

Intelligent diagnosis of bearing knock faults in internal combustion engines using vibration simulation

Jian Chen^{1,2*}, Robert Bond Randall²

1. Institute of Sound and Vibration Research, University of Southampton, Southampton SO17 1BJ, UK
2. School of Mechanical and Manufacturing Engineering, University of New South Wales, Sydney 2052 Australia

Telephone: +44-758-7456129

Email: jc2d14@soton.ac.uk; jian.chen1@unsw.edu.au

Abstract

Big-end bearing knock faults in IC engines can be considered as a real industrial case of a slider-crank mechanism including a joint with clearance and lubrication. In this paper, an Artificial Neural Network (ANN) based system was used to solve the problem of intelligent big-end bearing knock fault diagnosis in Internal Combustion (IC) engine. But when the ANN is used in machine condition monitoring, it is either unlikely or uneconomical to experience all different real faults to generate sufficient training data. Therefore, model based method should be a viable way to generate adequate data to train the networks for the intelligent big-end bearing fault diagnosis in IC engines. In order to evaluate and update the simulation model, experiments with normal bearing clearance and with different oversize bearing clearances were first carried out on the engine test rig. It was found that the relevant diagnostic information lies in the squared envelope of the vibration signals. Therefore, we only need build a proper simulation model to simulate the correct envelope signals rather than the raw vibration signals. As the important inputs of the simulation model, the inertia properties of the simulated engine components were also measured and studied. Next, we built an ANN-based bearing knock diagnosis system which consists of three phases: fault detection phase, fault localization phase and fault severity identification phase. Particularly, a saturating linear function is selected as the transfer function of the fault severity identification stage, so the networks can linearly classify the fault levels and the output is more in agreement with the reality in industry. Following the feature extraction and selection from the processed squared envelope signals, the networks were purely trained by the simulated data with normal bearing clearance and with different oversize bearing clearances. Finally the networks was tested by the real experimental data and it was demonstrated that the networks can successfully detect different bearing knock faults in real tests, and also classify the faults' location and severity levels.

Keywords: Intelligent diagnosis, internal combustion engine, journal bearing with oversize clearance, vibration simulation, Artificial Neural Network

1. Introduction

IC engines have severe operating conditions, involving high temperatures and high pressures, with large variations in internal forces within each cycle, and as such the wear mechanisms are complex. The journal bearings of IC engines are particularly vulnerable to these conditions, as opposed to those in rotating machines such as turbogenerators, because of the extreme variations in load. Therefore bearing damage, especially the big-end bearing damage, due to friction and wear accounts for and leads to a significant portion of Internal Combustion (IC) engine failure [1]. However, only few works have addressed the issue of big end bearing knock fault diagnosis based on vibration measurement. Moreover, traditional vibration based diagnostic techniques normally require engineers to analyse the vibration signal by means of their experience. Recently, Artificial Neural Networks (ANNs) based automated system has been successfully developed and applied to the diagnosis of rotating machinery. Therefore, following the proper signal processing method and feature

extraction/selection method, ANN should also be a potential solution to realize the intelligent diagnosis of the big end bearing faults in IC engines. The successful application of ANN is strongly depended on the adequate data for the networks' training. Both data-driving approach and model-based approach are the solutions to this issue. The former approach requires very long-term data record from one IC engine or a large number of identical engines, so it is a very expensive approach. Alternatively if a relevant simulation model is built up, different bearing knock faults can be easily simulated in the model, so sufficient training data can be generated as well. Beforehand only a small number of experiments need to be carried out on the engine test rig so as to evaluate the simulation model.

Big-end bearing knock faults in IC engines is a typical issue of a slider-crank mechanism with an oversize clearance joint and lubrication. Many researchers have addressed the dynamic modelling of the dry or lubricated evolute joints (journal bearing) with clearance in the slider-crank mechanism [2-6]. Even though all of the works are focussed on the modelling of the small end bearing with clearance, the analysis methods can also be extended to the modelling of the big end bearing with clearance in IC engines. Particularly, Daniel and Cavalca [6] analytically developed the kinematic model and hydrodynamic model for the evolute joint between piston and connecting rod with clearance and lubrication, and they used a numerical technique to iteratively compute the kinematic/kinetic equations and the lubrication equations, and finally solved the multi-Degree-Of-Freedom (DOF) problem raised by the non-negligible clearance.

This paper can be considered as the extension and a real application of modelling methods developed by Refs. [5, 6] and also our former work [7]. The critical issue for the model based intelligent condition monitoring is the model should create the signals with decent accurate diagnostic information, so normally a small number of experiments should be carried out on the engine test rig first. In this paper, after the advanced signal processing techniques were applied to process the experimental vibration signals, we found that the most useful diagnostic information lies in the squared envelopes of vibration signals. Therefore, even though many more sophisticated analytical modelling methods or finite element modelling methods have been developed for the journal bearing design in recent years, for the intelligent big-end bearing fault diagnosis, we only need to apply a proper modelling approach to generate accurate envelope signals rather than raw accelerations (or impact forces). The simulation model for the bearing knock diagnosis consists of the kinematic/kinetic part and the hydrodynamic part. The lubrication forces (bearing knock force) which is the interconnection between two models was numerically solved step by step. Based on the measured transfer functions between the bearing houses to the measurement point on the engine block, vibration signals were further calculated from the simulated bearing knock forces. As the important inputs of the simulation model, the parameters of the relevant engine components, such as the inertial properties, were also measured and calculated by some measurement techniques. Based on the experimental data, the amplitude and phase features of the Fourier coefficients of the squared envelopes were extracted. Particularly, the Genetic Algorithm (GA) was used select the best amplitude features as the inputs of the networks. Next, we used purely simulated data to train the three-phase network system, which consists of two MLP (Multi Layer Perceptron) networks for fault detection (MLP1) and fault severity identification (MLP2), and one PNN (Probabilistic Neural Network) for the fault localization. The selected amplitude features were used as the inputs of the MLP1 and MLP2, and the selected phase features were used as the inputs of the PNN. Finally the real experimental data were inputted into the intelligent diagnosis system to test its performance and it was demonstrated that the intelligent system got good results.

2. Principle of bearing knock

2.1 Kinematic/kinetic principle

If all the joints of piston-connection rod-crank system are perfect joints, the dynamics of the system is a classic signal-DOF issue. But if an oversize clearance is introduced into the big end bearing joint, the dynamic properties become more complicated and there are two pendulum motions in this mechanism system, one the the rocking motion of the connecting rod about the piston pin and the other is relative to the almost constant rotational speed of the crank pin. So the piston-connection rod-crank system become a multi-DOF system (as shown in Figure 1), and it can be separated into two subsystems, piston and connecting rod subsystem and crank subsystem. It is not difficult to build separate kinematic/kinetic equations for the two subsystems, but in order to solve this multi-DOF issue, the hydrodynamic lubrication equations which can introduce extra "constraints" into the system should be added into the system. Note that we only considered the oversize

clearance in the big end bearing joint in this paper, so we assumed both the piston pin (small end bearing) joint and crank main journal (joint) are perfect joints.

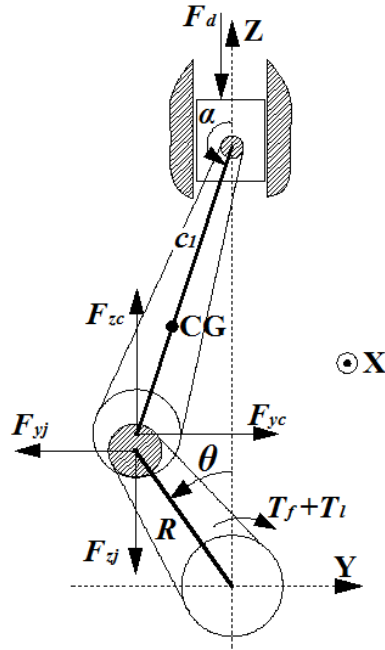


Figure 1. Force diagram of the big-end bearing joint with an oversize clearance

As shown in Figure 1, the whole system has three external forces/moments, F_d is the combustion force, T_f represents both friction torque and pumping torque of the engine, T_l represents the engine external load. We always consider the combustion pressure is evenly distributed on the top of the piston in this paper, so the combustion force F_d can be just integrated from the pressure about the piston top area. We will use subscript c and subscript j to represent the parameters associated with connecting rod (big end bearing) and crank journal in the following sections. For the connecting rod with a total length of c_c , we can measure the location of its centre of gravity (CG) and the detail of inertial property measurement of the connecting rod will be presented in Section 3.3. The distances from the centre of the small end to the centre of gravity (CG) of the connecting rod are c_1 . The crank throw radius is R (half stroke length). So for the piston and connecting rod subsystem, there are two DOFs, one is the translation motion of the piston in Z direction (z_p , \dot{z}_p and \ddot{z}_p denotes its displacement, velocity and acceleration respectively), and the other is the angular rotation of the connecting rod (α , $\dot{\alpha}$ and $\ddot{\alpha}$ denotes its angular displacement, velocity and acceleration respectively). It is not difficult to set up the kinematic equations of piston and connecting rod subsystem by Lagrangian mechanics [7],

$$\begin{bmatrix} m_p + m_c & -c_1 m_c \sin \alpha \\ -c_1 m_c \sin \alpha & I_c + c_1^2 m_c \end{bmatrix} \begin{bmatrix} \ddot{z}_p \\ \ddot{\alpha} \end{bmatrix} + \begin{bmatrix} 0 & -\dot{\alpha} c_1 m_c \cos \alpha \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{z}_p \\ \dot{\alpha} \end{bmatrix} = \begin{bmatrix} F_{zc} - F_g \\ -F_{yc} c \cos \alpha - F_{zc} c \sin \alpha \end{bmatrix} \quad (1)$$

where F_{yc} and F_{zc} are the lubrication forces acted on big end bearing in Y and Z directions respectively, m_p and m_c are the masses of the piston and the connecting rod respectively, and I_c is the moment of inertia of the connecting rod about its CG (the measurement of I_c will be introduced in Section 3.3).

The third DOF is presented in the crank subsystem, and it is the angular rotation of the crank (likewise we use θ , $\dot{\theta}$ and $\ddot{\theta}$ to denote the angular displacement, velocity and acceleration of the crank respectively). It is quite straightforward to build the kinematic equation of the crank subsystem based on the moment balance principle,

$$I_j \ddot{\theta} = F_{yj} R \cos \theta + F_{zj} R \sin \theta - T_l - T_f \quad (2)$$

where I_j is the polar moment of inertia of the crankshaft and flywheel (the detail of the measurement of I_j will be introduced in Section 3.3 as well), F_{yj} and F_{zj} are the lubrication forces acted on the crank journal in Y and

Z directions respectively. Note that the forces F_{yc} and F_{yj} are an action-reaction force pair, and same for the force pair F_{zc} and F_{zj} . It can be found that we need introduce lubrication equations to calculate the lubrication forces and eventually solve this multi-DOF issue.

2.2 Lubrication principle

The geometric representation of the big end bearing with clearance and lubrication is shown in Figure 2. There are two reference frames involved into the modelling, one is the fixed global XY frame, and the other is the intermediary MN frame, which is based on the position of maximum clearance (the direction of N axis is same as the direction from the bearing centre to the journal centre). β is the coordinate angle starting from the positive Y direction to the positive N direction. Note that different from normal bearing-housing fixed journal bearing, the load on journal/bearing in the big end journal bearing is a variable rotating load, therefore β cannot be simply considered as the general attitude angle, and it should be the combination of load rotation angle and attitude angle. The difference in radius ($R_c - R_j$) is the radial clearance c , R_c and R_j are the radii of the big end bearing and the crank journal respectively. e denotes the eccentricity at a certain instant, and it can be calculated by the locations of the bearing centre and the journal centre in the YZ frame at that moment.

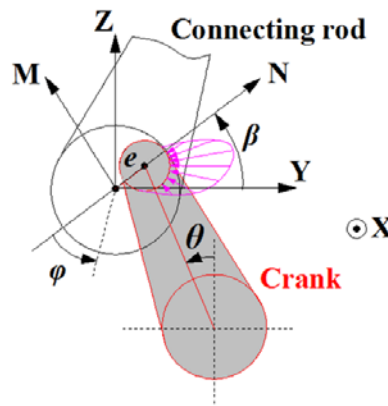


Figure 2. Geometric representation of the big end bearing with an oversize clearance and lubrication

The classic Reynolds equation can model the pressure distribution in the journal bearing. But it is very difficult to get the close form analytic solution of Reynolds equation, so two common assumptions were always introduced to simplify the solving of Reynolds equation, one is the journal bearing with an infinite length and the other is the journal bearing with an infinite short length, and the corresponding solution are called as Sommerfeld solution [8] and Ocvirk solution [9]. The ratio of bearing length to diameter for the majority of the big-end bearings in IC engines is about 0.5, so the Reynolds equation of the big end bearing can be simplified as the infinite short journal bearing. Based on the derivation in Ref [10, 11], we can write the pressure distribution solution for the big-end bearing as

$$p = 6\mu \left(x^2 - \frac{1}{4} L_c^2 \right) \frac{[\dot{\varepsilon} \cos \varphi + \varepsilon(\dot{\beta} - \bar{\omega}) \sin \varphi]}{c^2 (1 + \varepsilon \cos \varphi)^3} \quad (3)$$

where φ is the circumferential angle starting from the negative axis of N, μ is the viscosity, ε is the eccentricity ratio, equals to e/c , L_c is the length of the big-end bearing, x is the location on the axial direction of the bearing, and the zero point of X axis is at the centre of the bearing length (so the range of x is from $-0.5L_c$ to $0.5L_c$), $\bar{\omega}$ is the relative angular velocity of journal and equals to $0.5(\dot{\theta} - \dot{\beta})$.

Consequently, the lubrication force can be integrated from the pressure distribution in Eq. (3) about the angle φ ,

$$F_M = - \int_{-L_c/2}^{L_c/2} \int_{\varphi_1}^{\varphi_2} p R_c \cos \varphi d\varphi dx \quad (4)$$

$$F_N = - \int_{-L_c/2}^{L_c/2} \int_{\varphi_i}^{\varphi_o} p R_c \sin \varphi d\varphi dx \quad (5)$$

where φ_i is the start angle of the integration and φ_o is the end angle of the integration. Finally we can get the lubrication force components in Y and Z direction. The detail of force integration and the transformation from MN frame to YZ frame can be found in the Refs. [6, 7] and [11].

3. Experiments and signal processing

3.1 Experiments on engine test rig

The experiments were carried out on a four-cylinder four-stroke Toyota 3SFE petrol (gasoline) engine. The engine was connected to a hydraulic dynamometer which works as an external load. The engine was controlled to experience different operating conditions, at 3000rpm and at three output torques: 50Nm, 80Nm, 110Nm. The firing sequence is 1-3-4-2. Before all machining works, the engine was disassembled, the clearances of the big end bearing were measured by the engineers in an engine reconditioning workshop and this clearance was regarded as the normal clearance. Next, the journal of the crankshaft of cylinder 2 was ground twice to generate oversize clearance for two different stage experiments: two-time normal clearance (1st stage bearing knock fault) and four-time normal clearance (2nd stage bearing knock fault). The engineer in the workshop also advised on the two stages of oversize clearance for bearing knock faults to make sure the engine could operate safely, so the two stages of oversize clearance were selected as 0.2 and 0.4 mm for two-time normal clearance and four-time normal clearance, respectively.

One Bently Nevada 3300 proximity probe was used to pick up the once-per-rev tacho signal, which corresponds to the Top Dead Centre (TDC) of cylinder 1. A Kistler 6117B spark plug integrated pressure sensor was used to measure the cylinder pressure, and this result was used to update the cylinder pressure simulation model. For each cycle, there are two tacho signals corresponding to the TDCs in the firing stroke and expansion stroke. But only the tacho signal in the firing stroke is useful for the fault diagnostics (especially for identifying the localizations of faults). In the signal processing, the pressure signals in the cylinder 1 are close to the tacho signals in the firing strokes, so they were used to remove the tacho signals in the expansion strokes from the measured tacho signals. An Brüel & Kjær 4384 accelerometer, named as acc5 and shown in Figure 3 (extra accelerometers were also attached on the engine block for the diagnosis of misfires and piston slap faults, and the index number of each accelerometer just corresponds to its input channel number on the acquisition front-end) was attached adjacent to the big end bearing in the cylinder 2 and its measured vibration signal was used for the diagnosis of bearing knock faults. The sampling frequency of the recording system was set at 25600 samples/sec.

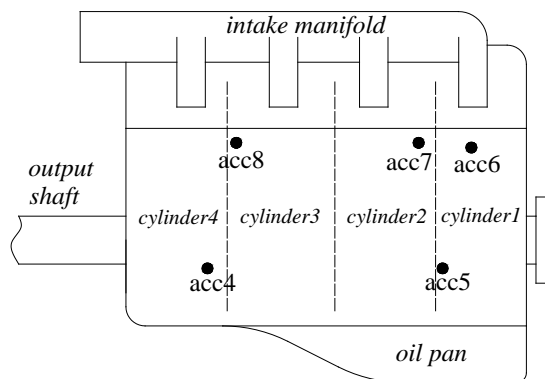


Figure 3. Accelerometer layout on the engine block

3.2 Signal processing

An example of raw vibration signal with bearing knock fault is shown in Figure 4. The former work [12] has found that owing to their characteristics of second order cyclostationarity, the diagnostic information in the signals for bearing knock fault should reside in the repetition frequency and pattern of the bursts. The envelope signals should be deterministic (actually first order cyclostationary) and therefore can be synchronously averaged. The envelope signal was also squared to represent it in power units (proportional to variance).

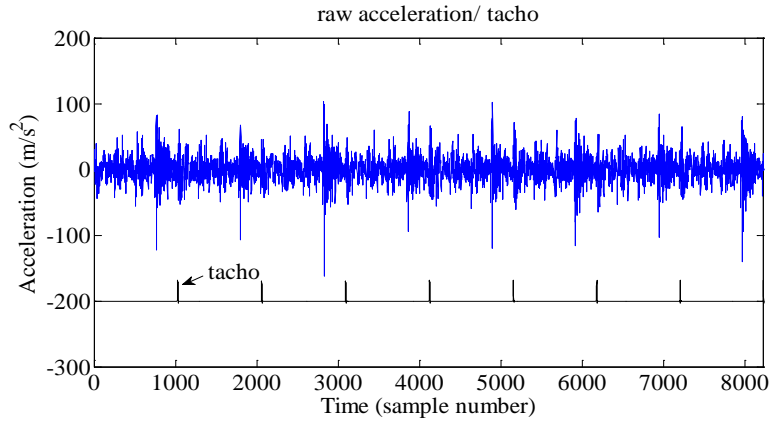
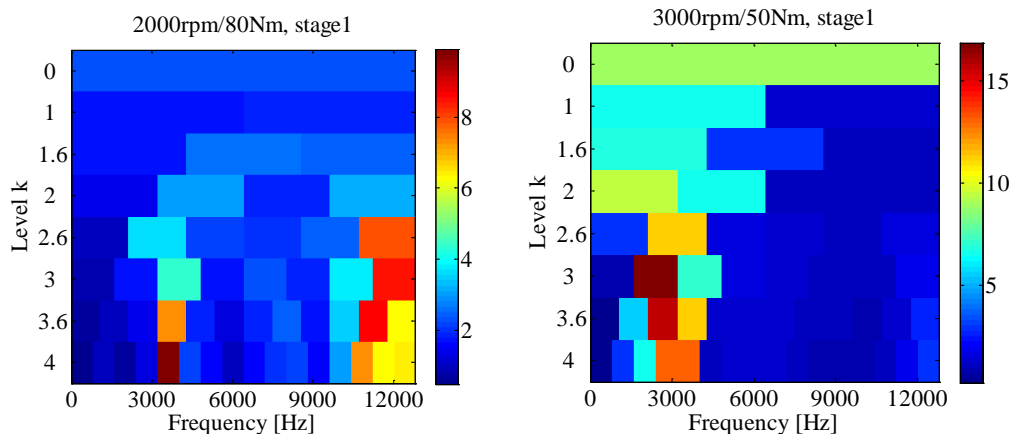


Figure 4. Raw vibration signal with 1st stage bearing knock fault at 3000rpm/110Nm

In order to find how the impulsiveness of the measured faults signals varies with frequency and remove some interfering noise, before performing envelope analysis, the “fast kurtogram” developed by Antoni [13] was used to provide some guidance for the selection of the optimum demodulation band for the bearing knock fault diagnosis. Compared to its application on rolling element bearing diagnosis, the situation for a big end bearing is more complicated. Firstly, the transfer function for the bearing knock is more complicated. The impact signal transmits from the crank journal to the main bearing, then to the external measurement point. Secondly, the oil lubrication of the big end bearing is subjected to a rotating dynamical load. Thirdly, because the measurement point is actually close to the main bearing between cylinders 1 and 2, the measurement signal is prone to the influence from the bearing knocks and combustion noises from the adjacent cylinder 1. However, it was found that the kurtogram was also useful in this case to find the frequency bands with highest impulsiveness. The range for bearing knock faults is wide (as shown in Figure 5), but mostly lie below 6400Hz. The cut-off frequency at 3200Hz is good for the bearing knock fault cases in the 1st stage, but when the cases in the 2nd stage were also considered (the kurtogram results for this stage are more scattered), it was found that the cut-off frequency set at 4400Hz is better for all cases in both stages. Therefore, allowing for the comparison of the whole range of bearing knock faults a common frequency band was found (400-4400Hz).



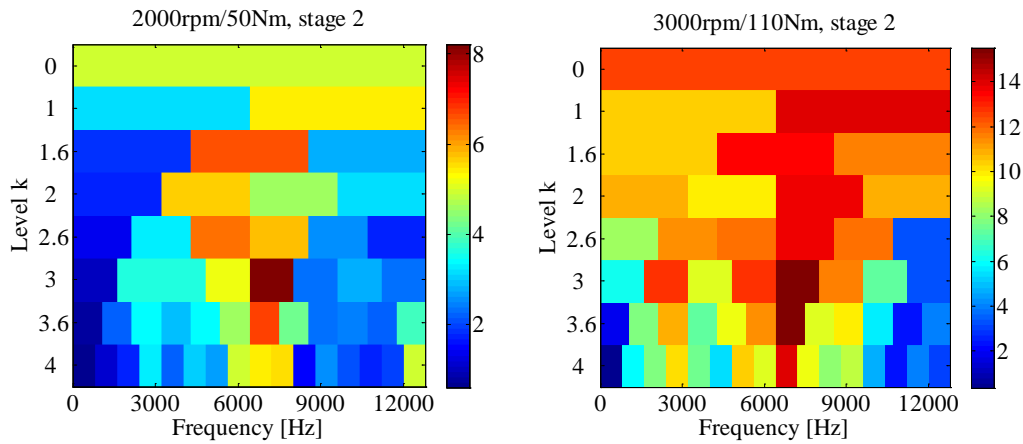
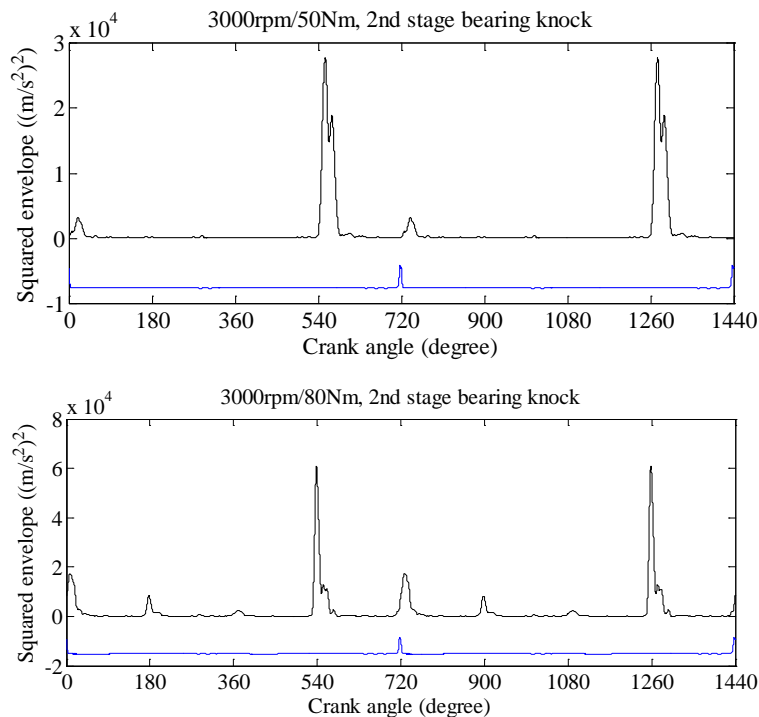


Figure 5. The “kurtogram” map for bearing knock faults

As mentioned above, the raw bearing knock fault signals contain some low frequency combustion noise and other random mechanical noise, but after the band pass filtering, it was found that the noise was essentially removed. It is worth mentioning that another combustion fault is engine knock, which also excites high frequency resonances. However, earlier research [14] has demonstrated that engine knock excites high frequency resonances with a typical range of 4-10 kHz. Therefore, the selected band pass filters should separate the engine combustion knocks from bearing knock fault. It should also be noted that other work [15] has studied piston slap, where a similar frequency range was used to generate envelope signals, but these were measured at different measurement points (eg, 6, 7, 8 in Figure 3) and measurement location is also a good separator of the different mechanical faults, as shown by results later in the paper, where some cases with piston slap faults were included in the “not bearing knock” category.

It is also worth pointing out that, for all the figures in the following section of this paper, the zero degree point in the crank angle corresponds to the TDC of the cylinder 1 in the firing stroke. The squared envelope signals at different load and same speed are shown in Figure 6, where it can be seen that the peak values of the squared envelope for bearing knock fault have no proportional relationship with the increase of load. That is because the bearing knock impact force is mainly attributed to the inertia force of the piston and connecting rod components, rather than the combustion force. Note that near TDC these forces act in opposite directions.



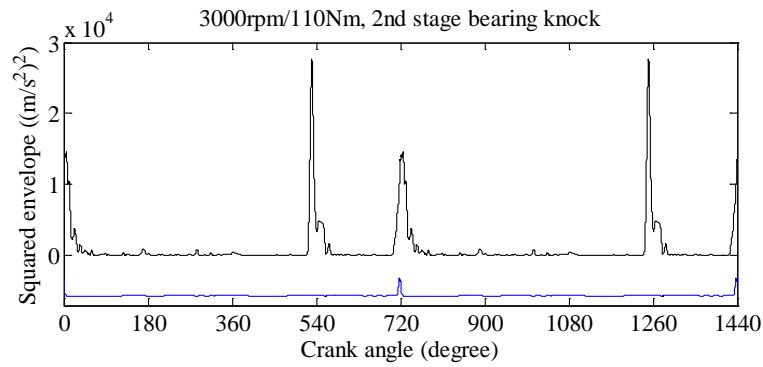


Figure 6. Squared envelopes for bearing knock faults at same speed but different load

An comparison example of the squared envelopes at the same speed/load condition but different fault stages is shown in Figure 7 and it can be found there is a substantial increase for the magnitude peak value (about 10 times from stage 1 to stage 2). The situation also happened for other speed/load conditions. The substantial increase implies the possibility for the fault severity identification from the squared envelope signals.

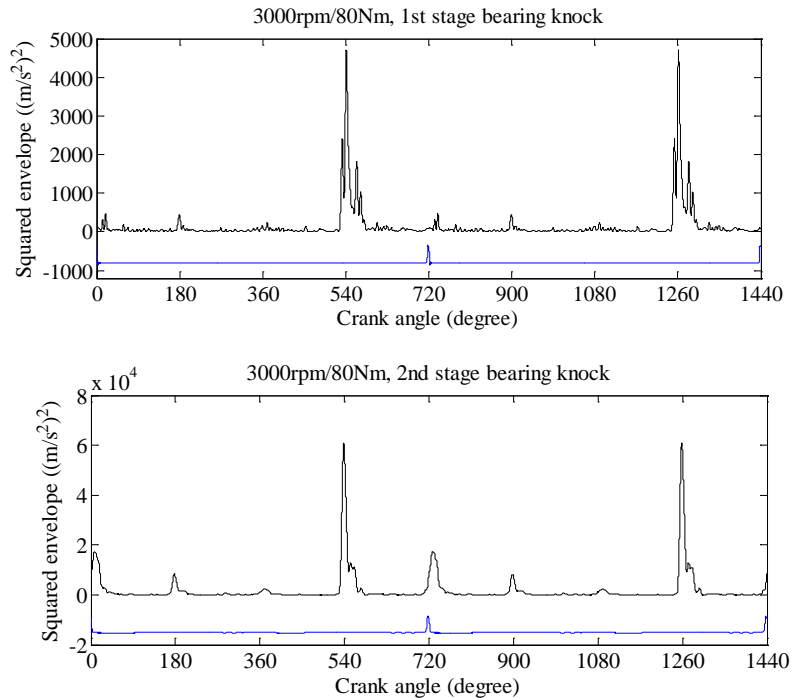


Figure 7. Envelope signals with different fault severity levels at 3000rpm/80Nm

3.3 Experiments on the kinematic components in the simulation model

Based on the analysis in Section 2, if we want to simulate the big-end bearing knock fault in IC engine (also for the comparison between the experiment and simulation), the parameters of the relevant engine components should be measured first. The parameters were measured from a disassembled engine which is identical to the test engine. Some parameters can be directly obtained from measurements, such as the length of the connecting rod and crank throw radius, but some parameters need further calculation from the measurements, such as inertia property and polar moment of inertia.

The vibration based mass line method is a suitable method to measure the inertia properties of connecting-rod. More detail of mass line method can be found in Refs. [16, 17]. The connecting rod was excited by a Brüel & Kjaer 8202 impact hammer. An Endevco 66A12 triaxial piezoelectric accelerometer was used to measure the accelerations. In order to get accurate final results by the mass line method, during the measurement, a couple of factors were specially considered:

- Suspensions were soft enough to get clear mass lines and the rigid modes and first bending mode could be consistently separated.
- Different from normal modal analysis experiments, the mass line method needs to measure the whole FRF matrix not just one column or one row of FRFs.
- The reference (excitation) points and response (measurement) points were the points with largest rigid body motion.
- Because the component was small, the size of the accelerometers was taken into account.
- In order to excite the points precisely, a very stiff rod (drift) was used. The top of the rod was hit by the hammer and end of the rod was located accurately on the structure and aligned. Therefore, the excitation location and direction could always be fixed.
- Different excitation DOFs, at least one in each of the X, Y, and Z directions, were selected.

The inertia property measurement of the connecting rod is shown in Figure 8. The Rigid Body Calculator in LMS Test.Lab was used to calculate the inertia properties from the measured FRFs.



Figure 8. The inertia property measurement of the connecting rod

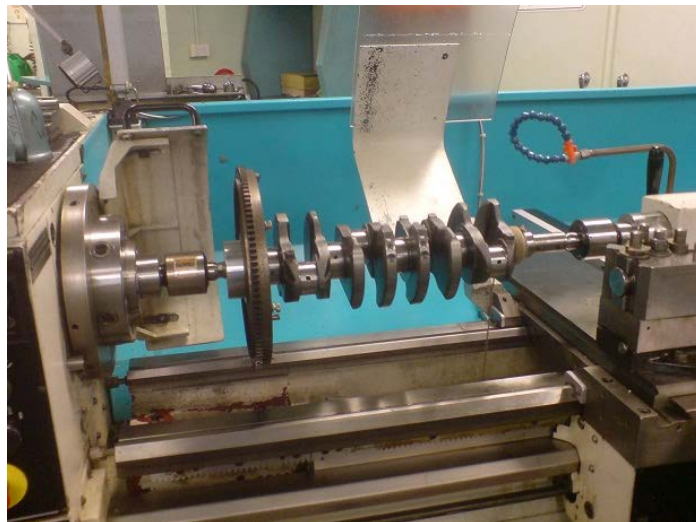


Figure 9. Polar moment of inertia measurement of the crankshaft

Figure 9 shows the measurement of polar moment of inertia of the crankshaft (with flywheel). In the measurement, the two ends of the crankshaft were suspended in two lathe centres. A string with a small mass at the end (mass m) was wound on the centre main journal of the crankshaft. After the mass was released from a certain height (h), the crankshaft was caused to rotate by the plumb movement of the mass. Since the length of the wound string was made $1.5h$, owing to the inertial effect, the crankshaft continued to rotate extra turns after the mass reached the floor. Finally the crankshaft stopped by virtue of the friction forces (moments) at the supports.

There are two assumptions for the experiment and the calculation of crankshaft polar moment of inertia: The energy dissipation per revolution was taken as constant and the CG of the crankshaft was coincident with its axis of rotation. During the measurement, the time interval of the whole plumb movement process was recorded as t . v denotes the velocity of the mass at the time t . Same as the definition in Section 2, I_j represents the polar moment of inertia of the crankshaft. ω is the angular velocity of the crankshaft at the time t . E is the energy per turn consumed by friction. r_{mb} is the radius of the middle journal of the crankshaft. While the mass was falling (interval t), the crankshaft rotated N turns. After the mass stopped, the crankshaft continued rotating an extra n turns.

In the interval t , the energy change of the mass is $mgh - 1/2mv^2$ and is equal to the maximum kinetic energy of the crankshaft and the energy consumed by the friction.

$$mgh - \frac{1}{2}mv^2 = \frac{1}{2}I_j\omega^2 + NE \quad (6)$$

After the mass hit the floor, the kinetic energy of the crankshaft was transformed into frictional energy in the extra n turns.

$$\frac{1}{2}I_j\omega^2 = nE \quad (7)$$

Substituting the Equation (7) into Equation (6) and eliminating variable E , one can get:

$$m\left(gh - \frac{v^2}{2}\right) = \frac{1}{2}I_j\omega^2\left(\frac{n+N}{n}\right) \quad (8)$$

and

$$m\left(gh - \frac{v^2}{2}\right) = \frac{1}{2}I_j \frac{v^2}{r_{mb}^2} \left(\frac{n+N}{n}\right) \quad (9)$$

The mass falling can be considered as a constant acceleration motion starting from zero, so substituting $v = 2h/t$ into Equation (9), the polar moment of inertia of the crankshaft can be calculated by

$$I_j = \frac{mt^2 r_{mb}^2}{2h} \left(g - \frac{2h}{t^2}\right) \left(\frac{n}{n+N}\right) \quad (10)$$

3. Simulation and result comparison

3.1. Simulation model

Based on the principle shown in Section 2, we built the kinematic/kinetic model and lubrication model for the big-end bearing with oversize clearance and lubrication, and two models can be related by a closed loop. The inputs to the kinematic/kinetic model are the two hydrodynamic forces in Y and Z directions (F_y and F_z , which are the outputs of lubrication model) and the outputs are the displacement parameters z_p , θ , α and the corresponding velocities. The inputs to the lubrication equations are the eccentricity ratio ε , the coordinate angle β and their derivatives, which can be calculated from the locations of the bearing centre and the journal centre in the kinematic/kinetic model. Now it is necessary to explain why we only applied the existing modelling methods (based on Refs [5-7]) to simulate the bearing knock faults in this paper. As studied in Section 3.2, the bearing knock fault diagnostic information lies in the squared enveloped signals (the interest lies in the total impact energy) rather than raw vibration signal. Many hydrodynamic (include friction mechanism or asperity lubrication), elasto-hydrodynamic, and thermo-hydrodynamic modelling methods have been recently developed for the mechanism design of journal bearings/revolute joints, so as to increase their performance, lifetime and so on, but for the bearing knock fault diagnosis, these modelling approaches may generate more accurate raw impact forces/vibration signals, but the processed squared envelope signals should have very small difference from those created by the model in this paper. This is because after the band pass filter and envelope processing, the nearby instantaneous samplings of the raw acceleration will be ‘‘mixed’’ and ‘‘wrapped up’’ together (this can be found from the comparison of raw signals in Figure 4 and the enveloped signals shown in Figures 6 and 7), and more importantly, the total energy under the envelope contour should be almost fixed for all methods.

Another important input of the simulation model, the combustion chamber pressure curves, corresponding to output loads of the engine at 50Nm, 80Nm and 110Nm, were simulated first and saved as look-up tables in the

model. For the compression and expansion process, the pressure can be solved by the classic polytropic equation. The pressure in the combustion process can be calculated by Wiebe's functions [18, 19]. The burn rate $w(\theta)$ and heat release $Q(\theta)$ as functions of crank angle are written as:

$$w(\theta) = \frac{6.908(m_v + 1)}{\theta_d} \left(\frac{\theta}{\theta_d} \right)^{m_v} e^{-6.908(\theta/\theta_d)^{m_v+1}} \quad (11)$$

$$Q(\theta) = w(\theta)r_{comb}m_{fuel}LHV \quad (12)$$

where m_v is Wiebe's combustion characteristic exponent, θ_d is the combustion duration in degrees about the crank angle, r_{comb} is the combustion efficiency, m_{fuel} is the fuel injection quantity which can be found from the look-up table of the engine fuel injection map, LHV is the lower heating value of the fuel, normally 43.9MJ/kg for petrol.

According to the thermodynamic relations among the pressure, chamber volume and heat release in the combustion process, the combustion pressure P_{comb} can be calculated step-by-step with finite angle increase:

$$P_{comb}(\theta + \Delta\theta) = \frac{\gamma - 1}{V(\theta)} \left(Q(\theta) + P_{comb}(\theta) \frac{(\gamma + 1)V(\theta) - \gamma V(\theta + \Delta\theta)}{\gamma - 1} \right) \quad (13)$$

The models were built in Simulink. As mentioned before, the loop of kinematic/kinetic model and lubrication model were iteratively solved step by step. The numerical integration in the kinematic/kinetic model was carried out by the ode45 Dortmund-prince approach. Variable time step mode was applied in the simulation, and the range of the time step is from 1e-10 to 2e-7 seconds. All the output signals have the same sampling ratio at 200 kHz. The initial parameters for the kinematic/kinetic model are shown in Table 1. Based on the initial conditions, the locations of the bearing centre and the journal centre in XY global frame were calculated first. Next the lubrication forces were calculated in the lubrication model and returned the next step input values to the kinematic/kinetic model, and then the simulation loop continued.

Table 1: Initial parameters of the kinematic/kinetic model

Paramters	Value	Unit
Initial displacement of the piston	0.177997	m
Initial velocity of the piston	0	m/s
Initial angular displacement of the connecting rod	pi	rad
Initial angular velocity of the connecting rod	-100.3/-66.9/-49.7	rad/s
Initial angular displacement of the crank	0	rad
Initial angular velocity velocity of the crank	315/210/156	rad/s

By changing the external load (as shown in Eq. (2)) and adjusting the clearance, the impact forces with different degrees of bearing knock at different operating conditions were simulated. As the analysis in Section 2.2, two forces in X and Y directions were obtained from the simulation, next their spectra were respectively multiplied by the corresponding measured Frequency Response Functions (FRFs) to get the impulse responses in the frequency domain. Afterwards, the simulated time series accelerations were obtained by inverse Fourier transform. During the FRF measurements, in accordance with Maxwell's theory of reciprocity, an hammer excitation force was applied to the measurement point on the engine block surface (acc 5 in Figure 3) and the acceleration was picked up on the crank journal in X and Y direction separately. After the FRFs from the big-end bearings in different cylinders to the location of acc5 measurement point were measured, the vibration signals due to the faults happening in different cylinders can be simulated as well.

3.2 Comparison of experiment and simulation

Since the target of the simulations here is to produce the decent accurate envelope signals rather than the raw acceleration vibration signals, the envelope signals of the simulated vibration were further calculated. Similar to the experiment results, it was found that amplitude in normal bearing clearance condition is quite small compared to that with 1st stage faults, so detecting the faults from normal clearance conditions should not be difficult based on the simulated data. The simulated squared envelopes with two stages of bearing knock faults

are shown in Figures 10 and 11. It can be seen that even through the peak value of the simulated amplitude with the 2nd stage bearing knock fault is smaller than that of the experiment, but the change ratios of the peak values (about 10 times each stage) from the 1st stage to the 2nd stage are adequate for the fault detection and the severity identification.

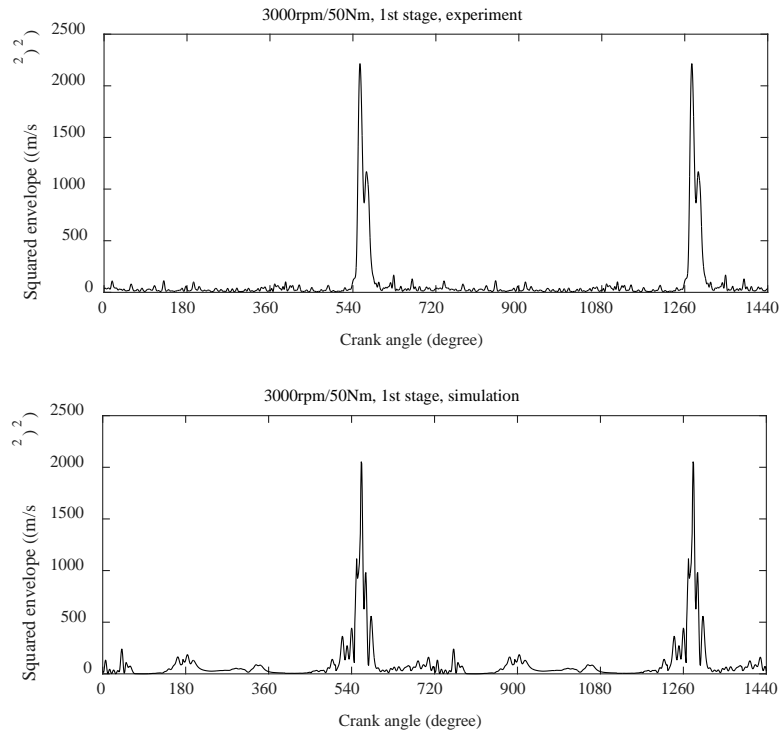


Figure 10. Experimental and simulated squared envelope signals with the 1st stage fault at 3000rpm/50Nm

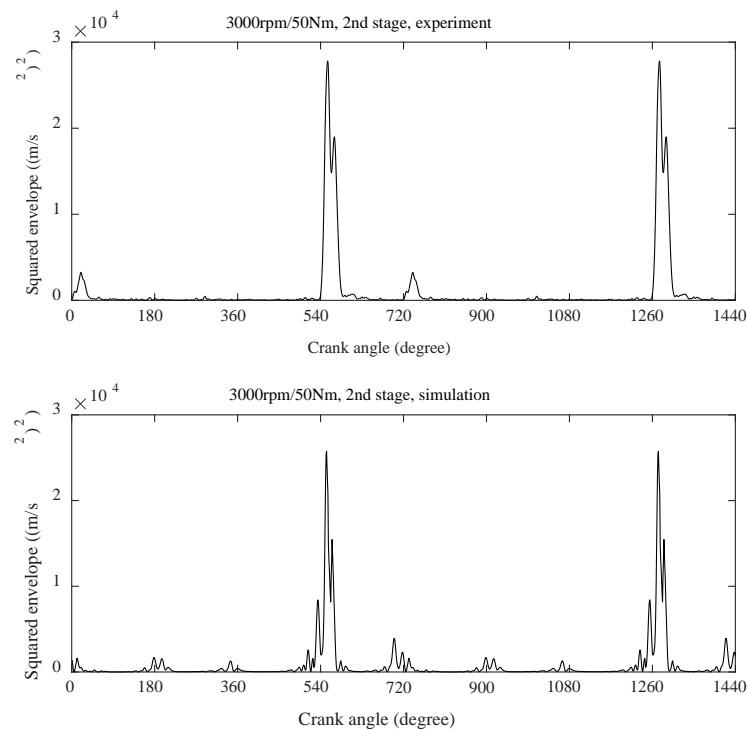


Figure 11. Experimental and simulated squared envelope signals with the 2nd stage fault at 3000rpm/50Nm

5. Intelligent diagnosis

As demonstrated by our former works in Refs. [20, 21], the amplitudes of the Fourier coefficients of the squared envelope signals were potential features to detect the bearing knock faults and also identify the severity level of the faults, and the phases of the Fourier coefficients were potential features to localize the faults in which cylinder. Therefore, separate neural networks were developed for the three phases of the diagnosis problem. The first phase is the bearing knock fault detection phase and the inputs of the networks are the selected amplitude features. The second phase is bearing knock fault localization phase and the inputs of the network are the selected phase features. Based on the detection and location results, the third phase network uses the selected amplitude features as inputs to identify the severity of the bearing knock faults. The success of the application of ANNs to the diagnostic problems also depends on which vectors are selected as the inputs to the networks. Due to the complication of vibration signal characteristics, the selection of best features from the processed signals is not straightforward for the bearing knock fault diagnosis; therefore an advanced selection technique needs to be used for the feature selection, especially for the selection of amplitude features.

5.1 Amplitude feature selection

There are two common advanced feature selection categories, one is filter-based methods and the other is wrapper methods. Filter based methods generally employ some quality criterion to score individual feature and provide a ranking (for example, the p -value of a t -test comparing two populations of interest with regard to the expression levels of the gene in the populations), and then based on this ranking order, subset of genes can be selected, normally manually setting a threshold. Wrapper methods use the inductive algorithm to evaluate the value of a given subset as a group, rather than individually [22]. Filter based methods requires less computing cost than wrapper methods, but the wrapper methods can qualify the important interactions among the features and can avoid the selection of redundant (highly-correlated) features. GA based feature selection is one of the typical wrapper methods. There should be some relevance among the candidate amplitude features for the bearing knock fault diagnosis. If highly-correlated amplitude features are selected together, the selected subset gives a poor representation of the entire feature set. Therefore, it is necessary to use GA based wrapper method [23] to find the best amplitude feature of mutually maximal dissimilarity, so as to increase the accuracy of the intelligent diagnosis system for the bearing knock faults.

In the GA approach, the solutions are encoded as a chromosome, where the individual candidate feature is coded as an individual binary digit (1/0). After a series of iterative computations on the chromosomes, the optimal solution is obtained. Two main operators can be used in the computation process: one is crossover and the other is mutation. Crossover is the basic operator for generating new chromosomes and it is a recombination process. In the process, new offspring are created, which share some common parts of both parents' genetic material. In addition, mutation is a background operator and acts as a safety net to recover good genetic material that may be lost through the action of selection and crossover [24]. The detail of GA can be found in [22, 24].

There are many new developments in the past twenty years, and a Mutual Information based GA approach [23] was used here to select the optimal features for the fault diagnosis. The fitness function of the GA algorithm is based on Max-Relevance and Min-Redundancy [24]. The maximum relevance condition is maximal $H(Y/x_i)$:

$$H(Y / x_i) = \frac{1}{|N|} \sum_{x_i \in N} I(Y, x_i) \quad (14)$$

In the cases of bearing knock faults, the Y is the detection target and the x_i are the amplitude features. The minimum redundancy condition is minimal $H(x_i/x_j)$:

$$H(x_i / x_j) = \frac{1}{|N|^2} \sum_{x_i, x_j \in N} I(x_i, x_j) \quad (15)$$

where N is the selected or desired feature subset, $|N|$ is the number of features in the subset, I is the mutual information of two variables m and n :

$$I(m, n) = \sum_{i,j} p(m_i, n_j) \log \frac{p(m_i, n_j)}{p(m_i)p(n_j)} \quad (16)$$

where $p(m, n)$ is the joint probabilistic distribution of m and n . $p(m)$ and $p(n)$ are marginal probabilities respectively.

The feature selection for the diagnostics of faults is a combinatorial optimization problem with the constraint that the feature index cannot repeat, and the introduction of a mutation operator will make the implementing of this constraint very difficult, so only the crossover operator was used. Based on the evaluation results of the fitness function (Max-Relevance and Min-Redundancy), the algorithm organizes the individual feature by its fitness values. During the crossover, each individual has a probability of reproduction that is given by its fitness value and more adapted individuals have more probability of participation. In order to take the individual fitness values into account, a random variable x_n (on which the selection criterion of crossover is based) was given by:

$$x_n = \frac{e^{ax} - 1}{e^a - 1} \quad (17)$$

where a is a positive and arbitrary constant and $x \in [0, 1]$ is a random variable with uniform distribution. The detail of the probability of reproduction of the algorithm can be obtained in [26]. Meanwhile if elite operators were applied, the best result of the current generation would be saved in the next generation and the training curve becomes monotonically descending, without oscillations. The amplitudes of the first forty harmonics of the processed envelope signals were considered as the candidate features for the fault detection and severity identification. In total, we had 37 real experimental cases, including 10 cases with different bearing knock faults and 27 cases with normal bearing clearance. The GA code is based on the program developed by the first author of Ref. [23, 26], the inputs of the program are the desired number N (in Eqs.(14, 15)) of selected features (we set it as 10), a matrix X , in which each column is a feature vector example (so a 37×40 matrix for the bearing knock fault diagnosis), and its respective target data y , which is a row vector (37 row in bearing knock fault diagnosis). The output vector represents of the indexes of the selected features that composes the optimum feature set. The elite value was set as 1. The maximal generation was set to 100. Finally, the best ten amplitude features were selected by the GA as the inputs of the MLPs of the intelligent diagnosis system, and they are: 2, 29, 14, 34, 16, 17, 9, 20, 8 and 18.

5.2 Phase feature selection

Based on the processed experimental data, when the bearing knock faults happen in a certain cylinder in different speed/load conditions, if the phases of the n th harmonic of the most cases are clustered, it implies that the phase of this harmonic is a potential feature for localizing the faults. Specifically, if the phases of the 1st harmonic of all cases are clustered, or near fixed, it means that there will be 90 degree difference for the faults in individual cylinders (60 degree difference for a six cylinder engine); therefore the fault localization can be directly and solely identified by the phases of the 1st harmonic. Moreover, if a fault occurs in the cylinder 1 or cylinder 3, the phase difference of the 1st harmonic is 90 degrees, and the phase difference of the 2nd harmonic is 180 degrees (90×2), so using the clustered phases of the 2nd harmonic can differentiate the faults in cylinder 1 from those in cylinder 3. On the other hand, if a fault occurs in cylinder 1 or cylinder 4, the phase difference of the 1st harmonic is 180 degrees, and the phase difference of the 2nd harmonic is 360 degrees (180×2), so using the clustered phases of the 2nd harmonic cannot differentiate the faults in cylinder 1 from those in cylinder 4. Ideally, the phase difference of the 2nd harmonic can only identify the location of the fault in 50% of cases. Note that the 4th harmonic represents the firing frequency, so the phases of the 4th harmonic and the $(4*n)$ th harmonics are useless to identify which cylinder has bearing knock faults. The phases of the first twenty harmonics of all experimental bearing knock fault cases are shown in Figure 12. It was found that the phases of higher harmonics are more scattered. So the conclusion is that the best phase features should be selected from the first ten harmonics rather than higher harmonics. It was also found that the phases of the 1st harmonic are near fixed, and the phases of the 2nd and 5th harmonics can also be added into the selected phase feature vector to make the fault severity identification more robust.

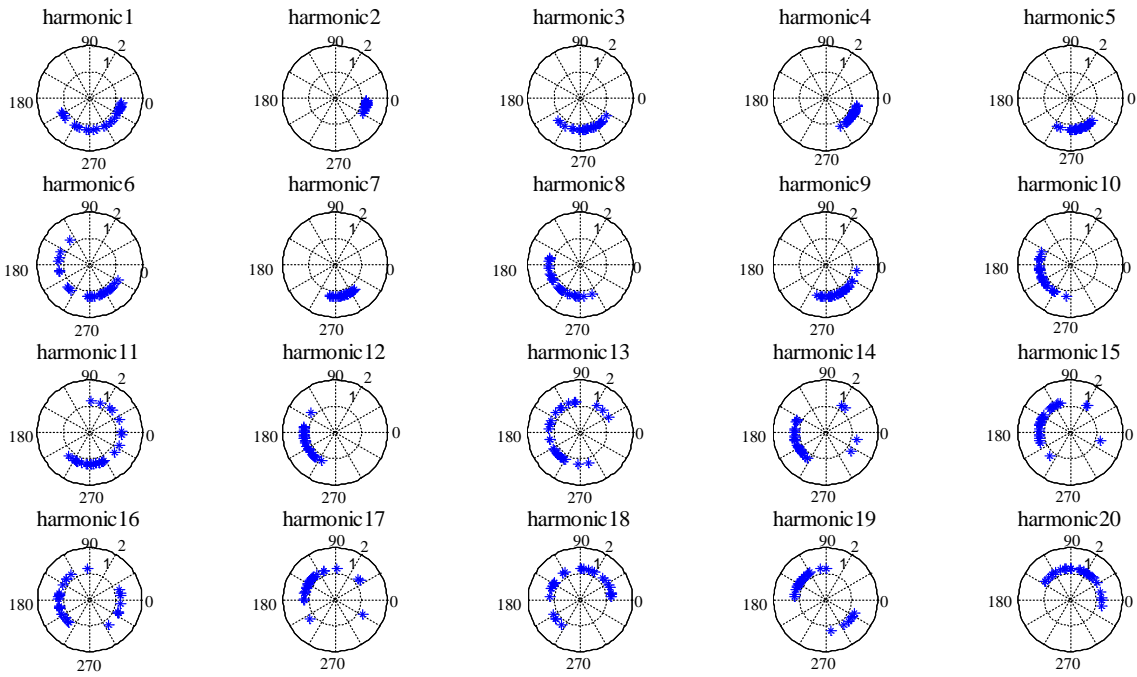


Figure 12. Phase distribution of the first twenty harmonics of bearing knock faults

5.3 ANN-based system

In each simulation, we simulated each bearing knock fault in each cylinder, at each load (50Nm, 80Nm, 110Nm at 3000rpm). The signals (with normal or oversize clearance) from the simulation model are deterministic, but the measurement signals in reality always have some deviations; in particular the transfer function from the bearing knock impact point to the measurement point may vary a little with operating condition (as shown in Figure 3, the measurement point is not very close to the big end bearing of cylinder 2, and the vibration is prone to contamination from other combustion and mechanical noises). Therefore, variations in the simulated signals were instituted to create representative cases for the ANN training. The standard deviation of the variations was set by analysing the normal conditions from experiment and was applied to the simulated enveloped signals. Two MLP networks were applied for the fault detection (MLP1) and fault severity identification (MLP2). A PNN was used to classify the locations of the faults. Because the outputs of the PNN can be any integer number, such as 1, 2, 3 and 4, so they can directly correspond to the number of the cylinder in which a fault happened. Both MLP and PNN are the typical feed-forward neural networks. In the feed-forward network, the artificial neurons are arranged in a layered configuration containing an input layer, usually one “hidden” layer, and an output layer. Neurons in the input layer introduce normalized or filtered values of each input into the network. Neurons in the hidden and output layers are connected to all of the neurons in the preceding layer. That activation signal then is passed through a transfer function to generate an output signal which is sent to processing neurons in the next layer. The configuration of the intelligent diagnosis system is shown in Figure 13.

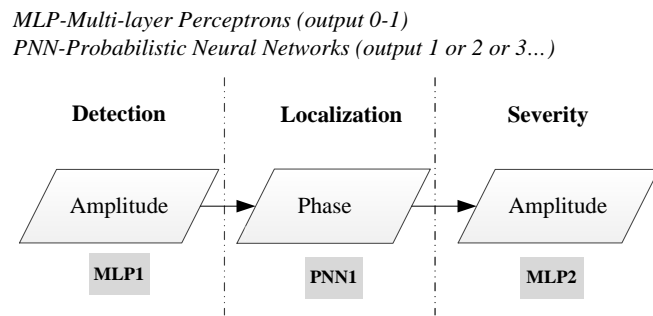


Figure 13. Structure of the three-stage ANN system

The configurations of two MLPs are similar, both have three layers: input, hidden and output. But the transfer functions of two MLPs are different. A log sigmoid function was used as the transfer function of MLP1, because it can efficiently classify all cases into two groups (converging to two boundary conditions); 0 for normal condition and 1 for bearing knock fault. In contrast to our former works [12], the saturating linear function, rather than nonlinear log sigmoid function, was introduced into MLP2 as its transfer function. The diagram of MLP2 is shown in Figure 14. The number of neurons in hidden layer was determined using a trial and error procedure (that shown in Figure 14 has 30 hidden neurons). I_w and b are respectively the weight and bias factors distributed to the individual elements of the input feature vectors. During the training stage, MLPs were led to a specific target output by adjusting the values of the weights and bias among the elements of the input vectors. This saturating linear function is linear in the interval $(-1,+1)$ and saturates outside this interval to -1 or $+1$, returning the value of x spread over the range 0 to 1. Thus the output of MLP2 can be any value linearly distributed from 0 to 1, for instance, 0.5 represents two-times normal clearance, 1 for four-time normal clearance and 0.75 for three-time normal clearance. Therefore, the output results for severity identification are more logically in agreement with the real situation. If the clearance could be larger than the largest value tested (in this case 4 times) the range of linear distribution could be extended to the largest allowable value.

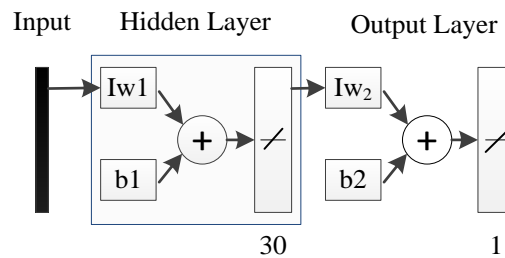


Figure 14. The diagram of the MLP2

MLP1 was trained purely by using the 120 simulated cases, and the case distribution is, 72 training cases with different bearing knock faults (12 cases for each severity level at each load condition, 18 cases for each cylinder) and 48 training cases with normal bearing clearance (16 cases for each load condition). MLP1 was tested by 37 real experimental cases, which includes 10 cases (case number from 1 to 10) with different bearing knock faults (including all three load conditions) and 27 cases (case number from 11 to 37) with normal bearing clearance. It is worth noting that the cases 32 to 37 are with oversize piston-cylinder wall clearance faults (piston slap faults). The MLP1 output results are shown in Figure 15. Particularly, the output of case 37 is 0.3034, but if the detection threshold is defined as 0.5, the MLP1 successfully classified the bearing knock fault cases from the noise contamination from other mechanical faults. Therefore we can get a conclusion that MLP1 100% accurately detected the bearing knock faults.

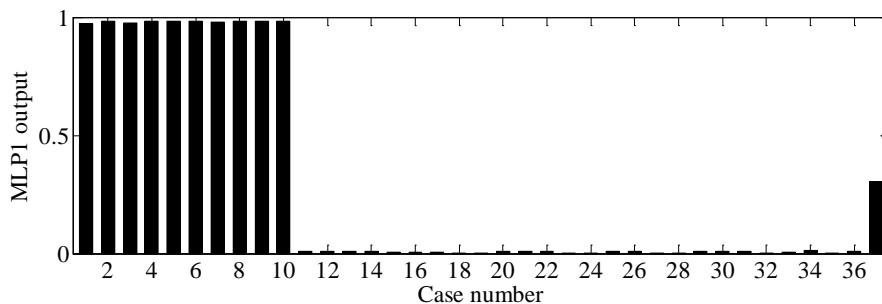


Figure 15. Output of MLP1 for fault detection

Because the fault cases have been selected by MLP1, the number of the training cases for the PNN were reduced to 72. 10 real experimental cases (all are in cylinder 2) were inputted the PNN for the test purpose. The test results show that all the output values from PNN are 2 (the cylinder number), therefore we can conclude that the PNN 100% correctly identified the location of the bearing knock faults. As mentioned before, in the MLP2, the output value of 0.5 means the 1st stage oversize bearing clearance faults and the output value 1 means the 2nd stage oversize bearing clearance faults. The final outputs of MLP2 is shown in Figure 16. Specially, we introduced an error value here to evaluate the performance of MLP2,

$$Error = \sum_{i=1}^N |(ANN(i) - VAL(i))| \quad (18)$$

where ANN denotes the output values from the MLP2 and VAL denotes the corresponding target values, I is the number of test cases. A lower fitness criterion means better MLP performance. Meanwhile, the values of fitness criterion also provides a guide for the trial and error process about the number of hidden neurons. It can be seen that the MLP2 also obtained good results for the fault severity identification. Moreover, due to the new saturating linear transfer function design, the MLP2 in this paper (error value 0.193) has better performance to identify the severity level in comparison with the MLP2) with log sigmoid transfer function (the error value (0.244) in reference [13].

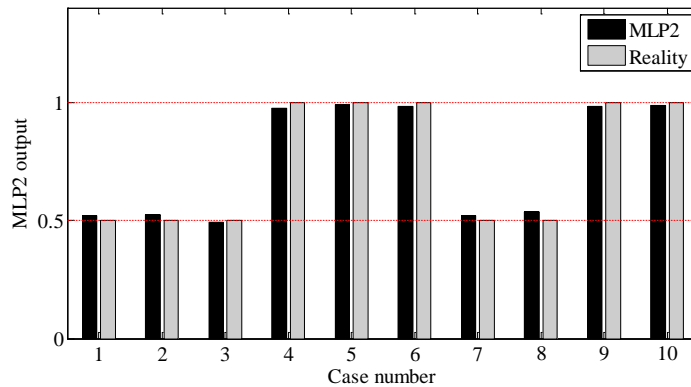


Figure 16. Output of MLP2 for fault severity identification

6. Conclusion

This paper built a model based intelligent diagnosis system for the bearing knock faults in IC engine, and the ANNs were used to implement the automated diagnosis. Based on the analysis on the experiment data from the engine test rig, it was found that the squared envelope signals, rather than the raw vibration signals, have useful information for the big-end bearing fault diagnosis. Therefore, based on existing modelling research, a proper modelling approach on the journal bearing with clearance and lubrication was selected to simulate different big end bearing knock faults and generate accurate squared envelope signals, and the abundant simulation data solved the issue of adequate data requirement of the ANNs' training. Meanwhile, as the important inputs of the simulation model, the inertia properties of the relevant engine components were also measured and calculated by some measurement techniques. The optimal amplitude features were extracted and selected (by the advanced algorithms) as the inputs of the detection stage and severity identification stage, and selected phase features were used to localize which cylinder had a fault. The results have demonstrated that the developed system could efficiently diagnose different bearing knock faults, and in particular could accurately identify the severity levels of the faults. Moreover, because the simulation models are based on the thermodynamic, dynamic and hydrodynamic principles of IC engines, the proposed intelligent diagnostic system can in principle be applied to any IC engine.

Acknowledgements

The authors would like to convey their special gratitude to the Australian Research Council and LMS International for sponsoring this research under Linkage Project LP0883486.

References

1. Taylor C.M.: Automobile engine tribology – design considerations for efficiency and durability, *Wear*, vol. 221, 1-8 (1998)
2. Flores P., Ambrósio J., Claro J. C. P., Lankarani H. M. and Koshy C. S.: Lubricated revolute joints in rigid multibody systems. *Nonlinear Dynamics*, vol.56, no.3, 277-295 (2009)
3. Flores P., Ambrósio J., Claro J. C. P., Lankarani H. M. and Koshy C. S.: A study on dynamics of mechanical systems including joints with clearance and lubrication. *Mechanism and Machine Theory*, vol. 41, no. 3, 247–261 (2006)

4. Flores P., Lankarani H. M., Ambrósio J. and Claro J. C. P.: Modeling lubricated revolute joints in multibody mechanical systems. *Proceedings of the Institution of Mechanical Engineers, Part K: Journal of Multi-body Dynamics*, vol.218, 183–190 (2004)
5. Machado M, Costa J, Seabra P, Flores P.: [The effect of the lubricated revolute joint parameters and hydrodynamic force models on the dynamic response of planar multibody systems](#). *Nonlinear Dynamics*, no.1-2, vol.69, pp.635–54 (2012)
6. Daniel G. B., Cavalca K. L.: Analysis of the dynamics of a slider-crank mechanism with hydrodynamic lubrication in the connecting rod-slider joint clearance. *Mechanism and Machine Theory*, vol.46, no.10, 1434–1452 (2011)
7. Chen, J., Randall, R., Feng, N., Peeters, B. and Van Der Auweraer, H.: Modelling and diagnosis of big-end bearing knock fault in internal combustion engines. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*. vol. 228, vo. 16, pp2973-2984 (2014)
8. Sommerfeld, A., “Zur hydrodynamischen Theorie der Schmiermittelreibung”, *Z. Math. Phys*, vol.50, pp.97- 155 (1904)
9. Dubois G. B. and Ocvirk F. W.: Analytical derivation and experimental evaluation of short bearing approximation for full journal bearings, NACA report 1157, Washington, USA (1953)
10. Hirani H., Athre K. and Biswas S.: Rapid and globally convergent method for dynamically loaded journal bearing design, *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology*, no.3, Vol. 212, pp. 207-214 (1998)
11. Frêne, J., Nicolas, D., Degneurce, B., Berthe, D. and Godet, M.: *Hydrodynamic lubrication - bearings and thrust bearings*, Elsevier, Amsterdam, Netherlands (1997)
12. Chen J., Randall R. B., Peeters B., Desmet W., and Van der Auweraer H.: Artificial neural network based fault diagnosis of IC engines. *Key Engineering Materials*, vol. 518, 47-56 (2012)
13. Antoni, J.: Fast computation of the kurtogram for the detection of transient faults. *Mechanical Systems and Signal Processing*, vol.21, no.1, 108-124 (2007)
14. Ren, Y.: Detection of knocking combustion in diesel engines by inverse filtering of structural vibration signals, PhD Thesis, UNSW, Australia (1999)
15. Chen J., Randall R. B.: Automated diagnosis system for mechanical faults in IC engines. *CM/MFPT 2013*, 17-20th June 2013, Kraków, Poland
16. J. Bretl, P. Conti, [Rigid body mass properties from test data](#), *Proceeding of the 5th Intentional Modal Analysis Conference, London, England, (1987) 655-659*.
17. E. Mucchi, G. Bottoni, R. Di Gregorio, [Indirect measurement of the inertia properties of a knee prosthesis through a simple frequency-domain technique](#), *Transactions of the ASME: Journal of Medical Devices*, 3 (4) (2009) 044501.1-045001.7.
18. Chen J., Randall R. B., Peeters B., Van der Auweraer H. and Desmet W.: Automated misfire diagnosis in engines using torsional vibration and block rotation. *COMADEM2012*, Huddersfield, UK, 18-20th June (2012)
19. Ghojel J. I.: Review of the development and applications of the Wiebe function: a tribute to the contribution of Ivan Wiebe to engine research. *International Journal of Engine Research*, vol.11 no.4, pp.297-312 (2010)
20. Chen J., Randall R. B.: Simulating bearing knock in an IC engine. the 20th International Congress on Sound and Vibration, 7-11th July 2013, Bangkok, Thailand
21. Chen J., Randall R. B., Peeters B., Desmet W. and Van der Auweraer H.: Neural network based diagnosis of mechanical faults in IC engines. *IMEchE Tenth International Conference on Vibrations in Rotating Machinery*, 11-13rd September 2012 London, UK
22. Guyon I. and Elisseeff A.: An introduction to variable and feature selection, *Journal of Machine Learning Research*, vol3, pp.1157-1182 (2003).
23. Ludwig O. and Nunes U.: Novel maximum-margin training algorithms for supervised neural networks, *IEEE Transaction on Neural Networks*, vol.21, no.6, pp.972-983 (2010)
24. Goldberg D. E.: *Genetic algorithms in search, optimization and machine learning*”, Addison Wesley, New York, USA (1989)
25. Ding C., Peng H.: Minimum redundancy feature selection from microarray gene expression data”. *Proceedings of the Computational Systems Bioinformatics*, pp. 523-528, 2003, Stanford, USA.
26. Ludwig O., Gonzalez P. C. and Lima A.: Optimization of ANN applied to non-linear system identification. *Proceedings of the 25th International Conference on Modelling, Identification, and Control*, 2006, Lanzarote, Spain.

