HSM2021-2021151

MACHINE DATA-BASED PREDICTION OF BLISK BLADE GEOMETRY CHARACTERISTICS

A. Ernst[1]*, M. Weigold[2]

[1]Rolls-Royce Deutschland Ltd & Co KG, Oberursel, Germany

[2]TU Darmstadt, Institute of Production Management, Technology and Machine Tools, Darmstadt, Germany

*Corresponding author; e-mail: alexander.ernst@rolls-royce.com

Abstract

The increasing availability of data recording solutions in the field of machining in combination with major developments in Machine Learning and Artificial Intelligence enable new approaches towards optimization in the industrial environment. In the aviation industry, critical components must fulfil extremely high quality standards. This requires a stable and error-free manufacturing process, as well as an extensive geometrical compliance, what is until now verified by long-lasting coordinate measuring machine (CMM) inspection. This publication shows how machine data analysis can contribute to reduce CMM measurement effort and thus decrease component cycle time. For this purpose, production machine data from an aircraft engine Inconel compressor blisk blade 5-axis milling operation was recorded and analysed by subsequent application of machine learning algorithms to predict the geometric measurement characteristics.

Keywords:
Quality; Machine Learning; Milling; Aviation Industry

1 INTRODUCTION AND MOTIVATION

In the production of each component in the machining industry, it is necessary to meet the required specifications of the technical drawings in order to ensure the desired functionality and durability. With regard to the geometric requirements, such as form and position tolerances, a measurement of the characteristics is always required in order to provide proof of compliance with the tolerances, which can often be carried out on a random basis. However, the higher the requirements for the component, the higher the percentage of measured components must be, which increases the measurement effort enormously. In the case of a compressor blisk (blade integrated disk) made of a Ni-based superalloy for a passenger aircraft engine, each individual blade of the blisk (Fig. 1) must be measured on the finished product using a coordinate measuring machine (CMM). Depending on the complexity and number of blades to be measured, this takes between six and eight hours. Thus, there is a high savings potential by reducing the conventional measurements.

To achieve this goal, it is necessary to evaluate data from the machining process in order to be able to predict the resulting geometry or the measured inspection characteristics. For this purpose, there are already analytical approaches that attempt to calculate all influencing variables by means of models in order to regress the geometry [Soori 2017] [Zeroudi 2012]. These

models are able to predict the geometrical deviation for simpler processes with certain boundary conditions, as not all influencing factors that occur in series production can be covered.

Fig. 1: High pressure compressor blisk; surfaces machined by the blade milling process are shown in red.

In addition, there are approaches that combine analytics with machine learning models in order to simplify the calculations. In this case, machining variables, like material removal rate, are calculated using a specially developed software and the resulting geometry is then predicted using machine learning models [Denkena 2016]. Similarly, in [Brecher 2019], calculations, such as determining the tool

speed, are carried out first before predicting the geometry using neural networks. [Klein 2020] does not calculate cutting parameters at all to regress geometrical characteristics with machine learning, although the basis is a honing process with a lower level of complexity than the considered milling process. These investigations show that machine learning models are generally capable of regressing geometries of machining processes. Therefore, this approach is investigated in the following for the prediction of geometrical characteristics of milled blisk blades. Due to the complexity of the process, the attempt is to avoid analytical models in this approach.

Apart from the machine data, a decisive point for the resulting final geometry is the milling tool. There are already various approaches to predict the wear condition of tools, for example [Wiesch 2019]. However, this is based on the fact that the tools have different initial states or exhibit different wear behaviour. Due to low manufacturing tolerances in tool production with close-meshed controls as well as a fixed number of machining operations per tool, the influencing factor of the tool can be regarded as constant in the investigated blisk manufacturing process and can therefore be left out of the equation.

2 MACHINING PROCESS

The blades of a blisk are milled out of a solid material disc. Most of the material removal is done by roughing processes, for which different strategies can be used [Klocke 2012], followed by a geometry-forming finishing process. In general, a helical path is followed from the tip to the bottom of the blade (Fig. 2), through simultaneous axes movements, to machine the final geometry with the tolerances shown in Tab. 1.

Previous investigations have shown that the influence of the rough machining is negligible, which means that only the finishing process is responsible for the geometry formation, which is therefore the core for the data evaluation in the following. The machining process is executed on a 5-axis milling machine with a Siemens 840D SolutionLine control and takes approximately one day in total. With the size of the manufactured blades, the machine inaccuracy is less than 2 μm.
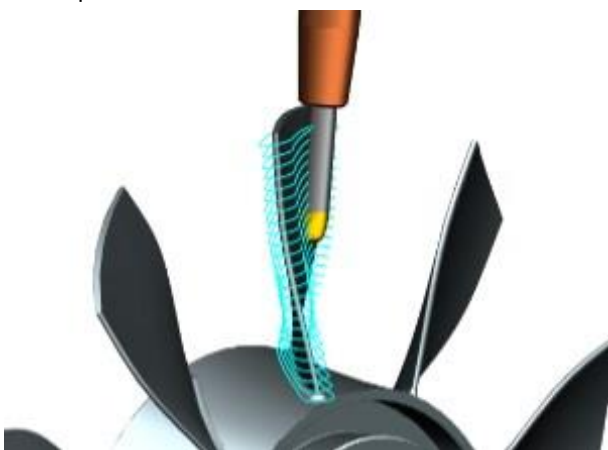

Fig. 2: Movements of the finishing milling tool [MMS 2019].

3 DATA RECORDING

In a research environment, often many sensors are attached to the tool, the workpiece or generally in the machining area in order to record data during the machining process and gain information about the tool or the process itself. However, this approach can only be implemented to a limited extent in industrial series production. Due to fixture changes, tool changes and the guarantee of a smooth process, additional sensors cannot be installed close to the machining process. For this reason, the approach to use machine data only is pursued, which can already be provided by default as a basis for data analysis. This means that no changes have to be applied to an already established process and thus no further influencing factor is added that could be a disruptive factor for industrial series production.

The Marposs Datalogger is used for data recording, which is installed on a laptop with connection to the machine via Ethernet. Using the S7COMM protocol, the following parameters can be recorded for each axis during the machining process:

- Actual position:
  Direct axis measurement system
- Target position:
  Setpoint of the interpolator
- Encoder position:
  Indirect axis measurement system
- Current consumption:
  Individually for each axis and the spindle
- Temperature compensation:
  Individually for each axis and adjusted during machining
- Overall compensation:
  Cross error compensation (changes due to tool position) plus temperature compensation

In addition, the length of the tool and the NC code lines are recorded. The corresponding controller variable names, which are necessary for the recording, are taken from [Siemens 2017]. In total a number of 74 parameters are recorded with a mean distance between two data points of 10 ms and saved as semi-structured data (csv files).

After the machining process, the finished blisk is measured using a tactile probe on a coordinate measuring machine. The measurement process consists of several radial and axial plane scans of each blade which are used to evaluate the geometric characteristic deviations, listed in Tab. 1 with the associated typical tolerances and the standard deviations of the datasets. As an example, the measured variables chord length and stagger angle are illustrated on a blade in Fig. 3. The chord length parameter can be measured with an accuracy of 1.6 μm as an example. A CAQ system is used as measurement data storage, whereby the data subsequently can be merged with the machine data for the following analysis (Fig. 4).
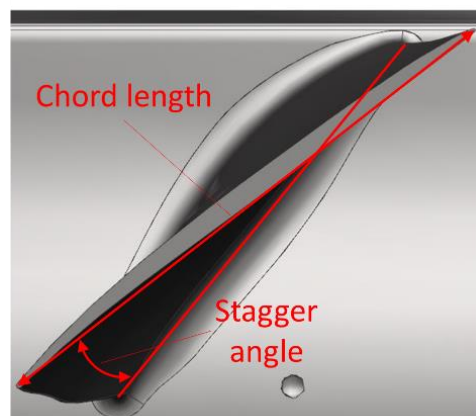

Fig. 3: Measured geometric characteristics examples from top view of the blade.

The recorded data all belong to the same component type with the identical raw material supplier and were all produced on the same machine. In total, the data of eleven blisks with 51 blades each are available, resulting in 561 datasets.
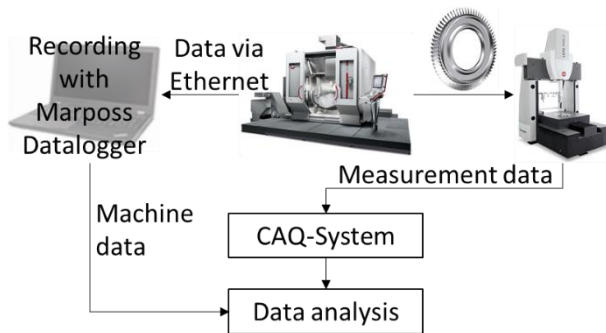

Fig. 4: Scheme of the data flow for model training and testing.

4 DATA PREPERATION

Due to the recording methodology and the process duration, the dataset of one component is split over several csv files. In the first step of data preparation, the corresponding data for one component must be identified and combined, followed by the marking of the finishing operations for further analysis. These two steps can be realised by means of unique NC code lines which occur in every recording. The final step of data slicing is to determine the associated data for the individual blade cycles. However, the slower data recording frequency of 10 ms compared to the machine's interpolation cycle of 2 ms poses a challenge. It is obvious that not every interpolation step of the machine can be recorded and consequently not every programmed line of the NC code, depending on the execution time. This is particularly the case with highly dynamic and fast machine movements, as here in the finishing process. Thus, special care must be taken to select code lines that are executed slowly enough to occur in each recording in order to perform an automated cutting process and thus ensure a constant basis for the following steps.
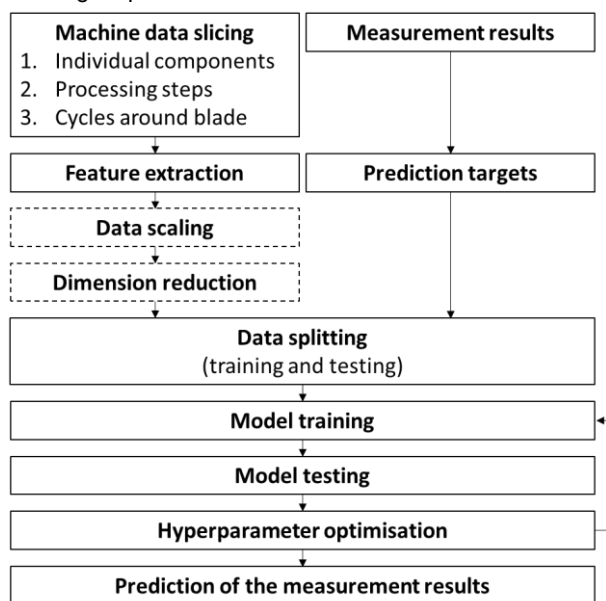

Fig. 5: Approach from data preparation to the optimisation of the regression models.

Subsequently, the deviations of the actual positions and the target positions as well as the encoder positions are calculated. From these results and the other recorded parameters, statistical values, such as maximum, minimum, median, mean, weighted average, standard deviation, variance, range, percentile, sum, skewness and kurtosis, are calculated afterwards and summarised into a feature dataset. Due to the recording rate, a frequency analysis of the data is not useful. The calculated features are scaled in the following with different methods to improve the model performance: scaling by minimum and maximum (MinMax), scaling by the absolute maximum (MaxAbs), standardisation (Standard), scaling by the range between the 25% and 75% of the data (Robust), normalisation and scaling by quantiles (Quantile). In the last step of data preparation, it is possible to reduce the number of feature dimensions, for which a principal component analysis (PCA) [Howley 2006] is used. The influence of a dimension reduction is explained below. Features and the associated targets result together in a complete dataset, which can then be split into the training and the test dataset. The entire process of data preparation up to the finished model is shown in Fig. 5.

Tab. 1: Measurement characteristics with corresponding tolerances and the measurement standard deviation of the used dataset.

Inspection characteristic	Upper tolerance [mm]	Lower tolerance [mm]	Standard deviation [mm]
Stacking axis	0.300	-0.300	0.071
Blade profile	0.100	-0.100	0.014
Chord length	0.300	-0.300	0.014
Stagger angel	0.400	-0.400	0.039
Bow	0.300	0.000	0.005
Blade thickness	0.200	-0.200	0.033
Leading edge A	0.200	-0.200	0.014
Leading edge B	0.170	-0.170	0.008
Leading edge C	0.150	-0.150	0.008

5 PREDICTION OF THE GEOMETRIC CHARACTERISTICS

5.1 Comparison of different regression models

The prepared data is divided in a ratio of 75/25 for the first tests. This results in a number of 420 datasets to train the following models:

- Multi-layer perceptron regressor (MLPR)
- Decision tree regressor (DTR)
- Support vector regressor (SVR)
- k-nearest neighbours regressor (KNNR)
- Gaussian process regressor (GPR)
- Partial least squares regressor (PLSR)
- Random forest regressor (RFR)

The hyperparameters of the models are all optimised using a randomised and a grid search with cross-validation. Tab. 2 shows the calculated mean absolute error (MAE) values over all geometrical characteristics for the optimised models using the scaled and non-scaled datasets. Among the four best results (marked in green) are the PLSR model with the non-scaled dataset, the GPR model with the MinMax-scaled dataset, the KNNR model and the RFR model each with the normalised dataset. Altogether, the RFR model produced the best results with a MAE of only 0.061 mm. As not only the MAE over all geometrical characteristics is decisive for a good regression result,

Tab. 2: Mean absolute error (calculated from all nine features from Tab. 1) of all optimised regression models and all different scaled datasets.

Model	MAE [mm]						
	Non-scaled	MinMax	MaxAbs	Standard	Robust	Normalised	Quantile
MLP	0.0100	0.0085	0.0113	0.0111	0.0239	0.0113	0.0086
DTR	0.0095	0.0095	0.0096	0.0099	0.0089	0.0089	0.0102
SVR	0.0164	0.0168	0.0162	0.0170	0.0168	0.0160	0.0169
KNNR	0.0066	0.0075	0.0079	0.0079	0.0135	0.0065	0.0075
GPR	0.0192	0.0077	0.0081	0.0560	0.4111	0.0103	0.0077
PLSR	0.0078	0.0078	0.0078	0.0078	0.0078	0.0085	0.0078
RFR	0.0064	0.0064	0.0064	0.0065	0.0064	0.0061	0.0065

further investigation of the predictions of the best models will be carried out. Thus, in the next step, the characteristics are considered individually and the standard deviation of the MAE as well as the maximum error (MaxE) are added to the evaluation (Tab. 3).

The largest MAEs and MaxEs are generated in the prediction of the stagger angle and the stacking axis. However, both characteristics are dependent on the raw material properties, especially the residual stresses released during the machining process. Since the material data are not included in the regression, this might result in larger errors. The RFR and the PLSR model show very good predictions regarding all other characteristics, resulting in MAEs in the single-digit micrometre range.

Tab. 3: MAE, MaxE and standard deviation of the MAE of the prediction for all geometry characteristics (all values in mm). The best result for each characteristic is marked green.

		KNNR	GPR	PLSR	RFR
Stacking axis	MAE	0.014	0.030	0.023	0.013
	STD	0.025	0.033	0.029	0.018
	MaxE	0.363	0.174	0.200	0.158
Blade profile	MAE	0.003	0.016	0.003	0.003
	STD	0.003	0.010	0.002	0.003
	MaxE	0.032	0.070	0.016	0.034
Chord length	MAE	0.006	0.017	0.006	0.006
	STD	0.005	0.014	0.005	0.005
	MaxE	0.036	0.068	0.038	0.036
Stagger angle	MAE	0.014	0.029	0.019	0.013
	STD	0.014	0.023	0.016	0.012
	MaxE	0.091	0.134	0.096	0.063
Bow	MAE	0.002	0.007	0.003	0.002
	STD	0.002	0.004	0.002	0.002
	MaxE	0.019	0.021	0.010	0.010
Blade thickness	MAE	0.006	0.017	0.005	0.006
	STD	0.007	0.016	0.004	0.006
	MaxE	0.082	0.102	0.034	0.079
Leading edge A	MAE	0.004	0.019	0.006	0.004
	STD	0.004	0.011	0.004	0.003
	MaxE	0.021	0.043	0.024	0.017
Leading edge B	MAE	0.004	0.019	0.005	0.004
	STD	0.003	0.014	0.003	0.003
	MaxE	0.018	0.056	0.015	0.016
Leading edge C	MAE	0.004	0.031	0.004	0.004
	STD	0.003	0.017	0.003	0.003
	MaxE	0.019	0.068	0.014	0.017

Also considering the MaxEs and the according tolerances from Tab. 1, the results are very accurate. The best algorithm overall, the RFR model, does not deliver the best results for all individual characteristics. For the determination of the blade profile and the blade thickness, the PLSR model generates the better predictions. The best total result can therefore be achieved by a combination of RFR and PLSR.

In order to understand the approach of the RFR model and to gain an insight into the decision-making process, the proportionally most important features can be read out. This analysis shows that in particular the comparisons of the encoder positions with the actual positions turn out to be most decisive, especially the features of the rotational axis of the swing table.

5.2 Influence of dimension reduction

To investigate the influence of a dimension reduction, the RFR model is used as a baseline, as it generally performs best. With the help of a PCA, the cumulative explained variances are calculated for all datasets, as shown in Fig. 7. This indicates the extent to which the variance of the complete dataset can be reflected by the reduced dimensions. Fig. 7 illustrate that for the non-scaled, normalised and robust-scaled dataset, a small number of dimensions, between five and 20, should be sufficient to achieve a good regression performance.
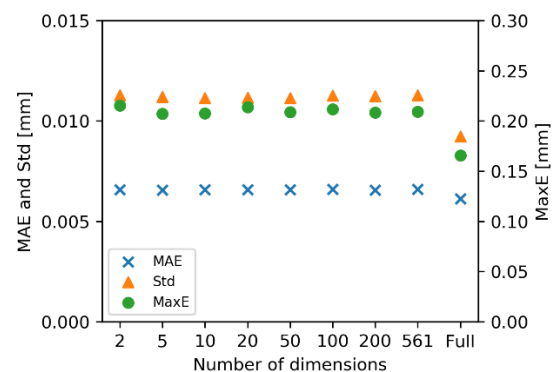

Fig. 6: Comparison of the RFR model predictions based on different datasets generated with varying numbers of dimensions (PCA), based on the normalised dataset.

To test this assumption, the RFR model is trained with the normalised dataset for different numbers of dimensions. The outcomes (Fig. 6) show that there is a negligible difference in performance for all used numbers of dimensions.
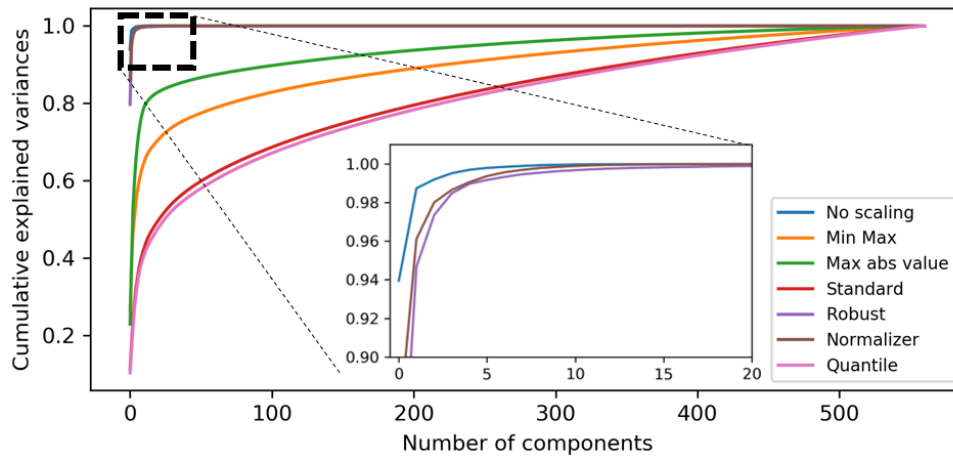
Fig. 7: Cumulative variance of all datasets as well as the original dataset for all possible PCA number of components.

However, the non-reduced dataset produces an even better result than all reduced datasets. Consequently, it can be stated that good results can already be achieved with a small number of dimensions, which are completely sufficient for quick analyses. However, if the geometric deviations are to be predicted as accurately as possible, especially to minimize the MaxE, it is essential to use the non-reduced dataset.

5.3 Influence of quantity of training data

Another influencing factor on the results of the regression models is the splitting of the data. To investigate this point, Fig. 8 shows the RFR results for different divisions of the normalised dataset. It becomes clear that a minimum of the prediction error is reached at a proportion of 75%, which corresponds to the already used 420 training datasets. Consequently, it can be stated that underfitting occurs below this value and overfitting occurs above it.
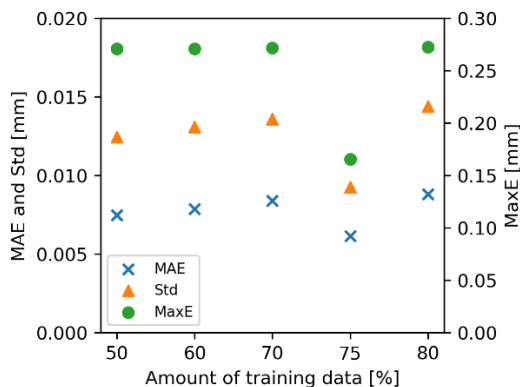

Fig. 8: Comparison of different train-test-split ratios for the RFR model with the normalised dataset.

6 SUMMARY AND OUTLOOK

This paper describes the prediction of blade geometry characteristics based on machine data from blisk series production, without the usage of additional sensors. The deviations of the machined geometry can be predicted very precisely with machine learning models. A combination of RFR and PLSR models is particularly suitable for the regression of conventional blade measurements, considering the right data selection and preparation. The results presented here provide a good baseline for future measurement reduction in blisk blade series production,

however, further steps are necessary to achieve the objective:

1. In addition to temperature compensation, the inclusion of temperature values
2. Improvement of the prediction of stacking axis and stagger angle
3. Proof of the prediction accuracy of the algorithms for a further manufactured number of pieces
4. Proof of the independence of the prediction from the raw material supplier
5. Proof of the independence of the prediction from the machine (series production on several identical machines)
6. Definition of confidence limits for the prediction
7. Fulfilment of the requirements in [SAE 2015] for the introduction of an alternative inspection frequency plan

As already mentioned in chapter 5.1, it is reasonable to assume that by including the raw material properties in the data analysis, the prediction errors of the stacking axis and the stagger angle could be reduced. This approach could also have a positive effect on the independence of the prediction from different raw material suppliers and therefore represents the next development step. In addition, other methods of data preparation will be investigated and the relevance of other machine parameters will be examined to improve the prediction as well as to generalise the method for further component types and machines.

In addition, the decision-making process of the RFR shows that further investigations need to pay closer attention to the rotational axis of the machine's swing table.

7 ACKNOWLEDGMENTS

8 REFERENCES

[Brecher 2019] C. Brecher, J. Ochel, et al. Merkmalsbasierte Qualitätsprädiktion durch maschinelles Lernen. Zeitschrift für wirtschaftlichen Fabrikbetrieb, 2019, 114 (11), pp. 784-787. DOI: 10.3139/104.112177.

[Denkena 2016] B. Denkena, M.-A. Dittrich, et al. Augmenting Milling Process Data for Shape Error Prediction. Procedia CIRP, 2016, 57, pp. 487-491. DOI: 10.1016/j.procir.2016.11.084.

[Howley 2006] T. Howley, M. G. Madden, et al. The effect of principal component analysis on machine learning accuracy with high-dimensional spectral data. Knowledge-Based System, 2009, 19 (5), pp. 262-270. DOI: 10.1016/j.knosys.2005.11.014.

[Klein 2020] S. Klein, S. Schorr, et al. Quality Prediction of Honed Bores with Machine Learning Based on Machining and Quality Data to Improve the Honing Process Control. Procedia CIRP, 2020, 93, pp. 1322-1327. DOI: 10.1016/j.procir.2020.03.055.

[Klocke 2012] F. Klocke, M. Zeis, et al. Technological and Economical Comparison of Roughing Strategies via Milling, EDM and ECM for Titanium- and Nickel-based Blisks. Procedia CIRP, 2012, 2, pp. 98-101, DOI: 10.1016/j.procir.2012.05.048.

[MMS 2019] Modern Machine Shop, Six Ways to Make 5-Axis Machining More Productive. Online available: www.mmsonline.com/articles/six-ways-to-make-5-axis-machining-more-productive, last updated: 8. November 2019, accessed: 18. May 2021.

[SAE 2015] SAE International, Requirements for Developing and Qualifying Alternate Inspection Frequency Plans, 2015, USA, Issue 2015-03.

[Siemens 2017] Siemens AG, SINUMERIK 840D sl NC-Variablen und Nahtstellensignale – Listenhandbuch, 2017.

[Soori 2017] M. Soori, B. Arezoo, et al. Accuracy analysis of tool deflection error modelling in prediction of milled surfaces by a virtual machining system. International Journal of Computer Applications in Technology, 2017, 55 (4), pp. 308-321. DOI: 10.1504/IJCAT.2017.086015.

[Wiesch 2019] M. Wiesch, A. Epple, et al. Tool Monitoring – A Scalable Learning Approach to Estimate Cutting Tool Conditions with Machine-Internal Data in Job Shop Production of a Milling Process. Proceedings of the 8th Congress of the German Academic Association for Production Technology (WGP), Aachen, 2018, pp. 101-111. DOI: 10.1007/978-3-030-03451-1_11.

[Zeroudi 2012] N. Zeroudi, M. Fontaine, Prediction of machined surface geometry based on analytical modelling of ball-end milling. Procedia CIRP, 2012, 1, pp. 108-113. DOI: 10.1016/j.procir.2012.04.017.