Design of a Physics-Based and Data-Driven Hybrid Model for Predictive Maintenance

Emiliano Traini(✉), Giulia Bruno(✉), and Franco Lombardi(✉)

Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Turin, Italy
{emiliano.traini,giulia.bruno, franco.lombardi}@polito.it

Abstract. The maintenance process is crucial in any system that is prone to failure or degradation, particularly in manufacturing operations. In fact, maintenance costs can reach up to 40% of the cost of production in certain industries. In the era of Industry 4.0, maintenance methods can maximize the use of components predicting the remaining useful life. These methods are identified as Predictive Maintenance and include several innovative technologies, such as IoT for deploying sensors that monitor machines and AI that provides the algorithms to interpret the data collected. The information generated from sensor data allows for more accurate predictions using statistical models that are sensitive to the peculiarities of an individual tool set on a particular machine and used by a certain operator. These models, unlike traditional methods based on physical laws, increase in efficiency as the data increases, and therefore are not efficient or usable when a sufficient bank of data is not available. This work proposes a hybrid model that, being based on both classical physics and data-drive models, demonstrates how it is possible to obtain a prediction method that estimates the state of the tool even in the absence of historical data and that increases its accuracy as such data increases. The proposed model is evaluated by using a public experimental milling dataset.

Keywords: Hybrid model · Predictive Maintenance · Tool condition monitoring · Machine learning · Milling

1 Introduction

Maintenance costs are estimated as a percentage of production costs that vary between 15%, for the manufacturing sector in general, and up to 40% for the metalworking industry [1]. With the proper implementation of Predictive Maintenance (PdM) strategies, these costs can be reduced by up to 30% [2], by automatizing part of monitoring activities, and by optimizing the decision of replacing the resources when strictly necessary. Furthermore, a PdM strategy can reduce the incidence of failures by up to 70%, allowing the productive time of systems to be increased by up to 30% [3]. Another important estimation is that PdM methods based on Machine Learning (ML) algorithms can reduce current maintenance costs by an additional 30%, increasing machine operating life and reducing downtime [4].

Published by Springer Nature Switzerland AG 2021
A. Dolgui et al. (Eds.): APMS 2021, IFIP AICT 634, pp. 536–543, 2021.
https://doi.org/10.1007/978-3-030-85914-5_57

In the era of Industry 4.0, several technologies, as Internet of Things (IoT) and Artificial Intelligence (AI), enable the real-time data collection required by high-performed PdM methods. In fact, they are based on real-time data collected by a sensors system and Big Data infrastructures. [5] Research has proposed several AI-based methods whose performance grows as the information possessed about the process under observation increases, but they are inoperable with no real-time information.

The performance of physics-based methods depends on (i) the number of variables (sources of variability) considered in the model, (ii) the complexity of the physical laws, and (iii) the estimation quality of few parameters with data offline generated by experiments and inspections. Contrarily, the accuracy of data-driven methods depends on the quantity and quality of historical data, which are difficult to replicate for research analyses [6]. The aim of this work is to propose a hybrid model that takes advantage of the strengths of both methods (physics-based and data-driven) while minimizing the effect of their weaknesses.

The rest of the paper is organized as follows: the second section describes the state of the art, the third one defines the proposed hybrid model, the fourth introduces the milling process case study, and, finally, the last section presents the conclusions and the ideas for future improvements.

2 State of the Art

Starting in 1950, preventive maintenance was introduced in order to limit the effects of a failure, which with the previous approach often led to downtime of the entire production process [7]. The first preventive maintenance methods were based only on time schedules, and for that they are called periodic maintenance. However, these approaches failed to predict abnormal failures and often led to unnecessary interventions.

Differently, it has been estimated that 99% of mechanical failures can be predicted with the help of specific indicators, on this basis was born the Condition Based Maintenance (CBM) [8]. It involves two main processes: diagnostic and prognostic [9]. The improvement of these methods is represented by PdM models, in which measurements on the machine are used in combination with process performance data measured by other devices. The use of such data jointly allows statistical models to analyze historical trends in order to predict the instant when the machine needs an intervention [10].

Prognostics methods can be categorized into data-driven, physics-based and hybrid approaches [11]. Despite the significant recent progress in the Model-Based (MB) and ML hybrid modeling domain, there are various challenges that throttle down the full-fledged growth of hybrid modeling: (i) there is no guidelines for selecting hybrid models, (ii) there are few benchmarks (problems and dataset) for evaluating and comparing hybrid models, (iii) training accurate models with low amount of data or labels, (iv) minimizing data collection costs, (v) solving the complexity due to geometric data formats as CAD files and imbalanced data [12].

3 Proposed Hybrid Model

The proposed method is a hybrid model between physics-based and data-driven approaches, and it aims to exploit the potential of each method.

The first step consists in training the physics-based model and data-driven model individually. In this way, the two methods generate estimations about wear levels (W) or Remaining useful Life (RUL) for each T-th run of a tool, called $W_{PB}(T)$ and $W_{ML}(T)$, respectively.

Then, always with the training set, for each run, the optimal weight $\omega(T)$ is calculated to generate the linear combinations of physics-based and data-driven predictions as stated in the following equation:

$$W = \omega W_{PB} + (1 - \omega) W_{ML}.$$

The weights are defined by choosing as objective to minimize the Root-Mean Square Error (RMSE) and the Root Relative Squared Error (RRSE) of the hybrid predictions (W).

Finally, the trained hybrid model is evaluated on the test set. The estimation of the wear level $W(T)$ during the T-th operation performed with the same tool is a weighted average of physics-based and data driven methods. However, the model estimates a dynamic weight $\omega := \omega(T, P)$ that depends on the number of operations performed by the same tool and on the process parameters P set by the operator.

The details on the two components of the hybrid model are reported in the following.

3.1 Physics-Based Component

Classical mechanics methods concerning the wear of rotating machine tools are based on the non-linear relationship between two main parameters: cutting speed (V_c) and RUL (T). A first equation was proposed by Taylor in 1906 and has the following form: $V_c \cdot T^{1/\beta} = C$ and $C = \alpha{\cdot}f^{-C/\beta} \cdot d^{-\gamma/\beta}$, where C is the cutting speed with which one minute of life is obtained, d is the depth of cut, f is the feed rate and α, β and γ are empirical constants. $1/\beta$ is an indicator of how much the tool life is affected by changes in cutting speed and empirical data has defined that $1/\beta \in [0.1; 0.15]$ for HS steel tools, $1/\beta \in [0.2; 0.25]$ for carbide tools and $1/\beta \in [0.6; 1]$ for ceramic tools [13].

The physics-based component of the hybrid model (called W_{PB}) is based on the extended Taylor equation for rotary tools, which includes all machining parameters. The equation has the following form: $T = \alpha_0 v_c^{\alpha_1} f^{\alpha_2} d^{\alpha_3} W_{PB}^{\alpha_4}$, where $[T] = [min]$ is the estimation of the useful life of the tool, $[v_c] = [mm/min]$ is the cutting speed, $[f] = [mm/rev]$ is the feed rate, $[d] = [mm]$ is the dept of cut. $[W_{PB}] = [mm]$ is the width of flank wear according to the physics-based method that can be measured in relation to the activity time T, and α_0, α_1, α_2, α_3 and α_4 are empirical constants to be estimated that will be called coefficients, since they represent the coefficients in the multiple regression model. The formula is the following:

$$W_{PB}(T) = e^{[\ln(T) - \ln(\alpha_0) - \alpha_1 \ln(v_c) - \alpha_2 \ln(f) - \alpha_3 \ln(d)]/\alpha_4}$$

The estimation of the coefficients can be done with a multiple linear regression by performing a logarithmic transformation, which can be written in matrix form.

$$Y = \begin{bmatrix} \ln(T_1) \\ \vdots \\ \ln(T_n) \end{bmatrix} = X\alpha = \begin{bmatrix} 1 & \ln(v_{c,1}) & \ln(f_1) & \ln(d_1) & \ln(W_{PB,1}) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \ln(v_{c,n}) & \ln(f_n) & \ln(d_n) & \ln(W_{PB,n}) \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix}$$

X is called sensitivity matrix [14] and α contains the coefficients that can be estimated with the method of least squares. The accuracy of the estimation of coefficients depends on the inverse of the matrix $X^T X$, the term on which the optimization criteria are based, as for example the one proposed by [15] that define the minimum data sample to be collected for training the method.

3.2 Data-Driven Component

In parallel, ML methods that can utilize the information contained in sensor measurements can be used to estimate tool wear W_{ML}. The framework used to estimate tool wear from sensor measurements is shown by [16] and [17] and it is based on a multi-step data pre-processing between the data acquisition and monitoring processes.

The first phase of data pre-processing, data cleaning, involves removing values that do not meet certain requirements and are inconsistent with requirements. The second phase is the outlier detection where outliers are the values that deviate strongly from other points in the sample. In the proposed framework, the distinction between searching for outliers within a sensor measurement and between measurements is considered, to recognize extreme signals that indicate process instability. After that, for the sensor measurement, considered as time series, the stationary window selection step is required, which allows the stationary phase of the machine to be selected using the Change Point Detection (CPD) technique.

When data are cleaned, the feature extraction phase follows, and it generates a complex set of predictors. This operation allows to reduce the computational load required by the algorithms to be implemented later and improves the speed of machine learning processes. The features obtained in the previous phase are normalized. Then, in the feature selection step, the large number of features extracted are reduced applying unsupervised methods described in [16].

Finally, a machine learning algorithm is applied to predict the tool wear W_{ML} based on the data preprocessed.

4 Case Study: Milling Process

4.1 The Milling Dataset

The analyses performed in this research work are based on the public Milling dataset, made available by the Prognostic Center of Excellence NASA-PCoE [18]. It contains the values recorded by six sensors throughout the life cycle of 16 tools (cases), under different working conditions identified with 8 scenarios, for a total of 170 machining operations (runs). Each case is characterized by three machining parameters, which follow the recommendations of the tool manufacturer, and by the type of material machined with fixed dimensions (483 mm $\times$ 178 mm $\times$ 51 mm). Except for the cutting speed, which remains unchanged at $200\,\mathrm{m/min}$, the other variables are dichotomous and in particular: the feed rate has been fixed at $0.25\,\mathrm{mm/s}$ or $0.5\,\mathrm{mm/s}$, the depth of cut has been fixed at $0.75\,\mathrm{mm}$ or $1.5\,\mathrm{mm}$, while the workpiece materials are cast iron or stainless steel 145.

While the sensors collected data continuously, both during operation and downtime, flank wear measurements (*VB*) were made periodically, removing the insert from the cutter and measuring the distance from the cutting edge to the end of the abrasive wear on the side face of the tool with a special microscope. Given the diversity of machining parameters, cycles with different activity time and number of machining operations were obtained for each tool.

4.2 Comparisons Between Hybrid Model and Single Ones

A comparison of the performance of tool wear prediction on this dataset is provided in [16]. Among different ML algorithms, Neural Network (NN) was the one with the best performance. For this reason, this model was chosen as the data-driven method to be used in the hybrid model.

In Fig. 1, the weights obtained with a training set composed by 10 tools are represented (ω values). In the later runs, the weights reflect the results obtained with the error analysis, in which there is an intermediate linear phase (from run 13 to run 19) in which the Taylor model is more accurate and therefore the weights tend to 1, while in the outer phases the data-driven model based on the NN algorithm prevails.

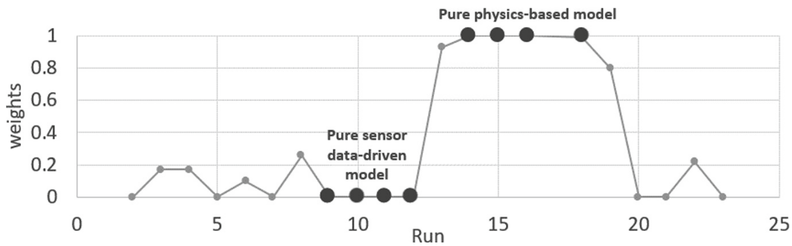

Fig. 1. Weights obtained on the training set.

Figure 2 shows the overall errors of the hybrid model and the single models. In terms of the RMSE, the hybrid model has a similar distribution to the NN model, slightly shifted downward and with one less outlier. While analyzing the distribution of the RRSE, the distribution of the hybrid model has a median value similar to the single models but with much less dispersion. It can be inferred that the hybrid model only slightly improves the overall accuracy of the predictions, while the main advantage of this approach is to obtain a more robust method whose goodness of predictions is less affected by the training data than the single models.

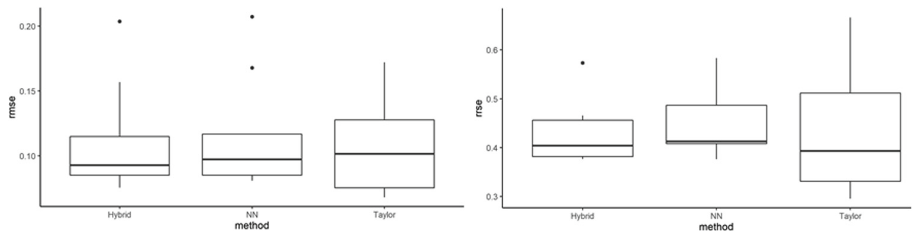

Fig. 2. Hybrid model performance compared to individual model performance.

As shown in Fig. 3, the performance of the hybrid and related single models was analyzed as the size of the training set varied. The error measured with the RRSE metric is always smaller with the hybrid model than with the single models, although the differences are more pronounced with larger training sets. While, analyzing the values of the metric RMSE, it was found that with a training set of small size (composed of 8/14 tools) is more accurate the Taylor model because there are not enough data to train the neural network, which has a much higher error than the physics-based method, and then the hybrid model has intermediate performance between them. While increasing the size of the training set the error of the NN model becomes very similar to that of the Taylor model and consequently the hybrid model obtains better performances than the single models. Moreover, it can be observed that with the hybrid model the increase in accuracy with the increase of the tools used in the training set is greater than that obtained by the single models.

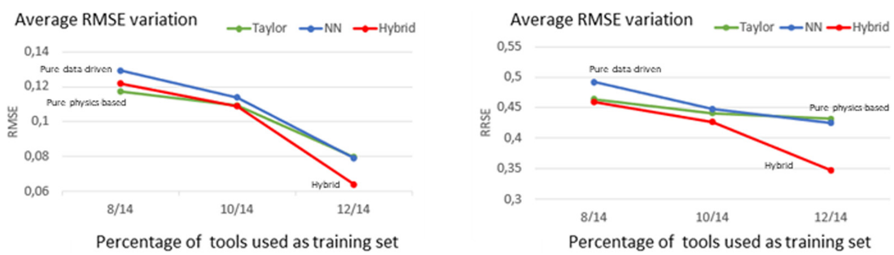

Fig. 3. Performance comparison of models for different training set size

4.3 Discussion of the Results

The results obtained is a primary formulation of a hybrid model applied on a real manufacturing case study, with the aim of monitoring the status of a CNC machine tool through a combination of an NN model with the extended Taylor law. Considering the two approaches applied on the case study, the physics-based methods result the most accurate and robust, in fact, the best overall prediction performance was obtained with the Taylor model. It has the advantage that it can be implemented even in the absence of sensors on board the machine and with few offline measurements to train the model. The main limitation of this method is that it can only be applied to wear phenomena for which a mathematical law describing the trend is known, i.e., only for common wear metrics. The category of data-driven methods based on NN model shows it potential during the last milling runs, i.e., with enough sensor data. This characteristic reflects the ability of such methods to explain high variable trends using the information provided by sensors. Limitations of data-driven methods are the need of installing many sensors on the machines for real-time monitoring, and then the high initial investment.

5 Conclusions and Future Works

The results show that the proposed hybrid approach, defined by the linear combination of a physics-based and a data-driven method, has the best performance than either single method. When single models achieve similar performance, the hybrid approach allows to significantly increase the overall accuracy and specially to obtain much more robust results. In addition, this approach can be used as an unique model to estimate tools that are monitored both offline of along each operation.

With further research, the approach can be validated on other case studies, even on different manufacturing processes and with different wear metrics. In turn, the hybrid model can be extended to predict the RUL of each tool after each run, in term of remaining runs or estimating the remaining time of usage. In addition, it is possible to implement a hybrid approach between a time series method (starting with a simple autoregressive one) and the two used in this work. Other improvements can be done to the hybrid model, such as considering other estimations as input nodes of the NN model.

Acknowledgments. This work has been funded by the Ministero dell'Istruzione, dell'Università e della Ricerca, Grant/Award Number: TESUN-83486178370409, finanziamento dipartimenti di eccellenza CAP. 1694 TIT. 232 ART. 6.

References

1. Agogino, A., Goebel, K.: Milling Data Set, NASA Ames Prognostics Data Repository. BEST lab, UC Berkeley. https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/
2. Heng, A., Zhang, S., Tan, A.C.C., Mathew, J.: Rotating machinery prognostics: state of the art, challenges and opportunities. Mech. Syst. Signal Process. **23**, 724–739 (2009)

3. Dos Santos, A.L.B., Duarte, M., Abrao, A.M., Machado, A.: An optimisation procedure to determine the coefficients of the extended Taylor's equation in machining. Int. J. Mach. Tools Manuf. **39**, 17–31 (1999)
4. Cline, B., Niculescu, R.S., Huffman, D., Deckel, B.: Predictive maintenance applications for machine learning, Orlando, FL, USA. IEEE (2017)
5. An, D., Kim, N.H., Choi, J.: Practical options for selecting data-driven or physics-based prognostics. Reliab. Eng. Syst. Saf. **133**, 223–236 (2015)
6. Johanssona, D., Hägglundb, S., Bushlya, V., Ståhla, J.: Assessment of commonly used tool life models in metal cutting. Procedia Manuf. **11**, 602–609 (2017)
7. Traini, E., Bruno, G., Lombardi, F.: Tool condition monitoring framework for predictive maintenance: a case study on milling process. Int. J. Prod. Res. (2020)
8. Traini, E., Bruno, G., D'Antonio, G., Lombardi, F.: Machine learning framework for predictive maintenance in milling. IFAC-PapersOnLine **52**, 177–182 (2019)
9. Trojan, F., Marçal, R.F.M.: Proposal of maintenance-types classification to clarify maintenance concepts in production and operations management. J. Bus. Econ. **8**, 560–572 (2017)
10. Sullivan, G.P., Pugh, R., Melendez, A.P., Hunt, W.D.: Operations & maintenance best practices - a guide to achieving operational efficiency (release 3.0). Pacific Northwest National Laboratory (2010)
11. Bloch, H.P., Geitner, F.K.: Machinery Failure Analysis and Troubleshooting. Butterworth-Heinemann (2012)
12. Hale, G.: InTech survey: predictive maintenance top technology challenge in 2007. InTech (2007)
13. Daily, J., Peterson, J.: Predictive maintenance: how big data analysis can improve maintenance. In: Richter, K., Walther, J. (eds.) Supply Chain Integration Challenges in Commercial Aerospace, pp. 267–278. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-46155-7_18
14. Arruda, J.R.F., Duarte, M.: Updating rotor-bearing finite element models using experimental frequency response functions. AIAA J. (1993)
15. Xu, L.D., Xu, E.L., Li, L.: Industry 4.0: state of the art and future trends. IJPR **56**, 2941–2962 (2018)
16. Lei, Y.: An improved exponential model for predicting remaining useful life of rolling element bearings. IEEE Trans. Ind. Electron. **62**, 7762–7773 (2015)
17. Liang, Z.: A study of tool life sensitivity to cutting speed. In: ASEE PEER (2002)
18. Liulys, K.: Machine learning application in predictive maintenance. IEEE (2019)
19. Kan, M.S., Tan, A.C.C., Mathew, J.: A review on prognostic techniques for non-stationary and non-linear rotating systems. Mech. Syst. Signal Process. **62–63**, 1–20 (2015)
20. Mobley, R.K.: An Introduction to Predictive Maintenance. Elsevier, Amsterdam (2002)
21. Zhang, Q., Tse, P.W., Wan, X., Xu, G.: Remaining useful life estimation for mechanical systems based on similarity of phase space trajectory. Expert Syst. Appl. **42**, 2353–2360 (2015)
22. Rai, R., Sahu, C.K.: Driven by data or derived through physics? A review of hybrid physics guided machine learning techniques with cyber-physical system (CPS) focus. IEEE Access **8**, 71050–71073 (2020)
23. Lei, Y., Lin, J., He, Z., Zuo, J.: A review on empirical mode decomposition in fault diagnosis of rotating machinery. Mech. Syst. Signal Process. **35**, 108–126 (2013)