Mechanical Systems and Signal Processing 138 (2020) 106611

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

journal homepage: www.elsevier.com/locate/ymssp

# Planetary gearbox spectral modeling based on the hybrid method of dynamics and LSTM

Ruo-Bin Sun [a,c], Zhi-Bo Yang [a,b,*], Liang-Dong Yang [a,b], Bai-Jie Qiao [a,b], Xue-Feng Chen [a,b], Konstantinos Gryllias [c,d]

[a] School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China
[b] State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China
[c] Department of Mechanical Engineering, KU Leuven, Celestijnenlaan 300, Leuven 3001, Belgium
[d] Dynamics of Mechanical and Mechatronic Systems, Flanders Make, Belgium

ARTICLE INFO

*Article history:*
Received 14 February 2019
Received in revised form 14 November 2019
Accepted 26 December 2019
Available online 7 January 2020

*Keywords:*
Planetary gearbox fault diagnosis
Spectral modeling
Long short-term memory network
Power spectral density estimation

ABSTRACT

It is the prerequisite for mechanical diagnosis to clarify the mapping relations between vibrations and damage conditions based on the dynamic model analyses and the systematic experiments. However, there is a certain deviation from the actual measurements in the existing planetary dynamic modeling methods which utilizes linear models or just considering simple non-linear factors. To solve the problem, the hybrid spectral modeling method based on the fault dynamics analysis and the long short-term memory (LSTM) network is proposed in this paper. Firstly, the response of several parts in planetary sets are simulated by the classical dynamic method. Then, as an exploratory attempt, an LSTM network is used to rectify the predicted spectral deviation caused by the over-simplification and parameter errors of the dynamic models. From the inputs of simulated spectral sequences of the several components, the LSTM successfully predicts the spectral sequence of the observation. Finally, by the overlapping weighted average, the power spectral density of the planetary system is estimated through the Welch method. The effectiveness and superiority of the proposed method are verified via the experimental data of a planetary gearbox test rig.



## 1. Introduction

Planetary gearboxes present strong load-carrying capacity and large transmission ratio within a relatively compact structure. Therefore, they are important components widely used in high-speed and heavy-duty machinery, such as helicopters, wind turbines, aero-engines and marine reducers. Fault diagnosis of planetary gearboxes is a rather complicated task due to the multiple gear pairs which mesh at the same frequency, the rotation of the carrier and the complex vibration transmission paths [1]. The classical mechanical fault diagnosis follows two main directions [2]: The first direction focuses in the exploration of the failure generation and the development mechanism by performing dynamic model simulations and at the extraction of accurate descriptions of the fault features; The second one targets to the analysis of vibration signals and the extraction of fault features/diagnostic indicators based on advanced signal processing methods, which can used for accurate diagnosis. In the frames of planetary gearbox diagnosis usually frequency domain methods are used, aiming at the

* Corresponding author.
*E-mail addresses:* phdapple@mail.xjtu.edu.cn (Z.-B. Yang), konstantinos.gryllias@kuleuven.be (K. Gryllias).

https://doi.org/10.1016/j.ymssp.2019.106611
0888-3270/

identification of fault feature frequencies. For example, McFadden [3] proposed a number of interpolation techniques and utilized synchronous averaging to enhance the rotation-related frequency components. Tong [4] et al. considered that the meshing frequency part in the gearbox may be modulated to the higher frequency band by the impacts of the fault, and proposed a multi-input single-output model to identify the resonance band. Moreover a high-order synchrosqueezing wavelet transform method was proposed in Ref. [5] to improve the frequency resolution of the planetary vibrations under a fast varying rotating speed. Furthermore in Ref. [6], a structured sparse representation method was proposed in order to separate the impact and the harmonic components, present in vibration signals captured on planetary gearboxes, and the diagnosis was carried out at a second step based on a frequency analysis of the impact components.

Till now, most of the planetary gearbox dynamic models are specially proposed for modal analysis and system identification, i.e. purely torsional lumped parameter models [7], translational-torsional lumped parameter models [8,9] and finite element models [10,11]. These models are well-designed for structure design and are over-simplified for characterization fault response. They often neglect some key factors that are distinctive to fixed-axis gearboxes. Firstly, there are multiple vibration transmission paths. The transmission paths of sun-planet and ring-planet meshing are different, thus their measured impact magnitudes in reality are dissimilar due to the different vibration attenuation. Secondly, the planets not only revolve on their axes, but also rotate with the carrier, which adds complicated amplitude modulation on observed vibration signals. Besides, the gear mesh includes many non-linear factors: backlash, tooth profile error, surface friction, etc. [12]. It is too complicated to take account of all these factors in the existing models. Simulation of a planetary dynamic model may give a clear description of the main impact waveform feature in the time domain, whereas there is a certain deviation from the measured vibrations in the frequency domain. Furthermore, due to the limitation of the sensors, the measured acceleration signals have a limited frequency bandwidth, but the main spectrum energy of the simulation response is beyond the effective measurement frequency band of the sensors. All these hinder to provide accurate fault prior knowledge for frequency domain diagnostic methods, and thus the simulated spectrum needs to be rectified to approximate the actual situation.

The popular artificial intelligence techniques, such as deep learning, promote the rapid development of intelligent fault diagnosis methods [13–15]. Various architectures of deep networks are applied to solve a certain diagnostic problem, e.g. sparse autoencoders [16,17], convolution neural networks [18,19], and recurrent neural networks [20]. This kind of methods can effectively handle and analyze massive data and directly extract fault features based on a pure data-driven strategy [21]. However, these methods are mainly black-box models and therefore the fault characteristics are not revealed from the mechanism perspective. Similarly, deep learning is used in identification modeling for control systems [22]. The parameters of the governing equations are determined through the training of the network.

The above-mentioned methods inspired us to use deep learning in order to rectify the deviation of planetary gearbox vibration simulations from reality, and give a more precise mapping relationship between the fault and the dynamic response in the frequency domain. Therefore a hybrid spectral modeling method based on the planetary gearbox fault dynamics and deep learning is proposed, focusing towards the correction/improvement of the dynamic model of planetary gearboxes and it use further for diagnostic purposes. By exploiting the strong nonlinear representation ability of the neural network, the deviation of the planetary gearbox dynamic model is rectified in the frequency domain. As an exploratory attempt, the problem is summarized as a sequence-to-sequence regression problem, where the inputs are simulated vibration signals of several planetary parts and the output is the real vibration sampled by a sensor. In this way, some dynamic phenomena and factors that are difficult to be modeled are taken into account. In consideration of a strong correlation in the Discrete Fourier Transform (DFT) sequences of the planetary vibration signals, the DFT sequences are rectified by a Long Short Term-Memory (LSTM) network structure [23], which is suitable for dealing with related events with relatively long intervals and delays in time series. Finally, through weighted averaging of the rectified DFT sequences, the Power Spectral Density (PSD) of the planetary gearbox is estimated by means of the Welch method [24]. The effectiveness of the proposed method is tested, evaluated and verified on a planetary gearbox test rig, and the comparison between simulated and experimental data reveals that the LSTM network is qualified to reduce the error of the dynamic model.

The rest of the paper is organised as following. The basic theory of planetary gearbox fault dynamic modeling and the Long Short Term-Memory (LSTM) networks are presented in Section 2. Moreover the hybrid spectral modeling method is presented in Section 3. The experimental set up, the application of the methodology and the results are analytically presented and analysed in Section 4. Finally, the paper closes with some conclusions in Section 5.

## 2. Basic theory

In this section, a pure torsional lumped parametric model of planetary gears is introduced. Several local failure modes of gears are simulated based on this model. Next, the architecture and the property of the LSTM networks are presented. A relatively simple dynamic model is chosen here specifically to fully demonstrate the ability of deep learning on simulated response modification.

### *2.1. Planetary gear fault dynamic model*

The model used in this study is a pure torsional lumped parametric model of one-stage planetary gear mechanism [7], in which only the torsional motion of each component is considered into modeling. The system consists of one sun ($s$), one ring

$(r)$, one carrier $(c)$, and $N$ planets $(p)$ and in total has $(3+N)$degrees of freedom, as shown in Fig. 1. The meshes of sun-planet and ring-planet pairs are modeled as linear springs and dampers, denoted by $k_{hn}$ and $c_{hn}$, where $h = s, r$ and $n = 1, 2 \ldots, N$. In this study, nonlinear factors of the gear mesh such as backlash and transfer errors are neglected in the dynamic modeling procedure. The torsional stiffness and the damping of parts are given by $k_{hu}$ and $c_{hu}$respectively, where $u = s, r, c$. Writing down the differential equations of the rigid body motion for each component, the following governing equation of the system can be obtained.

$$\text{Carrier}: \quad (I_c + \sum_{n=1}^{N} m_n r_c^2)/r_c^2 \ddot{u}_c + c_{cu}\dot{u}_c + k_{cu}u_c - \sum_{n=1}^{N}(c_{sn}\dot{\delta}_{sn} + c_{rn}\dot{\delta}_{rn})\cos\alpha - \sum_{n=1}^{N}(k_{sn}\delta_{sn} + k_{rn}\delta_{rn})\cos\alpha = T_c/r_c \tag{1}$$

$$\text{Ring}: \quad (I_r/r_r^2)\ddot{u}_r + c_{ru}\dot{u}_r + k_{ru}u_r + \sum_{n=1}^{N} c_{rn}\dot{\delta}_{rn} + \sum_{n=1}^{N} k_{rn}\delta_{rn} = T_r/r_r \tag{2}$$

$$\text{Sun}: \quad (I_s/r_s^2)\ddot{u}_s + c_{su}\dot{u}_s + k_{su}u_s + \sum_{n=1}^{N} c_{sn}\dot{\delta}_{sn} + \sum_{n=1}^{N} k_{sn}\delta_{sn} = T_s/r_s \tag{3}$$

$$n\text{th Planet}: \quad (I_n/r_n^2)\ddot{u}_n - c_{rn}\dot{\delta}_{rn} - k_{rn}\delta_{rn} + c_{sn}\dot{\delta}_{sn} + k_{sn}\delta_{sn} = 0 \tag{4}$$

where $u_h$ represents the torsional displacement of the corresponding components; $m_h$ and $I_h$ denote respectively the mass and the moment of inertia; $r_h$ is the radius of the base circle; $\alpha$ is the gear pressure angle; $T_h$ denotes the applied torque; $\delta_{sn} = u_s + u_n - u_c\cos\alpha$ and $\delta_{rn} = u_r - u_n - u_c\cos\alpha$ are the relative displacements, which project on the action line of the sun-planet and the ring-planet pairs.

The mesh stiffness is time-varying because of the alternating meshing of single-double teeth and the changing of the contact position. The deflection of a spur gear tooth could be regarded as a non-uniform cantilever beam and the stiffness could be estimated by the potential energy method [25]. The deformation caused by bending, shearing, compression and contact are taken into consideration, and the total effective mesh stiffness can be calculated by the following formula.

$$k_t = \sum_{i=1}^{2} \frac{1}{1/k_{h,i} + 1/k_{b1,i} + 1/k_{s1,i} + 1/k_{a1,i} + 1/k_{b2,i} + 1/k_{s2,i} + 1/k_{a2,i}} \tag{5}$$

where, $i = 1, 2$ represents the number of the gear pair that participates in meshing; and $k_h$, $k_b$, $k_s$, and $k_a$ denote respectively the Hertzian contact stiffness, the bending stiffness, the shearing stiffness, and the axial compressive stiffness of a pair of mating teeth, derived as in [26]:

$$\frac{1}{k_h} = \frac{4(1-v^2)}{\pi EL} \tag{6}$$

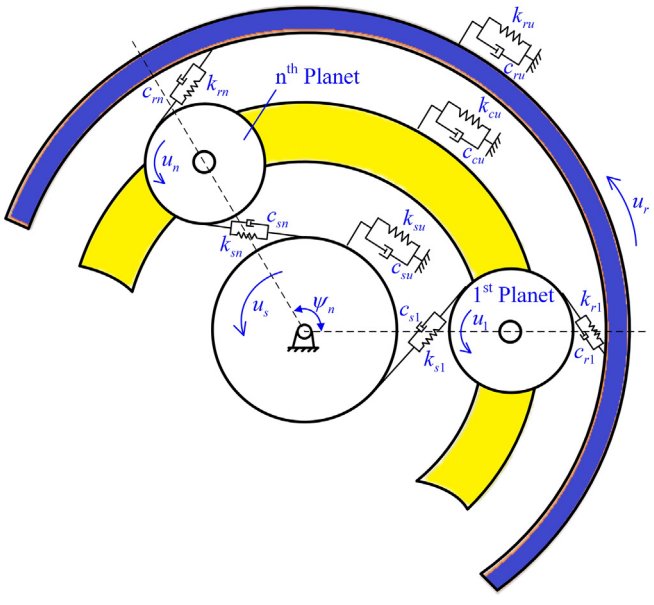


**Fig. 1.** The sketch of the pure torsional planetary sets model. The meshes are modeled as linear springs and dampers.

$$\frac{1}{k_b} = \int_0^d \frac{(x\cos\alpha_1 - h\sin\alpha_1)^2}{EI_x} dx \tag{7}$$

$$\frac{1}{k_s} = \int_0^d \frac{1.2\cos^2\alpha_1}{GA_x} dx \tag{8}$$

$$\frac{1}{k_a} = \int_0^d \frac{\sin^2\alpha_1}{EA_x} dx \tag{9}$$

where $E$, $G$, $v$ are respectively the elastic modulus, the shear modulus, and the poison ratio. $L$ denotes the width of the tooth. $I_x$ and $A_x$ are the area moment of inertia and the area of the section. $\alpha_1$ is the angle between the action line and the line perpendicular to the central line of the tooth, and the distance from the point of the applied force to the tooth root is denoted by $x$.

When a local fault occurs, the stiffness will be reduced on the corresponding defect tooth. Different failure modes affect the stiffness on different components. For example, a root crack changes the cross-section area of the tooth, and therefore the bending and the shearing stiffness will be reduced accordingly [27]. If a tooth is chipped, it can be modeled by the change of the tooth surface area, and thus the Hertzian contact stiffness decreases [28]. The pitting can be simulated by a shortening of the contact line of the mating teeth [29], and the most serious failure, a broken tooth, leads to a stiffness, total missing of the mating teeth. The mesh stiffness of a failure gear is first calculated and then is substituted into the motion equation. Thus, the vibration response of fault planetary sets can be simulated by solving the motion equation with numerical methods.

### 2.2. Concept of LSTM

LSTM network is a kind of advanced architecture of the Recurrent Neural Network (RNN), which is composed of LSTM units. The basic RNNs are networks with time recursive hidden layers. Each layer not only receives the input $x^{<t>}$ at the current time, but also takes information from the activation value $a^{<t-1>}$ of the last successive hidden layer [30]. Being different from the classical feed-forward neural network, an RNN utilizes its internal state as the memory to predict the output, which shares the features of inputs at different times. It is particularly suitable for time series modeling and could be used to solve sequence-to-sequence regression or classification problems.

However, the recursive structure may cause the gradient vanishing or the exploding problem during the training process by back-propagation. To address this issue, a more elaborate memory layer architecture, the LSTM, was proposed in [23]. The structure of the LSTM unit is described in Fig. 2. It is composed of a memory cell $c^{<t>}$, an input gate, an output gate, and a forget gate. The cell captures long-range connections, i.e., it remembers the activation value over arbitrary time intervals. Three gates adjust information flows into and out of the cell. More accurate mathematical expressions can be formulated as:

$$\tilde{c}^{<t>} = \tanh(W_c[a^{<t-1>}, x^{<t>}] + b_c) \tag{10}$$

$$\Gamma_u^{<t>} = \sigma(W_u[a^{<t-1>}, x^{<t>}] + b_u) \tag{11}$$

$$\Gamma_f^{<t>} = \sigma(W_f[a^{<t-1>}, x^{<t>}] + b_f) \tag{12}$$

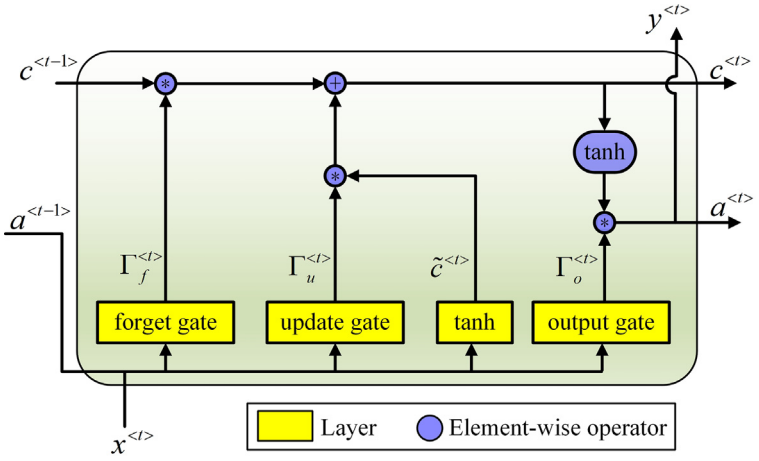


**Fig. 2.** The architecture of the LSTM unit, which consists of an input gate, an output gate, and a forget gate. The boxes indicate the layers of the unit and the circles indicate the element-wise operator.

$$\Gamma_o^{<t>} = \sigma(W_o[a^{<t-1>}, x^{<t>}] + b_o) \tag{13}$$

$$c^{<t>} = \Gamma_u \circledast \tilde{c}^{<t>} + \Gamma_f \circledast c^{<t-1>} \tag{14}$$

$$a^{<t>} = \Gamma_o \circledast \tanh(c^{<t>}) \tag{15}$$

where $\Gamma_u$, $\Gamma_f$, $\Gamma_o$ are the output values of the input, the forget, and the output gate. They determine patterns of the memory states, i.e., which part of the memory to be updated, discarded, and output. $a$ and $x$ are the input and the output value of the unit; the superscript $t$ denotes the time of the state. $\sigma$ and tanh represent the nonlinear activation functions: the sigmoid and the hyperbolic tangent. In addition, the mark $\circledast$ denotes the element-wise multiplication.

## 3. The hybrid spectral modeling method

Often mechanical fault diagnosis is carried out in the frequency domain. As mentioned before, the frequency spectra of the simulated response present a certain and often unacceptable deviation from the corresponding spectra of the observed vibration. As shown in Fig. 3 (a), the energy distribution of the amplitude spectrum of a simulated signal is mainly concentrated at the high-frequency band, that is 8–11 kHz, and the amplitudes of the first few orders of the mesh frequency are quite low. However, in the amplitude spectrum of a segment of an acceleration signal (Fig. 3 (b)), the energy is concentrated on the first few orders of the gear mesh frequencies. There are two reasons for this phenomenon. On the one hand, the dynamic model ignores too many factors and has a certain parameter error. On the other hand, the relatively narrow frequency response range of the accelerometer limits the sampling. For example, the response range of PCB 333B32 piezoelectric accelerometer used in the experiment is 0.5–3000 Hz (±5%). It can be found that only the information in the low-frequency band (< 3000 Hz) has engineering reference significance, the high-frequency band is in the range of the nonlinear frequency response of the specific sensor.

The deviation interferes with revealing the characteristic in the frequency domain of failure, thus can not guide diagnosis effectively. To address this problem, the dynamic model is rectified using a deep learning method, and therefore the complicated factors that have been neglected in the model could also be taken into consideration up to some extent, such as, the multiple different vibration transfer paths, the amplitude modulation caused by the rotational carrier, and the nonlinear gear mesh. The flowchart of the proposed hybrid spectral modeling method is plotted in Fig. 4.

The method consists of two major parts: the LSTM network training and the Power Spectral Density (PSD) estimation through the generalized Welch method [24]. Firstly, the network used to correct the dynamic model errors is trained. In essential, the spectrum correction of the simulation signal by the LSTM is a sequence-to-sequence regression problem. The training and testing data sets are acquired, respectively, from the simulation response of the planetary set fault dynamic model and the measured acceleration signals. Both the simulation and the measured signals are acquired under different operating conditions and various failure modes of the gear to ensure that the trained network has a certain generalization ability. The simulation vibrations of the different parts are regarded as the input data, and the measured signals of the test rig are used as the output sample labels. As explained before, only the low-frequency information of the measured data could be used for gearbox diagnosis, so the sample labels are firstly processed by a low-pass filter. Because the fault features are more

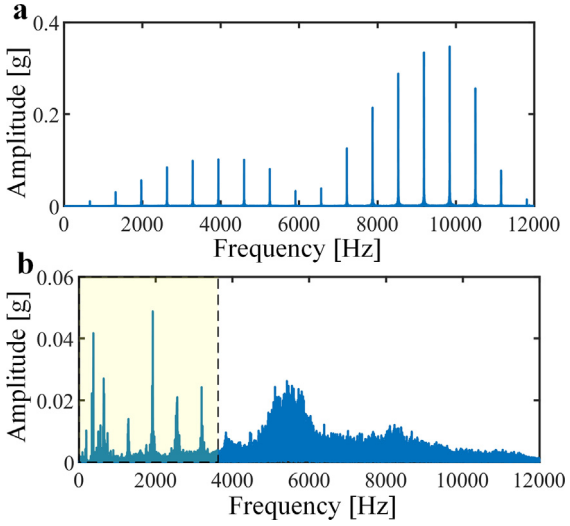


**Fig. 3.** (a) The amplitude spectrum of the simulated response of a healthy planetary gearbox model. (b) The amplitude spectrum of an observed acceleration signal on a test rig. The low frequency band is the effective measurement bandwidth.

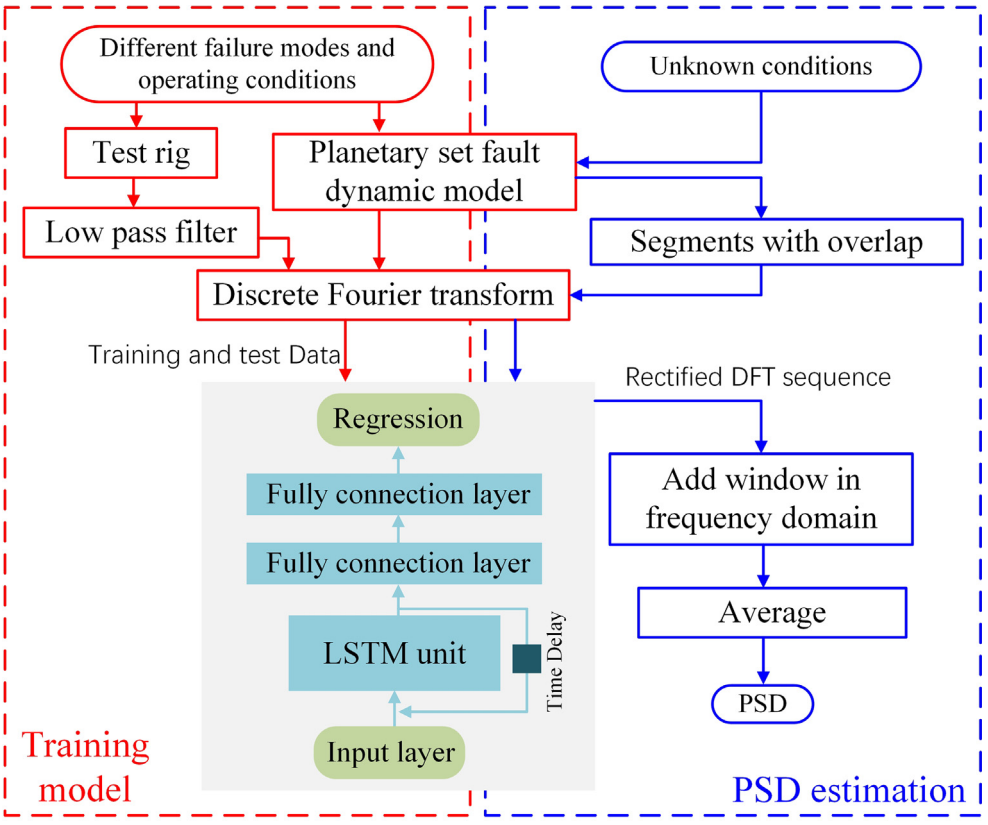


**Fig. 4.** The flowchart of the proposed method. The left part indicates the training of the LSTM network procedure; and the right part shows the PSD estimation procedure.

prominent in the frequency domain, and the labels still contain some noise, the training is carried out using the DFT sequences of the samples instead of the original waveform in the time domain. The LSTM has excellent properties that remember and share features between any interval range and, therefore, that could help to take advantage of the periodic structure of the DFT sequences; for instance, the peak of the mesh frequency and its higher orders, and the equal width side-bands. After training the network, one could estimate the spectral density under Welch's strategy of an unknown condition. That is to say, to average the modified periodograms of windowed overlapping segments, which is expressed as:

$$\hat{P}^i_{\mathrm{PER}}(\mathrm{e}^{\mathrm{j}\omega}) = \frac{1}{MU}\left|\sum_{n=0}^{M-1} x^i_N(n)w(n)\mathrm{e}^{-\mathrm{j}\omega n}\right|^2 = \frac{1}{U}\left|X^i_N(\mathrm{e}^{\mathrm{j}\omega}) * W(\mathrm{e}^{\mathrm{j}\omega})\right|^2 \tag{16}$$

$$\tilde{P}_{\mathrm{PER}}(\mathrm{e}^{\mathrm{j}\omega}) = \frac{1}{L}\sum_{i=1}^{L}\hat{P}^i_{\mathrm{PER}}(\mathrm{e}^{\mathrm{j}\omega}) \tag{17}$$

where $\hat{P}^i_{\mathrm{PER}}(\mathrm{e}^{\mathrm{j}\omega})$ is the Discrete Time Fourier Transform (DTFT) of the windowed $i$th segment of length $M$, which could also be calculated by the convolution of the DTFT sequences of the signal with the window function. The windows can improve the spectral distortion caused by the large side-lobes of truncated rectangular windows. $U = (1/M)\cdot\sum_{n=1}^{M-1} w^2(n)$ is the energy of the window function, which should be used in order to normalize the spectrum to get an unbiased estimation. The average $\tilde{P}_{\mathrm{PER}}$ of $L$ segment estimations reduces the variance of the estimation, thus the estimation performance can be improved. The procedures are described in details as follows.

*Step 1: Data preparation and pre-processing.* A fault dynamic model is prepared based on the geometry and the kinematics of the corresponding machine. The parameters of the model are determined based on the specific machine. Using numerical methods, the fault dynamic model is solved under different operating conditions. The calculated vibration responses of the simulations compose the input part of the original data set. Then vibration signals are captured over the corresponding machine under the same conditions with the simulations. The measured signals are low pass filtered and are termed as labels. Then the inputs and the labels are aligned through shifting by a certain time-delay which is obtained by a cross-correlation function. The data sets, simulated and captured under the same operating conditions, are truncated separately

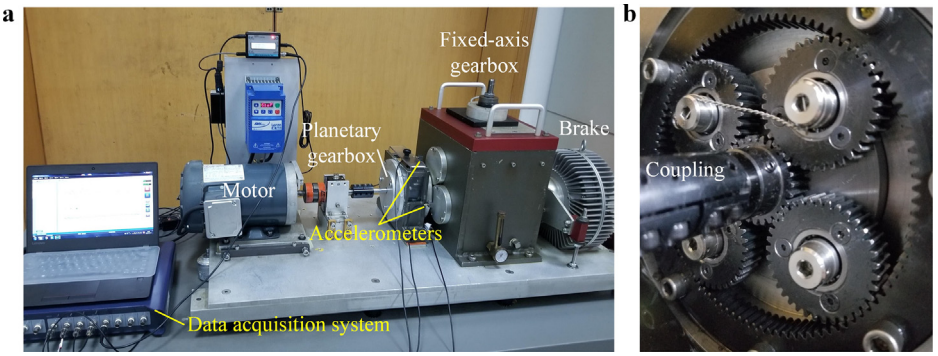


**Fig. 5.** (a) The SQI gearbox drive train diagnostic simulator, which consists of a motor, a planetary gearbox, a fixed-axis gearbox and a brake. (b) The single stage planetary set.

with the same length, and then the DFT is calculated. The sequences are centred and normalised and therefore have a mean value equal to 0 and a variance equal to 1. Due to the property of Fourier Transform only half of the DFT sequences are kept without loss of information. In this study the planetary gearbox of the test rig, presented in Fig. 5, is modeled. The simulated and the measured data sets are first preprocessed and then are divided randomly into two parts, the training set and the test set.

*Step 2: Preparation and training of the network.* Firstly the appropriate architecture of the deep network should be chosen. In this paper, the network consists of an input layer, an LSTM layer, two fully connected layers and a regression layer. At this step, the network parameters and the various settings, such as the number of nodes per layer, the cost function, the regularisation method and the optimization method, should be selected. Then the network is trained using a certain stopping rule.

*Step 3: Estimation of the spectral density for an unknown operating condition.* The final aim of the proposed method is the estimation of a rectified spectral density. For an unknown operating condition, the rectified spectral density could be estimated using the dynamic model and the trained network. The simulation response is firstly calculated by the dynamic model. Secondly, the response is divided into segments with length $M$ and a certain overlap, the modified DFT sequences are gained by predicting the output of simulated DFT sequences in the network. Besides, the modified DFT sequences are windowed in the frequency domain through convolution operation. Finally, the PSD function is estimated via averaging of several windowed sequences.

## 4. Experimental analysis and discussion

In order to test, evaluate and validate its effectiveness, the proposed methodology is applied on datasets simulated and captured on a experimental planetary gearbox test rig.

### *4.1. Description of the planetary gearbox test rig and the measurements*

The test rig is a planetary gearbox drive train diagnostic simulator manufactured by the Spectra Quest Inc. and is presented in Fig. 5 (a). It consists of a variable frequency controlled AC motor, a planetary gearbox, a fixed-axis gearbox and a magnetic brake. The original planetary gearbox has two stages. For the needs of this study, a coupling was installed to replace the first stage, thus the planetary gearbox was turned into a single-stage transmission, as shown in Fig. 5 (b). The stage used at the new configuration consists of four planets, a sun, a planet carrier and a fixed rig. The sun is the high-speed input and the planet carrier is the low-speed output. The accelerometers were placed on the bearing cover in both horizontal and vertical directions. The signals measured by the two different position sensors, have been analysed and were found to be very similar. Therefore, only the signals from the horizontal position were final used. The vibration signals were collected by a data acquisition system with a sampling frequency equal to 12.8 kHz. Therefore the Nyquist frequency was equal to 6.4 kHz, ensuring the coverage of the effective frequency band of the accelerometers used.

Four health cases of the gearbox have been considered in this study, including a healthy mode and three failure modes of the planet gears. The faulty cases include a root cracking, a chipped tooth and a broken tooth. The crack depth at the root was 3 mm while the chipped area was around 15% of the tooth surface. At the fourth case an entire tooth was missing. The three different failure modes of the planet as shown in Fig. 6. The physical parameters of the planetary set are listed in Table 1. A model of the gearbox has been prepared based on the physical parameters of the test rig, creating a kind of digital twin. The model was further used to generate the simulated vibration responses of the various components under the different input speeds and external loads. Three input speeds of the sun gear: 1200, 1500 and 1800 RPM and four equivalent external torques, applied on the carrier: 2.9, 7.3, and 11.7 N · m have been considered for this study. Totally the simulations were carried

out under 36 different conditions. Consistent with the simulations, the vibration signals were collected under the same 36 conditions. The data were then low pass filtered with a cutting frequency equal to 3000 Hz.

Five signal segments with a length equal to 3000 samples were randomly cut out from each condition to form a data set. The dataset was then split in three sets: the training set (70% of the samples), the validation set (15% of the samples) and the testing set (15% of the samples). As an example, the waveform and the amplitude spectra of a sample under the condition of a cracked planet, 1800 RPM sun rotational speed and 7.3 N · m external torque, is demonstrated in Fig. 7. It can be seen that the signals present periodic impulses in the time domain. Additionally the meshing frequency and its harmonics and some sideband components can be found in the frequency domain. The label signals were aligned utilizing a cross-correlation method and only the low-frequency band was retained.

Then the LSTM network was trained using the training data set. The numbers of nodes of were selected respectively equal to: 200 for the LSTM unit, 50 for the first fully connected layer and 30 for the second fully connected layer. The batch size of training was set equal to 20 and both the LSTM layer and the fully connected layers were initialized using the uniform distribution. The mean square error was used as the cost function in the training processing and the parameters were optimized by the adaptive moment estimation (ADAM) method [31]. The decay rates of the gradient and the squared gradient moving averages of the optimizer were respectively set equal to 0.9 and 0.999. Additionally the initial learning rate of the optimizer was set at 0.01. In order to prevent model from overfitting, a dropout regularization is considered at the second fully connected layer and the probability of the dropping out was set to 0.5. The distribution of the Root-Mean-Square Error (RMSE) with different frequency band in the test set is presented in Fig. 8. It can be found that the network has better rectification ability in the high-frequency part, as the label signals have been low-pass filtered, but on the other hand there is still a certain noise in the low-frequency part.

After training the network, the PSD could be estimated. Firstly, some estimation results under the training conditions are displayed in Fig. 9. It is clear that almost all the high-frequency components are inhibited, which is close to the actual measured vibrations. Moreover the estimated PSD retains some already known spectrum characteristics of planetary gearboxes. As a result, the first few harmonics of the meshing frequency can be clearly identified, and the fault characteristic frequency of a local fault on a planet can also be found in some sideband intervals, which are consistent with the feature frequency formula [32].

$$f_m = f_i \frac{Z_s Z_r}{(Z_s + Z_r)} \tag{18}$$

$$f_p = k\frac{f_m}{Z_p}, \quad (k = 1, 2) \tag{19}$$

where $f_m$ denotes the meshing frequency of gears; $Z_s$, $Z_r$, $Z_p$ represent respectively the teeth number of the sun, the ring and the planets; $f_p$ is the fault feature frequency of the planet. If the impacts of the sun-planet and the ring-planet pairs are similar, then $k = 2$, otherwise $k = 1$. Furthermore, an asymmetric sideband characteristic can be found, which is consistent with the conclusion of [33]. In order to give an intuitive comparison, Fig. 10 shows the PSD of the measured vibrations (labels) associated with Fig. 9. It can be found that measured spectra have more details near the meshing frequency of each order, and not surprisingly, actual vibrations have more noise components in the spectrum. The estimated spectra have a certain deviation from the actual vibration. On the other hand considering the highly random nature of mechanical vibration signals, the estimation provides the main features of the vibration in the frequency domain. Furthermore, it is clear that the rectified simulations have made great progress compared with the original simulations of the dynamic model.

To further verify the performance of the proposed method, the PSD of another special condition was estimated, as shown in Fig. 11. The condition include a chipped planet tooth, an input speed equal to 3000 RPM, and an external torque equal to 10 N · m, which had not been in the training set. In general, the RMSE value between the regression DFT sequences is 4.2, which is a little higher than the validation RMSE at the last iteration of the training. However, the first three harmonics of the meshing frequency and some sideband characteristics of the planet local fault can be clearly seen. It is therefore proven that the modeling method has a certain generalization ability. Similarly, the PSD function of the measured vibrations

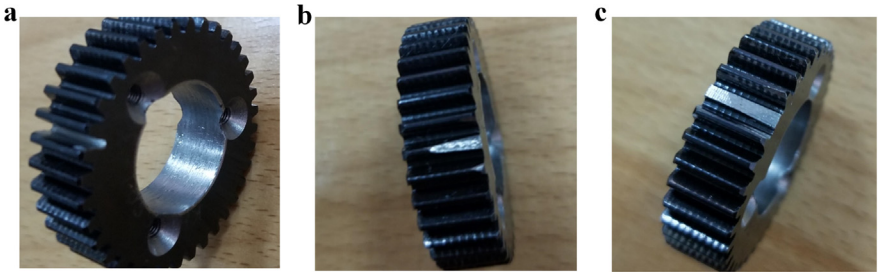


**Fig. 6.** Three different failure modes of the planet. (a) Root crack. (b) Chipped tooth. (c) Broken tooth.

**Table 1**
Physical parameters of the planetary gear set of the test rig.

| Parameters | Sun | Planet (4) | Ring | Carrier |
|---|---|---|---|---|
| Number of teeth | 28 | 36 | 100 | – |
| Module [mm] | 1 | 1 | 1 | – |
| Pressure angle [°] | 20 | 20 | 20 | – |
| Face width [mm] | 10 | 10 | 10 | – |
| Young's modulus [Pa] | $2.1 \times 10^{11}$ | $2.1 \times 10^{11}$ | $2.1 \times 10^{11}$ | – |
| Poisson's ratio | 0.3 | 0.3 | 0.3 | – |
| Mass [kg] | – | – | $9.86 \times 10^{-2}$ | – |
| Moment of inertia [kg · m$^2$] | $2.41 \times 10^{-6}$ | $1.60 \times 10^{-5}$ | $9.20 \times 10^{-3}$ | $4.99 \times 10^{-4}$ |
| Base circle [mm] | 13.2 | 16.9 | 47.0 | – |
| Torsional stiffness [N/m] | 0 | – | $1 \times 10^{9}$ | 0 |
| Torsional damping [N · s/m] | 0 | – | $1 \times 10^{3}$ | 0 |

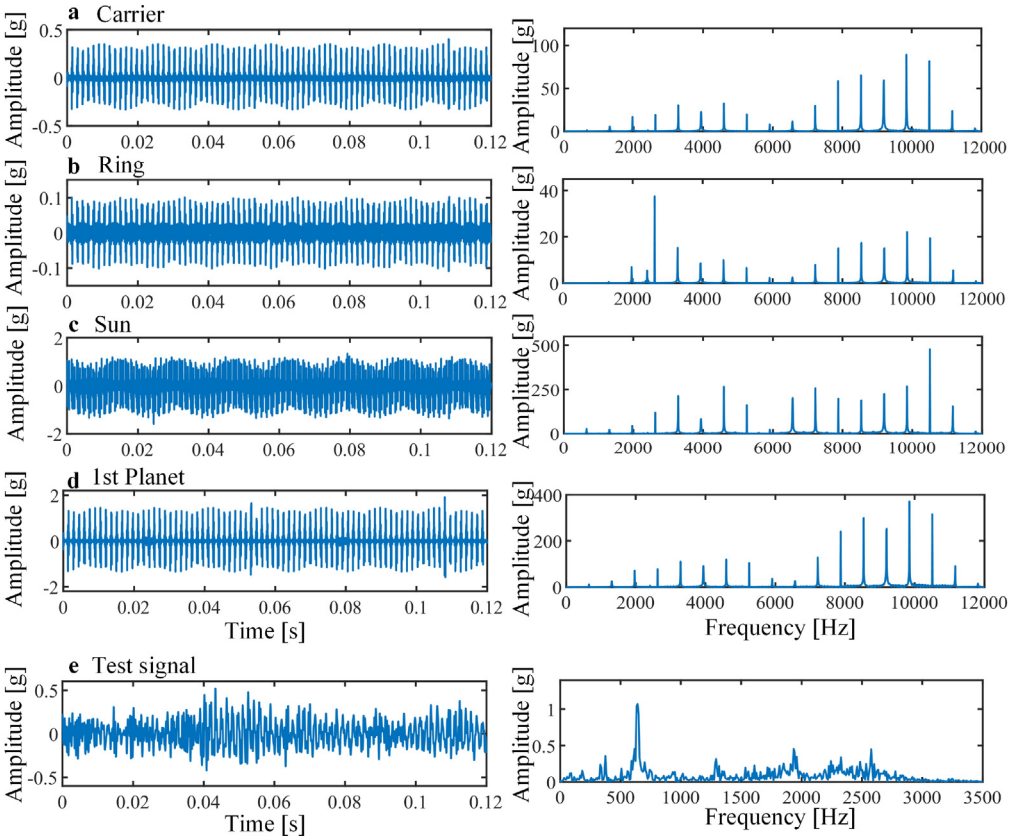


**Fig. 7.** The sample under the condition of a cracked planet, 1800 RPM input speed, and 7.3 N · m external torque. (a-d) Part of the input simulation signals and their amplitude spectrum. (Another three planet signals are omitted.) (e) The acquired vibration signal and its amplitude spectrum.

under the same condition is shown in Fig. 12. Generally speaking, the measured and the simulated vibration spectra are similar in appearance. On the other hand the measured spectrum present more details around the first harmonic of the meshing frequency. At the partially zoomed plot provided in Fig. 13, a number of sideband fault characteristics both around the first and the second harmonic of the meshing frequency can be identified. The simulated spectrum deviates from the actual vibration, but it still rectifies the error of the dynamic model to some extent.

### *4.2. Discussion*

Firstly, the effect of different kinds of networks on the vibration signal regression is analysed. For comparison, two other kinds of network architecture, a fully connected (FC) network and a convolution neural network (CNN), were also considered and trained. The FC network consists of three FC layers which respectively have 4000, 4000 and 1500 nodes. Moreover the

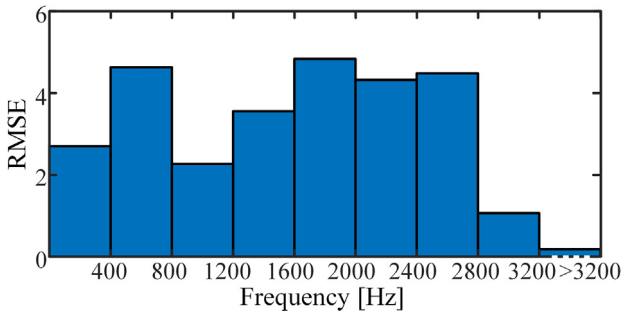


**Fig. 8.** The distribution of RMSE with different frequency band in the test set.

CNN network has four 2D convolution layers and an FC layer. The filter sizes of the convolution were set equal to $9 \times 7$, $5 \times 1$, $5 \times 1$, $5 \times 1$ and the FC layer has 1500 nodes. In both networks, the activation functions of the input and the hidden layers were set as the Rectified Linear Unit (ReLU) while for the output layer was the sigmoid function. To alleviate the effect of local optimal solutions and gain relative satisfying results, the hyperparameters were adjusted by trial and error, and all the models are repeatedly trained. The curves of the RMSE on the validation set during every training epoch are plotted in Fig. 14. It is illustrated that the predicted result of both the FC and the CNN have a larger bias than the LSTM, which further indicates that the memory mechanism of the LSTM unit is well adapted for regression DFT sequences of vibration signals.

Furthermore, a comparison was made to explain the prediction performance of different truncated sequence lengths. The LSTM network architecture and the training parameters were the same as the above experiment except for the sample length. The samples were divided into lengths of 2000, 3000, 4000 and 5000 samples respectively. Next, the regression networks were trained under each condition. The RMSE of different sample lengths on the validation set at the final iteration are illustrated in Fig. 15. It can be seen that the regression error increases, as the sample becomes longer. Obviously, it also takes more computing time as the sample length grows. On the other hand this does not mean that shorter samples should be applied. The regression is carried out in the frequency domain and therefore too short samples would make the spectral

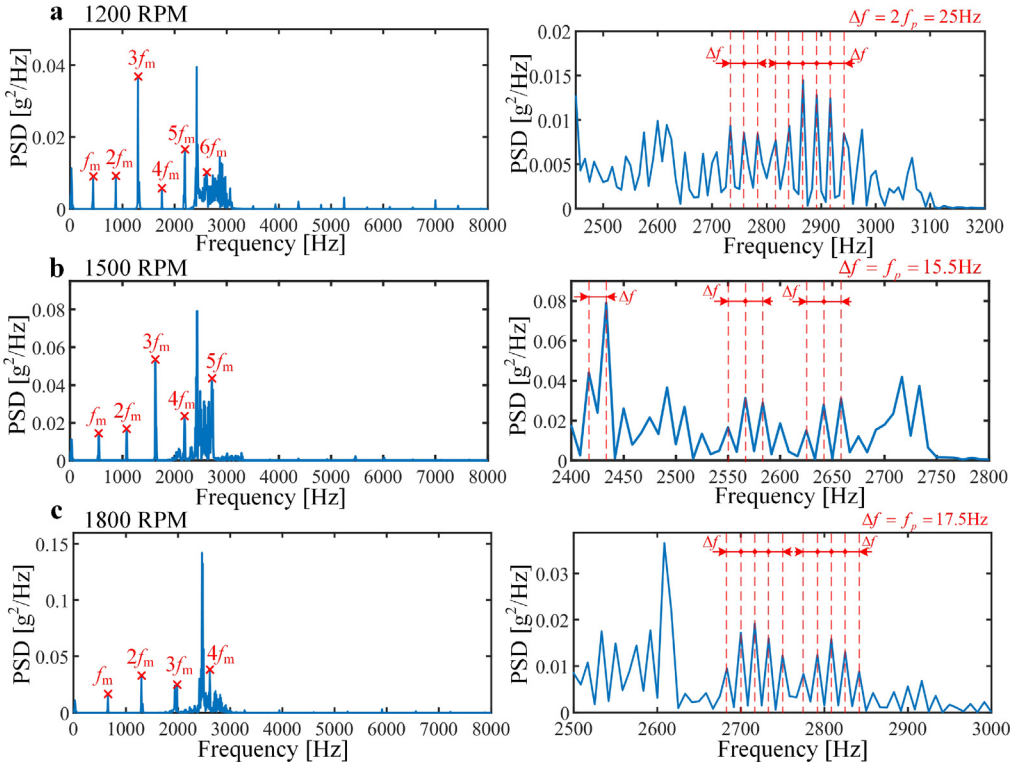


**Fig. 9.** The estimated PSD of some training conditons and their partially enlarged detail. The estimation conditions are: (a) chipped tooth, input speed = 1200 RPM, and torque = 7.3 N · m; (b) broken tooth, input speed = 1500 RPM, and torque = 11.7 N · m; (c) broken tooth, input speed = 1800 RPM, and torque = 11.7 N · m.

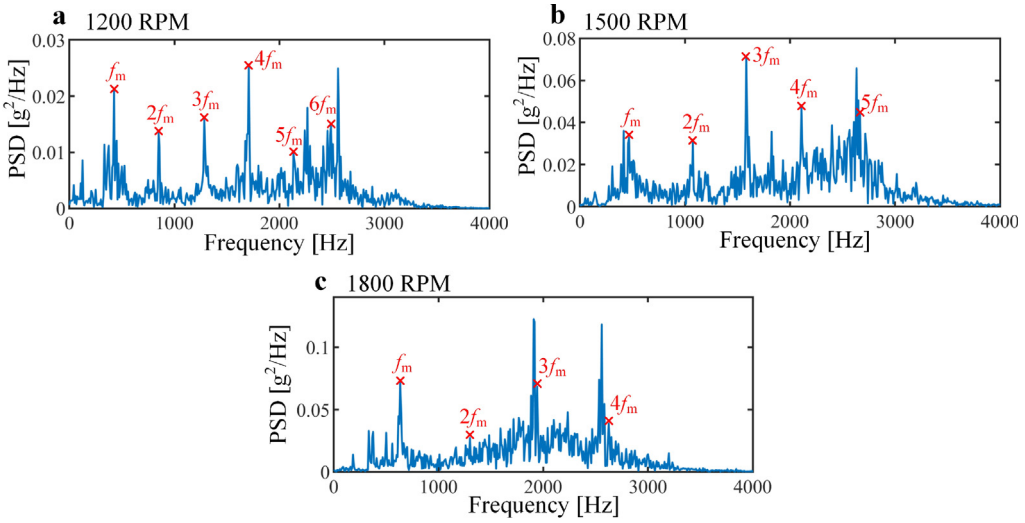


**Fig. 10.** The PSD of the measured vibration signals associated with the estimation in Fig. 9. (a) Chipped tooth, input speed = 1200 RPM, and torque = 7.3 N · m; (b) broken tooth, input speed = 1500 RPM, and torque = 11.7 N · m; (c) broken tooth, input speed = 1800 RPM, and torque = 11.7 N · m.

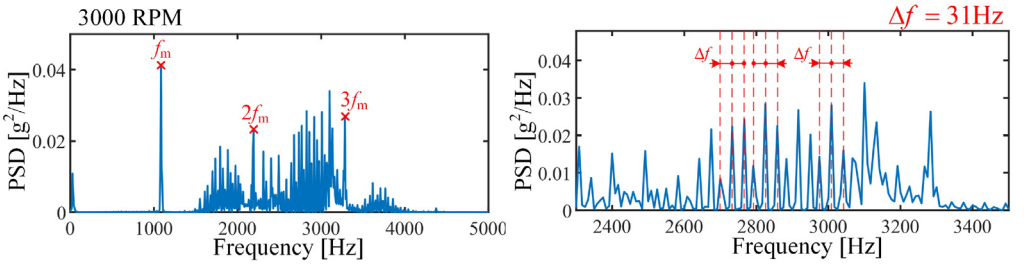


**Fig. 11.** The estimated PSD under condition of chipped tooth, input speed = 3000 RPM, and torque = 10 N · m.

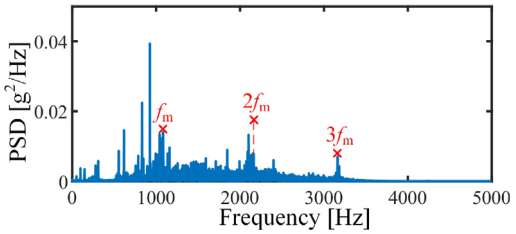


**Fig. 12.** The measured PSD under the same condition of chipped tooth, input speed = 3000 RPM, and torque = 10 N · m.

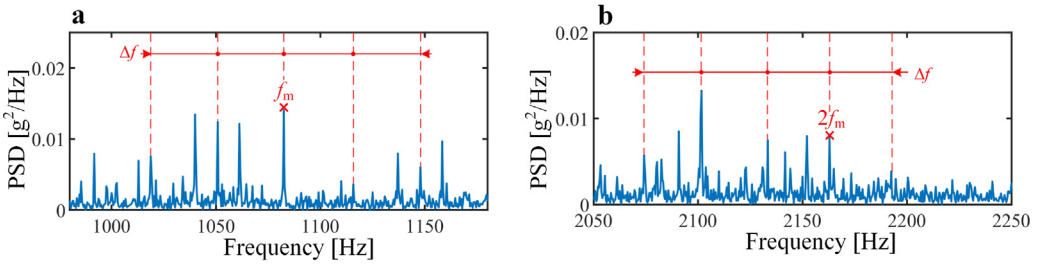


**Fig. 13.** The partial enlarged detail of the measured PSD under condition of chipped tooth, input speed = 3000 RPM, and torque = 10 N · m.

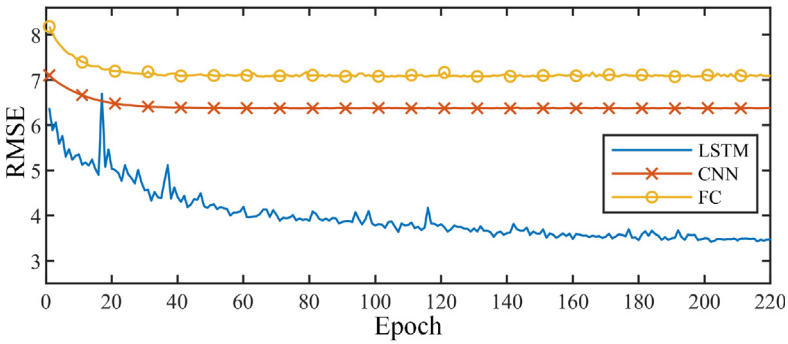


**Fig. 14.** The validation RMSE curve of every training epoch by the LSTM, CNN, and FC network structures.

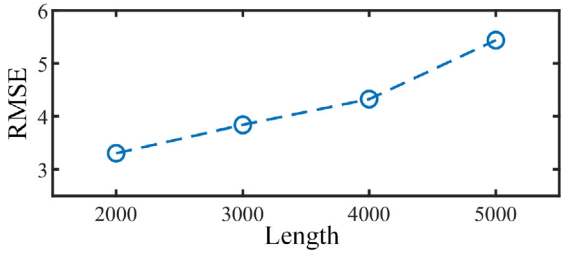


**Fig. 15.** The RMSE of different sample length on the validation set at the final iteration.

resolution too low to be used. As a result the length of the sample is a tradeoff between the regression accuracy and resolution. In this experiment, a moderate sample length was chosen.

## 5. Conclusion

In this paper a new method is proposed in order to improve the performance and the accuracy of dynamic simulation signals. Under this framework a dynamic model of the considered machinery is first prepared and a sequence of vibration measurements are realised under various speeds and loads. The measured data and the simulated data are fed as outputs and inputs at a machine learning technique. The machine learning model is then combined with the dynamic model, modeling in practice the unmodeled parts of the dynamic model, creating a new block which is the digital twin of the machine. The new block can be further used in order to simulate unmeasured data with a high precision. The general methodologies has been applied in the frames of this work on a one stage planetary gearbox. A planetary gearbox dynamic model was first prepared and data under various speeds and load were captured. An LSTM network was selected and trained. Three different types of networks have been tested and the LSTM unit seems to be, among the others, the more suitable for the regression of the signals' spectra. As a first attempt, the method fully demonstrates the great potential of deep learning in improving the performance of dynamic models. The validated digital twin can be then used for the analysis of the gear mechanisms, for the accurate description of fault characteristics and finally for model based condition monitoring. The authors are currently working on the extension and the testing of the methodology towards more complicated structures.

## Acknowledgements

We thank the supports given by the National Natural Science Foundation of China (No. 51875433), the Young Talent fund of University Association for Science and Technology in Shaanxi of China (No. 20170502), the Key Research and Development Program of Shanxi Province (Nos. 2017ZDCXL-GY-02-01 & 2017ZDCXL-GY-02-02), the Fundamental Research Funds for the Central Universities, and the China Scholarship Council.

## References

[1] Y. Lei, J. Lin, M.J. Zuo, Z. He, Condition monitoring and fault diagnosis of planetary gearboxes: a review, Measurement 48 (1) (2014) 292–305.
[2] M. Asefi, J. Nazarzadeh, A fast transient model for bearing fault analysis in induction machine drives, IEEE Sens. J. 19 (5) (2018) 1897–1904.
[3] P. McFadden, Window functions for the calculation of the time domain averages of the vibration of the individual planet gears and sun gear in an epicyclic gearbox, J. Vib. Acoust. 116 (2) (1994) 179–187.
[4] S. Tong, Y. Huang, Y. Jiang, Y. Weng, Z. Tong, N. Tang, F. Cong, The identification of gearbox vibration using the meshing impacts based demodulation technique, J. Sound Vib. 461 (2019) 114879.

[5] Y. Hu, X. Tu, F. Li, High-order synchrosqueezing wavelet transform and application to planetary gearbox fault diagnosis, Mech. Syst. Signal Process. 131 (2019) 126–151.
[6] R. Sun, Z. Yang, X. Chen, S. Tian, Y. Xie, Gear fault diagnosis based on the structured sparsity time-frequency analysis, Mech. Syst. Signal Process. 102 (2018) 346–363.
[7] A. Kahraman, Natural modes of planetary gear trains, J. Sound Vibr. 173 (1994) 125–130.
[8] J. Lin, R.G. Parker, Analytical characterization of the unique properties of planetary gear free vibration, J. Vibr. Acoust. 121 (3) (1999) 316–321.
[9] W. Liu, Z. Shuai, Y. Guo, D. Wang, Modal properties of a two-stage planetary gear system with sliding friction and elastic continuum ring gear, Mech. Mach. Theory 135 (2019) 251–270.
[10] R.G. Parker, V. Agashe, S.M. Vijayakar, Dynamic response of a planetary gear system using a finite element/contact mechanics model, J. Mech. Des. 122 (3) (2000) 304–310.
[11] Q. Thoret-Bauchet, P. Velex, M. Guingand, P. Casanova, Simulations of the dynamic response of planetary gears in the presence of localised tooth faults, Proc. Inst. Mech. Eng., Part C (2019), 0954406219846153.
[12] H. Ma, J. Zeng, R. Feng, X. Pang, B. Wen, An improved analytical method for mesh stiffness calculation of spur gears with tip relief, Mech. Mach. Theory 98 (2016) 64–80.
[13] J. Lee, F. Wu, W. Zhao, M. Ghaffari, L. Liao, D. Siegel, Prognostics and health management design for rotary machinery systems reviews, methodology and applications, Mech. Syst. Signal Process. 42 (1–2) (2014) 314–334.
[14] S.S. Udmale, S.K. Singh, S.G. Bhirud, A bearing data analysis based on kurtogram and deep learning sequence models, Measurement 145 (2019) 665–677.
[15] A. Diez-Olivan, J. Del Ser, D. Galar, B. Sierra, Data fusion and machine learning for industrial prognosis: trends and perspectives towards industry 4.0, Inf. Fusion 50 (2019) 92–111.
[16] F. Jia, Y. Lei, J. Lin, X. Zhou, N. Lu, Deep neural networks: a promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data, Mech. Syst. Signal Process. 72 (2016) 303–315.
[17] C. Lu, Z.Y. Wang, W.L. Qin, J. Ma, Fault diagnosis of rotary machinery components using a stacked denoising autoencoder-based health state identification, Signal Process. 130 (C) (2017) 377–388.
[18] J. Jiao, M. Zhao, J. Lin, J. Zhao, A multivariate encoder information based convolutional neural network for intelligent fault diagnosis of planetary gearboxes, Knowl.-Based Syst. 160 (2018) 237–250.
[19] B. Yang, R. Liu, E. Zio, Remaining useful life prediction based on a double-convolutional neural network architecture, IEEE Trans. Ind. Electron. 66 (12) (2019) 9521–9530.
[20] M. Yuan, Y. Wu, L. Lin, Fault diagnosis and remaining useful life estimation of aero engine using lstm neural network, in: Aircraft Utility Systems (AUS), IEEE International Conference on, IEEE, 2016, pp. 135–140.
[21] Z. Chen, K. Gryllias, W. Li, Intelligent fault diagnosis for rotary machinery using transferable convolutional neural network, IEEE Trans. Ind. Inf. (2019).
[22] A. Punjani, P. Abbeel, Deep learning helicopter dynamics models, in: Robotics and Automation (ICRA), 2015 IEEE International Conference on, 2015, pp. 3223–3230.
[23] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (8) (1997) 1735–1780.
[24] P. Welch, The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms, IEEE Trans. Audio Electroacoust. 15 (2) (1967) 70–73.
[25] D. Yang, J. Lin, Hertzian damping, tooth friction and bending elasticity in gear impact dynamics, J. Mech. Transm. Autom. Des. 109 (2) (1987) 189–196.
[26] F. Chaari, T. Fakhfakh, M. Haddar, Analytical modelling of spur gear tooth crack and influence on gearmesh stiffness, Eur. J. Mech.-A/Solids 28 (3) (2009) 461–468.
[27] Z. Chen, Y. Shao, Dynamic features of a planetary gear system with tooth crack under different sizes and inclination angles, J. Vib. Acoust. 135 (3) (2013) 031004.
[28] X. Tian, Dynamic Simulation for System Response of Gearbox Including Localized Gear Faults, University of Alberta Edmonton, Alberta, Canada, 2004 (Ph.D. thesis).
[29] Y. Lei, Z. Liu, D. Wang, X. Yang, H. Liu, J. Lin, A probability distribution model of tooth pits for evaluating time-varying mesh stiffness of pitting gears, Mech. Syst. Signal Process. 106 (2018) 355–366.
[30] K.-I. Funahashi, Y. Nakamura, Approximation of dynamical systems by continuous time recurrent neural networks, Neural Networks 6 (6) (1993) 801–806.
[31] D.P. Kingma, J. Ba, Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980 (2014).
[32] Z. Feng, M.J. Zuo, Vibration signal models for fault diagnosis of planetary gearboxes, J. Sound Vib. 331 (22) (2012) 4919–4939.
[33] P. McFadden, J. Smith, An explanation for the asymmetry of the modulation sidebands about the tooth meshing frequency in epicyclic gear vibration, Proc. Inst. Mech. Eng. Part C 199 (1) (1985) 65–70.