C. L., & Lau, D. (2024). Machine learning-assisted wood materials: Applications and future prospects. Extreme Mechanics Letters, 71, 102209. https://doi.org/10.1016/j.eml.2024.102209

[10] S.-W. Hwang, "Machine learning-driven research in wood science: from prediction to understanding through the framework of Wood Informatics," *Journal of Wood Science*, vol. 72, article no. 7, 2026, doi: 10.1186/s10086-026-02258-9.

[11] Purwanto, A. A., Amarta, Z., Rahmat, B., Widiyanto, W., & Fahrudin, M. (2026). Effects of depth of cut and feed rate on dimensional accuracy and surface roughness in CNC nesting of HMR panels. Jurnal Polimesin, 24(2), 368–372

[12] https://www.toolnation.com/cmt-19286011-spiral-cutter-6mm-shank-8mm.html, Accessed 20.06.2025

[13] https://www.hiteco.net/en/products/electrospindle-m.t.c.c8247, Accessed 20.06.2025

[14] https://3d-ing.si/wp-content/uploads/2019/03/Raptor3DX_EN_Brochure_181119_low.pdf, Accessed 20.06.2025

[15] https://pypi.org/project/smogn/, Accessed 20.06.2025

[16] https://scikit-learn.org/stable/, Accessed 20.06.2025