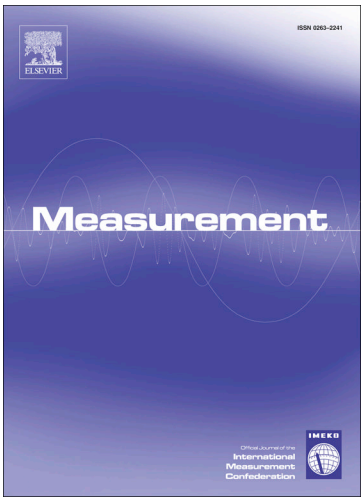


A new model for rolling element bearing defect size estimation

Aoyu Chen, Thomas R. Kurfess

PII: S0263-2241(17)30584-5
DOI: http://dx.doi.org/10.1016/j.measurement.2017.09.018
Reference: MEASUR 4968

To appear in: *Measurement*

Received Date: 16 May 2017
Revised Date: 18 July 2017
Accepted Date: 11 September 2017

Please cite this article as: A. Chen, T.R. Kurfess, A new model for rolling element bearing defect size estimation, *Measurement* (2017), doi: http://dx.doi.org/10.1016/j.measurement.2017.09.018

A new model for rolling element bearing defect size estimation

Aoyu Chen* and Thomas R. Kurfess
George W. Woodruff School of Mechanical Engineering
Georgia Institute of Technology
Atlanta, Georgia, USA
Corresponding author* at: Room 104, Love Building, Georgia Institute of Technology, GA 30332, USA
E-mail addresses: achen75@gatech.edu (Aoyu Chen), kurfess@gatech.edu (Thomas R. Kurfess).

Abstract

A new model based on the Hertzian contact theorem is proposed to estimate the size of a line spall defect located on the bearing's outer race. The entry point can be determined from the ball-race geometry relation, while the exit point can be identified from the time domain signal. Therefore, the defect size can be estimated from the vibration signal without requiring additional bearing load and stiffness measurements. Experiments were performed on a 3-axis CNC machine tool at speeds ranging from 500-3000rpm and three line spall defects were estimated using the proposed method. The proposed model is demonstrated to estimate defect size with minimal speed related error, offset error, and standard deviation.

Keywords: Bearing diagnostics; Condition based maintenance; Defect size estimation; Machine monitoring

1. Introduction

Condition Based Maintenance (CBM) is an effective method to prevent unnecessary maintenance cost and downtime resulting from unanticipated machine failure[1, 2]. Bearing diagnostics are a valuable tool for CBM, as they provide information related to a bearing's health in rotary machines including spindles [3, 4]. Recently, researchers have proposed that when a rolling element passes a line-spall defect on the raceway, a repeatable low-frequency entry event and a high-frequency exit event can be identified from the vibration response[5]. Based on dynamic models of the bearing system and experiment results, it has been proposed that the time difference between the entry and exit events can be used to estimate the size of the defect [6-8]. Calculating from the entry and exit events is a promising method to quantify the severity of a spall-like damage for a CBM algorithm. However, because the moment when the ball center pass through the entry edge is hard to determine in the vibration pattern, previous studies use indirect methods to estimate defect size [5, 9]. "Indirect methods" in this paper refer to estimation of the entry point without using information from the entry signal. In these indirect methods, the ball travel distance during the entry event is either ignored or calculated from bearing maximum load and stiffness. Thus, the estimation results in prior work suffer from speed related error, large offset error, and significant standard deviation. Therefore, a more comprehensive model that transfers the time information of the vibration signal into the defect size is critical to improve the estimation. A new defect size estimation model is proposed based on the Hertzian contact theorem. In the proposed model, the duration of the entry event is measured to locate the entry point in the vibration signal, based on the bearing geometry and material property. Therefore, measuring the bearing maximum load and stiffness is not required to calculate the defect size. Because bearing load and stiffness is difficult to determine and changes over time for an actual bearing system, the proposed method is a viable candidate for CBM.

In this paper, prior work is summarized in Section 2 and describes the vibration pattern due to a localized defect and estimation methods proposed by previous research. Section 3 introduces the new estimation model and the formula to estimate the defect size. The experimental setup and estimation results are

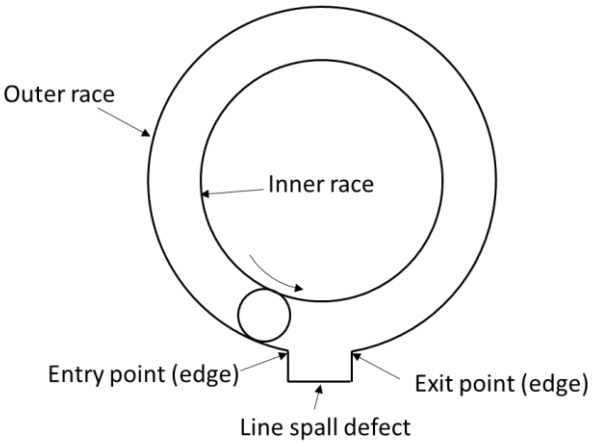
shown and discussed in Section 4 to validate the defect size estimation model. Finally, Section 5 discusses the summarizing conclusions.

2. Prior work

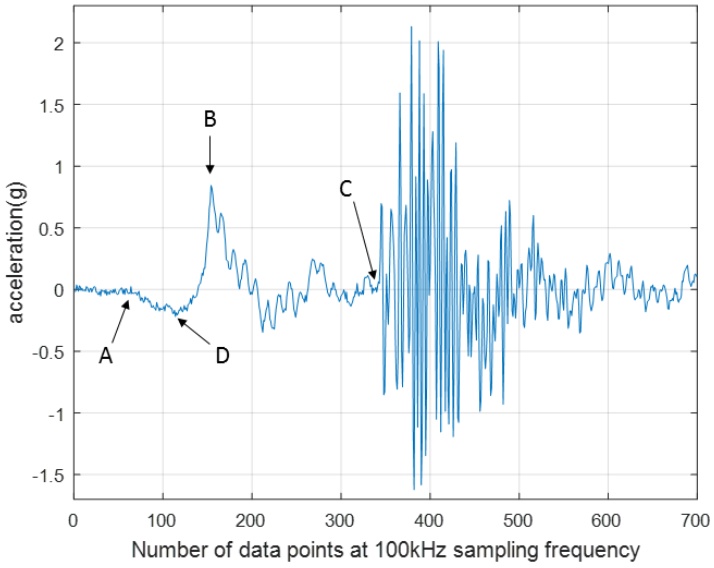
This section summarizes the pattern of the vibration response due to a localized defect on bearing races, and some defect size estimation methods provided by previous research.

2.1 The vibration pattern

The line-spall defect has a rectangular profile on bearing outer race as shown in Fig. 1(a). When the ball rolls into the defect zone, the destressing contact between the ball and both races causes the bearing system to vibrate [10]. A repeatable pattern in the vibration signal is shown in Fig. 1(b). The data in Fig. 1(b) comes from one of the experimental results in this research at 500 rpm and defect size 1.530mm. All the other data sets also show this similar pattern. At point A, the "destressing" process starts and then the vibration signal decreases to the local minimum at point D. Then, the vibration signal begins to increase and reaches the first peak at point B. Peak B is shown to be dominant and repeatable in the response pattern. Moazen Ahmadi proposed that peak B corresponds to the moment when the ball loses contact with both raceways [9]. After point B, the bearing system vibrates at a certain frequency until the ball collides with the exit edge and a high frequency response occurs at point C. In order to estimate the defect size, various models have been proposed to relate the vibration pattern to the actual ball center path [5, 9, 11, 12]. The moment when the ball center arrives the entry edge, which is defined as the entry point, must be determined to effectively estimate the defect size. However, according to the dynamic model proposed by Moazen Ahmadi, the entry point occurs between A and B and is difficult to isolate. Therefore, previous studies utilize simplifying assumptions or indirect methods to avoid locating the entry point. These estimation methods are introduced in the next section.



(a)



(b)

Fig. 1. Typical defective response due to a localized defect on the outer race.

2.2 Defect size estimation methods

Randall considered the local maximum in the entry signal (peak B) as the entry point and assumed that the ball-race exit collision occurs when the ball center is midway through the defect [5]. However, other work has shown these assumptions fail as defect size increases [12]. Prior work has shown the entry point starts earlier than the most dominant peak in the entry signal, which is between A and B [12]. Since the time between the entry point and B can occupy a significant portion of the total entry-to-exit time, Randall's estimation results were shown to be inaccurate and suffers from speed dependent error.

Smith proposed an alternative method to identify the entry point [11]. A consistent roll-off was shown to be present after point A in the entry event, and two methods were presented to estimate the entry point. In the first method, the zero-crossing point of the low-pass filtered acceleration gradient function is

considered as the entry point. However, under the influence of noise, the gradient function is not guaranteed to cross zero at an entry point. Therefore, this method is sensitive to noise and threshold identification. In the second method, the peak location of the low-pass filtered acceleration signal is considered as the entry point. However, this assumption only applies when the acceleration "hump" in the pre-impact signal coincides with the entry event.

Moazen Ahmadi studied the vibration pattern by the nonlinear multibody dynamic model of the bearing system and proposed a more accurate model [9, 13]. This model was based on the detection of the local maximum B in the entry event and the low-frequency exit event. However, the low-frequency exit event can be masked by the high-frequency exit signal and the detection is not always possible. Therefore, Moazen Ahmadi proposed another method using the time from points B to C. In both methods, the bearing maximum load and stiffness were used to calculate the time difference between A and B, and therefore entry point. However, the bearing load and stiffness is difficult to measure and changes over time. Therefore, the estimation result produces large standard deviation.

This paper proposes a new defect size estimation model, which is independent of the bearing load and stiffness measurement. The entry point is directly determined by the Hertzian contact theorem, the vibration signal and bearing geometry. The time from entry to impact is used to estimate the defect size. The new model is described in Section 3.

3. The defect size estimation model

This section describes the defect size estimation model proposed in this paper, which relates the vibration pattern to the defect size.

3.1 Bearing geometry relationship

The geometry relationship between the defect size d and the path traveled by the ball is shown in Fig. 2. O is the center of the bearing. The entry point lies on line OP, which means that it represents the moment when the ball center arrives the entry edge. Due to the contact deflection, two elliptical contact area exist between ball/inner and ball/outer race. When the deflected area arrives at the entry edge P, the ball center is at point A. Therefore, point A is located left to the entry point. The semi-major axis of the elliptical contact area between ball/outer race and ball/inner race is a_1 and a_2 respectively. The ball loses contact with both races at B. At this moment, the deflection between ball and races fully recovers and the distance from entry edge P to ball center B equals to the radius of the ball r. C is the ball center when the ball collides with the exit edge Q. Note that A, B and C in Fig. 2 correspond to A, B and C in the vibration signal as shown in Fig. 1. The angular positions of A and B relative to the entry point are defined as θ_1 and θ_2, respectively. The time from A to B is defined as the entry-to-peak time t_e, and the time from B to C is defined as the peak-to-impact time t_p. Both t_e and t_p can be extracted from the vibration pattern. The following will discuss how to estimate the defect size based on t_e, t_p and the geometry relationship shown in Fig. 2.

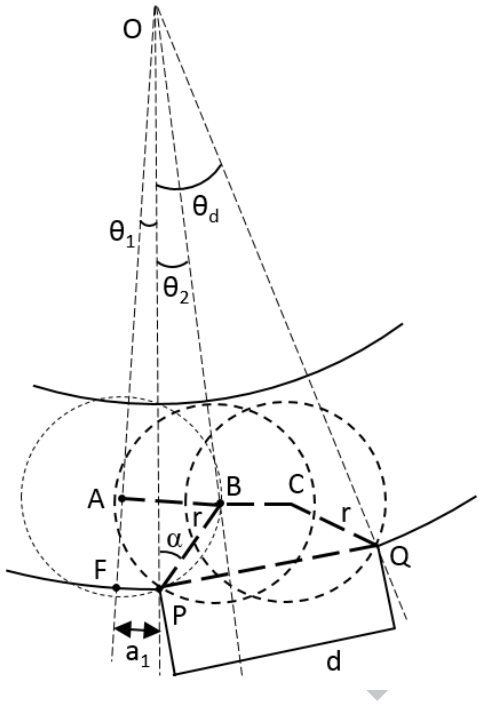


Fig. 2. The defect size estimation model.

3.2 Model derivation and discussions

At position A, given the initial contact force F_0, a_1 and a_2 can be derived from the Hertzian contact theorem as in Eq. (1) and (2). Within the speed range from 500-3000rpm, the centrifugal force is very small compared to the contact force, and therefore is not considered in this model.

$$_1 = \sqrt[3]{\frac{3F_0\left[\frac{1-v_1^2}{E_1}+\frac{1-v_2^2}{E_2}\right]}{4\left(\frac{1}{r}-\frac{1}{R_o}\right)}} \quad (1)$$

$$_2 = \sqrt[3]{\frac{3F_0\left[\frac{1-v_1^2}{E_1}+\frac{1-v_2^2}{E_2}\right]}{4\left(\frac{1}{r}-\frac{1}{R_i}\right)}} \quad (2)$$

Where E_1 and E_2 are the Young's moduli, v_1 and v_2 are the Poisson's ratios, R_o and R_i are the groove radius of the outer and inner races, respectively. Then, he deflection between ball/outer race and ball/inner race can be derived as:

$$\Delta x_1 = \ _1^2 \cdot \left(\frac{1}{r}-\frac{1}{R_0}\right) \quad (3)$$

$$\Delta x_2 = \ _2^2 \cdot \left(\frac{1}{r}-\frac{1}{R_i}\right) \quad (4)$$

If the distance from ball center to inner race center without deflection is R, θ_1 can be expressed as:

$$_1 = \frac{1}{R + r} \tag{5}$$

θ_2 can be solved in ΔOPB:

$$_2 = \frac{\sqrt{2(R + r)Rr\Delta x_0}}{R(R + r - \Delta x_0)} \tag{6}$$

Where Δx_0 is the total defection of the ball with both races. Both θ_1 and θ_2 are functions of the initial contact force F_0. Since the total deflection is small, θ_1 and θ_2 are also very small. Therefore, the velocity of the ball center from A to B can assumed to be constant. Thus,

$$t_e = \frac{_1 + \ _2}{\omega_c} \tag{7}$$

Where ω_c is the cage angular velocity, which depends on bearing geometry and spindle speed. Note that the ratio between θ_1 and θ_2 can be defined as a constant value η that only depends on the bearing geometry for small deflections. Then, the relationship between θ_2 and t_e can be derived as:

$$_2 = \frac{\omega_c \cdot t_e}{1 + \eta} \tag{8}$$

From position B, the ball is assumed to travel in the horizontal direction with a constant speed from B to C. Then, BC can be derived as:

$$BC = R \cdot \omega_c \cdot t_p \tag{9}$$

When the ball collides with the exit edge Q, the angle θ_d can be solved using the trigonometry relationship as shown in Fig. 2:

$$[OQ \cdot \cos \ _d - R \cos \ _2]^2 + [OQ \cdot \sin \ _d - R \sin \ _2 - R \cdot \omega_c \cdot t_p \ ]^2 = r^2 \tag{10}$$

Where OQ=R+r- Δx_0. Therefore, the center angle of the defect θ_d only depends on t_e and t_p, which can be extracted from the vibration signal. Then, θ_d can be solved numerically and the defect size d can be estimated using Eq. (11):

$$d = 2(R + r) \sin\frac{d}{2} \tag{11}$$

The proposed estimation model is experimentally validated as described in Section 4. A prerequisite for this model is that the ball collides with the exit edge after it leaves the entry edge. For small defects, this condition may not hold. The lower limit of the defect size valid for this model can be calculated from Eq. (10) when BC is zero. This condition corresponds to when the ball collides with the exit edge at the moment it loses contact with the entry edge. The lower defect size limit d_{min} only depends on the deflection between the ball and the raceways. When the contact force is larger, the deflection Δx_0 becomes larger accordingly, and then the lower size limit for this model will increase. When the defect size is smaller than d_{min}, the geometry relation changes and the proposed model will not hold in that situation.

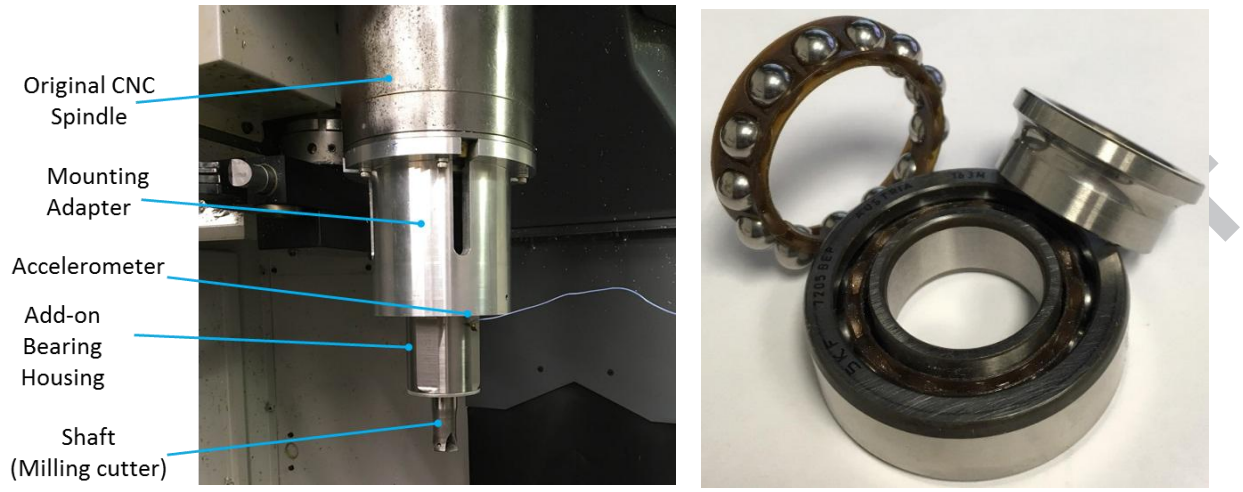
This model is not suitable for the inner race defect, because the ball path is regulated by the outer race after it loses contact with both races. Thus, for inner race defects, the ball does not travel "into" the defect as with the outer race case. Due to this reason, the ball path for the inner race defect is different from the outer race defect. Therefore, the scope of the proposed defect estimation model is limited to outer race defects.

4. Test equipment and experiment results

The test equipment, defect generation and experimental results are described in this section. The experimental setup is designed to validate the proposed bearing defect size estimation model described in Section 3. The bearing test system is mounted on the spindle of a 3-axis Okuma CNC mill. This setup provides operating conditions closer to an actual machine tool than the traditional bearing test rigs developed in previous bearing research [14, 15]. Three different defects are generated using Electric Discharge Machining (EDM) on the outer race. Results show that the new model provides estimation with error less than 10% for all three defects and small standard deviations.

4.1 Test equipment

Because frequent disassembly and reassembly of the spindle may result in the failure of the machine tool, the original spindle of the machine cannot be disassembled and seeded with defects. Therefore, an external bearing housing with the tested angular contact ball bearing is mounted onto the original machine spindle by a mounting adapter as shown in Fig. 3(a). Four flat planes perpendicular to each other are machined on the surface of the add-on housing to provide a robust support for the vibration sensor. The aluminum mounting adapter is fixed to the CNC and a setscrew is used to fix the outer race. The shaft of the add-on bearing system is also the milling cutter. Therefore, the add-on test system can be directly driven by the spindle of the 3-axis Okuma CNC mill. One set of two angular contact ball bearings is installed in the housing. The housing and the outer race are stationary, while the shaft and the inner race are rotating. The top bearing is seeded with defect and the outer race has a transition match for easy replacement. The outer races of the bearing set are supported by the housing shoulders and the inner races of the bearing set are supported by a spacer. The accelerometer is mounted on one of the flat planes outside the add-on bearing housing right above the seeded defect as shown in Fig. 3(a). The top cap is machined with a fine thread and thus the pre-load of the bearing can be adjusted by the screw on the top cap. This setup provides operating conditions closer to an actual machine tool than the traditional bearing test rigs developed in previous bearing work [14, 15]. In addition, the new setup has the capabilities to perform bearing diagnostics during machining operations. The tested ball bearing is SKF 7205BEP as shown in Fig. 3(b) and its geometry is listed in Table 1. Where OD, ID and BD are the bearing's outer race, inner race and ball diameter respectively.



(a) (b)

Fig. 3. (a) CNC-based bearing test setup (b) The tested bearing.

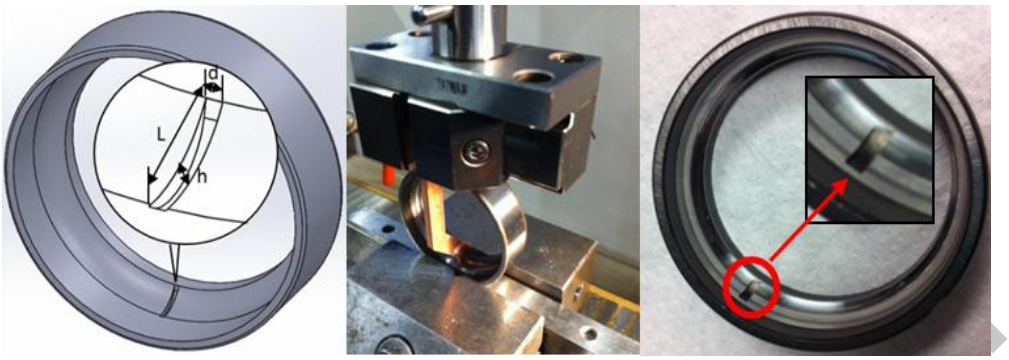
Table 1. Bearing geometry

Bearing type	OD(mm)	ID(mm)	BD(mm)	No. of balls	Contact angle
Angular contact ball bearing	52	25	8	13	40 °

The PCB 352a21 accelerometer is used to measure the vibration signature of the bearing system. The sensitivity of the accelerometer is 10mV/g. The signal is amplified by a Type 5134 Kistler power supply/coupler and collected by NI myDAQ with sampling rate of 100kHz.

4.2 Defect generation

For testing purposes, it is critical that consistent and measurable damage are generated on the test bearing surfaces to ensure the repeatability of the experiments as well as to provide a means by which the diagnostic systems may be calibrated. Therefore, a line spall, shown in Fig. 4(a), is used in this research. The defect is generated using a die sinking EDM on the bearing races, as shown in Fig. 4(b). The length of the defect along axis direction L is maximized, and therefore L is through the whole groove part of the outer race. The depth of the defect h is considered large compared with the defect width d, and therefore the ball cannot contact the bottom of the defect as it passes the defect zone. Thus, the values of L and h do not affect the defective signature. The primary focus of this research is the defect width d. The defect widths studied in this work are 0.794mm, 1.135mm and 1.530mm. Fig. 4(c) shows the 1.530mm width defect on the outer race.



(a) (b) (c)

Fig. 4. (a) A line-spall defect (b) Seed defect with EDM (c) 1.530mm outer race defect.

The defective bearing is tested under different spindle speeds and the vibration data is collected by the accelerometer mounted on the bearing housing right above the defect.

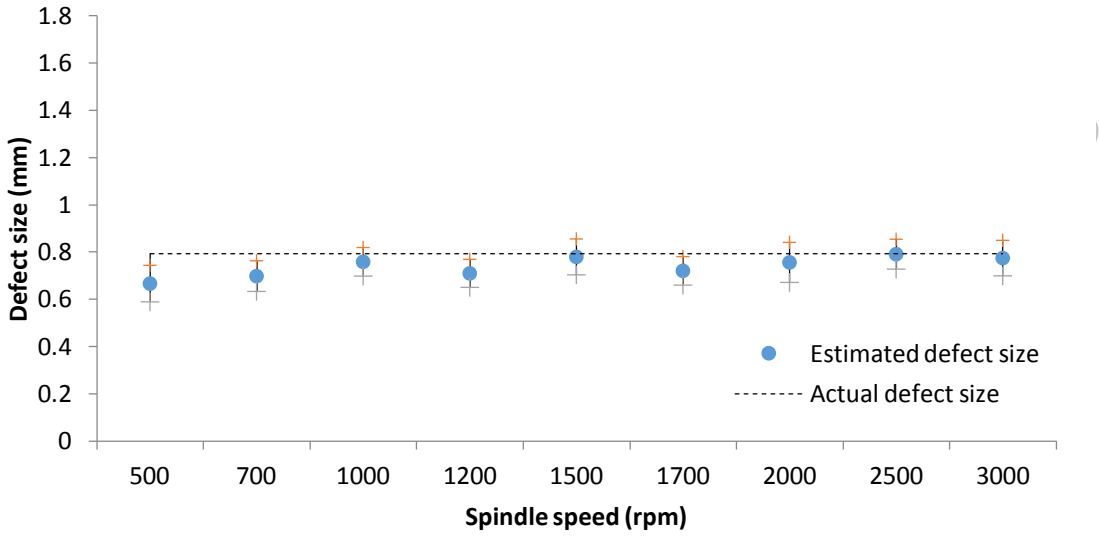
4.3 Experiment results and discussions

In this section, the defect size estimation results for the three defects at nine different spindle speeds from 500-3000rpm are shown in Fig. 5(a) through (c). Both t_e and t_p are extracted from the vibration signal and substituted in the estimation model derived in Section 3 to estimate the defect size. t_e and t_p are extracted with different signal processing techniques [16]. Due to the low signal-to-noise ratio of the entry signal, the Variational Mode Decomposition and empirical model based denoising procedures are applied to remove the high frequency noises from the entry signal to obtain t_e. t_p is extracted using cross-correlation and differential algorithms. Each data point in the plot is the averaged estimation result based on 100 impulsive responses in the vibration signals. The estimation model exhibits reasonable performance with a maximum absolute error of 0.054mm and the maximum relative error 7%, as listed in Table 2. Speed related errors and offset errors have been decreased a lot compare to previous work[5, 12]. The "offset error" measures the accuracy of the averaged estimation result compared to the true defect size. Therefore, the defect size estimation model as proposed in this paper is validated.

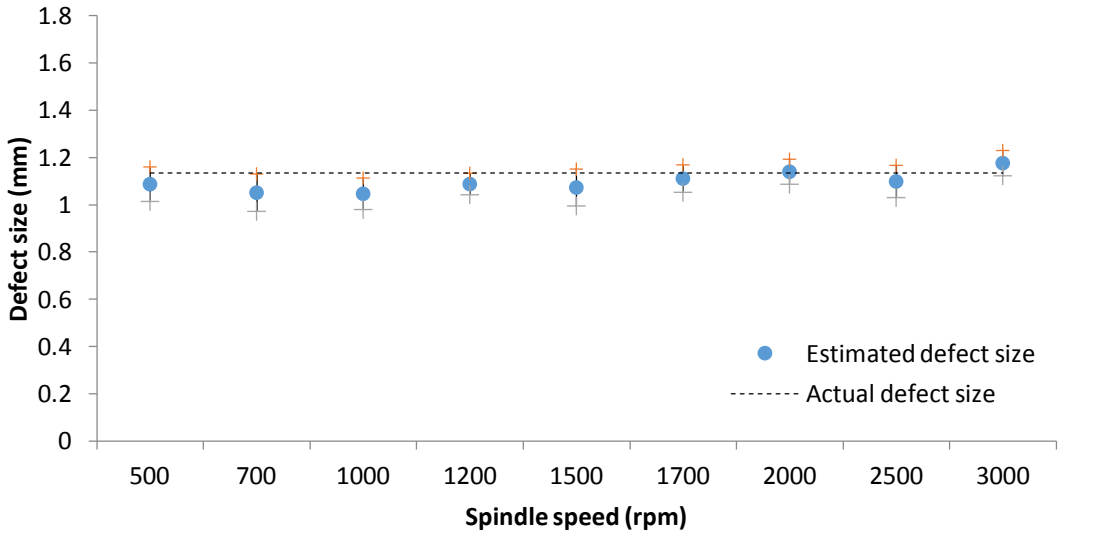
The standard deviation has been improved using this new model. In the previous work [12], only t_p is obtained from the vibration signal, and t_e is a constant value calculated based on maximum ball-race deflection. However, variations in bearing load when each ball passes the defect will affect both t_e and t_p. From the experiment results, a strong negative correlation exists between t_e and t_p. Thus, when the load increases, t_e will increase and t_p will decrease. Therefore, the standard deviation of t_p is larger than the standard deviation of t_e+t_p. Using the new model, both t_e and t_p are measured from the vibration signal, and the variations in t_p due to the variation of load is compensated by the measurement of t_e. Therefore, the new model yields smaller standard deviation than in previous work.

At speeds lower than 500rpm, due to the small signal-to-noise ratio between the entry signal and the low frequency noise, the estimation result is smaller than the true defect size. Signal processing methods for lower speeds is recommended for future work. At speeds larger than 3000rpm, uncertainty level is high due to the short time interval from entry to exit. Therefore, for speeds above 3000rpm, the sampling frequency should be increased. In addition, the bearing sliding issue becomes more prominent at higher speeds, and therefore the estimation result will be larger than the true defect size. Estimation models

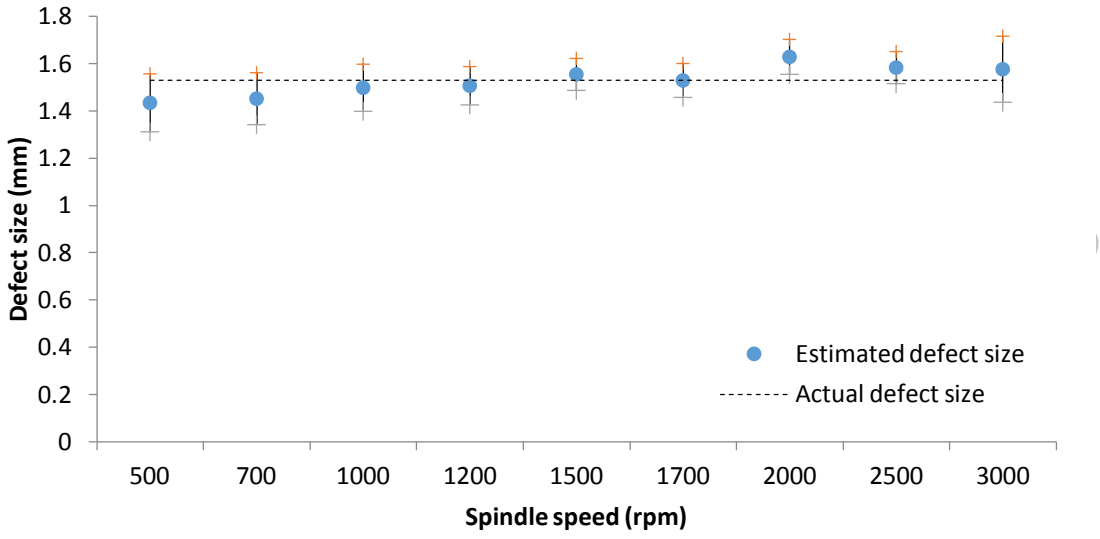
considering bearing sliding can help improve the estimation for high speed implementation and is recommended for future work.



(a)



(b)



(c)

Fig. 5. Defect size estimation results. (a) 0.794mm (b) 1.135mm (c) 1.530mm.

Table 2. Average absolute and relative error of three different estimation models.

Defect size (mm)	Absolute error (mm)	Relative error
0.794	0.054	7%
1.135	0.048	4%
1.530	0.050	3%

5. Conclusions

This paper proposes a new model for rolling element bearing defect size estimation on the outer race. In this model, the entry point when the ball center arrives the entry edge can be calculated from the ball-race deflection based on the Hertzian contact theorem and the ball-race geometry relationship. The entry-to-peak time (A to B) and the peak-to-impact time (B to C) are used to estimate defect size. All the parameters required for estimation can be obtained from the vibration signal and is therefore free from bearing load/stiffness measurement. Experiments were performed on a CNC machine tool with an add-on bearing test system with three EDM defects tested at speeds of 500-3000rpm. Results show that the new model provide accurate estimation with less speed related error and standard deviation. The maximum absolute error is 0.054mm, which is only 7% of the 0.794mm defect.

References

[1] F. Camci, R.B. Chinnam, Health-state estimation and prognostics in machining processes, IEEE Transactions on Automation Science and Engineering, 7 (2010) 581-597.

[2] A.K. Jardine, D. Lin, D. Banjevic, A review on machinery diagnostics and prognostics implementing condition-based maintenance, Mechanical systems and signal processing, 20 (2006) 1483-1510.

[3] T.R. Kurfess, S. Billington, S.Y. Liang, Advanced diagnostic and prognostic techniques for rolling element bearings, Condition Monitoring and Control for Intelligent Manufacturing, Springer, 2006, pp. 137-165.
[4] N. Tandon, A comparison of some vibration parameters for the condition monitoring of rolling element bearings, Measurement, 12 (1994) 285-289.
[5] N. Sawalhi, R. Randall, Vibration response of spalled rolling element bearings: Observations, simulations and signal processing techniques to track the spall size, Mechanical Systems and Signal Processing, 25 (2011) 846-870.
[6] L. Niu, H. Cao, Z. He, Y. Li, Dynamic modeling and vibration response simulation for high speed rolling ball bearings with localized surface defects in raceways, Journal of Manufacturing Science and Engineering, 136 (2014) 041015.
[7] M. Behzad, A.R. Bastami, D. Mba, A new model for estimating vibrations generated in the defective rolling element bearings, Journal of Vibration and Acoustics, 133 (2011) 041011.
[8] S. Khanam, N. Tandon, J. Dutt, Fault size estimation in the outer race of ball bearing using discrete wavelet transform of the vibration signal, Procedia Technology, 14 (2014) 12-19.
[9] A.M. Ahmadi, C.Q. Howard, D. Petersen, The path of rolling elements in defective bearings: Observations, analysis and methods to estimate spall size, Journal of Sound and Vibration, 366 (2016) 277-292.
[10] A. Moazenahmadi, D. Petersen, C. Howard, N. Sawalhi, Defect size estimation and analysis of the path of rolling elements in defective bearings with respect to the operational speed, INTER-NOISE and NOISE-CON Congress and Conference Proceedings, Institute of Noise Control Engineering, 2014, pp. 4242-4251.
[11] W.A. Smith, C. Hu, R.B. Randall, Z. Peng, Vibration-Based Spall Size Tracking in Rolling Element Bearings, Proceedings of the 9th IFToMM International Conference on Rotor Dynamics, Springer, 2015, pp. 587-597.
[12] A. Moazen-Ahmadi, C.Q. Howard, A defect size estimation method based on operational speed and path of rolling elements in defective bearings, Journal of Sound and Vibration, 385 (2016) 138-148.
[13] A.M. Ahmadi, D. Petersen, C. Howard, A nonlinear dynamic vibration model of defective bearings–The importance of modelling the finite size of rolling elements, Mechanical Systems and Signal Processing, 52 (2015) 309-326.
[14] J.-Y. Shen, C.-W. Tseng, I. Shen, Vibration of rotating disk/spindle systems with flexible housing/stator assemblies, Journal of Sound and Vibration, 271 (2004) 725-756.
[15] R.B. Randall, J. Antoni, Rolling element bearing diagnostics—a tutorial, Mechanical Systems and Signal Processing, 25 (2011) 485-520.
[16] A. Chen, MACHINE TOOL SPINDLE BEARING DIAGNOSTICS UNDER OPERATING CONDITIONS, Georgia Institute of Technology, 2017.

Highlights:

- Entry signal is used to find the moment when the ball center arrives entry edge.
- All the parameters for the estimation can be obtained from the vibration signal.
- Experiments were performed on a CNC machine tool rather than test rigs.