Development of Machine Learning prediction models for their integration in a Digital Twin for a tapered roller bearing production line

J Domínguez1*, A Esteban2, J A Romeo2, F Cebrián2, S Santo Domingo2 and J J Aguilar1

1 Department of Design and Manufacturing Engineering, University of Zaragoza. María de Luna 3, 50018 Zaragoza, Spain
2 Fersa Bearings S.A., C/ Bari 18, Polígono Industrial PLAZA, Zaragoza 50197, Spain

*Corresponding author: jdoco@unizar.es

Abstract: The aim of this work is to develop the prediction models that are integrated in a digital twin for a tapered roller bearing multi-stage production line. The manufacturing process consists of rings machining and component assembly processes, including intense quality control. This work proposes the use of machine learning techniques for a 4-step strategy which consists of: a data analysis, the development of one prediction model for the manufacture of the double outer ring, one model for the two inners rings, and finally their integration in the digital twin. The strategy is validated with real data. Several regression techniques are tested and the selected model is the exponential regression due to its better performance when compared with other algorithms. Once incorporated in the digital twin, the developed models can predict the process behaviour under potential changes so determining the optimum operating conditions can be fairly facilitated; as well as predict the final bearing setting under different machining conditions.

Keywords: Industry 4.0, Digital twin, Machine learning, Tapered roller bearing production.

1. Introduction

Bearing manufacturing requires high precision production processes. It is necessary to guarantee not only the composition and hardness of the material but also the geometric micrometric dimensions [1, 2]. Manufacturing errors will cause problems related to surface roughness, waviness, misaligned races, and out-of-size rolling elements [3]. They all affect fatigue, which is the predominant mode of bearing failure [4]. Tapered roller bearings can accommodate high radial loads as well as high axial loads. They have four main components: inner ring, outer ring, rollers, and cage.

The quality of the product is an important strategic factor of competitiveness of the European manufacturing industry in the global market. In that context, process control, automation and optimization are key to reach the highest quality at competitive cost [5, 6]. Process control ensures quality and reduce scrap and rework but needs dimensional measurements. Automation is key to measure at competitive cost, using machines with enough accuracy and cycle time, and automatic adjustment of the manufacturing process in real time. Current data collection and inspection technologies allow data to be collected online along the process chain and can significantly increase quality control and improvement in current dynamic and modifiable environments. The real challenge companies face is the problem of synthesizing such heterogeneous data to fully understand the correlations between variables throughout the process stages of a multi-stage system. This would be...
aimed at achieving the generation of zero defects, at the single process level, and the propagation of zero defects, at the system level, through proactive process control [7].

With the development of intelligent machining, optimization of machining process configuration during actual production has become more accessible. The grinding machine control system can receive as feedback information the relevant process data, including the result of the measurement of the machined part in the post-process verification. Developing a robust and accurate Digital Twin (DT) requires a sufficient Internet of Things (IoT) network that efficiently records live data from the process [8]. Big Data tools are also required in order to process the information recorded by the IoT network. By using these tools, it is possible to know the state of the process as well as to achieve a deep understanding of it, what permits to propose changes and substantial improvements [9]. Finally, by applying machine learning on the recorded data, DTs can predict the process behaviour under possible changes or incidents [10, 11], becoming a powerful tool for efficiency and performance improvement.

This work is conducted as part of a project whose aim is to develop a DT for a tapered roller bearing production line and to apply it in order to achieve a noticeable improvement on its performance. To achieve this improvement, the DT determines the optimum operation conditions under different line conditions. This article focuses on the machine learning techniques applied to develop the prediction models that integrate the DT. The work described in this article also shows how these models can help improving the performance of the production line even before integrating them in the DT.

2. Bearing and production line description

2.1. Bearing overview

Tapered roller bearings manufactured in this production line consist of 3 basic components which are a double outer ring and two dimensionally different inner rings, “A” and “B”. These 3 components have a direct impact on the bearing setting, which is the final critical dimension that must be strictly controlled once the bearing is mounted. The setting is defined as the axial clearance within the bearing.

On their own, both outer and inner rings have critical dimensions that have a direct effect on the setting and that must be controlled as well. The critical dimension for both cones is the axial location of the reference surface of the inner ring with respect to the reference surface of a master outer ring, as shown in figure 1. From now on this dimension is referenced as $T_{i,inv}$. Regarding the outer ring, its critical dimension is the distance between its two raceways, and it is designed as $T_e$. $T_e$ and $T_{i,inv}$ are henceforth understood as deviations from the nominal dimension shown in figure 1 under the same naming.

![Figure 1. Critical dimension of inner and outer rings.](image)

These three elements (both cones and the outer ring) are classified into 20 groups depending on their critical dimensions ($T_{i,inv}$ or $T_e$). Each of these groups covers a range of 4 μm and 8 μm of $T_{i,inv}$ and $T_e$, respectively, meaning a total allowance of 160 μm for the double cup and 80 μm for each cone. By using this classification, the three components can be matched to achieve an optimum setting.

Components are rarely classified in the outermost groups (1-3 and 18-20), so the objective of models is not improving the process capability, but rather manufacturing optimum components to be matched.
2.2. Line overview
The production line is divided into three sections. Two of them are parallel to each other and the third one, at the end of the line, is the result of the junction of the other two.

The parallel sections are dedicated to the grinding processes of the outer ring and the two inner rings, respectively. These sections are formed of several stations such as dimensional controls, super finish processes, washing stations and the grinding station itself. The two grinding stations are the core of both parallel sections since it is in these stations where critical dimensions $T_{inv}$ and $T_e$ are adjusted.

The junction of the parallel sections is a storage and matching area where outer and inner rings are classified into 20 groups and then matched to form the bearing. If the station finds the necessary pairs to form the set in the storage area, the received ring is matched instead of stored. However, if the station cannot find the necessary pairs in the storage area, the received ring is stored in this area until appropriate pairs are received. To match inner and outer rings they must belong to similar groups ($G \pm 2$).

From this point to the end of the line, the set is dimensionally verified, greased, oiled, and packaged. It is also worthwhile to mention that the inner ring grinding station is the bottleneck of the whole process.

3. Models scope and objectives
As previously mentioned, critical dimensions of inner and outer rings are obtained in the grinding stations. This is the main reason to focus the scope of the prediction models on these two machines.

By training prediction models, grinding stations can automatically suggest or even adopt operation parameters to obtain optimum inner and outer rings regarding their critical dimensions. Optimum rings are understood as those that can be matched immediately after reaching the storage and matching area. That is, if there are inner rings of group $G$ accumulated in the storage area, the optimum outer ring belongs to group $G \pm 2$ so it can be matched immediately after reaching the station.

Once they are integrated in the digital twin, prediction models consider as inputs both the machine state and the piece requirements to predict the necessary operation parameters to obtain the desired critical dimensions. Figure 2 shows how prediction models are included in the production line.

![Figure 2. Integration of prediction models in the production line. In blue: Offline processes of data analysis, training, and validation. In black: Online production process.](image)

Given the line configuration, training proper prediction models to operate grinding stations would have the following effects:

- Maximizing matching efficiency. This means to perform pairing of three elements of the same group most times. This also leads to a narrower setting allowance after the matching.
- Reducing the storage area. The current storage area has room for 220 inner rings and 110 outer rings, which implies an area of more than 10 m².
- Reducing the cycle time by performing immediate matchings instead of storing components.
• Generating a line simulator or digital twin. Predicting the performance of grinding stations makes it possible to integrate a complete line simulator with which perform experiments and predictions about production patterns, machine adjustment, etc.

4. Data analysis
The line IoT infrastructure makes it possible to acquire every machine parameter as well as rings parameters such as temperature or dimensions. After uploading data to an internal database, data analysis can be performed (figure 2). The total amount of data available to perform this study is equivalent to the grinding of 40.511 outer rings, 56.915 inner rings A, and 57.143 inner rings B.

The aims of this study are to know which machine and ring parameters must be considered to train the prediction models; to evaluate the consistency of predictions once the model is trained; and to compare experimental relations between parameters with theoretic relations.

Three types of parameters can be defined before the data analysis:

• Control variables (CVs): Those machine parameters with significant influence on the critical dimension (T_inv or T_e) and which can be modified to obtain optimum rings.

• Machine factors: Those machine parameters with significant influence on the critical dimension (T_inv or T_e) and which cannot be modified. Along with control variables, these parameters constitute the state of the machine. The state array can be denoted as x.

• Critical dimension: Is the output of the model. Its value depends on the state of the machine. This dimension is T_e or T_inv, and it can be denoted as y.

Under these definitions, models can be understood as a function that relates the machine state with the critical dimension of the ring, as shown by equation (1):

\[
y = f(\text{Control variables}, \text{Machine factors}) = f(x)
\]

Therefore, by knowing the ring requirements (T_e or T_inv), the model can be used to predict which machine state is needed to accomplish these requirements. That is, once integrated in the digital twin, models can predict CVs values to obtain the optimum ring, y, under certain machine factors (figure 2).

To study the influence of each parameter of interest on the output T_e or T_inv, data were split into datasets formed of all machine states that meet three conditions:

• Every parameter must remain constant along the state, except for the parameter of interest.

• The parameter of interest must have a different value in every state.

• Every state included in the dataset must correspond to the same wheel and shift. When the wheel is changed or the machine reset, every parameter reference changes.

It can be assumed that changes in T_e and T_inv under these conditions are only consequence of the changes in the parameter of interest. Equation (2) and equation (3) show datasets notation.

\[
X_{pi}^R = (x_{p_{i1}}, \ldots, x_{p_{i,j}}, \ldots , x_{p_{i,N}})
\]

\[
x_{p_{ij}}^R = (p_{i,j}, p_{i,j+1}, \ldots, p_M)
\]

\[
p_i = (p_{i,1}, \ldots, p_{i,j}, \ldots, p_{i,N})
\]

where: p_i: array containing all the N values of parameter p_i within the same wheel and shift; x_{p_{ij}}^R is the state of the machine that includes the j^{th} value of parameter p_i. All the other M – 1 parameters that form the state remain constant over the whole X_{pi}^R dataset; X_{pi}^R is the dataset formed by all the states x_{p_{ij}}^R of the grinding machine of R ring (outer -OR-, inner -IR-). Notice that many X_{pi}^R states from different wheels and shifts are used to study the influence of each parameter of interest on the output T_e or T_inv.
To assure enough consistency, only states $x_{R_{Pi}}^R$ greater than 25 observations that pass a Kolmogorov-Smirnov normality test have been kept in the analysis.

### 4.1. Outer ring

Five parameters out of eighty-nine from the outer ring grinding machine were found to have a significant influence on the critical dimension $T_e$. Table 1 shows the classification of the five parameters into factors and control variables. We do not consider interactions between parameters during data analysis since single effects are enough to identify which parameters have the highest impact on the output.

| Variable name | Description                               | Type         |
|---------------|-------------------------------------------|--------------|
| R120          | Radial adjustment of the grinding wheel   | Control variable |
| R280          | Radial position of the wheel during grinding | Factor          |
| R121          | Axial adjustment of the grinding wheel    | Control variable |
| R218          | Wheel diameter                            | Factor          |
| R115          | Dressing compensation                     | Factor          |
| $T_e$         | Distance between raceways                 | Output          |

As an example, figure 3(a) shows the output $T_e$ under the machine states $X_{R_{120}}^{OR}$. $T_e$ has been characterized as $T_e \sim N(\mu, \sigma^2)$ with $\sigma = 11.1 \mu m$.

![Figure 3. (a) Output $T_e$ under states $X_{R_{120}}^{OR}$](image)

Table 2 shows the correlations found between the parameters listed in table 1 as well as the theoretic known relations. Table 2 also shows the mean $R^2_{Di}$ coefficient, which is calculated from all considered datasets, $X_{pi}^{OR}$. We use $R^2_{Di}$ to determinate which parameters have to be included in the model. Only correlations with $R^2_{Di} > 0.4$ were kept.

| Relation                                             | Theoretic value | Experimental value | $R^2_{Di}$ | #Datasets |
|------------------------------------------------------|-----------------|--------------------|------------|-----------|
| Radial adjustment of the wheel–Wheel consumption (R218) | 0.5             | 0.5                | 1          | 51        |
| Axial adjustment of the grinding wheel - $T_e$        | 2.5             | 2.8                | 0.62       | 406       |
| Radial adjustment of the grinding wheel - $T_e$      | -               | 3                  | 0.58       | 140       |

### 4.2. Inner rings

Grinding cycles of inner ring A and B are different from each other due to their dimensional differences. In fact, inner ring B must be processed in two different cycles. However, in spite of these differences,
parameters that have been found to have the most significative influences on $T_{i,inv}$ are common to both grinding processes (inner ring A and B). These common parameters are shown in Table 3.

| Variable name | Description                                      | Type     |
|---------------|--------------------------------------------------|----------|
| L51           | Radial adjustment of the grinding wheel          | Control variable |
| L52           | Axial adjustment of the grinding wheel           | Control variable |
| R35           | Reference for radial adjustment of the wheel     | Factor   |
| R75           | Reference for axial adjustment of the wheel      | Factor   |
| R322          | Wheel diameter                                   | Factor   |
| $T_{i,inv}$   | Distance between surfaces of inner and master outer ring | Output |

A positive implication of finding common parameters is that the same prediction model can be used for both inner rings with no loss in accuracy as proved in section 6.

As an example, figure 3(b) shows the output $T_{i,inv}$ under the machine state $X_{L51}$. $T_{i,inv}$ has been characterized as $T_{i,inv} \sim N(\mu, \sigma^2)$ with $\sigma = 5.3 \mu m$. Table 4 shows the correlations with $R^2_p > 0.4$ found between the parameters listed in table 3 as well as the theoretic known relations.

| Relation                        | Theoretic value | Cone type | Experimental value | $R^2_p$ | #Datasets |
|---------------------------------|-----------------|-----------|--------------------|---------|-----------|
| Radial adjustment of the        | 0.25            | A         | 0.23               | 0.69     | 153       |
| grinding wheel – $T_{i,inv}$    |                 | B         | 0.23               | 0.66     | 119       |
| Axial adjustment of the         | -               | A         | 0.1                | 0.62     | 197       |
| grinding wheel – $T_{i,inv}$    |                 | B         | 0.08               | 0.55     | 117       |

5. Data normalization

As previously mentioned, resetting the grinding machines or changing their wheels entail a manual adjustment of the machines. This adjustment produces a noticeable change in the global references of parameters. To train models with data from different wheels and shifts, it is necessary to perform a normalization of data; this is, a change on its reference. This normalization consists on using incremental data with respect to the last change of wheel. This way, models predictions are also given with respect to the last change of wheel, which is known. Figure 4 shows the control variable R120 both before and after applying normalization of its reference.

Figure 4. (a) Non-normalized data. (b) Normalized data.

Changes of wheel occur every ~700 and ~15,000 pieces in outer and inner ring grinding machines, respectively. This low frequency of changes in inner ring machine makes it unnecessary to perform data normalization on it while it is still necessary on the outer ring machine. The first stable grinding
operations that act as a reference for successive values are considered as the first group greater than 15 grindings with constant parameters after changing the wheel. This consideration avoids using values from the high variability calibration process that takes place after every change of wheel as references.

### Table 5. Validation RMSE obtained with different models.

| Model Type         | Validation mean RMSE over 45 experiments (μm) |
|--------------------|-----------------------------------------------|
|                    | Outer ring (31300 data) | Inner ring (31500 data) |
| Exponential GPR    | 7.29                           | 3.57                  |
| Rational Quadratic GPR | 7.44                         | 3.61                  |
| Matern 5/2 GPR     | 7.42                           | 3.62                  |
| Fine tree          | 8.18                           | 3.66                  |

6. Trained models

The selected model is the exponential regression due to its better performance when compared to other 12 algorithms such as decision trees, SVM or other gaussian process regressions (GPR). Table 3 shows RMSE (Root Mean Squared Error) obtained from the four algorithms with best performance after evaluating them over 45 experiments. Each experiment consists in training and validating the algorithm over data from a unique wheel in the case of outer rings and over subsets of 700 data in the case of inner rings. Cross-validation with 5 folds has been used.

6.1. Outer ring model

Since control variable R120 and factor R280 are tightly related variables \(R^2 = 1\), Table 2), keeping both is redundant. To train the simplest model, only control variable R120 has been kept.

Once a model is trained and validated with the same dataset, it must be tested on new data unknown to the model. Figure 5(a) shows the response of the selected exponential GPR model when it is trained and validated with 5,000 data and tested with 3,500 new data. Table 4 summarizes the RMSE obtained in every data interval shown in figure 5(a). For the sake of clarity, \(T_e\) dispersion has been removed.

Considering that \(3\sigma_{Te} \approx 33 \mu m\) (dispersion inherent to the grinding machine process), models with a testing RMSE between 20 and 25 μm are considered accurate and robust enough to be reliable.

6.2. Inner ring model

The main difference of inner ring model with respect to the outer ring model is that, as previously mentioned, data are not normalized. Training, validation, and testing has been done under the same conditions. That is, testing data is unknown to the model and \(T_{inv}\) dispersion has been removed.

Figure 5(b) shows predictions with the 5,000 training and validation data and with the 3,500 test data. Note that \(3\sigma_{T_{inv}} \approx 16 \mu m\) while \(RMSE_{IR} < 6 \mu m\) when tested with up to 3,500 data (table 6).

7. Conclusions

From the results, it can be concluded that the developed models reproduce the actual grinding process behaviour. A validation with real data is carried out and the results have confirmed the high variability of the process over time and the need to normalize the values of the outer ring at each change of grinding wheel. The exponential regression technique has been chosen because of its better performance compared to other regression algorithms. The results show that for the outer ring model, with a dispersion inherent to the process of \(3\sigma_{Te} \approx 33 \mu m\), a RMSE < 26 μm is obtained. In the case of the inner ring, with a dispersion inherent to the process of \(3\sigma_{T_{inv}} \approx 16 \mu m\), a \(RMSE_{IR} < 6 \mu m\) is obtained. These models predict the grinding behavior, so once integrated in the DT, it can be used to determinate the optimum operating conditions and the final bearing setting under different machining conditions.
Figure 5. (a) Response $T_e$ predicted (yellow dots) by outer ring model on training (blue line) and testing (red line) data. (b) Responses $T_{i,inv}$ predicted by inner ring model.

Table 6. RMSE obtained from the validation and testing of the trained models.

| Model    | RMSE Training data (μm) | RMSE Testing data (μm) |
|----------|-------------------------|------------------------|
|          | 350 data                | 700 data               | 1400 data | 2100 data | 2800 data | 3500 data |
| Outer ring| 11.3                    | 20.43                  | 18.69     | 23.67     | 22.69     | 22.49     | 25.68     |
| Inner ring| 6.94                    | 5.55                   | 5.33      | 4.66      | 4.51      | 4.86      | 5.29      |

Acknowledgements
This research was funded by the Ministerio de Ciencia, Innovación y Universidades and the CDTI (Centro de Desarrollo Tecnológico Industrial) with project number CIEN IDI-20181145 and by Aragon Government through the Research Activity Grant for research groups recognized by the Aragon Government (T56_17R Manufacturing Engineering and Advanced Metrology Group).

References
[1] El-Thalji I and Jantunen E 2015 A summary of fault modelling and predictive health monitoring of rolling element bearings *Mechanical Systems and Signal Processing* 60–61 pp 252–272
[2] Liu J, Shao Y and Qin X 2017 Dynamic simulation for a tapered roller bearing considering a localized surface fault on the rib of the inner race *Proceedings of the Institution of Mechanical Engineers, Part K: Journal of Multi-body Dynamics* Institution of Mechanical Engineers 231 (4) pp 670–683
[3] Brosed FJ, Victor Zaera A, Padilla E, Cebrían F and Aguilar J J 2018 In-process measurement for the process control of the real-time manufacturing of tapered roller bearings *Materials* 11 (8) p 1371
[4] Sadeghi F, Jalalahmadi B, Slack T S, Raje N, Arakere N K 2009 A review of rolling contact fatigue *ASME Transactions on Tribology* 131 p 041403-1-15
[5] D’Addona D M, Bracco F, Bettoni A, Nishino N, Carpanzano N and Bruzzone A A 2018 Adaptive automation and human factors in manufacturing: An experimental assessment for a cognitive approach *CIRP Annals* 67 (1) pp 455-458
[6] Finlay B, Rackwitz N, Conerney B, Warren E, Erdmann D, Stoddard, K, Weber A and Scanlon T 2018 Requirements for First-Time-Right Response in Advanced Manufacturing. 29th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC) (Saratoga Springs, NY) p 186
[7] Colledani M, Fischer A, Iung B, Lanza G, Schmitt R, Tolio T and Vancza J 2014 Design and management of manufacturing systems for production quality *CIRP Annals - Manufacturing Technology* 63 (2) pp 773-796
[8] Ren S, Zhang Y, Liu Y, Sakao T, Huisingsh D and Almeida C M V B 2019 Review of A comprehensive review of big data analytics throughout product lifecycle to support sustainable
smart manufacturing: A framework, challenges and future research directions *Journal of Cleaner Production* **210** pp 1343–1365

[9] Wan J, Li X, Dai H-N, Kusiak A, Martinez-García M and Li D 2021 Artificial-Intelligence-Driven Customized Manufacturing Factory: Key Technologies, Applications, and Challenges in *Proceedings of the IEEE* **109** (4) p 377

[10] Kritzinger W, Karner M, Traar G, Henjes J and Sihn W 2018 Digital Twin in manufacturing: A categorical literature review and classification *IFAC-PapersOnLine* **51** (11) pp 1016–1022

[11] Tao F, Cheng J, Qi Q, Zhang M, Zhang H and Sui F 2018 Digital twin-driven product design, manufacturing and service with big data *The International Journal of Advanced Manufacturing Technology* **94** (9-12) pp 3563–3576