Bearing degradation prediction based on support vector regression

S Darwis¹, N Hajarisman¹, S Suladi¹ and A Widodo²

¹ Department of Statistics, Universitas Islam Bandung, Bandung, Indonesia
² Department of Mechanical Engineering, University of Diponegoro Semarang, Semarang, Indonesia

*std.darwis@gmail.com

Abstract. Predicting bearing degradation before reaching the state of risk of accident is one important issues in power generation insurance. This paper proposes a method based on support vector regression to achieve the goal. The method is applied on PRONOSTIA dataset which is an experimental platform dedicated to test methods related to bearing health assessment. The results show that the method can effectively model the evolution of the bearing degradation.

1. Introduction

The delay time defines the failure process of an asset as two-stage process. The first one is the normal operating stage from new to the point that a defect has been identified. The second one is defined as the failure delay time from the point of defect identification to failure. Optimal inspection interval can be identified with appropriate modeling of the durations of these two stages. The delay time is random and the initial point of defect identification is very important to the setup of an appropriate inspection interval [1]. The first instance where the presence of a defect is called the initial point u of the defect and time h to failure from u is called the delay time of the defect. Wang shows that the failure process is a marked Poisson process with the delay time as the marker, and the failure process is a nonhomogeneous Poisson process [1]. Wang and Zhang used statistical process control to model the identification of the initiation point of a random defect [2]. A point considered as the initial point of the failure if one point outside the action line [2]. Wang and Zhang proposed adaptive statistical process control based on autoregressive model. The adaptive Shewhart chart overcomes the drawback of adaptive moving charts, also not very sensitive to small changes in the data [2].

There are two major approaches in modeling aggregate loss: the individual risk model and the collective risk model [3]. Individual risk assumes there are in independent loss. Distribution of the aggregate loss can be obtained through the convolution method. The collective risk model treats the aggregate loss as having a compound distribution with the primary distribution being the claim frequency and the secondary distribution being the claim severity.

Predicting the bearing degradation process before reaching failure is extremely important in mechanical system [4]. Feature extraction is the process of transforming the vibration data collected from running equipment to relevant information of bearing health condition. The prediction of degradation process can be classified into model-based and data-driven methods. The model-based methods predict the degradation process using the physical model. Data-driven methods are derived directly from monitoring data of the sensor system installed in the equipment and have been increasingly
used to bearing future life time estimation. Sensor-based degradation signals measure the accumulation of damage of a mechanical system using sensor technology [5]. Vibration analysis is intended to ensure the safety of the installation by avoiding significant damage by trigging an alarm when the vibration level reached excessive values for the functioning.

In traditional regression analysis, via some statistical criteria such as error sum of squares the parameters can be estimated. In support vector regression context, the criterion is defined as ignoring observations which error less than. Amiri et al., used support vector regression to analyze the time series data [6]. The model produced by support vector regression only depends on a subset of the training data. Instead of minimizing the residual sum of squares, support vector regression attempts to minimize the generalization error bound so as to achieve generalized performance [7]. Support vector regression has shown excellent forecast performance due to its particular design of minimizing structural risk. It uses support vector machine to predict a continuous variable while other regression models minimize the error between the predicted and the actual value. Support vector regression tries to classify all the prediction line in two types, ones that pass through the error boundary and ones that do not. The line that pass is considered for potential support vector to predict the value of unknown.

PRONOSTIA is an experimental platform for testing and validating bearing degradation methods. Its aims to provide real data related to bearing degradation test. The monitoring is ensured by gathering two types of signal: temperature, and horizontal and vertical acceleration. Depending on various factors, the degradation may be different for distinct bearing. Interaction with other parts of the rotating system produces irregular pattern [8].

This paper explores the applicability of support vector regression (SVR) in predicting the bearing degradation process. The FEMTO dataset was used for training, validating and predicting using SVR. Four hundred data was sampled from dataset, 300 used as training and validating, and 100 used for prediction. The aims of the study are to investigate the feasibility, to evaluate the performance, and reliability of SVR in bearing degradation analysis.

The remainder of the paper is organized as follows: section 2 discusses the methodology of support vector regression and FEMTO bearing degradation dataset. Section 3 presents the results and discussion of bearing vibration data analysis. Conclusion is presented in section 4.

2. Method

The loss function is defined as ignoring observation which error is less than [6]. The linear support vector regression can be expressed as linear combination of training pattern $x_i$

$$f(x) = \sum_{i=1}^{t} (\alpha_i - \alpha_i^*) \langle x_i, x \rangle + b$$

(1)

The Lagrange multipliers $\alpha_i$ and $\alpha_i^*$ represent solutions to the quadratic optimization

$$\min \frac{1}{2} \sum_{i,j=1}^{t} (\alpha_i - \alpha_i^*) \langle x_i, x_j \rangle - e \sum_{i=1}^{t} (\alpha_i + \alpha_i^*) + \sum_{i=1}^{t} y_i (\alpha_i - \alpha_i^*)$$

subject to $\sum_{i=1}^{t} (\alpha_i - \alpha_i^*) = 0$, $\alpha, \alpha^* \in [0, C]$ (2)

Only the nonzero values of $\alpha - \alpha_i^*$ are useful in forecasting the regression line. The solution to the support vector regression involves only the inner product of the observations [9]. The parameters are nonzero for the support vectors in the solution. Training observations that are far from $x$ will play no role in prediction. The radial kernel has very local behaviour, only nearby training observation have an effect on the prediction.

PRONOSTIA is an experimental platform for testing and validating bearing degradation methods [8]. The monitoring is ensured by gathering two types of signal: temperature, and acceleration. Three different treatments were considered: 1) 1800 rpm & 4000 N, 2) 1650 rpm & 4200 N, 3) 1500 rpm &
5000 N. Six run-to-failure datasets are provided in order to build degradation model and used to estimate 11 bearing vibration data. Table 1 summarizes the bearing degradation experiment, and table 2 shows the characteristics of FEMTO bearing experiment.

### Table 1. Load, learning and testing FEMTO datasets.

| Load (N) | Learning | Testing |
|----------|----------|---------|
| 1800-4000 | 1-1, 2-1, 3-1 | 1-3, 2-3, 3-3 |
| 1650-4200 | 1-2, 2-2, 3-2 | |
| 1500-5000 | |

### Table 2. Characteristics of experiments of the FEMTO dataset.

| Bearing | Date       | Duration (h.m.s) | # files |
|---------|------------|------------------|---------|
| 1-1     | 01.12.2010 | 7.47.00          | 3269    |
| 1-2     | 06.04.2011 | 2.25.00          | 1015    |
| 2-1     | 06.05.2011 | 2.31.40          | 1062    |
| 2-2     | 17.06.2011 | 2.12.40          | 797     |
| 3-1     | 07.04.2011 | 1.25.40          | 604     |
| 3-2     | 28.06.2011 | 4.32.40          | 1637    |
| 1-3     | 17.11.2010 | 5.00.10          | 1802    |
| 1-4     | 07.12.2010 | 3.09.40          | 1327    |
| 1-5     | 13.04.2011 | 6.23.30          | 2685    |
| 1-6     | 14.04.2011 | 6.23.29          | 2685    |
| 1-7     | 15.04.2011 | 4.10.11          | 1752    |
| 2-3     | 19.05.2011 | 3.20.10          | 1202    |
| 2-4     | 26.05.2011 | 1.41.50          | 713     |
| 2-5     | 27.05.2011 | 5.33.30          | 2337    |
| 2-6     | 07.04.2011 | 1.35.10          | 572     |
| 2-7     | 08.06.2011 | 0.28.30          | 200     |
| 3-3     | 08.04.2011 | 0.58.30          | 410     |

PRONOSTIA is composed of three main parts: rotating, degradation generation and measurement part [8]. The experiment is carried out by applying loads on the bearing exceeding the loads allowed by the catalog in order to accelerate the degradation. The referenced bearing was NSK 6804 DD which can operate at maximum speed of 13 000 rpm and a load limit of 4 000 N. PRONOSTIA is composed of two main treatments: speed variation and load generation [10]. Two accelerometers are mounted, horizontally and vertically on the housing of the bearing to pick up horizontal and vertical accelerations. The sensors are connected to a data acquisition to provide the monitoring bearing condition. The vibration data provided by the accelerometers are collected every one second, the sampling frequency is 25600 Hz. The duration of the experiments is (hours, minutes): 1.25, 2.25, 3.10, 4.35, 5.00, 6.25, 7.50. Vertical and horizontal accelerations collected from each bearing are used for feature identification. During the experiments, new bearing is used until it is failed (Benkedjouh et al., 2013).

### 3. Results and discussions

The support vector regression was used to estimate the initial point of bearing degradation using the action line. The time when one observation outside of action line considered as the initial point of degradation. Figure 1 shows three hundred bearing vibration used to train the support vector regression
of vibration time series. Some peaks amplitudes shown at the beginning and the following periods the data shows an increasing. Figure 2 shows the histogram of the support vector regression with the number of support vector was $n_{sv} = 130$. The coefficients are in the range of $-0.2$ to $0.2$.

Figure 3 shows the actual vibration and the prediction of the support vector regression. The correlation of the prediction with the actual data is high and the rmse is acceptable. The model is acceptable for forecasting the state of the bearing. The model catches the dynamic of time series bearing vibration data. Figure 4 shows the action line obtained from support vector regression in training phase. The initial point of bearing degradation is estimated at $t = 24$.

![Figure 1](image1.png)  
**Figure 1.** Three hundred bearing vibration data used for training support vector regression model. Some peak vibration shows at the beginning of the data, and the vibration increases shows a trend.

![Figure 2](image2.png)  
**Figure 2.** Histogram coefficient SVR, $n_{sv} = 130$, $C = 1000$, epsilon = .01, gamma =1.
Figure 3. Training support vector regression, --+ is the actual vibration, --- is the prediction. The model catches very well the dynamic of bearing vibration time series. The model is acceptable and can used as forecast of bearing degradation.

Figure 4. Estimation of initial point of the degradation. The action line is \( y = 2.321 \), and the initial point of degradation is estimated at time \( t = 24 \).

4. Conclusion
This paper explores the application of support vector regression in the prediction of initial point of bearing degradation. The support vector regression is a smoothing technique and has some similarity with other smoothing technique such as spline smoothing. In this study, the support vector regression was used to estimate the action line for estimating the initial point of bearing degradation.
Acknowledgments
This research is supported by LPPM-Unisba under grant number 002/B.04/LPPM/I/2019.

References
[1] Wang W 2008 Delay Time Modeling University of Salford, anchester UK
[2] Wang W and Zhang W 2008 Early Defect Identification: Application of Statistical Process Control Journal of Quality in Maintenance Engineering 14(3) pp 225-236
[3] Tse Y K 2009 Nonlife Actuarial Models: Theory, Methods and Evaluation Cambridge University Press
[4] Dong S, Yin S, Tang B, Chen L and Luo T 2014 Bearing Degradation Process Prediction Based on the Support Vector Machine and Markov Model Shock and Vibration Hidawi Publishing Company Corporation
[5] Zhou Z, Du Y and Dong S 2011 DNA–Ag nanoclusters as fluorescence probe for turn-on aptamer sensor of small molecules Biosensors and Bioelectronics 28(1) 33-37
[6] Amiri S, van Rosen D and Zwanzig S 2009 The SVM Approach for Box-Jenkins Models REVSTAT Statistical Journal 7(1) 23-26
[7] Basak D, Pal S, Patranabis D C 2007 Support Vector Regression Neural Information Processing – Letters and Reviews 11(10)
[8] Nectoux P, Gouriveau R, Medjaher K, Ramasso E, Morello B, Zerhouni B and Varnier C 2012 PRONOSTIA: An Experimental Platform for Bearing Accelerated Degradation Test IEEE International Conference on PHM Denver, Co, USA
[9] James G, Whitten D, Hastie T and Tibshirani R 2013 An Introduction to Statistical Learning with R Springer
[10] Benkedjouh T, Medjaher K, Zerhouni N and Rechak S 2013 Remaining Useful Life Estimation Based on Nonlinear Feature Reduction and Support Vector Regression Engineering Applications of Artificial Intelligent