ICIC Express Letters
Volume **12**, Number **11**, November **2018**

ICIC International ©2018 ISSN 1881-803X
pp. 1145–1153

A MODEL-DRIVEN PREDICTIVE ANALYTICS APPROACH FOR MACHINING TIME USING HISTORICAL MACHINE-MONITORING DATA

Prita Meilanitasari[1] and Seung-Jun Shin[2,*]

[1]Graduate School of Management of Technology
Hanyang University
222, Wangsimni-ro, Seongdong-gu, Seoul 04763, Korea
prita.meilanitasari@gmail.com

[2]Graduate School of Management of Technology
Pukyong National University
365, Sinseon-ro, Nam-gu, Busan 48547, Korea
*Corresponding author: sjshin22@gmail.com

Received April 2018; accepted July 2018

Abstract. *Machine-monitoring data represent the reality on machine shops where uncertainties exist due to variety of machining conditions. Hence, predicting machining time using the machine-specific monitoring data in individual machine tools enables to reduce the uncertainties, thereby building up precise process planning and execution. However, previous research focused on creating time prediction models at the product or machining feature level and, did not work properly in different forms of products or changes in geometry of machining features. To overcome this limitation, it is necessary to create granular prediction models at the tool path level for gaining accurate time anticipation through creating and composing these granular models along with sequential tool movements. For such purpose, this study presents a model-driven methodology for machining time prediction using historical machine-monitoring data. The methodology consists of: 1) data processing, 2) predictive analytics modeling, and 3) model verification and validation. A case study is demonstrated to show the effectiveness of the proposed methodology.*

Keywords: Machine-monitoring data, Machining time, Predictive analytics, Machine learning

1. **Introduction.** The machine system parameters provide a large amount of data and information about the machine's condition, maintenance requirement, quality and efficiency. Those parameters are being designed to meet the requirements of high-speed and high-precision production which are essential in this era of smart manufacturing [1]. To support the smart manufacturing technology, machine-monitoring is important. Machine-monitoring gives the ability of measuring the machine's performance in real time. Machine-monitoring obviously outputs the measured datasets called machine-monitoring data. The machine-monitoring data collected are valuable because they can be used to make data-driven prediction for efficient shop floor operations [2]. Machine-monitoring also can be used to improve manufacturing performance by supplying accurate productivity metrics to improve operations and make better decisions. Data gathered in this study come from machine-monitoring dedicated to executing a Numerical Control (NC) program in machine tools [3].

Machine-monitoring data, which records the time span consumed for fabricating a part by the execution of an NC program, can be used to accurately estimate machining time. The machining time largely relies on the characteristics of the process parameters to

DOI: 10.24507/icicel.12.11.1145

be decided in process planning. In milling machine tools, standard machining time is determined by such process parameters as feed rate, cutting depth, cutting speed [4]. As automation and mechanization increasingly require faster response to orders, predicting machining time becomes more critical for deciding accurate delivery time, manufacturing cost, and production/process planning [5].

In the previous literature, Zhou et al. [6] minimized total completion time in the hierarchical uniform machine. They use total completion of machining time to make a manufacturing scheduling. Pfeiffer et al. [7] estimated manufacturing lead time. They use statistical and simulation methods to make prediction by discrete event simulation model data. Both researches focused on how to minimize machining time at the production-level. Coelho et al. [5] presented a practical mechanistic method for milling time estimation. They developed a software that estimated real machining time more accurately for free-form geometries. However, their study is only effective in linear tool path interpolation (G01), but not in circular interpolations (G02/G03). Monreal and Rodriguez [8] discussed the influences of the tool path strategy on the cycle time of high speed milling operation by constructing a mechanic model for zig-zag tool path in the pocketing operations only. Previous studies have contributed to providing good prediction models for machining time; however, their models have been made within the product or machining-feature level. These prediction models at product and feature levels accurately work for designated products or sets of machining features. However, when the geometry of a product or the geometric parameters of a certain machining feature change, their prediction may lose the accuracy due to the dependency of those models with such high-level stratification.

The differences of our study with the previous studies are to: 1) use the historical machine-monitoring data, which have been collected and accumulated from real machines' operations, and 2) generate granular prediction models that can anticipate machining time at tool path levels. The purpose study of this research is to propose a model-driven methodology for machining time prediction using historical machine monitoring data. The application of the proposed methodology can make better accuracy in time prediction through the decomposition of such models into tool-path levels and thus reduce uncertainties in predictive process planning. This paper is structured as follows: Section 2 proposes the methodology; Section 3 explains a case study; Section 4 discusses the case study; Section 5 remarks conclusions.

2. **Methodology.** This section describes the proposed methodology for machining time prediction based on predictive analytics with the use of historical machine-monitoring data. Our goal is to predict machining time at the tool path level on a milling machine along with sequential tool movements. For clear understanding, we use an example of three

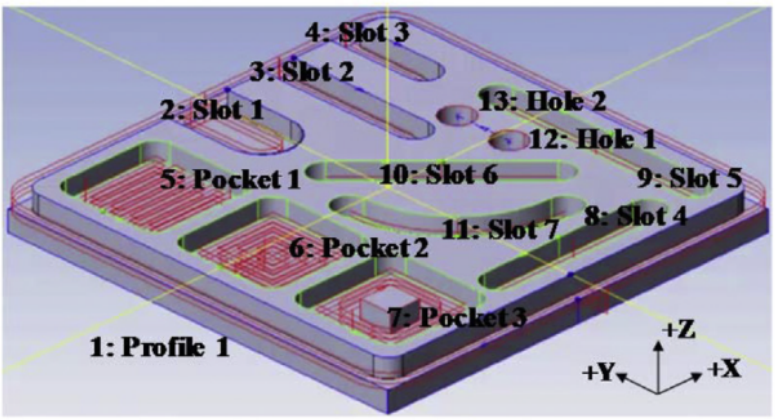

FIGURE 1. Machining part [9]

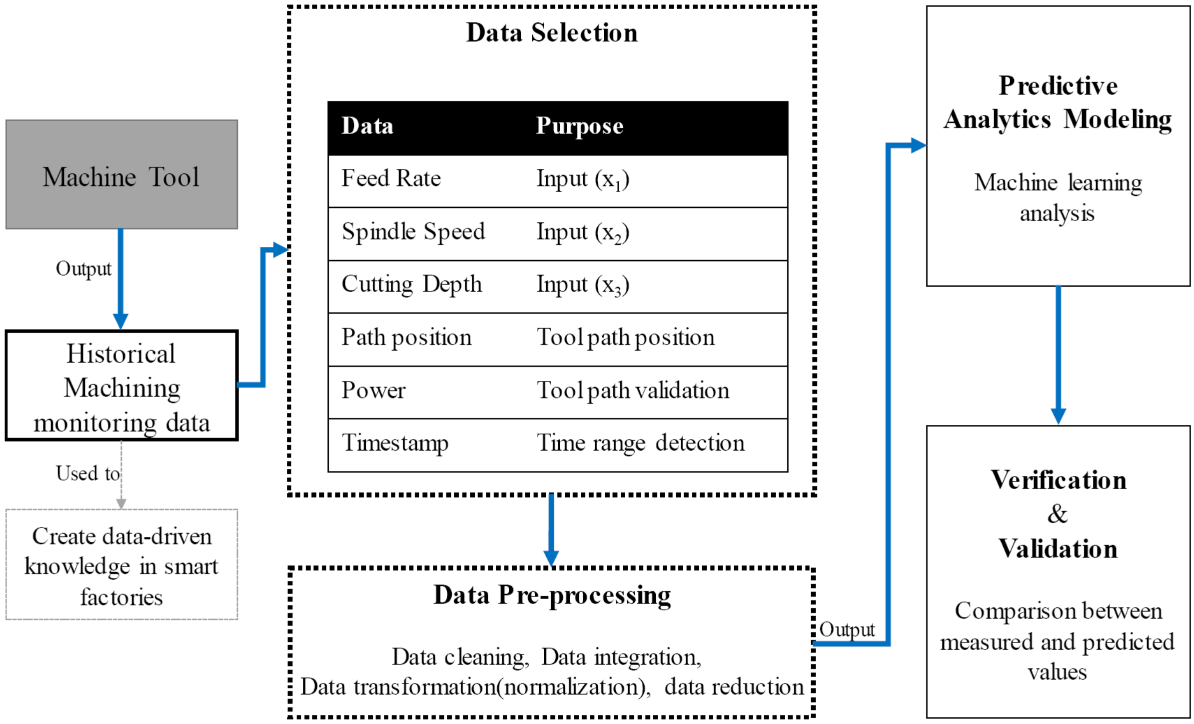

FIGURE 2. A predictive analytics modeling methodology

out of thirteen machining features and their associated roughing operations as shown in Figure 1. Figure 2 shows the methodology that consists of data processing, predictive analytics modeling, and verification and validation. The following subsections explain the details of methodology.

2.1. **Data processing.** Data processing deals with data selection and data pre-processing. The data selection identifies the data parameters that should be extracted from historical machine-monitoring data. Data selection can be made by finding the relationship between input and output parameters at a given machining condition. In this paper, we identify Machining Configuration (MC) parameters as well as input and output parameters in data selection. MC parameters specify a certain machining condition where a prediction model can be applied. That means an identical set of MC parameters uses the same prediction model, but a different set of them should use a different model due to their different machining conditions [9]. MC parameters can be identified as a set of a machining operation, command, trajectory and tool path, as presented in Table 1. 'Machining operation' can be normally obtained from process plan data, whereas 'command', 'trajectory' and 'tool path' from NC program data.

Regarding input parameters, we use three main process parameters (feed rate, spindle speed, and cutting depth). It comes from that these parameters are controllable factors

TABLE 1. Machine Configuration (MC) parameters [9]

Parameter	Examples
Operation	Contouring, pocketing, slotting, drilling
Command	Rapid (G00), linear feed (G01), clockwise-circular (G02), counterclockwise-circular (G03)
Trajectory	X-direction, Y-direction, feed, back, step over, approach, retract, circular
Tool path	Approach, Yfeed, Xfeed, Cfeed, Feed, Step, Back, Retract

and significantly affect machining time [9]. The output parameter is set to be machining time. Additionally, we extract path position (x, y, z) parameters to detect an accurate position of each machining feature, and power parameter to detect the stroke count and to check the validity of position, and timestamp parameter to decide the starting and finish time at each tool path.

Data pre-processing is the process of transiting "dirty" and "untidy" into "clean" and "tidy" datasets [10]. It comprises data cleaning, data integration, data transformation and reduction. Raw machine-monitoring data unintentionally contain missing, erroneous or inconsistent data due to lack in certain attributes of interest, errors in data transmission, faults in technological limitations of measurements, or human errors [10]. Hence, data cleaning is necessary for reducing the uncertainty in data through handling such missing, erroneous or inconsistent data. In this study, the missing data found are neglected via deletion methods if it is detected that the rest of data at the same MC parameter set are enough.

Data integration combines various data sources into a consistent dataset to make it simple and complete. In this study, for example, we combine several datasets acquired from three different MC sets into one dataset to make a prediction model for 'approach' tool path, which totally depends on machine tool's performance not on feed rate, spindle speed and cutting depth. Data transformation adjusts data values into the designated format, scale or unit to provide more suitability for predictive analytics modeling. Here, the input parameters (feed rate, spindle speed, and cutting depth) and the output parameter (time) are scaled-down to the -1-to-1 range from their original values using z-score normalization. Data reduction removes redundant datasets with satisfaction of data integrity.

2.2. **Predictive analytics modeling.** Once the data processing is complete, input-and-output datasets are prepared to generate predictive models via machine learning techniques, which learns from training datasets so that it outputs predictive values. Various machine learning techniques like regression and artificial neural network are available. Equation (1) expresses the 2^{nd}-order polynomial regression-based model that figures out the relationships between the input (x: feed rate, spindle speed, and cutting depth) and the output (y_t: time-range on each stroke) parameters. This equation predicts the time-range for each stroke, i.e., single tool movement instructed by an NC block, on a set of MC parameters.

$$y_t = \alpha_0 + \sum_{i=1}^{n} \alpha_i x_i + \sum_{j=1}^{n} \alpha_j x_j^2 + \varepsilon \tag{1}$$

where α: coefficient, n: number of x variables, ε: error.

Because roughing operations generally contain multiple strokes correspondent to individual cutting depth layers, Equation (1) can be adjusted to Equation (2) that expresses the total time-range for each set of MC parameters, which multiplies the number of strokes (s) given for removing a machining feature.

$$y_{total} = s\left(\alpha_0 + \sum_{i=1}^{n} \alpha_i x_i + \sum_{j=1}^{n} \alpha_j x_j^2 + \varepsilon\right) \tag{2}$$

2.3. **Verification and validation.** Verification and validation is necessary to measure and compare predicted machining time values with real ones (the values actually observed). For such purpose, we can calculate Root Mean Square Error (RMSE), as expressed in Equation (3). RMSE measures the differences between individual measured and predicted machining time values. Prediction models make better performance to predict correctly as RMSE values are closer to 0. We can also calculate Total Relative Error

(TRE), which measures the total difference between measured and predicted machining time.

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{n}} \tag{3}$$

where $\hat{y}_i$: predicted value, y_i: real value, n: the number of pairwise datasets.

3. Case Study.

This section describes a case study to demonstrate the feasibility and effectiveness of our methodology.

3.1. Experimental setup.

The objective of this case study is to predict the machining time consumed during tool path movements. We use the machining part as shown in Figure 1 and only deal with the three machining features (Profile 1, Pocket 1 and Slot 2) for problem simplification. Figure 3 shows the configuration of tool path movements in Profile 1, Pocket 1 and Slot 2. Table 2 presents the sequence of tool path movements for the three machining features. We perform actual machining to gather real datasets. Table 3 presents a list of process parameters for twelve trials. The experimental environments are: Mori Seiki NVD 1500DCG for a machine tool, Fanuc 0i for a numerical controller, cold finish mild steel 1018 for workpiece material, $10.16\text{cm} * 10.16\text{cm} * 1.27\text{cm}$ for workpiece geometry, solid carbide flat end mill for a cutting tool, 8mm diameter and 4 numbers of flutes for cutting tool geometry, and systems insights high speed power meter for power

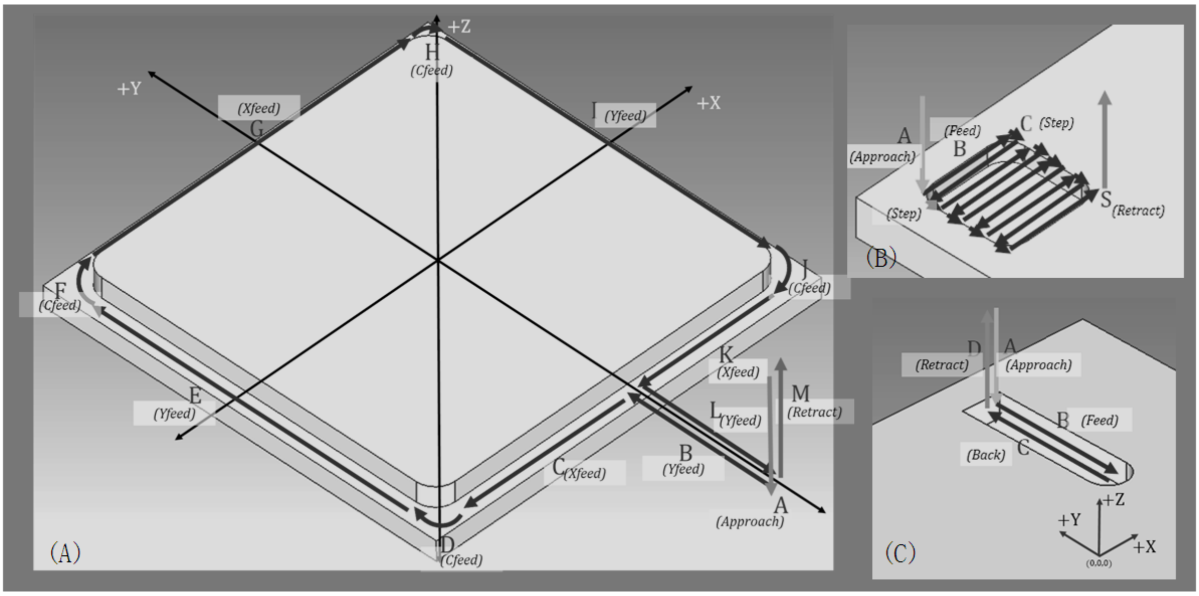

FIGURE 3. Tool path movement for (A) Profile 1; (B) Pocket 1; (C) Slot 2

TABLE 2. Tool path movement sequence on Profile 1, Pocket 1, and Slot 2

Machining feature	Tool path movement (in sequence)
Profile 1	Approach (A) → Yfeed (B) → Xfeed (C) → Cfeed (D) → Yfeed (E) → Cfeed (F) → Xfeed (G) → Cfeed (H) → Yfeed (I) → Cfeed (J) → Xfeed (K) → Yfeed (L) → Retract (M)
Pocket 1	Approach (A) → Feed (B) → Step (C) → Feed (D) → Step (E) → Feed (F) → Step (G) → Feed (H) → Step (I) → Feed (J) → Step (K) → Feed (L) → Step (M) → Feed (N) → Step (O) → Feed (P) → Step (Q) → Feed (R) → Retract (S)
Slot 2	Approach (A) → Feed (B) → Back (C) → Retract (D)

TABLE 3. A list of process parameters

Process Parameter \ Trial	1	2	3	4	5	6	7	8	9	10	11	12
Feed rate (mm/tooth)	0.0005	0.0005	0.0005	0.0009	0.0005	0.0007	0.0007	0.0007	0.0007	0.0003	0.0006	0.0005
Spindle speed (RPM)	1500	2000	1750	1750	1750	1500	2000	2000	1750	1750	1750	1750
Cutting depth (mm)	1.5	1.5	1	1	2	1	1	2	1.5	1.5	1.5	1.5

measurement. Cutting width is fixed as the diameter of the cutting tool (8mm). The MC parameters (operation, command, trajectory, and tool path) vary depending on the machining configurations given by individual NC blocks.

3.2. **Data acquisition.** During actual machining, machine-monitoring data are gathered through an MTConnect agent, i.e., software that receives and stores a time series of data samples or events and acts as a bridge between a machine and a client application [12]. Together, the process planning and NC programming data corresponded to the machine-monitoring data are obtained. The process planning data include the technical description for the experimental setup described in Section 3.1. The NC programming data suitable for the numerical controller (Fanuc 0i) are created based on the process planning data. Here, operation sequence describes sequential tool path movements along with the execution of NC blocks. In the technical manners described in Section 2, data pre-processing is performed to make clean and tidy training datasets.

3.3. **Modeling.** Machining time-predictive models are created using the training datasets gathered by the technical descriptions in Section 2. The number of predictive models depends on the types of tool path movements on individual machining features. In Table 2, for example, Profile 1 consists of seven predictive models: 'App', 'Profile1-Yfeed-1', 'Profile1-XFeed-1', 'Profile1-Cfeed', 'Profile1-Yfeed-2', 'Profile1-XFeed-2', and 'Ret' while Pocket 1 comprises four models: 'App', 'Pocket1-Feed', 'Pocket1-Step', and 'Ret'. Here, the predictive models associated with 'App' and 'Ret' are identically applied in both Profile 1 and Pocket 1 because they do not relate to machining features. Meanwhile, the rest of the predictive models except those of 'App' and 'Ret' need to be differentially made because their machining time largely depends on the types of machining features and tool path movements. In other words, Equation (2) needs to be individually applied with regard to each set of MC parameters. The following equations are the 2nd order polynomial regression models for machining time prediction in Profile 1, Pocket 1, and Slot 2. It is noted that we apply a −1-to-1 normalization to adjusting data scales for disperse data distributions in feed rate, spindle speed and cutting depth.

$$\{\text{G01}\}:\quad y_{t(\text{App})} = 0.711 - 0.022x_1 + 0.0061x_1^2 + 0.0417x_2 - 0.055x_2^2 - 0.233x_3 + 0.265x_3^2$$

$$\begin{aligned}
\{\text{G01}\}:\quad y_{t(\text{Profile1-Yfeed-1})} &= 0.205 - 0.0628x_1 + 0.035x_1^2 - 0.0089x_2 + 0.00164x_2^2 \\
&\quad - 0.00076x_3 + 0.00208x_3^2 \\
\{\text{G01}\}:\quad y_{t(\text{Profile1-Xfeed-1})} &= 0.679 - 0.720x_1 + 0.402x_1^2 - 0.142x_2 + 0.0611x_2^2 \\
&\quad + 0.0195x_3 - 0.011x_3^2 \\
\{\text{G02}\}:\quad y_{t(\text{Profile1-Cfeed})} &= 0.275 - 0.161x_1 + 0.0919x_1^2 - 0.0362x_2 + 0.0182x_2^2 \\
&\quad + 0.00826x_3 - 0.00716x_3^2 \\
\{\text{G01}\}:\quad y_{t(\text{Profile1-Yfeed-2})} &= 1.043 - 0.763x_1 + 0.261x_1^2 - 0.188x_2 + 0.0377x_2^2 \\
&\quad + 0.00952x_3 - 0.00617x_3^2 \\
\{\text{G01}\}:\quad y_{t(\text{Profile1-Xfeed-2})} &= 1.048 - 0.776x_1 + 0.271x_1^2 - 0.205x_2 + 0.0624x_2^2 \\
&\quad + 0.0088x_3 - 0.00482x_3^2 \\
\{\text{G01}\}:\quad y_{t(\text{Pocket1-Feed})} &= 0.902 - 0.812x_1 + 0.415x_1^2 - 0.169x_2 + 0.0670x_2^2
\end{aligned}$$

$$+0.0411x_3 - 0.0313x_3^2$$

$$\{\text{G01}\}: \quad y_{t(\text{Pocket1-Step})} = 0.25 - 0.0777x_1 + 0.0658x_1^2 - 0.0126x_2 + 0.00888x_2^2 + 0.00772x_3 - 0.00663x_3^2$$

$$\{\text{G01}\}: \quad y_{t(\text{Slot2-Feed})} = 1.220 - 1.219x_1 + 0.314x_1^2 - 0.356x_2 + 0.0611x_2^2 - 0.0526x_3 + 0.0592x_3^2$$

$$\{\text{G01}\}: \quad y_{t(\text{Slot2-Back})} = 1.191 - 1.0874x_1 + 0.153x_1^2 - 0.344x_2 + 0.0573x_2^2 - 0.0461x_3 - 0.0546x_3^2$$

$$\{\text{G01}\}: \quad y_{t(\text{Ret})} = 0.437 - 0.000037x_1 + 0.0000309x_1^2 + 0.0000714x_2 - 0.000072x_2^2 + 0.00034x_3 - 0.0004x_3^2$$

where x_1: feed rate, x_2: spindle speed, x_3: cutting depth, y_t: predicted machining time.

3.4. **Model verification and validation.** Model verification and validation is made by comparing predicted machining time results with measured ones. Figure 4 visualizes the predicted vs. measured machining time on the twelve trials with regard to individual machining features. Here, the predicted machining time values are derived from the application of the polynomial regression models above. Table 4 presents the result of verification and validation. This table shows that the RMSE values on the twelve trials score under 1.762 sec and the TRE values do absolute 29.9 percent on the individual machining features at each trial.

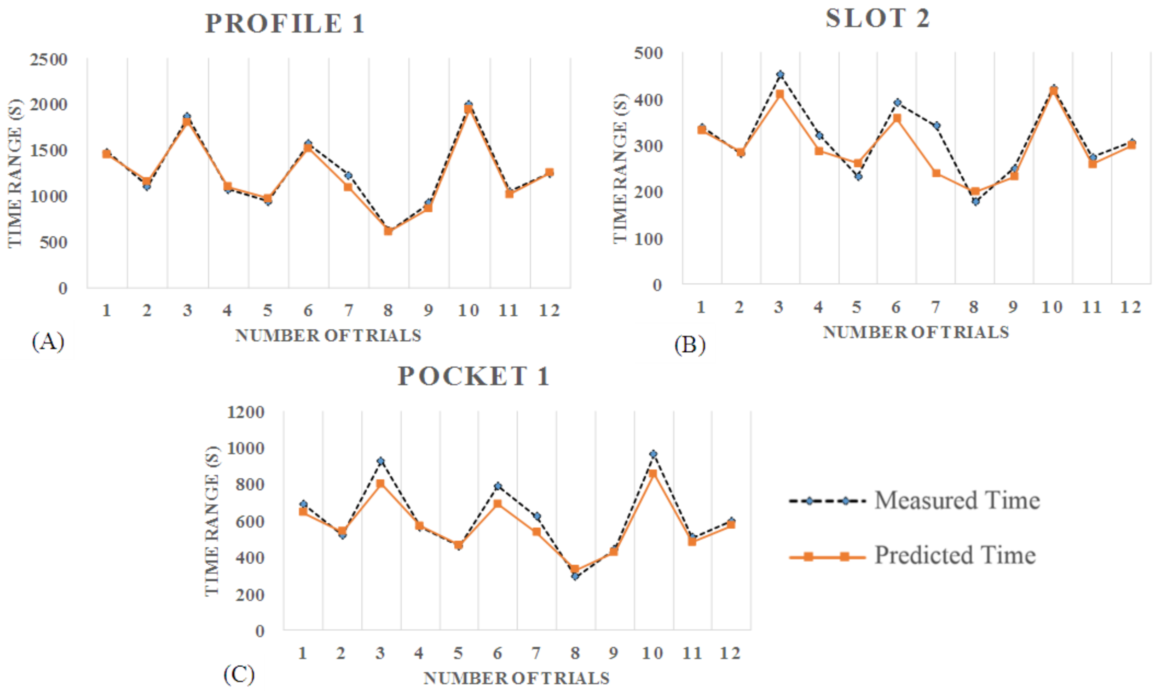

FIGURE 4. Total measured time vs total predicted time on (A) Profile 1; (B) Slot 2; (C) Pocket 1

TABLE 4. RMSE and TRE for measured vs. predicted machining time

Trial		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
RMSE (sec)		0.239	1.611	0.720	0.677	0.795	0.209	0.847	0.525	0.874	0.650	1.762	0.711
TRE (%)	Profile 1	−1.85	5.07	−3.08	2.49	3.61	−3.06	−11.23	−0.84	−6.52	−2.86	−3.24	0.43
	Pocket 1	−6.53	3.61	−13.63	0.9	0.99	−12.52	−14.13	12.72	−3.64	−11.14	−5.03	−4.32
	Slot 2	−2.43	0.82	−9.52	−10.8	11.8	−8.8	−29.9	12.3	−7.13	−1.46	−5.5	−2.41

4. **Discussion.**

(1) Prediction accuracy: The average value of TRE in Table 4 scores (−1.76%), (−4.39%), and (−4.42%) on Profile 1, Pocket 1, and Slot 2, respectively. These results show that our predictive models are acceptable to calculate the predicted machining time values close to the actual measured values. However, Slot 2 has the highest TRE, which means the predictive models for Slot 2 may have the lowest significance when they are used to predict machining time. It comes from the lack of data due to shorter tool path movements than them of Profile 1, Pocket 1, and thus datasets are not enough to create acceptable models. The same case was shown in trial 7. Trial 7 contains the lowest size of datasets, compared with the other trials. This high TRE gives an implication of the influence of the size of data being used on prediction. It is conjectured that our methodology needs to consider sufficient data sizes when performing machine learning analysis.

(2) Data integration: our methodology starts from the assumption where every set of MC parameters has its own operation sequence and different position and thus it has at least one predictive model. Despite the uniqueness, we integrate separate approach and retract models in different machining features into identical approach and retract models. By using data integration, we can share common data for the approach and retract models and put together into single models. The data integration will show an efficiency when calculation burdens increase due to the increase in the number of NC blocks or predictive models.

5. **Conclusion.** The purpose of the present work is to propose a model-driven predictive analytics approach for machining time using historical machine-monitoring data. Based on the proposed approach, we can create granular prediction models for gaining accurate machining time anticipation on the tool path level along with tool path movements. Using the real and historical machine-monitoring data, we generate accurate machining time predictive models using polynomial regression technique. Results on the applied approach show that our approach can be acceptably used to increase the accuracy of machining time prediction thereby enabling precise production and process planning. However, we only focused on machining time prediction on a single machine within the boundary of the given process planning. We do not deal with machining time optimization, which eventually contributes to increasing production efficiency by providing optimal process parameters for time minimization. The future works include: production time prediction on a production line where various machines exist and run together in complex circumstances, and machining time optimization with minimum energy uses through deriving optimal process parameters, which relates to multi-objective optimization problems.

Acknowledgment. This work was supported by a Research Grant of Pukyong National University (C-D-2017-1264).

REFERENCES

[1] J. P. Hung, Y. L. Lai, T. L. Luo and H. H. Hsiao, Prediction of the dynamic characteristics of a milling machine using the integrated model of machine frame and spindle unit, *International Journal of Mechanical and Mechatronics Engineering*, vol.6, no.7, 2012.

[2] M. Albert, *Modern Machine Shop*, https://www.mmsonline.com/articles/getting-started-with-machine-monitoring, [Access: 20-2-2018], 2016.

[3] S. J. Shin, J. Y. Woo and S. Rachuri, Predictive analytics model for power consumption in manufacturing, *Procedia CIRP*, vol.15, pp.153-158, 2014.

[4] Machining time, *The Great Soviet Encyclopedia*, 3rd Edition (1970-1979), The Gale Group, Inc., https://encyclopedia2.thefreedictionary.com/Machining+Time, [Access: 13-3-2018], 2010.

[5] R. T. Coelho, A. F. de Souza, A. R. Roger, A. M. Y. Rigatti and A. A. de Lima Ribeiro, Mechanistic approach to predict real machining time for milling free-form geometries applying high feed rate, *International Journal Advanced Manufacturing Technology*, vol.46, pp.1103-1111, 2010.

[6] H. Zhou, Y. Jiang, P. Zhou, M. Ji and Y. Zhao, Total completion time minimization scheduling on two hierarchical uniform machines, *Theoretical Computer Science*, vol.702, pp.65-76, 2017.

[7] A. Pfeiffer, D. Gyulai, B. Kadar and L. Monostori, Manufacturing lead time estimation with the combination of simulation and statistical learning methods, *Procedia CIRP*, vol.41, pp.75-80, 2016.

[8] M. Monreal and C. A. Rodriguez, Influence of tool path strategy on the cycle time of high-speed milling, *Computer Aided Design*, vol.35, pp.395-401, 2003.

[9] S. J. Shin, J. Y. Woo and S. Rachuri, Energy efficiency of milling machining: Component modeling and online optimization of cutting parameters, *Journal of Cleaner Production*, vol.161, pp.12-29, 2017.

[10] B. Malley, D. Ramazzotti and J. T.-Y. Wu, Chapter 12: Data pre-processing, in *Secondary Analysis of Electronic Health Records*, Springer International Publishing, Cham, Switzerland, 2016.

[11] C. M. Salgado, C. Azevedo, H. Proenca and S. M. Vieira, Chapter 13: Missing data, in *Secondary Analysis of Electronic Health Records*, Springer International Publishing, Cham, Switzerland, 2016.

[12] B. E. Lee, J. Michaloski, F. Proctor, S. Venkatesh and N. Bengtsson, MTConnect-based Kaizen for machine tool processes, *Proc. of ASME 2010 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference*, Montreal, Canada, 2010.