Chapter 15
Time Series Modelling of Non-stationary Vibration Signals for Gearbox Fault Diagnosis

Yuejian Chen, Xihui Liang, and Ming J. Zuo

Abstract Gearboxes often operate under variable operating conditions, which lead to non-stationary vibration. Vibration signal analysis is a widely used condition monitoring technique. Time series model-based methods have been developed for the study of non-stationary vibration signals, and subsequently, for fault diagnosis of gearboxes under variable operating conditions. This chapter presented the latest methodologies for gearbox fault diagnosis using time series model-based methods. The main contents include widely used time-variant models, parameter estimation and model structure selection methods, model validation criteria, and fault diagnostic schemes based on either model residual signals or model parameters. Illustrative examples are provided to show the applications of model residual-based fault diagnosis methods on an experimental dataset collected from a laboratory gearbox test rig. Future research topics are pointed out at the end.

Keywords Gearboxes · Non-stationary · Time series models · Fault diagnosis · Vibration analysis

15.1 Introduction

Gearbox fault diagnosis refers to fault detection, fault mode identification, and severity assessment, which are critical for the prevention of sudden failures of gearboxes, enabling condition-based maintenance, and thus minimizing downtime and/or maintenance costs. Vibration analysis is the most widely used technique for gearbox fault diagnosis.

Y. Chen · M. J. Zuo (✉)
Department of Mechanical Engineering, University of Alberta, Edmonton, Alberta T6G 19, Canada
e-mail: ming.zuo@ualberta.ca; mzuo@ualberta.ca

X. Liang
Department of Mechanical Engineering, University of Manitoba, Winnipeg, MB R3T 5V6, Canada

K. B. Misra (ed.), *Handbook of Advanced Performability Engineering*,
https://doi.org/10.1007/978-3-030-55732-4_15

In many industrial applications, gearboxes operate under variable speed conditions. For instance, wind turbine gearboxes run under variable speed conditions due to the randomness of wind [1]. The gearbox that drives the fans of demand ventilation systems operates under variable speed conditions to reduce operating costs [2]. In railway systems, gearboxes experience run-up and coast-down conditions. The varying speeds modulate the amplitude and frequency of vibration signals. Therefore, the vibration signals become non-stationary. Effective non-stationary signal analysis tools are needed for gearbox fault diagnosis.

Time series model-based methods (TSMBMs) were initially employed in the structural health monitoring (SHM) field and have now drawn increased attention for gearbox fault diagnosis [3–5]. TSMBMs use time series models to model the vibration signals that are generated by gearbox systems. How to identify a time series model can be regarded as a response-only system identification problem.

Modelling non-stationary vibration signals need time-variant time series models that are realized by configuring the parameters of time-invariant models to be time variant. In this chapter, we will describe four widely used time-variant time series models. They are categorized based on how the parameters of time-invariant models are configured as time variant.

A time series model is generally composed of autoregressive (AR) and moving average (MA) terms. The MA terms are often ignored because (1) the AR terms can approximate the MA terms and (2) the consideration of MA terms makes the model identification more complex. Thus, in this chapter, we describe time-variant time series models that are composed of the AR terms only. Meanwhile, for simplicity, this chapter is limited to time series models for a single-channel vibration signal.

The materials in two of our earlier journal papers [3, 4] have been summarized and included in this chapter. Note that this chapter has also described other methods [5–10] to provide a comprehensive introduction of the latest TSMBMs.

The rest of this chapter is organized as follows: Sect. 15.2 introduces four time-variant time series models; Sect. 15.3 presents parameter estimation and model structure selection methods for the identification of time-variant time series models as well as the criteria for model validation; Sect. 15.4 describes two schemes (i.e. model residual-based scheme and model parameter-based scheme) for fault diagnosis; Sect. 15.5 presents the applications of the model residual-based fault diagnostic scheme on an experimental dataset collected from a laboratory gearbox test rig; conclusion remarks are drawn in Sect. 15.6.

15.2 Time Series Models for Non-stationary Vibration Signals

In this section, we present four time-variant AR models for representing non-stationary vibration signals. The first one is the periodic AR (PAR) model [6]. The PAR has AR parameters varying periodically with a specified period T. The PAR model has the following difference equation

$$y_t = \sum_{i=1}^{n_a} a_i(t) y_{t-i} + \varepsilon_t \tag{15.1}$$

where y_t and $y_{t\text{-}i}$ denote the vibration at time t and t-i, respectively; n_a is the AR model order; a_i stands for the AR parameters, which are periodic with the same period T; and ε_t is a zero-mean Gaussian white noise at time t. The PAR model is useful for representing non-stationary vibration signals with periodic time-varying characteristics. Wyłomańska et al. [6] used the PAR model for fault diagnosis of a gearbox in a bucket-wheel excavator that is a heavy-duty mining machine subjected to cyclic load/speed variation due to the digging/excavating process. However, the PAR model may not be applicable for representing non-stationary vibration signals collected under non-periodic variable speed conditions, such as the random variable speed condition that winds turbine gearbox experience.

The second one is the adaptive AR model with its model parameters adaptively (recursively) adjusted by recursive parameter estimation methods [7]. The adaptive evolution of model parameters enables the AR model time variant, and thus, it can track the non-stationary characteristics of vibration signals. Zhan et al. [7] and Shao et al. [8] used the adaptive AR model for fault diagnosis of fixed-axis gearboxes. The adaptive AR model requires a proper tuning of the convergence rate for its recursive parameter estimation algorithm. Too high of a convergence rate results in overfitting, and too low of a convergence rate causes underfitting.

The third one is the functional series time-dependent AR (FS-TAR) model [5, 10]. The FS-TAR model has a model difference equation the same as Eq. (15.1), but $a_i(t)$ is no longer periodic. Instead, $a_i(t)$ is represented by a function of time that is expanded in functional series (basis expansion). Therefore, the FS-TAR model is not limited to applications with cyclic load/speed variations. Reported basis functions include discrete cosine transform functions, Legendre polynomials, Harr functions, normalized B-spines, etc.[5]. Take the Legendre polynomials basis as an example. The dependency $a_i(t)$ is of the form:

$$a_i(t) = \sum_{j=1}^{p} a_{i,j} t^j \tag{15.2}$$

where $a_{i,j}$ stands for the AR parameters of projection and p specifies the order of functional spaces. The FS-TAR model has been widely used in the SHM field, such as a pick-and-place mechanism [10].

The last one is the functional pooled AR (FP-AR) model [3, 4]. The FP-AR model has the following model difference equation [3, 4, 11]

$$y_t = \sum_{i=1}^{n_a} a_i(k_t) y_{t-i} + \varepsilon_t \tag{15.3}$$

where k_t denotes the operating condition at time t and a_i is a function of k_t. We can see that the FP-AR model has the same model structure as the FS-TAR model, but with its AR parameters dependent on operating condition variable k_t. In the case when k_t is a vector, the FP-AR model is extended to the vector functional pooled autoregressive (VFP-AR) model in which the AR parameters are functions of a vector. Traditionally, the FP-AR models were identified to represent the vibrations under different levels of operating conditions [12–14]. It has recently been shown that the FP-AR model can be used to represent non-stationary signals which have a continuous time-varying spectrum [3, 4]. Chen et al. [3] presented an FP-AR model-based method for tooth crack fault detection of fixed-axis gearboxes under variable speed conditions. Chen et al. [4] presented a VFP-AR model-based method for tooth crack severity assessment of fixed-axis gearboxes under random variable speed conditions.

15.3 Model Identification and Validation

Model identification refers to the estimation of time series models based on the vibration data records y_t (for $t = 1, 2, \ldots, N$, where N is the number of data points). The identification of time series models includes parameter estimation and model structure selection. Model structure selection refers to the selection of lagged terms and/or functional basis. Once a time series model is identified, the model needs to be validated to ensure its modelling accuracy. In this section, we will describe the most widely employed methods for parameter estimation, model structure selection, and model validation.

15.3.1 Parameter Estimation Methods

Typical parameter estimation methods include the least squares (LS) and maximum likelihood (ML) [5]. The LS estimator can be used for PAR, FS-TAR, and FP-AR models. The LS estimator of the model parameter vector $\boldsymbol{\theta}$ is based on minimizing the squared summation of residuals $\mathbf{e}$

$$\hat{\boldsymbol{\theta}} = \text{argmin}\{\|\mathbf{e}\|\} = \text{argmin}\left\{\|\mathbf{y} - \boldsymbol{\Phi}^T\boldsymbol{\theta}\|\right\} \tag{15.4}$$

where $\|\bullet\|$ denotes the l_2 norm, $\mathbf{y} = [y_1, \ldots, y_N]^T$ denotes the observed time series, and $\boldsymbol{\Phi}$ is a hat matrix that is constructed from k_t and/or $y_{t-1}, \ldots, y_{t-n_a}$, depending on the time series model structure. The residual $\mathbf{e}$ means the one-step-ahead prediction error. The above minimization problem yields the solution expressed as

$$\hat{\boldsymbol{\theta}} = \left[\boldsymbol{\Phi}^T\boldsymbol{\Phi}\right]^{-1}\boldsymbol{\Phi}^T\mathbf{y} \tag{15.5}$$

The recursive least squares (RLS) estimator [5, 15] computes the model parameter vector $\boldsymbol{\theta}$ recursively by making use of the new data record at a given time instant t. When RLS is employed for estimating the AR model, we can realize an adaptive AR model (as introduced in Sect. 15.2) to represent non-stationary vibration signals. Readers may refer to Refs [5, 15] for more details about RLS estimators.

The maximum likelihood estimator can be used for PAR, FS-TAR, and FP-AR models. The ML estimator of the model parameter vector $\boldsymbol{\theta}$ is based on maximizing the log-likelihood function given as follows [5]

$$\hat{\boldsymbol{\theta}} = \text{argmax} L(\boldsymbol{\theta};\ \mathbf{e}|\mathbf{y}) \tag{15.6}$$

$$L(\boldsymbol{\theta};\ \mathbf{e}|\mathbf{y}) = -\frac{N}{2}\ln 2\pi - \frac{1}{2}\sum_{t=1}^{N}\left(\ln\sigma_t^2 + \frac{\varepsilon^2[t,\boldsymbol{\theta}]}{\sigma_t^2}\right) \tag{15.7}$$

where σ_t is the standard deviation of the residual, which is also dependent on the model parameter vector $\boldsymbol{\theta}$. The standard deviation σ_t is estimated directly from the residual sequence $\mathbf{e}$ using the sample standard deviation formula [5]. In Eq. (15.7), it is assumed that ε_t follows the zero-mean Gaussian distribution with a standard deviation σ_t.

Both LS and ML estimators are asymptotically Gaussian distributed with mean coinciding to the true value. The ML estimator, however, achieves lower model parameter estimation variance than the LS estimator [5]. On the other hand, the ML estimator has a higher computational cost than the LS method.

15.3.2 Model Structure Selection

Typical methods for model structure selection include various information criteria and regularization. The most widely used criteria are the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). The AIC is of the form

$$\text{AIC} = N\ln(\text{RSS}/N) + 2d \tag{15.8}$$

where d is the number of model parameters and RSS is the training residual sum of squares. The BIC is similar to the AIC, with a different penalty for the number of parameters as follows

$$\text{BIC} = N\ln(\text{RSS}/N) + d\ln(N) \tag{15.9}$$

Both AIC and BIC penalize the model structural complexity and thus avoid overfitting. Based on minimizing these criteria, model structure selection becomes an

integer optimization problem. Such an optimization problem can be solved via backward and/or forward regression, genetic algorithm, particle swarm algorithm, etc. It is important to note that these criteria-based methods require the assumption of consecutive AR set (and identical sets of functional spaces for FS-TAR and FP-AR) for the aforementioned four time-variant AR models to simplify the model structure selection procedure [14]. Without such an assumption, the integer optimization problem will have 2^d (d is usually greater than 100) different solutions, which is computationally impossible to find the global minimiser.

The regularization (e.g., l-1 norm)-based method has recently been adopted for model structure selection for the FP-AR model [3]. The regularization-based methods are free from the assumptions of consecutive AR set and identical sets of functional spaces, and, therefore, achieve higher modelling accuracy [3, 4]. With an initial (sufficient large) n_a and high dimension functional spaces, the least absolute shrinkage and selection operator (LASSO) estimator is given as follows,

$$\hat{\boldsymbol{\theta}} = \text{argmin}\left\{\|\mathbf{y} - \boldsymbol{\Phi}^T\boldsymbol{\theta}\|\right\} + \lambda|\boldsymbol{\theta}| \tag{15.10}$$

where $\lambda \geq 0$ is a tuning parameter and $|\bullet|$ denotes the l-1norm. The selection of λ is critical. When $\lambda = 0$, the LASSO will reduce to the LS estimator. Too large λ value will force too many coefficients to zero, whereas too small λ value will force a limited number of coefficients to zero. The λ can be selected by either the K-fold cross-validation [3] or validation set approach [4]. Other regularizations, such as l-2 norm and elastic net [15], are also options for time series model structure selection.

15.3.3 Model Validation

Model validation is mainly based on a validation signal. Upon the identification of a time series model, the inverse filter is constructed and then applied to process the validation signal. We refer to the residual obtained from the validation signal as 'residual-of-validation'. Model accuracy can be judged by the mean squared error (MSE) of the residual-of-validation, the randomness of the residual-of-validation, and the frozen-time spectrum [3, 5]. First, a model with a lower MSE of the residual-of-validation is more accurate in modelling the baseline vibration than those with a higher MSE [36, Sect. 7]. Second, the more random the residual-of-validation is, the more accurate the model is. Ljung–Box test [16] can be conducted to quantify the randomness. Last, the frozen-time spectrum $S(f, t)$ of time series models can be obtained and compared with the non-parametric spectrum (e.g. short-time Fourier transform) of non-stationary signals. An accurate time series model should give a parametric spectrum in good agreement with the non-parametric one.

15.4 Time Series Model-Based Fault Diagnosis

Fault diagnosis may be based on either model residuals or model parameters. In this section, we introduce both model residual-based methods and model parameter-based methods.

15.4.1 Model Residual-Based Method

For fault detection, the model residual-based method relies on the identification of a baseline time series model. Figure 15.1 shows the schematic of the model residual-based method for fault detection [3]. A time series model is identified to represent the baseline vibration signals. Then, the vibration signals collected under future unknown health state are processed by an inverse filter constructed from the baseline model. Any changes in the residual signals may indicate the occurrence of a fault. Researchers have examined the whiteness [8], variance [17, 18], Gaussionality [7], and impulsiveness [3, 8] to quantify the changes in residual signals.

For severity assessment or fault mode identification, the model residual-based method relies on the identification of time series models under each severity level or fault mode. Figure 15.2 shows the schematic of the model residual-based method for severity assessment [4, 19]. The presented scheme can also be used for fault mode identification by changing the fault severity states to fault modes. During the training phase, the training signals collected under each fault severity level and a wide range of the speed variation are used. Under each fault severity level, a time series model is identified to represent the vibration signals of that state. We refer these time series models as state models. During the testing phase, these trained state models are used for severity assessment. Vibration signal y_t, along with necessary operating condition variables is collected under an unknown health state of the gearbox. Afterwards, the inverse filters from each of the state models are applied to process the vibration signals collected under the unknown health state and to obtain residuals of the state models. The final health state is classified as the state with an inverse filter that gives minimal residual MSE. Note that in industrial applications, it is not easy to obtain the signals under known health states. This is the major challenge associated with this model residual-based method for severity assessment or fault mode identification.

15.4.2 Model Parameter-Based Method

The model parameter-based method requires the identification of a model during the testing stage. For fault detection, the model parameter-based method is based on comparing the parameters of the current model with the parameters of the baseline time series model. Figure 15.3 shows the schematic of the model parameter-based

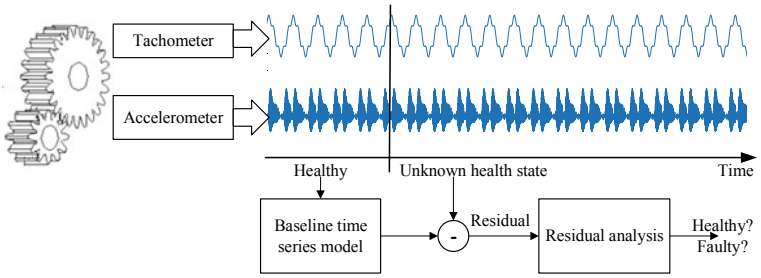

Fig. 15.1 Model residual-based fault detection method [3]

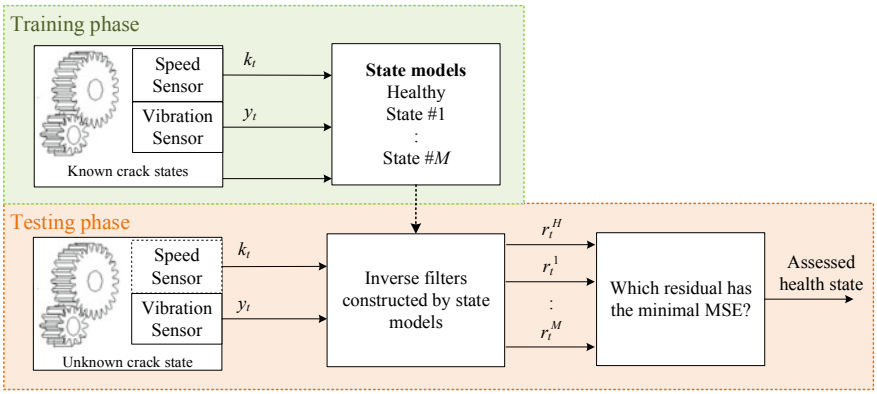

Fig. 15.2 Model residual-based severity assessment or mode identification method [4]

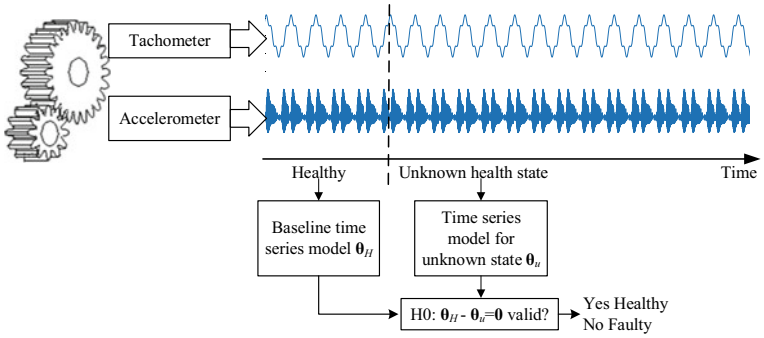

Fig. 15.3 Model parameter-based fault detection method [21]

method for fault detection [20, 21]. During the training stage, a time series model is identified to represent the baseline vibration signals. During the testing stage, another time series model is identified to represent the current vibration signals from

an unknown health state. Fault detection is based on testing statistical differences between the model parameters under the healthy state $\boldsymbol{\theta}_H$ and the model parameters under the unknown health state $\boldsymbol{\theta}_u$ through the following hypotheses [20].

$$\text{H0} : \boldsymbol{\theta}_H - \boldsymbol{\theta}_u = 0; \text{H1}; \boldsymbol{\theta}_H - \boldsymbol{\theta}_u \neq 0 \tag{15.11}$$

where H0 denotes the null hypothesis and H1 denotes the alternative hypothesis. If H0 is valid, then the unknown health state is deemed healthy. Otherwise, the unknown health state is detected as faulty.

For severity assessment or fault mode identification, the model parameter-based method is based on comparing the parameters of the current model with the parameters of the trained state models [10]. Figure 15.4 shows the schematic of the model parameter-based method for severity assessment or fault mode identification. During the training stage, the state time series model is identified for each severity level or each fault mode. During the testing stage, another time series model is identified to represent the current vibration signals from an unknown health state. Severity assessment or fault mode identification is based on testing statistical differences between the model parameters obtained from the training stage ($\boldsymbol{\theta}_H$, $\boldsymbol{\theta}_1$, …, $\boldsymbol{\theta}_M$) and the model parameters under the unknown health state $\boldsymbol{\theta}_u$ through hypotheses similar to Eq. (15.11). The final health state is classified as the state with a valid null hypothesis. Since the model parameter-based method requires the identification of a model during the testing stage, it is generally not suitable for online applications.

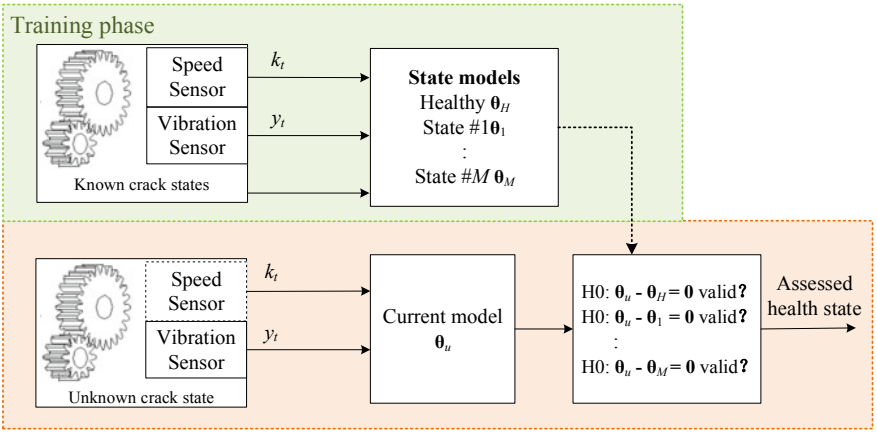

Fig. 15.4 Schematic of the model parameter-based severity assessment method [11]

15.5 Applications of Model Residual-Based Methods

This section presents the applications of the two model residual-based methods [3, 4] for fault detection and severity assessment of gear tooth crack. Section 15.5.1 will briefly describe the experimental dataset. Section 15.5.2 will present the application of the sparse FP-AR model residual-based fault detection method [3] on vibration signals different from the applications presented in ref.[3]. Section 15.5.3 will summarize the application of the sparse VFP-AR model residual-based fault detection method, as reported in Ref.[4].

15.5.1 Experimental Dataset

The experimental dataset was collected at the University of Pretoria, South Africa [22, 23]. Readers can refer to refs [22, 23]. for the detailed experimental setup. Two vibration sensors were equipped on this test rig. One vibration sensor is single axial and is labelled as #7. The other vibration sensor is a triaxial accelerometer. The experimental dataset contains 100 data files from a healthy gearbox and 1400 data files from a run-to-failure experiment with 50% initial crack and duration of around 21 days of continuous running. Each data file contains data collected within 20 s. The sampling frequency was $f_s = 25.6$ kHz. These vibration signals were further low passed using an FIR filter with a cut-off frequency of $f_c = 1.6$ kHz and then downsampled from $f_s = 25.6$ kHz to $f_s = 3.2$ kHz. When collecting each data file, an electrical motor drove the transmission train such that the rotating speed of the target gearbox followed a sinusoidal-like profile with a period equals to 10 s. The alternator generated a load torque positively correlated to the speed.

15.5.2 FP-AR Model-Based Fault Detection

In this subsection, we present the application of the sparse FP-AR model residual-based method [3] for gear tooth crack fault detection. In ref.[3], the method was applied to the vibration signal collected from the sensor labelled as #7. They did not analyze the vibration data collected from the triaxial accelerator. In this subsection, we are to apply the method reported in ref [3] to the vibration signal collected from the x-direction of the triaxial accelerometer (with a sensitivity of 100 mV/g). In other words, a one-dimensional vibration data series from a different sensor of the same test rig will be used to assess the effectiveness of the sparse FP-AR model residual-based method [3].

The following configurations are the same as used in Ref. [3]: The signals from a zebra-tape shaft encoder are used in this subsection to obtain the rotating speed

information. Training and validation data are arbitrarily selected from the 100 baseline data files. The training data were 7.5 s in length, which was truncated from the length of 20 s. The validation data also have a length of 7.5 s. During the sparse FP-AR modelling, the initial set of functional spaces was configured as $\{1, \omega_t, \omega_t{}^2, \ldots, \omega_t{}^7\}$, where ω_t is the rotating speed. The candidate set for λ was configured as $[0, 1 \times 10^{-8}, 1 \times 10^{-7}, \ldots, 1 \times 10^{-1}]$.

The following results are obtained when we apply the sparse FP-AR model residual-based method to the vibration signal collected from the x-direction of the triaxial accelerometer: The initial n_a was determined as $\{1, 2, \ldots, 63\}$ by BIC. When $\lambda = 0$, we achieved the minimum CVMSE. By increasing λ, the CVMSE will get bigger. We need to use a larger λ value that does not have too big a CVMSE. When λ increases to 1×10^{-5}, the CVMSE is still within one standard deviation of the minimum CVMSE [3]. Therefore, $\lambda = 1 \times 10^{-5}$ is chosen. Table 15.1 lists the modelling performance of the identified sparse FP-AR. From the table, we can see that the residual-of-validation of the sparse FP-AR model has a p-value of 0.1509 from the Ljung–Box test. This p-value is significantly higher than the p-value of 0 of the validation signal. This means that the randomness of the residual-of-validation is much higher than the original validation signal.

Figure 15.5 shows the non-parametric spectrum (a) and the frozen-time spectrums (b) obtained from the sparse FP-AR model. The frozen-time spectrum obtained from the sparse FP-AR model aligns well with the non-parametric spectrum by tracking the time-varying spectral contents. Given the above model validation criteria,

Table 15.1 Modelling performance of the sparse FP-AR model. Algorithms were coded in MATLAB 2019a and implemented on a desktop with two Intel 2.4 GHz processors and 16 GB of RAM

Ljung–Box test, p-value, of validation signal	Ljung–Box test, p-value, of residual-of-validation	CPU time in training, (min)	CPU time in testing, (s)
0	0.1509	48	0.4

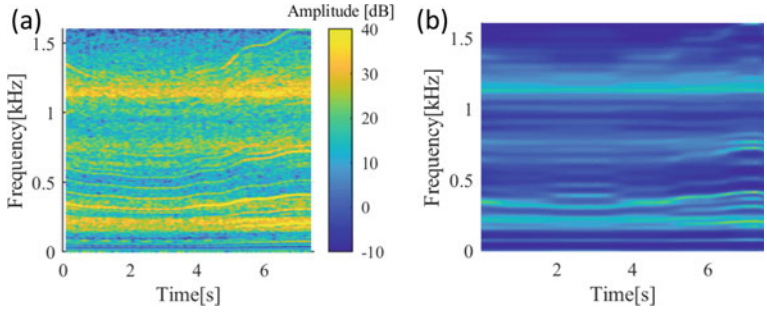

Fig. 15.5 Time–frequency spectrums: **a** non-parametric spectrum of the validation signal obtained by MATLAB spectrogram function; **b** Frozen-time spectrum obtained from the sparse FP-AR model. Z-axis scales are the same

we can conclude that the sparse FP-AR model has good modelling accuracy when representing the baseline vibration signal collected from the *x*-direction of the triaxial accelerometer. The sparse FP-AR model used about 48 min in the training stage and 0.4 s in the testing stage. Since this training process is completed offline, the length of time required is not very critical [3]. The computational time in the testing stage is more critical as it determines whether the TSMBM based on the sparse FP-AR model is practically useful or not [3]. In many applications where incipient faults do not immediately lead to a catastrophic failure of the gearbox system, updating the fault detection information every second is acceptable and thus requiring about 0.4 s in testing is acceptable [23].

The identified sparse FP-AR model was used for detecting the gear tooth crack faults. Each of the 99 baseline data files (the 100 baseline data files exclude the one used for training the sparse FP-AR model) and 1400 run-to-failure data files are truncated to have the speed profile the same as the validation signal. Figure 15.6 shows the normalized periodic modulation intensity (NPMI) [3] calculated from both the raw data and the residuals obtained from the identified sparse FP-AR model for each data file. The NPMI is the periodic modulation intensity (PMI) value of the residual divided by the PMI of the residual of baseline vibration, where the PMI represents the energy ratio between tooth crack-induced impulses and other components. Figure 15.6a shows the baseline case, whereas Fig. 15.6b shows the damaged case. It is clear that for the NPMI obtained from the raw signals, they are of similar magnitude for both the healthy data files and faulty data files. On the other hand, the NPMIs from the residuals of the sparse FP-AR model for the faulty data files are obviously higher than those for the healthy data files. This means that using the solid blue plots, we are able to detect the faults.

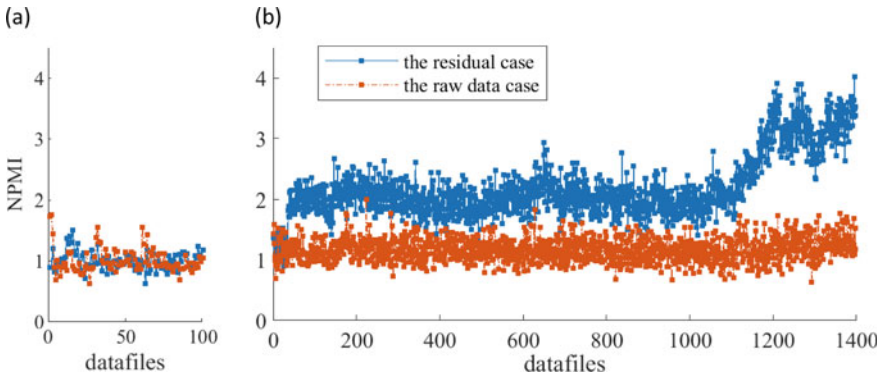

Fig. 15.6 Normalized PMI for detecting the tooth crack fault. **a** Healthy data files; **b** faulty data files

15.5.3 VFP-AR Model-Based Severity Assessment

In this subsection, we summarize the application of the sparse VFP-AR model-based method for gear tooth crack severity assessment, as reported in ref.[4].

Four discrete health states, namely, healthy (H), initial crack (F1), intermediate crack (F2), and missing tooth (F3) were considered. The F1 state corresponds to the gear with the initial 50% tooth crack. The F2 state corresponds to the gear that had run 17 days after the initial 50% tooth crack. The F3 state corresponds to the gear that ran right before the end of the run-to-failure experiment.

Under each health state, training, validation, and testing signals were prepared and preprocessed. In total, 43 data files under each health state were used in which one served for training (training signal), two for validation (one for model identification and the other for measuring modelling accuracy), and the rest 40 for testing severity assessment performance. The inverse filter constructed by the baseline sparse FP-AR model identified in Sect. 15.5.2 was used to preprocess these signals and to obtain residual signals.

For the training and validation signals, the first half (10 s) of the vibration signal in a data file was used. Such a vibration signal experienced a full cycle of the speed variation. On the other hand, the 40 segments of testing signals only lasted 5 s with a starting point p_s randomly sampled from [0, 0.25, 0.5, …, 10]s.

For the identification of the sparse VFP-AR models, Legendre polynomial basis functions were used for $\mathbf{G}(\omega_t)$ and the refined B-splines for $\mathbf{G}(\theta_t)$ [4] where θ_t is the rotating phase. The $\mathbf{G}(\omega_t)$ was configured as $\{1, \omega_t, {\omega_t}^2, \ldots, {\omega_t}^7\}$. As for the refined B-splines, r was configured as 3 and K as 40 (i.e. the number of teeth 37 plus $r = 3$). Two parameters k and n were further determined by estimating a sparse VFP-AR model with a small $n_a = 5$ and examining the occurrence of periodic B-spline bases. The k and n were determined as (24, 1) for F1 state; (24, 2) for F2 state; (24, 3) for F3 state. Since the vibration signal under the H state did not have crack induced impulses, its corresponding sparse VFP-AR model did not need to consider the phase. In other words, the sparse VFP-AR model for the H state reduced to a sparse FP-AR model. Afterwards, the n_a was to be determined after obtaining the k and n via the validation set approach. The n_a was determined to be (50, 40, 35, 35) for health states (H, F1, F2, F3), respectively.

Upon the identification of both the sparse VFP-AR models, the inverse filter was constructed and then applied to process the validation signals for measuring the modelling accuracy. Table 15.2 lists the MSE and the randomness of both residuals and the validation signals (i.e. the residual of the baseline sparse FP-AR model). The p-values from the Ljung–Box tests were reported, which means the probability of being random. From this table, we can see that the sparse VFP-AR models return a residual with reduced MSE and a higher probability of being random compared with the validation signals.

Computational costs were evaluated and listed in Table 15.2 as well. The training data points were 32,000. For four sparse VFP-AR models, the time required for training was around 17.9 h. Since this training process was completed offline, the

Table 15.2 Modelling accuracy and computational cost of sparse VFP-AR models [4]

Health state	H	F1	F2	F3
MSE of residual(normalized m/s^2)	0.881	1.172	1.114	1.191
MSE of validation signal (normalized m/s^2)	1.008	1.347	1.271	1.369
Randomness of residual (p-value)	1.55×10^{-5}	1.56×10^{-5}	0.0034	6.93×10^{-5}
Randomness of validation signal (p-value)	0	0	0	0
Time in training (h)	1.5	4.9	5.2	6.3
Time in testing (s)	2.2	2.4	2.4	2.6

length of time needed was not very critical [3, 4]. The computational time in the testing stage is more critical to determine whether the method is practical or not [3]. The inverse filter constructed by the four sparse VFP-AR models used less than 9.6 s. Again, in many applications where incipient faults do not immediately lead to a catastrophic failure of the gearbox system, updating the fault detection information every 10 s is acceptable.

Figure 15.7 shows the non-parametric spectrum (a, b, c, d) of the validation signals as well as the frozen-time spectrums of the sparse VFP-AR models (e, f, g, h). We can see the tooth crack-induced impulses as vertical lines in these spectrums. The vertical lines in the frozen-time spectrums of sparse VFP-AR models behave discretely, which are in good agreement with the discrete lines in non-parametric spectrums. The tooth crack-induced impulses can be represented using sparse VFP-AR models.

Figure 15.7 shows the non-parametric spectrum (a, b, c, d) of the validation signals as well as the frozen-time spectrums of the sparse VFP-AR models (e, f, g, h). We can see the tooth crack-induced impulses as vertical lines in these spectrums. The vertical lines in the frozen-time spectrums of sparse VFP-AR models behave discretely, which are in good agreement with the discrete lines in non-parametric spectrums. The tooth crack-induced impulses can be represented using sparse VFP-AR models.

The sparse VFP-AR models under known health state were applied to testing signals for the severity assessment. For each testing signal, four model residuals were obtained, and their MSE values were calculated. The health state was classified as the state with an inverse filter that gave minimal residual MSE. Figure 15.8 shows the classification results when processing the 40 testing signals under each health state. The classification accuracy was reported to be 93.8%. The results showed the effectiveness of the sparse VFP-AR model-based method.

15.6 Summary and Conclusion

This chapter presented the latest methodologies related to the time series model-based techniques for gearbox fault diagnosis. We described four most widely used time-variant time series models, typical parameter estimation and model structure

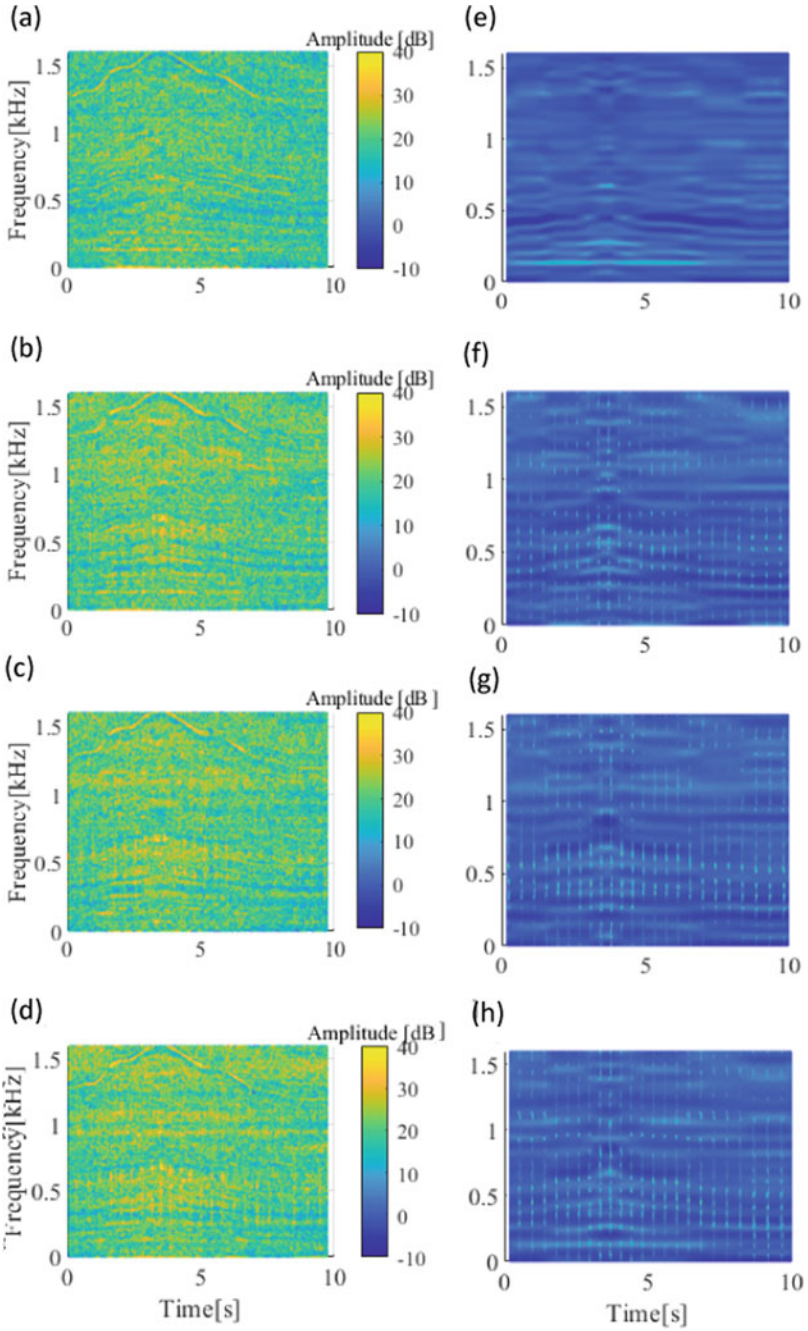

Fig. 15.7 Time–frequency spectrums: **a** ~**d** STFT spectrum of the validation signal; **e** ~ **h** Frozen-time spectrum of sparse VFP-AR models. Z-axis scales are the same for all spectrums [4]

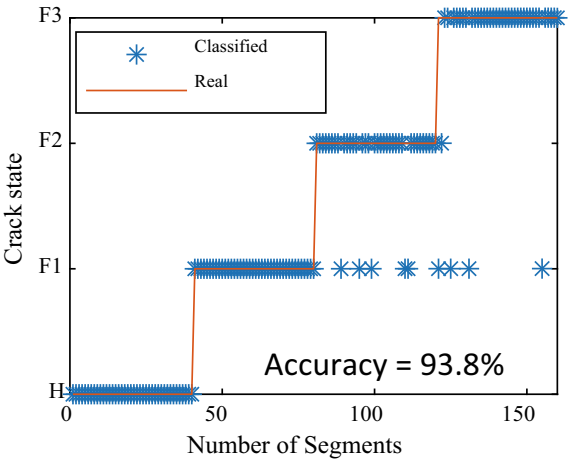

Fig. 15.8 Classification results [4]

selection methods for model identification, model validation criteria, and fault diagnosis schemes based on either model residual or model parameters. Finally, this chapter gave two examples to illustrate the applications of the model residual-based fault diagnosis method on a lab gearbox. The following aspects may be further investigated in future studies: (a) various regularization techniques, such as l-2 norm and elastic net, for the structure selection of time-variant time series models; (b) the consideration of more than two operating condition variables in a VFP-AR model, such as temperature, rotating speed, and load torque; and (c) the account of uncertainties of operating condition variables when identifying a time-variant time series model.

Acknowledgements This research is supported by the Natural Science and Engineering Research Council of Canada, Canada [grant number RGPIN-2015-04897, RGPIN-2019-05361]; Future Energy Systems under Canada First Research Excellent Fund [grant number FES-T11-P01, FES-T14-P02]; University of Manitoba Research Grants Program (URGP); Sadler Graduate Scholarship in Mechanical Engineering, Canada; and China Scholarship Council, China [grant number 201506840098].

References

1. Carlin, P. W., Laxson, A. S., & Muljadi, E. B. (2003). The history and state of the art of variable-speed wind turbine technology. *Wind Energy., 6*(2), 129–159.
2. McBain, J., & Timusk, M. (2009). Fault detection in variable speed machinery: Statistical parameterization. *Journal of Sound and Vibration, 327*(3), 623–646.
3. Chen, Y., Liang, X., & Zuo, M. J. (2019). Sparse time series modeling of the baseline vibration from a gearbox under time-varying speed condition. *Mech Syst Signal Process., 1*(134), 106342.
4. Chen, Y., Schmidt, S., Heyns, P. S., & Zuo, M. J. (2020). A time series model-based method for gear tooth crack detection and severity assessment under random speed variation. *Mechanical System and Signal Processing, 9,* 1–32.

5. Spiridonakos, M. D., & Fassois, S. D. (2014). Non-stationary random vibration modelling and analysis via functional series time-dependent ARMA (FS-TARMA) models—A critical survey. *Mechanical System and Signal Processing, 47*(1–2), 175–224.
6. Wyłomańska, A., Obuchowski, J., Zimroz, R., Hurd, H. (2017). Periodic autoregressive modeling of vibration time series from planetary gearbox used in bucket wheel excavator. In: Chaari, F., Leśkow, J., Napolitano, A., Sanchez-Ramirez, A., (Eds.), *Cyclostationarity: Theory and Methods* [Internet]. Springer International Publishing; 2014 [cited 2017 Mar 31]. pp. 171–86. (Lecture Notes in Mechanical Engineering). Available from: https://link.springer.com/chapter/10.1007/978-3-319-04187-2_12.
7. Zhan, Y., & Mechefske, C. K. (2007). Robust detection of gearbox deterioration using compromised autoregressive modeling and Kolmogorov-Smirnov test statistic—Part I: Compromised autoregressive modeling with the aid of hypothesis tests and simulation analysis. *Mechanical System and Signal Processing, 21*(5), 1953–1982.
8. Shao, Y., & Mechefske, C. K. (2009). Gearbox vibration monitoring using extended Kalman filters and hypothesis tests. *Journal of Sound and Vibration, 325*(3), 629–648.
9. Spiridonakos, M.D., Fassois, S.D. (2014). Adaptable functional series TARMA models for non-stationary signal representation and their application to mechanical random vibration modeling. *Signal Process*, *96*, Part A, 63–79.
10. Spiridonakos, M. D., & Fassois, S. D. (2013). An FS-TAR based method for vibration-response-based fault diagnosis in stochastic time-varying structures: Experimental application to a pick-and-place mechanism. *Mechanical Systems and Signal Processing, 38*(1), 206–222.
11. Sakellariou, J.S., Fassois, S.D. (2007). A functional pooling framework for the identification of systems under multiple operating conditions. In: *2007 Mediterranean Conference on Control Automation*, pp. 1–6.
12. Kopsaftopoulos, F., Nardari, R., Li, Y.-H., & Chang, F.-K. (2018). A stochastic global identification framework for aerospace structures operating under varying flight states. *Mechanical Systems and Signal Processing, 1*(98), 425–447.
13. Sakellariou, J. S., & Fassois, S. D. (2016). Functionally Pooled models for the global identification of stochastic systems under different pseudo-static operating conditions. *Mechanical Systems and Signal Processing, 1*(72–73), 785–807.
14. Kopsaftopoulos, F. P., & Fassois, S. D. (2013). A functional model based statistical time series method for vibration based damage detection, localization, and magnitude estimation. *Mechanical Systems and Signal Processing, 39*(1), 143–161.
15. Hastie, T., Tibshirani, R., Friedman, J.H. (2009). *The Elements of Statistical Learning: data mining, inference, and prediction.* [Internet]. 2nd ed. 2009 [cited 2020 Feb 18]. Available from: https://web.stanford.edu/~hastie/ElemStatLearn/.
16. Ljung, G. M., & Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika, 65*(2), 297–303.
17. Heyns, T., Godsill, S. J., de Villiers, J. P., & Heyns, P. S. (2012). Statistical gear health analysis which is robust to fluctuating loads and operating speeds. *Mechanical Systems and Signal Processing, 27,* 651–666.
18. Yang, M., & Makis, V. (2010). ARX model-based gearbox fault detection and localization under varying load conditions. *Journal of Sound and Vibration, 329*(24), 5209–5221.
19. Li, G., McDonald, G.L., Zhao, Q. (2017). Sinusoidal synthesis based adaptive tracking for rotating machinery fault detection. *Mechanical Systems and Signal Processing*, *83*(Supplement C), 356–370.
20. Hios, J. D., & Fassois, S. D. (2014). A global statistical model based approach for vibration response-only damage detection under various temperatures: A proof-of-concept study. *Mechanical Systems and Signal Processing, 49*(1), 77–94.
21. Aravanis, T.-CI., Sakellariou, J.S., Fassois, S.D. (2019). A stochastic Functional Model based method for random vibration based robust fault detection under variable non–measurable operating conditions with application to railway vehicle suspensions. *Journal of Sound and Vibration*, 115006.

22. Schmidt, S., Heyns, P. S., & de Villiers, J. P. (2018a). A novelty detection diagnostic methodology for gearboxes operating under fluctuating operating conditions using probabilistic techniques. *Mechanical Systems and Signal Processing, 1*(100), 152–166.
23. Chen, Y., Liang, X., & Zuo, M. J. (2020). An improved singular value decomposition-based method for gear tooth crack detection and severity assessment. *Journal of Sound and Vibration, 3*(468), 115068.