Health Index Development for Planetary Gearboxes

by

Weixuan Tang

A thesis submitted in partial fulfilment of the requirements for the degree of

Master of Science

in

Engineering Management

Department of Mechanical Engineering

University of Alberta

ABSTRACT

Planetary gearboxes are widely used in machineries such as wind turbines and helicopters. To maximize their effectiveness over their lifecycle, condition monitoring is often used, and proper health indexes can be developed utilizing condition monitoring data. Health index (HI) development for planetary gearboxes contains two important parts: input feature selection and HI smoothing procedure. Input feature selection is to select the best combination of features as the HI modeling input that provides the highest HI prediction accuracy. HI smoothing procedure is to further improve the modeled HI to get an even higher HI prediction accuracy.

A reported method uses a feedforward neural network (FFNN) to develop an HI for a type of electric motor. The FFNN is to find the relationship between condition monitoring data and the HI. The reported method uses a fixed stepsize following sequential ordering to select the optimal input features. In addition, the reported method reports an HI dynamic smoothing procedure to further smooth the modeled HI in order to get a higher accuracy of HI prediction. This thesis investigates in-depth the reported method and finds that there are two aspects that are unclear and deficient. These two aspects are thus investigated and the suggestions are provided to address the shortcomings. The findings of this thesis are listed as follows:

1) The impact of the combinations of the input features in the FFNN-based HI model is investigated. A feature selection method is proposed to find the optimal subset of features.

2) The impact of both the window size parameter and the maxdrop parameter in the reported HI dynamic smoothing procedure is investigated. An improved HI dynamic smoothing procedure using the optimized window size parameter and the optimized maxdrop parameter is proposed.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisor Dr. Ming J. Zuo for his patient supervision and support through all my study and research. I have gained solid academic training and have achieved this work. My sincere thanks also go to the members of my examining committee: Dr. Zhigang Tian and Dr. Basel Alsayyed Ahmed.

I would like to thank all the members of Reliability Research Lab for their kind help during my M.Sc. study.

This thesis would not have been possible without the support of my parents and my girlfriend. I take this opportunity to express my thanks to all of them for their understanding and consistent encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

ANN	Artificial Neural Network
BA	Backpropagation Algorithm
BFP	Boiler Feed Pump
CBM	Condition-Based Maintenance
DGA	Dissolved Gas Analysis
FFNN	Feed-Forward Neural Network
FR	Fisher's Ratio
FT	Fourier Transform
GA	Genetic Algorithm
GMF	Gear Meshing Frequency
HE	Heat Exchanger
HI	Health Index
HIF	Health Index Factor
MR	Monotonicity Ratio
PM	Preventive Maintenance
PT	Power Transformer

RMS	Root Mean Square
RMSE	Root Mean Squared Error
RTF	Run-To-Failure
RUL	Remaining Useful Life
WTG	Wind Turbine Gearbox

Chapter 1

Introduction

This chapter consist of three sections. Section 1.1 and 1.2 introduces background and planetary gearboxes, respectively. Section 1.3 presents the research objective.

1.1 Background

Engineering assets are widely used to serve the needs of humankind. These assets include energy infrastructure such as power generation stations, wind farms, power transmission and distribution networks, and chemical refineries. They include transportation infrastructure such as highways, railways, bridges, and pipelines. They also include transportation vehicles such as automobiles, trucks, rail cars, airplanes, and space shuttles. All such assets require proper design, manufacture, construction, operation, and maintenance to ensure their ability to deliver the services our society needs. To maximize the effectiveness of these assets over their life cycles, condition-based maintenance (CBM) strategy and tactics have been showing more and more benefits.

Traditionally, the role of maintenance is to fix failed parts in an engineering asset. Maintenance activities were treated as the reactive tasks of repair works or parts replacement triggered by failures [1]. This type of maintenance is called corrective maintenance and it has been employed in industries for many decades. As its name indicates, corrective maintenance is used only when an actual failure has occurred in an engineering asset. If a critical part in an engineering asset actually fails, this will generate

unplanned downtime. This may cause huge disruption to the needed services and have major economic losses.

To prevent failures of critical engineering assets, another type of maintenance called preventive maintenance (PM) has been developed to address the shortcomings of corrective maintenance. PM is a scheduled maintenance considering running hours of equipment [1]. Usually a constant time or usage interval is used to schedule maintenance activities under this strategy. If this interval is too short, too frequent preventive maintenances are performed. Though this would reduce unexpected failures, it also results in extra time and resources spent on such maintenance activities and increases the operation and maintenance costs of such assets. Due to the uncertainties in the potential failure times of critical components of engineering assets, it is difficult to select the best PM interval.

Due to the technological advances in recent decades, CBM strategy has been used more and more widely. The advances in sensors, computing, and data processing have made accurate nondestructive evaluation of the health status of running engineering assets possible. Once such health status is obtained, appropriate operating and maintenance decisions can be made to ensure reliable operation of engineering assets with minimal operating and maintenance costs. This strategy is called CBM as maintenance decisions are made based on the assessed health condition of the assets.

Many condition monitoring approaches have been developed to assess the health status of various engineering assets. Vibration analysis has been widely used for engineering systems with moving parts including engines, pumps, gearboxes, and bearings [2]. This method uses the vibration data collected from such systems to assess the health condition

of critical moving parts inside these systems. Thermographic analysis which processes the collected temperature data has been used to assess the health condition of parts in plain journal bearing systems [3]. Ultrasonic testing which replies the bounced ultrasonic echoes from pipeline walls has been used to measure cracks and corrosions that may be experienced by buried metal pipelines [4]. Such techniques may be used continuously or periodically to monitor the health status of running engineering assets without interrupting their usual operations. With the health status of the running engineering assets obtained with such condition monitoring techniques, the operation and maintenance decisions that are the best suited to address such health conditions can be made. Unlike the PM strategy, CBM is a condition-driven or data-driven method considering condition monitoring data collected by sensors. With sensory data, we are able to optimize maintenance actions and make better maintenance decisions.

A critical step for effective CBM decisions is the accurate assessment of the health status of the running engineering asset. In recent literature, various health indexes (HIs) have been developed to describe the health status of various assets [5]. What health index (HI) to use and how it is formulated depend on the target asset, its failure modes, and the types of data collected. Aiming to develop accurate health indexes, we next introduce the target engineering asset to be studied in this thesis research.

1.2 Planetary Gearboxes

Planetary gearboxes are widely used in engineering systems such as helicopters, wind turbines, and transmissions in various industrial machinery because of their high compactness, high torque-to-weight ratio, and high transmission ratio [6]. Planetary

gearboxes are also commonly used in mining machineries. See Figure1.1 for some of such applications. Mining machineries are used in mines and oilfields to refine ore, oil sand, petroleum, etc. Planetary gearboxes are critical components of mining machineries. The reliable operation of such planetary gearboxes directly affects the efficiency of mining. According to [7], gearboxes cause the highest downtime in a rotating machinery. To make contribution to mining industry development, this thesis selects planetary gearboxes as the target asset for their health index development.

Figure 1.1 Heavy industrial applications of planetary gearbox [8]

The structure of a 1-stage planetary gear set with four planet gears is shown in Figure 1.2. The key parts of a planetary gear set includes a ring gear, a sun gear, a number of planet gears, and a carrier [6], as shown in Figure1.2. The number of planet gears is dependent on specific applications. The planet gears which are fixed on the carrier may rotate around

their own axes. The carrier may rotate around its own axis carrying all planet gears with its rotation. The sun gear and the ring gear may rotate around their own axes. The planet gears mesh simultaneously with the sun gear and the ring gear. All rotating axes are supported by bearings. Depending which axis is non-rotating, planet gearboxes may be divided into the following categories with their typical applications [6]:

(1) Planetary gearboxes with standstill ring gear

(2) Planetary gearboxes with standstill sun gear

(3) Planetary gearboxes without any standstill gear

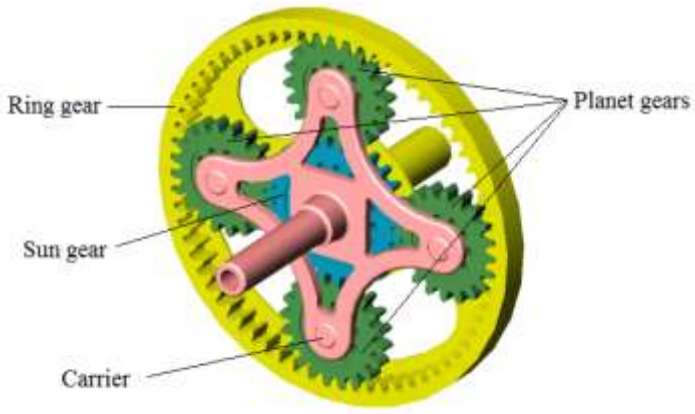

Figure 1.2 Structure of a planetary gear set with 4 planet gears [8]

The planetary gearbox used in mining machinery are of type (1) described above. In mining planetary gearboxes, the ring gear is fixed, while the sun gear, the planet gears, and the carrier are rotating [9]. The typical used bearings in mining planetary gearboxes are tapered roller bearings or needle roller bearings for planet gears, tapered roller bearings for sun gear, and ball bearings for output shaft [10]. The usual number of planet gears in mining

machinery are 3 or 4. In this thesis research, we will use a planetary gearbox with 4 planet gears structure produced by SpectraQuest Inc. as our study object.

Mining machinery are usually located in underground. The operating condition and operating environment may be humid, muddy, high-pressure, etc. The key components of planetary gearboxes such as gears and bearings are subject to fault types such as pitting, crack, and wear [6]. A failure of such a planetary gearbox may lead to unexpected shutdown of the whole mining machinery, leading to major economic losses and even human casualties may occur due to the unexpected shutdown. Condition monitoring, degradation analysis, and health index development of planetary gearboxes aim to help engineers to prevent such accidents and optimize the CBM actions of planetary gearboxes.

The usual data collected from mining gearboxes include vibration signals, acoustic emissions, temperature, etc. Such data need to be used effectively for accurate assessment of the health index of these gearboxes. In this thesis, we develop the health index based on the available vibration signals collected in a lab planetary gearbox.

1.3 Research Objectives

As described earlier in this chapter, this research focuses on development of more accurate health indexes for planetary gearboxes used in mining machinery utilizing vibration data collected from running gearboxes.

This thesis is organized as follows. Chapter 2 will review the works reported in the literature on health index development for engineering assets including gearboxes (and planetary gearboxes). Detailed research objectives will be outlined based on this literature review in Chapter 2. Chapter 3 will introduce the fundamental knowledge of artificial

neural network (ANN) and genetic algorithm (GA), since we will use ANN to develop HI and we will use GA for feature selection. Chapter 4 mainly describes the run-to-failure experimental data collected from a lab planetary gearbox that will be used for illustration of the development of HI. It also describes the data preprocessing which includes feature extraction, dimension reduction of the vibration signals collected from the lab planetary gearbox that have been reported in [11], [12]. After the dimension reduction, the features selected via dimension reduction will be used as the input for our research investigations in this thesis. Chapter 5 describes the investigations and numerical experiments conducted in the process of developing a more accurate HI for the planetary gearbox using a Feed-Forward Neural Network (FFNN). Detailed FFNN will be introduced in Chapter 3. We will compare three different feature selection strategies in order to evaluate which strategy performs the best for HI development. After developing an HI, a dynamic smoothing procedure as reported in [13] is then used and modified to smooth the developed HI in order to decrease the fluctuation of the developed HI. Furthermore, Chapter 5 will also discuss the optimal parameter selection in the reported dynamic smoothing procedure in [13]. Chapter 6, the final chapter, draws our summary and conclusions. Possible future works will also be outlined.

Chapter 2

Literature Review

As mentioned in Section 1.3, health index (HI) development of planetary gearboxes based on condition monitoring data is the focus of this thesis. This chapter first provides a general review of reported approaches for HI modeling based on condition monitoring data. The approaches reviewed are not limited to the applications to planetary gearboxes. Then this chapter provides a detailed review of reported HI related approaches for planetary gearboxes. The pros and cons of the best reported approach that can be applied on planetary gearboxes will be identified. The detailed scope and objectives of this study will be provided at the end of this chapter based on the detailed literature review.

2.1 HI Modeling Based on Condition Monitoring Data

Health index (HI) is a term that can be divided into "health" and "index". The definition of "health" in this thesis is the state of an engineering asset which represents the ability of the asset to meet its designed and required functions [14]. The definition of "index" in this thesis is a number on a scale usually between 0 (the worst health) and 1 (the best health) that can stand for the health of an engineering asset [14]. The HI takes its highest value when an asset is brand new and it generally decreases monotonically as the asset is used and is deteriorating. Therefore, the HI in this thesis aims to reflect the comprehensive health condition of such engineering assets.

Many features have been utilized to track the health condition of a component or a system in use [15]. The features mentioned here are the features or statistical values extracted from raw sensory data. For example, the kurtosis, a common statistical measure, of a number of data points may be a feature. A few commonly used features will be introduced and defined later in Section 2.2. The features can reflect the health condition of the component or the health condition of the whole system from one specific aspect. A reliable feature is expected to vary along with the change of health conditions of the target component or the target system.

In this thesis, we treat the HI as a comprehensive index that combines all the available health indicators together to reflect comprehensive health condition of an engineering asset. Theoretically, the value of an HI is expected to decrease from the perfect value of one eventually to zero throughout the life cycle of an engineering asset.

In order to develop a reliable HI for engineering assets, HI modeling is a crucial task. This section reviews reported approaches in HI modeling. The reported approaches can be divided into two categories: the weighted summation category and the linear regression category. The reported work on these two categories is summarized next.

Jahromi *et al.* [16] constructed an HI using a linearly weighted summation model with fixed weights where different features have different weights. The features are generated by sensory data of power transformers (PTs). In [16], total 24 types of data are utilized to measure the health condition of such PTs. 24 types of data (such as dissolved gas analysis (DGA), load history, power factor, etc.) were used in [16]. One Health Index Factor (HIF) is used to quantitatively measure the health condition of the PT based on one of the above-

mentioned data types. Each HIF can take integer values of 4, 3, 2, 1, and 0 corresponding to five condition levels of a type of measurement. Thresholds in each HIF are determined by expert experience and recommended technical standards. Then, HI is formulated as the weighted summation of all the HIFs. Equation (2.1) shows the final linear weighted summation equation [16].

$$HI = \frac{100\% \times \sum_{j=1}^{n} K_j HIF_j}{\sum_{j=1}^{n} 4K_j} \tag{2.1}$$

where K_j and HIF_j are weights and integer value of health condition level mentioned above, respectively. The formulated HI then can reflects the comprehensive health condition of the PT.

Chen *et al.* [17] used the same categorized approach to develop an HI by a weighted summation model for critical assets (e.g. PTs, Wind Turbine Gearboxes (WTGs), Boiler Feed Pumps (BFPs), and Heat Exchangers (HEs)) in power generation plants. The differences in [17] to the approach used in [16] are as follows:

1) Chen *et al.* [17] added dynamic weights in the HI linear combination equation. The dynamic weights used to guarantee that once an exact failure occurs in the asset, the HI should be sensitive to reflect the occurred failure. They defined the HI to a range in [10,0], wherein 10 represents the best condition (i.e., newly installed) and 0 represents the worst. They categorized the HI into four regions, Good ($HI \in [10,8]$), Monitor ($HI \in [7.99,6]$), Action Required ($HI \in [5.99,3]$), and Poor ($HI \in [2.99,0]$). The dynamic weights in the approach of Chen *et al.* can

satisfy the following two principles: (1) when an HIF equals to zero, the HI must be in Poor region; (2) when an HIF equals to one, the HI must be in the Action Required region [17].

2) Unlike Jahromi *et al.* [16] took integer values of 4, 3, 2, 1, and 0 corresponding to five condition levels of a type of measurement for HIFs, Chen *et al.* [17] took the integer values of 3, 2, 1, and 0 corresponding to four condition levels of a type of measurement for HIFs.

3) Jahromi *et al.* [16] used an HI which belongs to the range of [1,0]. Instead, Chen *et al.* [17] used $HI \in [10,0]$ as required by the collaborating company.

4) For HI development of PTs, Chen *et al.* selected 5 types of data with high priority based on the recommendations of the collaborating company's engineers. These 5 types of data are DGA, fluid analysis (also called oil quality), power factor, insulation resistance, and winding resistance.

We can see details of the modified weighted HI model proposed by Chen *et al.* [17] in Equation (2.2):

$$HI = \frac{10 \times \sum_{i=1}^{n} K_{di}(HIF_i) K_{ri} HIF_i}{3 \times \sum_{i=1}^{n} K_{di}(HIF_i) K_{ri}} \tag{2.2}$$

where $K_{di}(HIF_i)$ is the dynamic weight for the i^{th} *HIF*, K_{ri} is the relative weight for the i^{th} *HIF*, and n is the number of *HIF*s. The dynamic weight K_{di} varies when HIF_i value varies. However, the relative weight K_{ri} remains fixed for each HIF_i no matter what values HIF_i

takes. The relative weights are used to measure the importance of each HIF among all HIFs. These K_r weights are determined by subject-matter experts in the collaborating company [17]. Chen *et al.* [17] also determined the HIFs by expert experience and recommended technical standards.

Chen *et al.* [17] did some works beyond PTs. They also developed HIs for WTGs, BFPs, and HEs. The general HI models for WTGs, BFPs, and HEs are same as the HI model of PTs as shown in Equation (2.2). While, it is worth noting that the types of data selected to generate HIFs for WTGs, BFPs, and HEs are different. In WTG HI model, Chen *et al.* [17] considered two types of data (vibration signals and oil debris). They extracted two features (FM4 and NA4) from vibration signals, and they used one feature called oil debris mass from oil debris data. In addition, Chen *et al.* [17] used three features (vibration analysis, efficiency analysis, and oil analysis) in their BFP HI model. Furthermore, Chen *et al.* [17] employed the differential pressure of fluid of HE and the effectiveness of HE as two features in their HE HI model.

In summary, we can see that the first category of methods uses the weighted summation model to develop the HI. Multiple sub-HIs may be obtained by multiple types of condition monitoring data. Weights were given directly to each sub-HI based on expert knowledge. The comprehensive HI is integrated by these sub-HIs using the weighted summation model.

Using the second category of linear regression, Wang *et al.* [18] proposed a comprehensive heath indicator by fusing multi-dimensional features that were extracted from raw sensory data. Wang *et al.* [18] set the comprehensive health indicator in the range of [1,0]. The model they used is a linear regression model as shown in Equation (2.3).

$$y = \alpha + \boldsymbol{\beta}^T \mathbf{x} + \varepsilon = \alpha + \sum_{i=1}^{N} \beta_i x_i + \varepsilon \tag{2.3}$$

where $\mathbf{x} = (x_1, x_2, ..., x_N)$ is the N dimensional feature vector, y is a scalar denoting the comprehensive health indicator, $(\alpha, \boldsymbol{\beta}) = (\alpha, \beta_1, \beta_2, ..., \beta_N)$ is $N+1$ model parameters, and ε is the noise term, a scalar.

In addition, Wang *et al*. [18] did some works in sensor selection. They used 21 sensors simultaneously and conducted physical experiments. They selected data from two subsets of sensors, one subset has 3 sensors and the other subset has 7 sensors. The subset with 7 sensors is picked since it has a better HI prediction accuracy. However, they ignored other possible combinations of sensors which may result in even better HI prediction accuracy.

Riad *et al*. [19] used the same linear regression model as shown in Equation (2.3) to construct the HI of turbofan engines. The dataset they used is the turbofan engine degradation simulation dataset provided by NASA. They utilized an HI in the range of [1,0] (1 means completely heathy and 0 means completely failed). Unlike Wang *et al*. [18], Riad *et al*. [19] chose the smoothed features extracted by raw sensor readings as their HI model input. Simple moving average was used to smooth the extracted features. In addition, Riad *et al*. [19] also smoothed the formulated HI through a third-order polynomial curve fitting. The smoothed HI was utilized as the input to a multi-layer perceptron ANN to estimate the health condition. It is worth noting that total 21 sensors are available in their dataset, and they chose 14 sensors based on their intuitive degradation trend. But Riad *et al*. did not consider any optimal combinations of sensors in their HI model, and they did not minimize the dimensions of their model input.

Yang *et al.* [13] developed an HI for electric motors based on features extracted from raw condition monitoring data. They treated HI development as a regression problem and used an FFNN to find the relationship between the extracted features and HIs. Chapter 3 will introduce details about the FFNN. The FFNN-based HI for this specific type of motor is trained using assumed true HI and condition monitoring data collected from Run-To-Failure (RTF) experiments on 10 identical motors. They started 10 RTF experiments for these 10 identical motors simultaneously, but the failure time for these 10 motors are different. An HI is then modeled for each of these 10 motors based on its own RTF experimental dataset. For each motor, the inputs of the FFNN-based HI are the features extracted from 11 channels of condition monitoring data collected from this motor, whereas the output is the HI for this motor. The averaged HI of these 10 HIs is treated as the HI for this type of motor, and the Remaining Useful Life (RUL) for this type of motor is then determined. The assumed true HI is in the range of [1,0] which covers the whole degradation process and has a linear degradation trend with respect to time (i.e. if a motor has a life cycle of T hours, then its HI = 1 at its age = 0, HI = 0 at its age = T, and HI = 1-t/T at its age = t). The performance is measured by root mean squared error (RMSE) between the developed HI and the true HI. Chapter 3 will illustrate details about RMSE. Yang *et al.* [13] did some works in feature selection, they tried top 5, top 10, all the way to top 50 features as inputs. The reported HI modeling approach in [13] has a great flexibility in either choosing the modeling algorithm for HI prediction or in designing their corresponding smoothing strategies based on the assumption of HI degradation trend.

In summary, the second category of methods uses the linear regression model to develop HI. Using the linear regression model can address the uncertainty of the weights

determined by ambiguous expert knowledge and has a great flexibility in adding or deleting candidate inputs.

Compared with the first category of methods, the second category of methods is more applicable to planetary gearboxes. To be more precise, the method for motors developed in [13] can be applied to planetary gearboxes due to its great flexibility in either choosing the modeling algorithm for HI prediction or in designing their corresponding smoothing strategies based on the assumption of HI degradation trend. More detailed comments on [13] will be provided in the following section.

2.2 HI Related Approaches for Planetary Gearboxes

In order to improve availability of planetary gearboxes, multiple condition monitoring data such as vibration signal, ultrasonic data, and thermographic data are often collected. Among all these condition monitoring data types, the vibration signal is the most commonly used for planetary gearboxes. The vibration signal usually changes when a fault occurs in a planetary gearbox. Features can reflect such changes or faults in a planetary gearbox. In addition, the most widely used features for planetary gearboxes are the statistical features extracted from vibration signals [20].

In 1962, Kenney *et al.* [21] proposed a feature called root mean square (RMS), defined in Equation (2.4). The RMS may be the most commonly used feature in vibration monitoring and trending the overall condition of machines over time. It is defined as the square root of the average of the squares of the measurements in the data series.

$$RMS = \sqrt{\frac{\sum_{i=1}^{N} x_i^2}{N}} \tag{2.4}$$

where N is the number of data points in the series and x_i is the i^{th} measurement (i=1,2,…N). RMS has been used to represent the power of the data series.

Samuel *et al.* [22] proposed a feature called kurtosis, expressed in Equation (2.5). Kurtosis is actually the ratio between the fourth central moment and the squared value of the second central moment of the data series. It provides a measure of the peakedness of the data series.

$$Kurtosis = \frac{N\sum_{i=1}^{N}(x_i - \bar{x})^4}{\left[\sum_{i=1}^{N}(x_i - \bar{x})^2\right]^2} \tag{2.5}$$

where x_i is the i^{th} value in the sample and $\bar{x}$ is the sample mean.

The two features above are widely used in condition monitoring of planetary gearboxes. Beyond these two features, there are a large number of other features reported in the existing literature. Hoseini *et al.* [11] summarized 213 features for planetary gearboxes. These 213 features are comprehensive and exhaustive. Hoseini *et al.* [11] divided these 213 features into 3 categories including: 1) general system features, 2) gearbox specific features, and 3) frequency domain indicators.

In the category of general system features, there are twenty three features that are not only limited to planetary gearboxes, but also useful for generic systems. For example, RMS and kurtosis as mentioned in Equation (2.4) and Equation (2.5), respectively.

In the category of gearbox specific features, there are twelve features that have been developed specifically for condition monitoring of gearboxes. These type of features are defined based on the residual signal r or the difference signal d. The residual signal is obtained as a vibration signal filtering out the shaft frequency, the Gear Meshing Frequency (GMF), their harmonics, and the first order sidebands [17]. The GMF of a gear can be simply defined as times gear teeth are meshing with each other and is calculated by the number of teeth of a gear multiplied by the rotational speed of this gear [11]. A harmonic is a wave with a frequency that is a positive integer multiple of the frequency of the original wave [23]. The shaft frequency is the shaft speed (rpm) divided by 60. A sideband is a band of frequencies higher than or lower than the frequency that is the result of the modulation process [24]. Modulation is the process of varying one or more properties of a periodic waveform, with a modulating signal that typically contains information to be transmitted [25]. A difference signal is defined as a vibration signal filtering out the first order sidebands about the GMFs [11]. Two examples of the gearbox specific features are FM4 and NA4, defined in Equation (2.6) and Equation (2.7) [26], respectively.

$$FM4 = \frac{N\sum_{i=1}^{N}(d_i - \bar{d})^4}{\left[\sum_{i=1}^{N}(d_i - \bar{d})^2\right]^2} \tag{2.6}$$

where d_i is the i^{th} value of the difference signal in the sample and $\bar{d}$ is the mean value of the difference signal.

$$NA4 = \frac{\frac{1}{N}\sum_{i=1}^{N}(r_i - \bar{r})^4}{\left\{\frac{1}{M}\sum_{j=1}^{M}\left[\frac{1}{N}\sum_{i=1}^{N}(r_{ij} - \bar{r}_j)^2\right]\right\}^2} \tag{2.7}$$

where r_i is the i^{th} value of the residual signal in the sample, $\bar{r}$ is the mean value of the residual signal, and M is the number of previous data readings.

In the category of frequency domain features, the well-known Fourier Transform (FT) must be performed on the time domain vibration data collected by sensors [27]. FT decomposes a time domain signal into a summation of sinusoidal waveforms at specific frequencies with specific amplitudes [27]. FT converts the signal from time domain to frequency domain. A vibration spectrum is the frequency domain representation of a signal [28]. The vibration spectrum of a healthy gearbox usually has a dominant component at the GMF. When a localized gear tooth fault occurs, say a tooth crack is present, the amplitudes of the sidebands around the GMF and the GMF's harmonics increase [11]. In the category of frequency domain features, there are 178 features that are extracted based on the GMF, its harmonics, and the sidebands around them. For example, sideband index as defined in Equation (2.8) [11]:

$$Sideband\ Index = \frac{1}{2}\left[Amp(sb1) + Amp(sb2)\right] \tag{2.8}$$

where *Amp* is the amplitude, *sb1* is the dominant sideband about GMF of the 1st stage planetary gearbox, *sb2* is the dominant sideband about GMF of the 2nd stage planetary gearbox. Details about the 1st stage planetary gearbox and the 2nd stage planetary gearbox will be introduced in Chapter 4.

From the above, we can see that a large number of features are available. We do not expect all these features have the same capability in revealing the fault degradation for a specific planetary gearbox. These available features will be selected via feature selection process in HI modeling.

In [13], Yang *et al.* modeled an HI for electric motors based on their extracted statistical features. They firstly selected top 50 features as the candidate dimensions of the input with the best ability in revealing degradation trend from all their extracted features. They did some works in the selection of features, but they only used fixed stepsize following top-down ranking in their approach. They tried top 5, top 10, all the way to top 50 features with a fixed stepsize of 5 as inputs in their HI modelling. However, other possible combinations of features may perform better compared to their proposed fixed stepsize feature selection. Potential application and improvement of the approach [13] will be investigated in this thesis research.

2.3 Summary

This chapter reviews HI development based on condition monitoring data. Section 2.1 reviewed and categorized the existing approaches for HI modeling using condition monitoring data. Section 2.2 introduced the HI related works for planetary gearboxes.

This thesis aims to develop an HI using sensory vibration signals collected from the planetary gearboxes. Yang *et al.* [13] reported an HI modeling method which has great potential application on planetary gearboxes, for motors. In their HI modeling, an FFNN is used. Detailed fundamental knowledge about the FFNN will be illustrated in Chapter 3. In the HI modeling, Yang *et al.* tried top 5, top 10, all the way to top 50 features as inputs.

However, they ignored other possible feature combinations. In addition, Yang *et al.* [13] reported a smoothing algorithm called HI dynamic smoothing procedure to smooth the modeled HI. This thesis will introduce and discuss the reported HI dynamic smoothing procedure in Chapter 5. In their reported HI dynamic smoothing procedure, Yang *et al.* used a fixed window size of 5 and a fixed max drop parameter of 0.1. However, they ignored other possible combinations of parameters that may produce an HI with higher accuracy.

To address the shortcomings in [13], this thesis proposes a feature selection strategy using Genetic Algorithm (GA) [29] to select the best feature combination that produces an HI with the highest accuracy. Detailed fundamental knowledge about GA will be introduced in Chapter 3. In addition, proper smoothing procedure needs to be applied on the developed HI in order to decrease fluctuations of the developed HI. This thesis will improve the reported HI dynamic smoothing procedure [13] in order to best fit an HI with the lowest fluctuation and the highest accuracy. The developed HI is then used to reflect the comprehensive health condition of the target planetary gearbox. More details about HI modeling and smoothing will be discussed in Chapter 5.

Chapter 3

Fundamental Knowledge

As mentioned in Chapter 2, HI development of planetary gearboxes based on condition monitoring data is the focus of this thesis. We will develop an HI for planetary gearboxes based on the reported work of Yang *et al.* in [13]. Yang *et al.* developed an HI for electric motors using a Feed-Forward Neural Network (FFNN). FFNN is one type of Artificial Neural Networks (ANNs). An ANN is a supervised learning algorithm that can learn and model the relationship between the input and the output. Besides, this thesis will also use the FFNN to develop the HI for planetary gearboxes, so we will introduce both ANN and FFNN in this chapter. In addition, we do not expect all the inputs are able to reveal monotonic degradation trend that producing the single outputted monotonic HI. The feature selection process is important, since it selects capable inputs in revealing monotonic HI. In [13], Yang *et al.* used top-ranked 5, top-ranked 10, top-ranked 15, all the way to top-ranked 50 features with a fixed stepsize of 5 in their feature selection process. However, they ignored other possible combinations of features. This thesis will use a more advanced method called Genetic Algorithm (GA) in selecting the optimal combination of features as the FFNN-based HI model input. A GA is a widely used optimization algorithm based on bio-inspired operators [30]. We will introduce the GA in detail in Section 3.3. Based on above-mentioned illustrations, we conclude that two crucial issues about our HI modeling that need to be considered are as follows:

1) The algorithm how to model the HI using condition monitoring data and,

2) How to optimize the dimension or combination of the input of the FFNN-based HI model.

As we mentioned in 1), we will use an ANN to model the HI. In 2), we will use a GA to optimize the dimension or combination of the input of the FFNN-based HI model. Besides, the way we model the HI is a regression. Therefore, the fundamental knowledge on regression, ANN, and GA is needed. This chapter aims to provide such fundamental knowledge for ease of reference in later chapters.

Section 3.1 introduces the regression problem in HI modeling. Section 3.2 introduces the basics of the ANN. Section 3.3 introduces the fundamental knowledge of the GA. Section 3.4 gives summary of this chapter.

3.1 Regression Problems

The descriptions of regression models in this section are based on [31]. Suppose we have n data points. Each data point consists of an input vector $\mathbf{x}_t \in R^n$, t is the time index, an output scalar $y_t \in R^1$. The relationship f is what we want to map from the input to the output, which can be expressed in Equation (3.1):

$$y_t = f(\mathbf{x}_t) \tag{3.1}$$

Generally, regression analysis is a set of statistical analyses for evaluating the relationships among different variables. It includes techniques for modeling and analyzing variables, when the focus is on the relationship between a dependent variable and one or more independent variables [31].

Many techniques for executing regression analysis have been developed, and can be divided into two categories [31]. First category of such techniques for regression analysis is parametric, in that the regression function is defined in terms of a finite number of unknown parameters that are obtained from the data analysis. Second category of the techniques is nonparametric, which refers to techniques that allow the regression function to lie in a specified set of functions with infinite dimensions [31]. In this thesis, we use the nonparametric type technique called ANN for regression analysis in HI modeling for planetary gearboxes. While, for ease of understanding regression, we introduce the regression model using a parametric type regression called linear regression in this section.

Regression models which involve the following parameters and variables: 1) The unknown parameters $\boldsymbol{\beta}$, which is a vector of scalars. 2) The independent variables, $\mathbf{X}$, which is a vector. 3) The dependent variable, Y, which is often a scalar. A regression model relates Y to a function of $\mathbf{X}$ and $\boldsymbol{\beta}$ can be shown in Equation (3.2):

$$Y \approx f(\mathbf{X}, \boldsymbol{\beta}) \tag{3.2}$$

We can see from Equation (3.2) that the regression analysis is to find the unknown parameters $\boldsymbol{\beta}$ that will minimize the difference between the measured and predicted values of the dependent variable Y.

We use linear regression models as examples to show how the regression model is to be determined. In a 1-dimensional linear regression model, the specification of this type of model is that the dependent variable y is a linear combination of a single dependent variable x. Equation (3.3) shows an example of such a simple linear regression for modeling n data

points with one dependent variable y, independent variable x, and two parameters β_0 and β_1:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i,\ i = 1,\ldots,n \tag{3.3}$$

In addition, we can represent a multiple linear regression by adding another variable z as follows:

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 z_i + \varepsilon_i,\ i = 1,\ldots,n \tag{3.4}$$

The regression Equation (3.4) is still a linear regression because the powers of the independent variable on the right hand side is 1.

Returning our attention to the simple linear regression model with only one independent variable, we can also represent the simple linear regression model as follows:

$$y_i^* = \beta_0^* + \beta_1^* x_i \tag{3.5}$$

where y_i is the i^{th} measurement of the dependent variable y when the independent variable x takes the i^{th} value x_i, y_i^* is the predicted value of the dependent variable y at the specified independent variable value x_i.

The error $e_i = y_i - y_i^*$ is the difference between the predicted value y_i^* and the true value y_i. Theoretically, if the difference e_i is lower, the result of the regression is better. The performance measure called the root mean squared error (RMSE) is often used to evaluate the accuracy of regression modeling. Equation (3.6) shows the RMSE.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i}^{n}(y_i - y_i^*)^2} \tag{3.6}$$

where y_i^* is the model predicted value, y_i is the real value.

In this thesis, RMSE is employed as the performance measure in HI development. Given the predicted HI (HI_t) by our HI model and the assumed real HI (HI_t^*) at the time index t, the RMSE in this thesis following Equation (3.6) above is expressed as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{t}^{n}(HI_t - HI_t^*)^2} \tag{3.7}$$

The assumed real HI in this thesis is the same assumed real HI in [13]. Details about the assumed real HI will be introduced in Chapter 5.

The above discussions on regression are parametric. Besides, the regression models introduced above are mathematical expressions. A parameter vector **β**, a vector **X**, and a scalar Y are used in these mathematical expressions. The scalar Y is a linear combination of the vector **X** by using the parameter vector **β**.

The ANN as mentioned earlier in this chapter uses a supervised learning algorithm that can learn and model complex relationship between the input and the output. In addition, an ANN is usually considered as a nonparametric regression [32]. The ANN-based method in HI modeling is used in the best reported HI work [13]. We will also use the ANN in our HI modeling for planetary gearboxes. Therefore, the basics of ANN will be introduced in the following section.

3.2 Basics of ANN

3.2.1 Structure of ANN

An ANN consists of interconnected neurons. Each neuron stands for a mapping, usually with multi-inputs and a single output. Figure 3.1 shows the structure of a single neuron. The function f at the output side of a specific neuron is called the activation function. The activation function of a neuron defines the output of that neuron given inputs [29]. The activation may be in a few forms, for example, the logistic function shown in Equation (3.8) and the TanH function shown in Equation (3.9) [11].

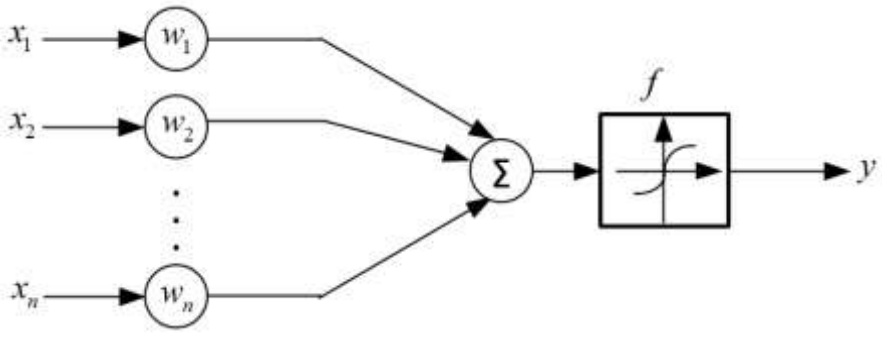

Figure 3.1 A single neuron [29]

$$f = \frac{1}{1+e^{-\mathbf{x}}} \tag{3.8}$$

$$f = \frac{(e^{\mathbf{x}} - e^{-\mathbf{x}})}{(e^{\mathbf{x}} + e^{-\mathbf{x}})} \tag{3.9}$$

where e is a constant approximately equal to 2.7183, $\mathbf{x} = [x_1, x_2, x_3, x_4, ...x_n]^{\mathrm{T}} \in R^n$ is the input vector, and f is the output value of the activation function.

If the activation function is the identity activation function (also called linear activation function), then the output of the single neuron in Figure 3.1 will be:

$$y = \sum_{i=1}^{n} w_i x_i = \mathbf{x}^{\mathrm{T}} \mathbf{w} \tag{3.10}$$

where the input vector $\mathbf{x} = [x_1, x_2, x_3, x_4, ...x_n]^{\mathrm{T}} \in R^n$, its corresponding weight vector $\mathbf{w} = [w_1, w_2, w_3, w_4, ...w_n]^{\mathrm{T}} \in R^n$, and the single output y, namely a scalar.

An FFNN is one type of ANNs. This thesis uses an FFNN to model the HI for planetary gearboxes. In a typical FFNN, the neurons are connected with each other in different layers. The data of the FFNN transmit from one layer to its succeeding layer in a specific direction. Therefore, the neuron in the succeeding layer receives data only from the neurons in the previous layer. The first layer is called the input layer, the last layer is called the output layer, and the layers between the input layer and output layer are called hidden layers and there may be multiple hidden layers. Figure 3.2 shows a structure of an FFNN having only 1 hidden layer [29]. There are *n* inputs, one output, and *l* neurons in the input layer, the output layer, and the hidden layer, respectively.

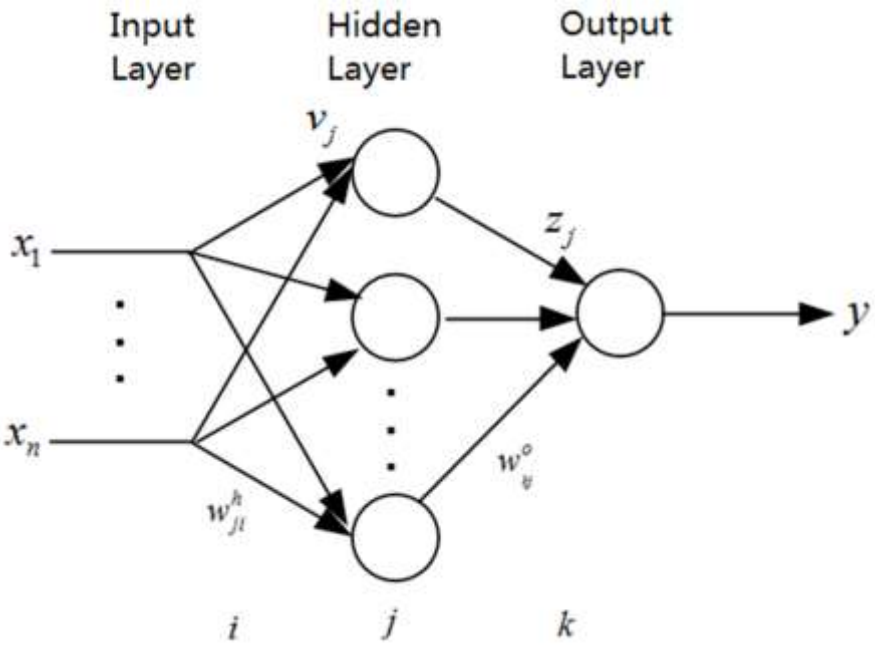

Figure 3.2 An FFNN having three layers [29]

The input vector $\mathbf{x}=[x_1,x_2,x_3,x_4,...x_n]$, input to the hidden layer: v_j $(j=1,2,3,4,...,l)$, the single output y, hidden layer to output: z_j $(j=1,2,3,4,...,l)$, the interconnected weights from the input layer to the hidden layer: w^h_{ji} $(j=1,2,3,4,...,l;i=1,2,3,4,...,n)$, the interconnected weights from the hidden layer to the output layer: w^o_{kj} $(j=1,2,3,4,...,l;k=1)$, the activation functions: f^h_j $(j=1,2,3,4,...,l)$, and f^o. The following equations show how the neural network calculates its output using the given input, the weights, and the activation functions.

$$v_j=\sum_{i=1}^{n} w^h_{ji}x_i \tag{3.11}$$

$$z_j=f^h_j(v_j) \tag{3.12}$$

$$y = f^o(\sum_{j=1}^{l} w_j^o z_j) \tag{3.13}$$

$$y = f^o(\sum_{j=1}^{l} w_j^o z_j) = f^o(\sum_{j=1}^{l} w_j^o f_j^h(v_j)) = f^o(\sum_{j=1}^{l} w_j^o f_j^h(\sum_{i=1}^{n} w_{ji}^h x_i)) \tag{3.14}$$

where y is the single output, namely, a scalar. Firstly, we consider a single training data point $(\mathbf{x}_d, y_d)$, $\mathbf{x}_d \in R^n$, and $y_d \in R^1$. The weights w_{ji}^h $(j = 1,2,3,4,...,l; i = 1,2,3,4,...,n)$ and w_{kj}^o $(j = 1,2,3,4,...,l; k = 1)$ need to be determined. Thus, we have the objective function to be minimized as follows:

$$\text{Minimize} \quad E(w, \mathbf{x}_d) = \frac{1}{2}(y_d - y)^2 \tag{3.15}$$

where y is the output calculated using the input data $\mathbf{x}_d \in R^n$ and the weights will be optimized, y_d is a scalar that describes the real output values.

The weights of neurons in the FFNN can be adjusted and optimized by a training algorithm in order to best fit the FFNN model. The training algorithm will be introduced in the following section.

3.2.2 The Training Algorithm

The Adaline which denotes adaptive linear element is the single neuron and its corresponding training algorithm [29]. Figure 3.3 which is based on Figure 3.1 shows the Adaline. The e is the error between the real y_d and predicted y mentioned in Equation (3.15).

As to the training of a whole network including many neurons, the Backpropagation Algorithm (BA) is a widely used training algorithm. The BA is gradient-based optimization algorithm that exploits the chain rule [33]. The chain rule is a formula for computing the derivative of the composition of two or more functions [33]. We use the BA to minimize the e, and the BA will be introduced in this section.

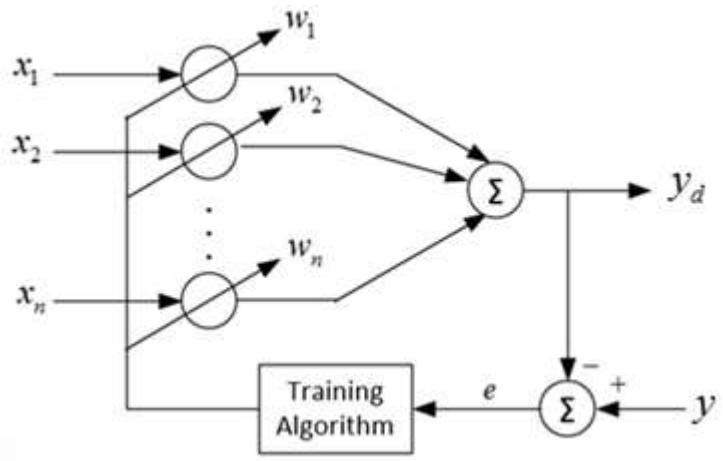

Figure 3.3 Adaline [29]

In this section, we briefly introduce a fixed stepsize gradient method to illustrate how a BA works in an FFNN. Suppose we have a function $f(\mathbf{x})$ to be minimized, this function could be our error function shown in Equation (3.15) for the whole network. The gradient vector $\nabla f(\mathbf{x})$ at point $\mathbf{x}$ stands for the direction of maximum rate in maximizing the $f(\mathbf{x})$ [29]. The direction $-\nabla f(\mathbf{x})$ at point $\mathbf{x}$ then stands for the direction of maximum rate in minimizing the $f(\mathbf{x})$. Then we can represent this fixed stepsize gradient algorithm as the iterative equation in Equation (3.16). As we repeat the iterative process step by step as Equation (3.16) shows, we will get a minimized value of $f(\mathbf{x})$ in the end.

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \alpha_k \nabla f(\mathbf{x}^{(k)}),\ k = 0, 1, 2, 3, \ldots \tag{3.16}$$

where α_k is called the fixed stepsize [29]. It is also called the learning rate in neural network training terminology.

We can also apply the iterative process mentioned in Equation (3.16) to optimize the weights of neurons in an FFNN. The weights of neurons in the FFNN are optimized using BA. We can directly choose an FFNN model with BA in the Machine Learning and Deep Learning Toolbox of Matlab.

This thesis uses an FFNN in HI modeling. The input of our FFNN-based HI model are the extracted features mentioned in Chapter 2, and the single output is the modeled HI. Besides, the error between the modeled HI and the assumed true HI to be minimized by updating the weights of neurons in the FFNN. The objective is to minimize the RMSE between the predicted HI and the assumed true HI shown in Equation (3.17) which is based on Equation (3.7):

$$\text{Minimize} \quad RMSE(\mathbf{x}) = \sqrt{\frac{1}{n}\sum_{t}^{n}(HI_t(\mathbf{x}) - HI_t^*)^2} \tag{3.17}$$

where the predicted HI (HI_t) by our HI model and the assumed true HI (HI_t^*) at the time index t, $\mathbf{x}$ is the input vector.

In the objective function shown in Equation (3.17), the gradient vector is ∇RMSE(x), the fixed stepsize (learning rate of the neural network training) is α_k, thus, the fixed stepsize gradient algorithm for our proposed FFNN-based HI model can be expressed in Equation (3.18):

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \alpha_k \nabla RMSE(\mathbf{x}^{(k)}),\ k = 0, 1, 2, 3, \ldots \tag{3.18}$$

3.3 GA for Integer Variable Optimizations

In this thesis, the optimization of the dimensions and the combinations in the input of our proposed FFNN-based HI model is an integer variable optimization problem. Thus, this section will introduce the integer variable optimization and the type of integer variable optimization algorithm called GA used in the selection of the input in our FFNN-based HI model.

An integer variable optimization is a mathematic optimization in which some or all of the variables are restricted to be integers [34]. The case we considered in our FFNN-based HI model is an integer variable optimization with all of the variables to be integers. Thus, a general mathematical expression of an integer variable optimization with all of the variables to be integers can be expressed in Equation (3.19):

$$\begin{aligned} &\text{Minimize} \quad f(\mathbf{x}) \\ &\text{Subject to} \quad \mathbf{x} \in Z^n \end{aligned} \tag{3.19}$$

where $f(\mathbf{x})$ is an arbitrary function, $\mathbf{x}$ is a vector with all of the dimensions to be integers.

We can also apply the integer variable optimization in our proposed FFNN-based HI model based on Equation (3.18) and Equation (3.19), thus, Equation (3.20) shows the integer variable optimization in the feature selection of the input in our proposed FFNN-based HI model:

$$\text{Minimize} \quad RMSE(\mathbf{x}) = \sqrt{\frac{1}{n}\sum_{t}^{n}\left[HI_t(\mathbf{x}) - HI_t^*\right]^2} \qquad (3.20)$$
$$\text{Subject to} \quad \mathbf{x} \in Z^{40}$$

where the predicted HI (HI_t) by our HI model and the assumed true HI (HI_t^*) at the time index t, $\mathbf{x}$ is a 40-dimension vector that contains the 40 candidate features selected by the preliminary feature selection. The preliminary feature selection will be introduced in Chapter 4. If a specific feature is selected, we will mark its corresponding position as "1" in the $\mathbf{x}$; otherwise, we will mark its corresponding position as "0" in the $\mathbf{x}$. In our model, the vector $\mathbf{x}$ actually contains binary integer variables only.

A GA is a typical method in solving integer variable optimizations. This thesis uses the GA to minimize the RMSE between the predicted HI and the assumed true HI by adjusting the dimensions and combinations of variables in the input of the FFNN-based HI model.

The root of the GA is in the principles of genetics. Figure 3.4 shows the general procedure to implement a GA [29]. In Figure 3.4, the P(0) called initial population and it is an initial set of individuals which contains the points in feasible region [29]. Each individual, an input vector is a solution to the optimization problem we want to solve. Furthermore, each dimension of this input vector is called a gene, it may be 1 or 0. All these genes are jointed into a string to form a solution, also called a chromosome in the GA terminology. M(k) is the mating pool which is formed by the P(k) using a randomized procedure [29]. Crossover and mutation are two operations that belongs to evolution procedure. In the evolution procedure, the GA mimic the evolution of animals and plants in nature [29]. The crossover is the one of the most significant phase in a GA [29]. For each pair of parents to be mated, a crossover point is chosen randomly within the chromosome. Mutation operation is to

form a new offspring with some of its genes can be changed with a low random probability, for example the mutation probability p=0.001 [29]. The mutation can also be implied that some of the bits in the bit string can be flipped. An example of the crossover and an example of the mutation are shown in Figure 3.5 and Figure 3.6, respectively.

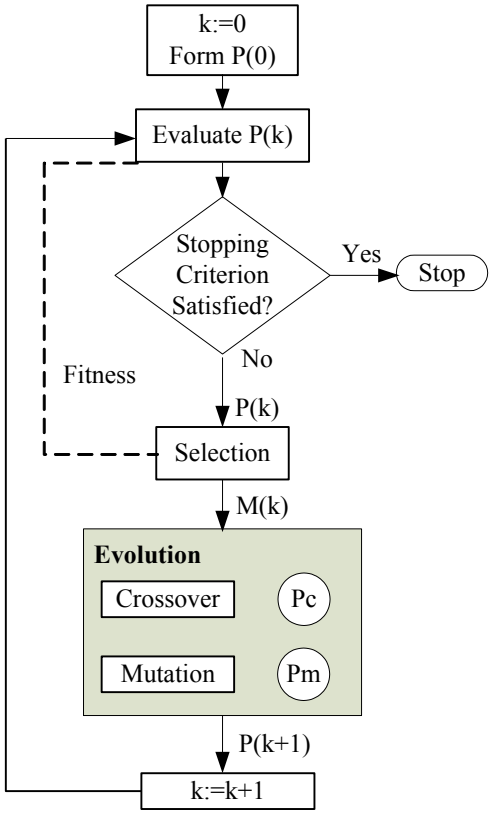

Figure 3.4 Flowchart for the GA [29]

The general procedure of the GA shown in Figure 3.4 is as follows: 1) Set k=0, and generate an initial population P(0). 2) Evaluate P(k). 3) If the stopping criterion is satisfied, then stop the procedure of the GA. 4) If the stopping criterion is not satisfied, then select M(k) from P(k). 5) Evolve M(k) to form P(k+1). 6) Set k=k+1, then go to step 2). The details of the GA can be found in [29].

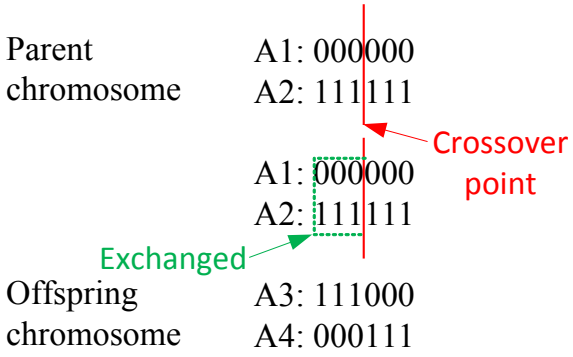

Figure 3.5 Crossover operation

Before mutation A1: 111000

After mutation A2: 110111

Figure 3.6 Mutation operation

This thesis proposes a GA-based feature selection method in determining optimized combination of features in the input of the FFNN-based HI model. In our proposed GA-based feature selection method, the decision variables are the extracted features, namely, the dimensions of the input of the FFNN-based HI model. We briefly illustrate the idea of the GA-based feature selection method in the following paragraph.

In our proposed GA-based feature selection method, we first generated an initial population which contains a set of individuals. An individual consists of multiple points ranging from 1-2^{40}. (We used a 40-bit binary string to represent any feature combination subset. Each bit corresponds to a feature. If a specific feature is selected, we marked its corresponding bit as "1" in our binary string; otherwise, we marked its corresponding bit as "0" in the binary

string). The 40-bit represents candidate 40 features with good capability in revealing degradation trend. Chapter 4 will introduce how we get the candidate 40 features in details. An HI will be modeled by the FFNN using the initial population. Subsequently, the RMSE is evaluated between the predicted HI and the assumed true HI. The objective function is to minimize the evaluated RMSE by adding or deleting the input features. Minimizing the RMSE is an iterative process that is implemented by the GA. We can directly use the GA function in the Matlab. We will introduce details of our proposed GA-based feature selection method in Chapter 5.

3.4 Summary

This chapter first introduced the regression problem since we treated the process modeling the HI as a regression problem. Subsequently, we illustrated the basics of the ANN, since this thesis uses a type of ANN called FFNN to formulate the HI. In addition, this chapter presented the GA for integer variable optimization, since this thesis will apply the GA in the selection of the input of our proposed FFNN-based HI model.

To addressing the shortcomings in [13] and to developing an HI for planetary gearboxes, this thesis uses a Run-To-Failure (RTF) data set of a lab planetary gearbox conducted by Reliability Research Lab (RRL) of University of Alberta (UofA). The following chapter, namely, Chapter 4 will introduce the details of the RTF data set.

Chapter 4

Data Description for HI Development for Planetary Gearboxes

This thesis focuses on HI development for planetary gearboxes, and the FFNN-based HI model is used in HI development for planetary gearboxes. The features extracted from the vibration signals collected from a lab planetary gearbox will be used as the input of our proposed FFNN-based HI model for planetary gearboxes. In addition, one target of this thesis is to find an optimal combination of the input features of the FFNN-based HI model with the highest HI prediction accuracy. In order to reach this goal, the details about the experiment data of the lab planetary gearbox are needed. The works of the physical experiment and data preprocessing that will be described in this chapter were completed by the Reliability Research Lab (RRL), University of Alberta (UofA) in 2010 [35].

Section 4.1 introduces the test rig. Section 4.2 introduces the Run-To-Failure (RTF) experiments conducted and the raw data collected. Section 4.3 introduces the feature calculation and the preliminary feature selection conducted earlier for degradation trend prediction. Finally the feature sets selected earlier in [36] that will be used as the input in Chapter 5 will be summarized.

4.1 Introduction of the Test Rig

The RTF experiment data collected in 2010 from a 2-stage lab planetary gearbox test rig as shown in Figure 4.1 are used in this thesis. This test rig consists of a 20 HP drive motor, a 1-stage bevel gearbox, a 2-stage planetary gearbox, two speed-up gearboxes, and a 40 HP load motor [35]. The 2nd-stage planetary gearbox was the focus of this experiment. All the gears in the 2-stage planetary gearbox are standard spur gears without tooth profile modification. In the test rig, six sensors including two low sensitivity vibration sensors, two high sensitivity vibration sensors, and two acoustic emission sensors were used to collect the vibration signals and the acoustic emissions [35]. A low sensitivity sensors, a high sensitivity sensors, and an acoustic emission sensors were mounted on the casing of the 1st stage of the planetary gearbox, and the other three sensors were mounted on the casing of the 2nd stage of the planetary gearbox. The locations of these sensors are shown in Figure 4.2. In addition to these 6 sensors, the lubrication system shown in Figure 4.1 was also able to collect metal scan data automatically.

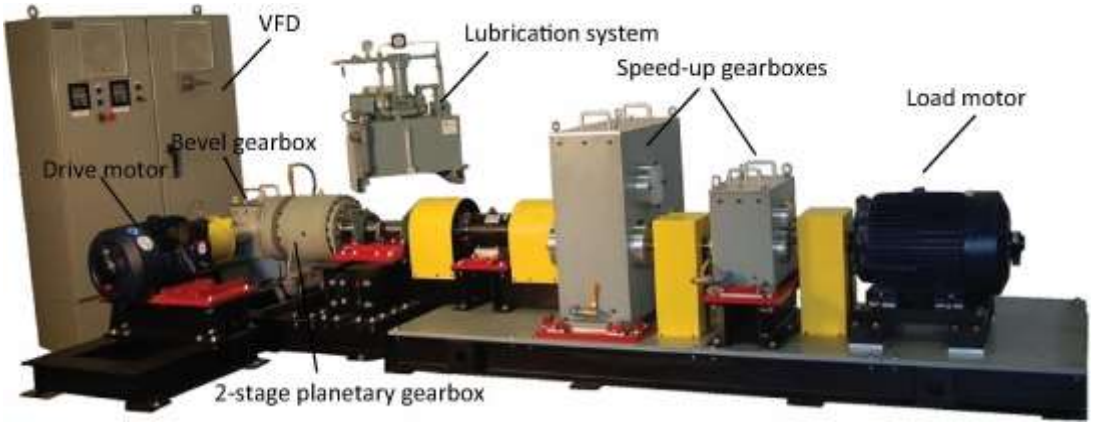

Figure 4.1 Planetary gearbox test rig [35]

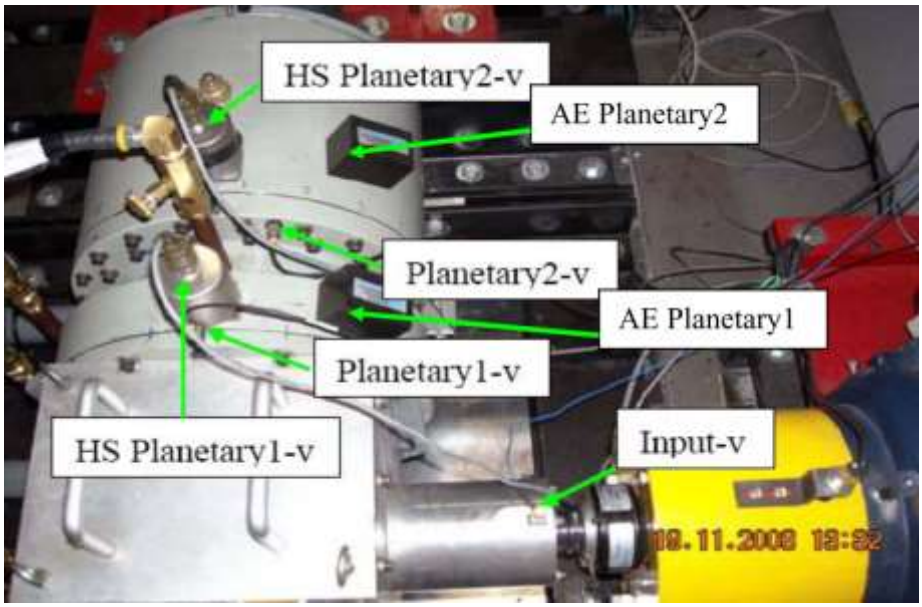

Figure 4.2 Sensor locations for the planetary gearbox test rig [35]

4.2 The RTF Experiments and Data Collection

The RTF experiments of the 2-stage lab planetary gearbox were conducted using the test rig shown in Figure 4.1. The gears on the 2nd stage of the planetary gearbox were allowed to naturally damaged in the test rig during these RTF experiments.

These RTF experiments consisted of 19 runs and lasted for 772 hours in total during the time from Sept. 11, 2009 to Aug. 18, 2010 as documented in [35]. The test rig was first ran at a driving motor rotating speed of 1200 rpm and a load of 10 k lb-in for 8 hours and then stopped for its first open inspection of the internal condition of the gearbox. The measurements taken during such open inspections will be given later in this section. This was the so-called run #1. At the end of the run #1, they did not see any damage on the gears on the 2nd stage of the planetary gearbox. They then decided to increase the load level to 19 k lb-in at the same driving motor rotating speed of 1200 rpm for another 8 hours. This

was the run #2. When the test rig was opened for inspection at the end of the run #2, there was still no visible damage to the gears on the 2nd stage of the planetary gearbox. Therefore, they increased the load level to 20 k lb-in during run #3 and run #4 and later to 25 k lb-in during run #5. At the end of run #5, they found some pits on the sun gear of the 2nd stage of the planetary gearbox. Thus, they decided to keep the load level at 25 k lb-in for all subsequent runs. Eventually, after 19 runs of the experiments, the gears in the 2nd stage gearbox were considered to have failed. A summary of these 19 runs of experiments is given in Table 4.1.

Table 4.1 Summary of the RTF experiment [35]

Run #	Driving motor speed (rpm)	Duration (Hour: Minutes)	Load (k lb-in)
1	1200	8:22	10
2	1200	8:40	19
3	1200	16:53	19
4	1200	32	20
5	1200	32	20
6	1200	34	20
7	1200	32	20
8	1200	32	25
9	1200	48	25
10	1200	48	25
11	1200	54	25
12	1200	48	25
13	1200	48	25
14	1200	48	25
15	1200	56	25
16	1200	56	25
17	1200	58	25
18	1200	51	25
19	1200	31	25

At each open inspection of the planetary gearbox, the gears of the 2nd stage of the planetary gearbox were taken out, cleaned, and checked for damage. The gears were weighted and

their photos were taken. Details about the weights and the photos of the gears can be found in [37]. Decisions to move on to the next run or not were up to the visual observations of the gear tooth health conditions. The last RTF experiment was stopped when the tooth mass of the sun gear of the 2nd stage of the planetary gearbox lost 50% of its mass [35]. In other words, the whole RTF experiment end up with 50% tooth mass loss of the sun gear on the 2nd stage of the planetary gearbox.

Descriptions of the RTF experiments and all data including vibration signals, acoustic emissions, photos, weights, and metal scan collected during all these RTF experiments were documented in [1]. This thesis treats all of these 19 runs of the RTF experiment as a single stream of data ignoring the open inspection breaks between each two neighbor runs, which means that we use a nonstop time-series to represent the whole RTF experiment. Besides, this thesis only uses the vibration signals to develop the HI for planetary gearboxes. Other data such as acoustic emissions, weight, and photo data will not be used.

For collection of the vibration data, a 5-minute time span of vibration signals were collected every two hours during these RTF experiments. The sampling frequency f_s used during the first 4 runs was 10 KHz. After 21 hours in run #4, the sampling frequency was changed from 10 KHz to 5 KHz. All vibration signals collected for run #5 to run #19 used the sampling frequency of 5 KHz. There are 1.5e6 data points (for f_s = 5KHz) or 3e6 data points (for f_s = 10KHz) in each 5-minute time span. Details about the raw data can be found in [35].

4.3 Feature Calculation and Preliminary Feature Selection

Based on the collected raw vibration signals mentioned in the last section, Hoseini *et al.* [11] calculated 852 (213 individual features times 4 vibration sensors) features. Each 5-minute time span was divided into 8 equal time segments, except one 5-minute time span from each of run #5, run #14, and run #16 which are divided into 7 segments. There are 3629 time segments in total for the whole RTF experiment. The 213 individual features calculated for each time segment of each sensor were categorized into 3 categories as mentioned in Section 2.2 including: 1) general system features (two examples including RMS and kurtosis were introduced in Section 2.2), 2) gearbox specific features (two examples including FM4 and NA4 were introduced in Section 2.2), and 3) frequency domain indicators (one example called sideband index was introduced in Section 2.2). Details of all these 213 features can be found in [11]. Each time segment contains 852 dimensions, namely, the calculated 852 features because there is data from 4 sensors for each time segment. In other words, RMS, kurtosis, FM4, NA4, sideband index, etc. of the vibration signals from all these four vibration sensors are calculated in each time segment. The calculated 852 features in the 3629 time segments cover the whole degradation process of the target planetary gearbox.

Since it was not expected for all these calculated 852 features for each time segment to have the same capability in revealing the degradation trend inside the gearbox, a proper preliminary feature selection process was carried out and documented in [2].

Zhao *et al.* [36] described the evaluation and ranking of these 852 features calculated from the RTF experiment data. They were interested in the features that were able to reflect the

monotonic degradation trend. They used a measure called monotonicity ratio (MR) to select the features that are sensitive to the degradation trend. The MR is in a range of [1,0]. The MR of a feature equal to 1 means that this feature perfectly reflects the monotonic degradation trend of this planetary gearbox. The MR of a feature equal to 0 means this feature is unable to reflect the monotonic degradation trend of this planetary gearbox at all. Features having higher MRs are more sensitive to the degradation trend. More details about the MR can be found in [36]. In this thesis, this process of selecting good features based on the measure of MR as used in [2] is called the preliminary feature selection.

Zhao *et al.* [36] selected 40 top-ranked features among all these 852 features mentioned earlier according to their corresponding MRs. Table 4.2 lists 4 of these 40 top-ranked features. This thesis uses these 40 top-ranked features as the input of our proposed FFNN-based HI model for planetary gearboxes.

Table 4.2 The 40 top-ranked features for RTF data [36]

Ranking No.	Feature Name	MR
6	LS2_kurtosis	0.5898
15	HS1_RMS	0.4825
17	LS1_max value	0.4800
37	HS2_FM4	0.3435

Here we briefly explain the 4 highlighted features listed in Table 4.2. LS1 and LS2 stand for the low sensitivity vibration sensor #1 and #2 as shown in Figure 4.2, respectively. HS1 and HS2 stand for the high sensitivity vibration sensor #1 and #2, respectively. Kurtosis (feature #6), RMS (feature #15), and FM4 (feature #37) were calculated from LS2, HS1,

and HS2, respectively. Their definitions are given in Equation (2.5), Equation (2.4), and Equation (2.6), respectively. Equation (4.1) shows a feature called max value [11]:

$$Max\ value = \max(x_i) \tag{4.1}$$

where x_i is the i^{th} value in the sample. In Table 4.2, max value (feature #17) was calculated from LS1. For definitions of other features included in the 40 top-ranked features, please refer to [36].

4.4 Summary

This chapter firstly introduced the RTF experiment test rig. Then the RTF experiments and data collection process were introduced. This chapter also illustrated the feature calculation and the preliminary selection as the feature calculation provides the input of the preliminary feature selection. Then the preliminary feature selection provides the input of our FFNN-based HI model to be described later in this thesis. The proposed FFNN-based HI model for planetary gearboxes using the candidate input, namely, the result of the preliminary selection will be introduced in Chapter 5.

Chapter 5

HI Development for Planetary Gearboxes

This chapter investigates the two aspects of Yang's method [13] which has vagueness and shortcomings. The RTF experiment data collected by former members of the RRL [35] will be used for the investigations. Section 5.1 introduces Yang's method and describes the two aspects to be investigated. Section 5.2 investigates the impact of the combinations of the input features in the FFNN-based HI model. Section 5.3 investigates the impact of the combinations of the window size parameter and the maxdrop parameter in the reported HI dynamic smoothing procedure in Yang's method. Section 5.4 shows comparisons between our improved method and Yang's method based on the results of data analyses. Then Section 5.5 summarizes the results of the investigations.

5.1 Introduction to Yang's Method

Yang *et al*. [13] developed an HI for electric motors. They used the RTF experiment data of a type of motor in their study. The information about the RTF experiment data of the motors will be described in this paragraph. They started 10 RTF experiments for 10 identical motors simultaneously, but the failure times for these 10 motors are different. They collected 11 channels of raw condition monitoring data including three-phase current data, three-phase voltage data, vibration signals, load data, speeds of the motor, temperature data, and acoustic emissions. Details of the RTF experiment data of the motors can be found in [13].

In the analysis of the data collected, Yang *et al.* [13] used the following assumption. For example, the RTF experiment for one of the 10 motors lasted 170 hours. The true HI was assumed to be decreasing linearly from the perfect value of 1 at time t = 0 to the lowest value of 0 at the time of 170 hours. That is, if a motor has a life of T hours, then its HI = 1 at its age = 0, HI = 0 at its age = T, and HI = 1-t/T at its age = t.

With the above assumption, Yang *et al.* [13] treated the HI development as a regression problem and used an FFNN to find the relationship between some extracted features and the HI. The FFNN-based HI for this specific type of motor is trained using the assumed true HI trend described above and the condition monitoring data collected from the RTF experiments on 10 identical motors as described in the earlier paragraphs. An HI was then modeled for each of these 10 motors based on its own RTF experimental dataset. For each motor, the inputs of the FFNN-based HI were the features extracted from the 11 channels of condition monitoring data collected from this motor as mentioned in the previous paragraph, whereas the output was the HI for this motor. The averaged HI of these 10 HIs was treated as the HI for this type of motor and the Remaining Useful Life (RUL) for this type of motor was then determined. The performance was measured by RMSE as defined in Equation (3.7) between the developed HI and the assumed true HI.

Figure 5.1 shows the framework of Yang's method. They used the features extracted from the 11 channels of the raw condition monitoring data as mentioned above and can be found in [13]. Then they extracted features such as RMS, kurtosis, etc. as we mentioned in Chapter 2. They did not mention how many features were extracted in total, but they mentioned that a term called Fisher's Ratio (FR) was used to preliminarily select the top-

ranked 50 features with the highest degradation reflection ability. The FR measures the degradation reflection ability of a feature X_j and is expressed in Equation (5.1) [13]:

$$FR(X_j) = \frac{(m_{j(1)} - m_{j(2)})^2}{\delta_{j(1)}^2 - \delta_{j(2)}^2} \tag{5.1}$$

where $m_{j(c)}$ and $\delta_{j(c)}^2$ are the mean and the variance of the feature X_j, respectively, within class c which is the health condition class, for c = 1, 2 (1 means healthy and 2 means faulty). The larger the FR is, the better the feature is able to reflect degradation.

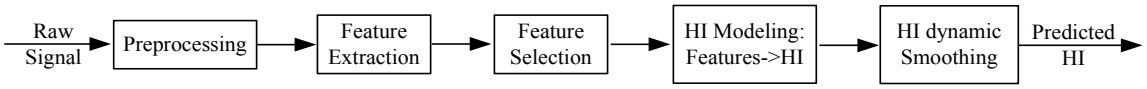

Figure 5.1 Procedure for HI development in Yang's method [13]

Yang *et al.* [13] did some work in feature selection. They tried the top-ranked 5, the top-ranked 10, all the way to the top-ranked 50 features following the FR ranking as the inputs of their reported HI model. However, they ignored other possible feature combinations. In addition, they used a fixed window size of 5 and a fixed maxdrop of 0.1 in their reported HI dynamic smoothing procedure. The maxdrop is a parameter in their HI dynamic smoothing procedure, the details about the maxdrop will be provided in Section 5.3. The HI dynamic smoothing procedure was used to smooth the modeled HI to get a smoothed HI with a lower RMSE, namely, a higher accuracy. The detailed HI dynamic smoothing procedure will be introduced in Section 5.3. However, they neglected other possible window sizes and maxdrops in their reported HI dynamic smoothing procedure.

To sum up, there are two important aspects with shortcomings that need to be addressed in Yang's method [13]: 1) the selection of the features of the FFNN-based HI model, 2) the selection of the two parameters (the window size parameter and the maxdrop parameter) in the reported HI dynamic smoothing procedure in Yang's method. The following sections will investigate these two aspects in detail.

To address the shortcomings in [13], this thesis proposes a GA-based method in selecting the best feature combination that models an HI with the lowest RMSE. With the modeled HI, this thesis proposes to use our improved HI dynamic smoothing procedure based on the reported HI dynamic smoothing procedure in [13] to further smooth the modeled HI in order to get an even lower RMSE. The final developed HI via HI dynamic smoothing is used to reflect the comprehensive health condition of the target planetary gearbox. Unlike Yang *et al.* used 10 identical motors to develop HIs and then determined the RUL using the averaged HI for this type of motor, we will not do any work in RUL prediction since we only have one planetary gearbox test rig. This thesis only focuses on the HI development for planetary gearboxes. The following sections will illustrate our proposed improvements over Yang's method in detail.

5.2 Selection of the Features for HI Modeling of Planetary Gearboxes

5.2.1 Using Yang's Method in HI Modeling for Planetary Gearboxes

Yang *et al.* [13] used an FFNN with three layers (10 neurons in the hidden layer and a single output layer neuron) using Machine Learning and Deep Learning toolbox in Matlab to develop the HI for the motors. In training the FFNN model, Yang *et al.* selected a BA in Matlab. The BA has been illustrated in Chapter 3. This thesis uses the same FFNN with the same structure and the same training algorithm as Yang *et al.* did in [13] to develop HI for planetary gearboxes. The data to be used in this thesis is not the motor data from Yang *et al*. but from the RTF experiments done by the RRL described in Chapter 4.

In order to select the best feature combination as the input that develop an HI with the lowest RMSE and the highest accuracy, Yang *et al.* used the top-ranked 5, the top-ranked 10, the top-ranked 15 all the way to the top-ranked 50 features with a fixed stepsize of 5 following the top-down ranking in terms of the FR in [13] with their motor data. In this thesis, using the gearbox RTF data, we will first use the same stepsize of 5 [13] in terms of the MR following the top-down ranking in input feature selection. The MR as mentioned in Chapter 4 also has the ability to measure whether a feature has a good degradation reflection. In this section, we use the top-ranked 5, the top-ranked 10, all the way to the top-ranked 40 features with a stepsize of 5 as the input of our FFNN-based HI model for planetary gearboxes.

The data we use in this thesis is the RTF experiment data conducted by the RRL in UofA [35] as introduced in Chapter 4. We use the features extracted from 3608 time segments of the RTF experiment dataset in our HI modeling. These 3608 time segments will be divided into 2 sub-datasets with 1804 time segments in each, we use the features of the first 4 time segments in every 8 time segments for training the FFNN and we use the features in the remaining 4 time segments for testing the FFNN. As mentioned in Chapter 4, every 8 time segments were collected in a 5-minute time span every 2 hours, thus, the health condition of the planetary gearbox in every 8 time segments should remain constant or change only slightly. The way we separate the RTF dataset into 2 sub-datasets is to artificially create 2 datasets in the FFNN modeling. We have done data analysis using the feature selection method with the fixed stepsize of 5 reported in [13] based on the top-ranked 40 features with the highest MR (we labeled them following the top-down MR ranking: [#1 ~ #40]). Table 5.1 lists the RMSE values of using the top-ranked 5, the top-ranked 10, the top-ranked 15, all the way to the top-ranked 40 features with the stepsize of 5. From this table, we can see that the best feature combination with a stepsize of 5 is the feature subset containing the top-ranked 25 features and its corresponding RMSE of 0.0261. Figure 5.2 a) shows graphically the results in Table 5.1. Figure 5.2 b) shows the trend of the developed HI using these selected top-ranked 25 features.

Table 5.1 Comparisons in the feature selection using the stepsize of 5

Feature subsets	RMSE values
Top-ranked 5	0.0947
Top-ranked 10	0.0454
Top-ranked 15	0.0344
Top-ranked 20	0.0344
Top-ranked 25	0.0261
Top-ranked 30	0.0277
Top-ranked 35	0.0368
Top-ranked 40	0.0343

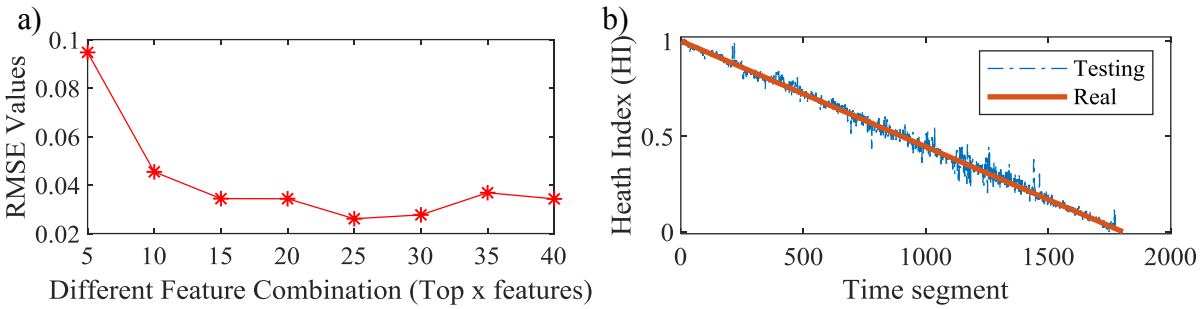

Figure 5.2 a) RMSE vs. top features (stepsize of 5); b) HI vs. time (using top 25 features as the input of the FFNN)

5.2.2 Selection of Features with a Finer Fixed Stepsize of 1

After repeating Yang's method using the fixed stepsize of 5 in the HI modeling for planetary gearboxes as described in the previous section, we take consideration that if we relax the restriction of stepsize of 5, a better feature combination for HI development with a lower RMSE may be obtained. Therefore, we have used a fixed stepsize of 1 instead of the reported fixed stepsize of 5 in this section. Based on this assumption, we have done data analysis using the feature selection method with the fixed stepsize of 1. Table 5.2 lists the RMSE values of using the top-ranked 1, the top-ranked 2, the top-ranked 3, all the way to the top-ranked 40 features with the stepsize of 1. From this table, we can see that the

best feature combination with the stepsize of 1 is the top-ranked 34 features and its corresponding RMSE is 0.0247, which is lower than that with the stepsize of 5 (with the RMSE=0.0261).

Figure 5.3 a) shows graphically the results listed in Table 5.2. Figure 5.3 b) shows the HI v.s. time using the selected feature subset that contains the top-ranked 34 features. We find that using the stepsize of 1 is better than using the reported stepsize of 5 in our HI modeling, which matches our assumption.

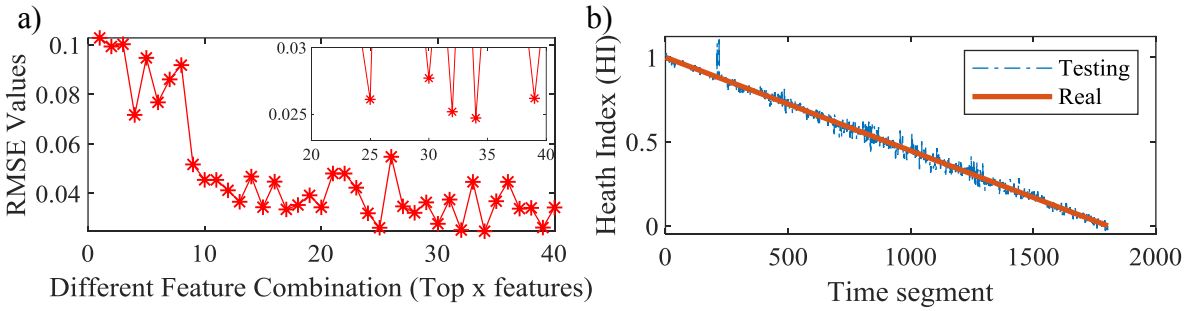

Figure 5.3 a) RMSE vs. top x features (stepsize of 1); b) HI vs. time (using the top-ranked 34 features as the input of the FFNN-based HI model)

Table 5.2 Comparisons in the feature selection using the stepsize of 1

Feature subsets	RMSE values
Top-ranked 1	0.1028
Top-ranked 2	0.0993
Top-ranked 3	0.1002
Top-ranked 4	0.0717
Top-ranked 5	0.0947
Top-ranked 6	0.0768
Top-ranked 7	0.0860
Top-ranked 8	0.0918
Top-ranked 9	0.0517
Top-ranked 10	0.0454
Top-ranked 11	0.0454
Top-ranked 12	0.0412
Top-ranked 13	0.0366
Top-ranked 14	0.0468
Top-ranked 15	0.0344
Top-ranked 16	0.0447
Top-ranked 17	0.0335
Top-ranked 18	0.0353
Top-ranked 19	0.0392
Top-ranked 20	0.0344
Top-ranked 21	0.0480
Top-ranked 22	0.0480
Top-ranked 23	0.0423
Top-ranked 24	0.0320
Top-ranked 25	0.0261
Top-ranked 26	0.0548
Top-ranked 27	0.0347
Top-ranked 28	0.0321
Top-ranked 29	0.0364
Top-ranked 30	0.0277
Top-ranked 31	0.0375
Top-ranked 32	0.0252
Top-ranked 33	0.0446
Top-ranked 34	0.0247
Top-ranked 35	0.0368
Top-ranked 36	0.0447
Top-ranked 37	0.0338
Top-ranked 38	0.0341
Top-ranked 39	0.0262
Top-ranked 40	0.0343

5.2.3 Selection of Features Using the Proposed GA-based Method

If we select features from the top-ranked 40 features without using the strict ranking measured by MR value, a further optimized feature combination may be obtained. This thesis uses a GA-based method to further select the subset of features from the top-ranked 40 features. Table 5.3 shows the details of the proposed GA-based method. The feature subset used to model the HI with the lowest RMSE is chosen to be the feature subset in the HI modeling. We have done data analysis using the proposed GA-based method.

Table 5.3 The proposed GA-based feature selection method

GA-based Method in Feature Selection:
1: Generate an initial population which contains N number of points. Each point ranges from 1-2^{40}. (We use a 40-bit binary string to represent any feature combination subset. Each bit corresponds to a feature. If a specific feature is selected, we mark the bit as "1" in our binary string; otherwise, we mark the bit as "0" in the binary string.). Evaluate the RMSE values for each point in the current population.
2: Based on the RMSE values of the individuals in the current population, a mating pool consisting of N individuals will be formed by randomly selecting members from the current population. An individual may be selected more than once. The fitter individuals have a higher probability being selected.
3: Reproduce the population by evolution operation. The members in the mating pool are called parents. The evolution operation includes two sub-operations, namely, crossover operation and mutation operation.
4: Evaluate the RMSE values of the new population. If the max number of generations is met, stop; otherwise, go to step 2.

The GA-based method results in the following subset including 19 features: [#1, #2, #3, #4, #6, #7, #9, #10, #13, #18, #19, #20, #21, #24, #29, #33, #34, #36, #38], and its

corresponding RMSE is 0.0219, which is much lower than that of the previous two feature selection methods with the fixed stepsize of 5 (with the RMSE=0.0261) and the fixed stepsize of 1 (with the RMSE=0.0247), respectively. Table 5.4 lists these 19 winning features, the detailed definitions and equations for these winning features can be found in [11]. As we mentioned in Chapter 4, LS1, LS2, HS1, and HS2 denotes the four vibration sensors. Figure 5.4 shows the HI trend using the subset of features selected by the GA-based method is much smoother than using the previous two feature selection methods.

Table 5.4 The 19 winning features selected by the proposed GA-based method [11]

Top-ranked #	Feature name
1	LS1_m2_k-1_s-2
2	HS1_m2_k-1_s-2
3	LS2_m2_k-1_s-2
4	HS2_m2_k-1_s-2
6	LS2_kurtosis
7	LS2_coefficient of kurtosis
9	LS2_mean frequency
10	LS1_mean frequency
13	LS2_FM0
18	LS2_m3_k2_s2
19	HS1_average absolute value
20	HS1_stage1_sb_L5
21	LS2_sb level factor
24	LS1_stage2_sb_R5
29	LS1_energy ratio
33	LS2_stage1_sb_L5
34	HS2_stage1_sb_L6
36	HS2_sb level factor
38	HS2_M6A

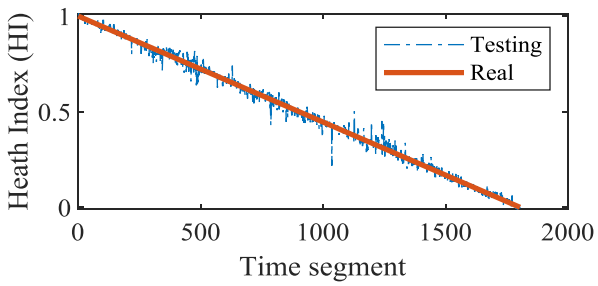

Figure 5.4 HI vs. time (features selected by GA as the input of the FFNN)

5.2.4 Investigations on Additional Data Separations

This thesis used 3608 time segments of the RTF experiment data documented in [35] to model the HI. As described in Chapter 4, the vibration signals of each 5-minute segment were divided into 8 time segments. We first chose to use the features of the first 4 segments in every 8 time segments for training the FFNN, and we used the features in the remaining 4 time segments for testing the FFNN in the previous sections. This is the original data separation. We are curious about the impact on the way in choosing the training data set and the testing data set in the FFNN modeling. Therefore, we do further investigations on two additional data separations as follows:

1) There are 3608 time segments in total. We use the 1st ½ time segments (time segments 1-1804) for training the FFNN, the remaining ½ time segments (time segments 1805-3608) for testing the FFNN. This is the additional data separation #1. This additional data separation is to test whether it can get accurate HI predictions using the first half lifetime for training, trying to extrapolate far into the future for the remaining half lifetime.

2) In 1), the 1st ½ time segments are used for training the FFNN, the remaining ½ segments are used for testing the FFNN. Here, we randomly select ½ time segments from the total 3608 time segments for training the FFNN, and the remaining ½ time segments are used for testing the FFNN. The time segments in each set (either the training set or the testing set) are arranged in chronological order. This is the additional data separation #2. This additional data separation is to test whether it can get accurate HI predictions using the randomly selected half lifetime data for training the HI model and using the remaining half lifetime data, which is also random, for testing the HI model. The expected RMSE value using the randomly selected half lifetime data should be higher than the RMSE value using the original data separation as the uncertainty of using the randomly selected half lifetime data is higher than using the original data separation. The motivation to try this additional data separation is to test whether the HI modeling is still effective when the number of time segments in each data collection period is different.

For the additional data separation #1, we have done data analysis using the GA-based feature selection method in the HI modeling. It is worth mentioning that the second ½ time segments are used for testing the HI model, thus the HI in Figure 5.5 starts from 0.5 instead of 1.

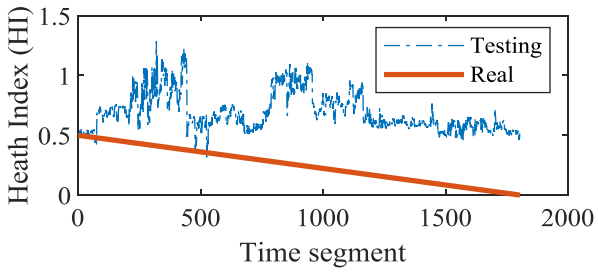

Figure 5.5 HI vs. time (using additional data separation strategy #1)

The additional data separation strategy #1 results in the following subset including 20 features:

Ω={#1,#2,#4,#8,#9,#11,#15,#16,#17,#18,#19,#21,#23,#24,#25,#26,#29,#30,#32,#37}

and its corresponding RMSE is 0.4749, which is as 20 times higher as the RMSE using the original data separation (with RMSE=0.0219). Table 5.5 lists all these 20 winning features, the detailed definitions and equations for these winning features can be easily found in [11]. Figure 5.5 shows that the HI fluctuates a lot in the whole degradation process. We do not see any monotonic trend between the HI and the time segment in Figure 5.5. This means that using the first half lifetime for training, trying to extrapolate far into the future for the remaining half lifetime is difficult to get accurate HI predictions. Therefore, we could not consider the additional data separation #1 in the HI modeling.

Table 5.5 The 20 winning features selected by the proposed GA-based method [11]

Top-ranked #	Feature name
1	LS1_m2_k-1_s-2
2	HS1_m2_k-1_s-2
4	HS2_m2_k-1_s-2
8	LS1_m3_k2_s2
9	LS2_mean frequency
11	HS1_mean frequency
15	HS1_RMS
16	HS1_STD
17	LS1_max value
18	LS2_m3_k2_s2
19	HS1_average absolute value
21	LS2_sb level factor
23	HS1_clearance factor
24	LS1_stage2_sb_R5
25	HS2_m3_k2_s2
26	LS2_stage1_sb_L5
29	LS1_energy ratio
30	HS2_mean frequency
32	LS2_coefficient of variation
37	HS2_FM4

For additional data separation #2, we have also done data analysis using the GA-based feature selection method in the HI modeling. The additional data separation strategy #2 results in the following subset including 18 features: Ω={#1,#4,#7,#8,#9,#14,#18,#19,#20,#21,#25,#26,#29,#31,#32,#34,#36,#39}, and its corresponding RMSE is 0.0352, which is slightly higher than the RMSE using the original data separation (with RMSE=0.0219). Table 5.6 lists all these 20 winning features, the detailed definitions and equations for these winning features can be easily found in [11].

Table 5.6 The 18 winning features selected by the proposed GA-based method [11]

Top-ranked #	Feature name
1	LS1_m2_k-1_s-2
4	HS2_m2_k-1_s-2
7	LS2_coefficient of kurtosis
8	LS1_m3_k2_s2
9	LS2_mean frequency
14	HS1_variance
18	LS2_m3_k2_s2
19	HS1_average absolute value
20	HS1_stage1_sb_L5
21	LS2_sb level factor
25	HS2_m3_k2_s2
26	LS1_stage2_sb_L5
29	LS1_energy ratio
31	HS1_m2_k-1_s2
32	LS2_coefficient of variation
34	HS2_stage1_sb_L6
36	HS2_sb level factor
39	HS2_stage1_sb_L3

Figure 5.6 shows that the modeled HI using the additional data separation #2, the HI experiences a clear monotonic degradation trend with respect to the time segment. This means that using the randomly selected half lifetime data for training the HI model and using the remaining half lifetime data, which is also random, to model the HI can get accurate HI predictions. The RMSE value of using this additional data separation #2 is higher than using the original data separation, which matches our assumptions. While the RMSE value of using the additional data separation #2 is acceptable. In other words, the additional data separation #2 is acceptable.

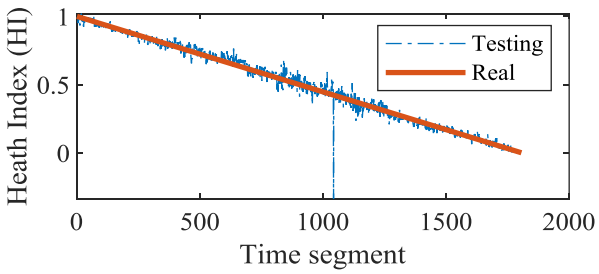

Figure 5.6 HI vs. time (using additional data separation strategy #2)

In summary, we do not consider the additional data separation #1 in our HI modeling since the HI developed could not reflect any monotonic trend of the HI. The additional data separation strategy #2 is acceptable, which matches our assumption. However, the original data separation shows a better performance in the HI modeling, thus, this thesis uses the original data separation in HI development.

5.2.5 Summary

In Section 5.2, we have investigated the impact of the combinations of the input features in the FFNN-based HI modeling. Three variations including the fixed stepsize of 5, the improved fixed stepsize of 1, and the proposed GA-based method for input feature selection are employed to test the accuracy of HI modeling. The results of data analyses show that the GA-based method outperformed its counterparts. Based on the data analyses, using the proposed GA-based method to model the HI with the lowest RMSE is recommended for the HI modeling for planetary gearboxes.

5.3 HI Smoothing for Planetary Gearboxes

As we can see from Figure 5.2 b), 5.3 b), and 5.4 in the previous sections, the originally modeled HIs fluctuate all the way during the whole degradation process. Thus, a proper HI

smoothing procedure is needed. Yang *et al.* [13] reported an HI smoothing procedure called the HI dynamic smoothing procedure in order to decrease the fluctuations of the originally modeled HI. Figure 5.7 shows the detailed procedures of the reported HI dynamic smoothing procedure in Yang's method.

In the reported HI dynamic smoothing procedure, three requirements including "Monotonicity", "Gradualness", and "Consistency" need to be satisfied. The "Monotonicity" means that the health condition of the engineering system should degrade monotonically if there are no maintenance actions on this engineering system [13]. The "Gradualness" means the health condition of the engineering system degrades gradually without a sudden big drop [13]. The "Consistency" means that the health condition of the engineering system should remain constant or change only slightly in a short period of time. More details about the reported HI dynamic smoothing procedure can be found in [13]. Theoretically, the reported HI dynamic smoothing procedure can be used to decrease the fluctuations of the originally modeled HI. We will investigate the reported HI dynamic smoothing procedure in the following sections.

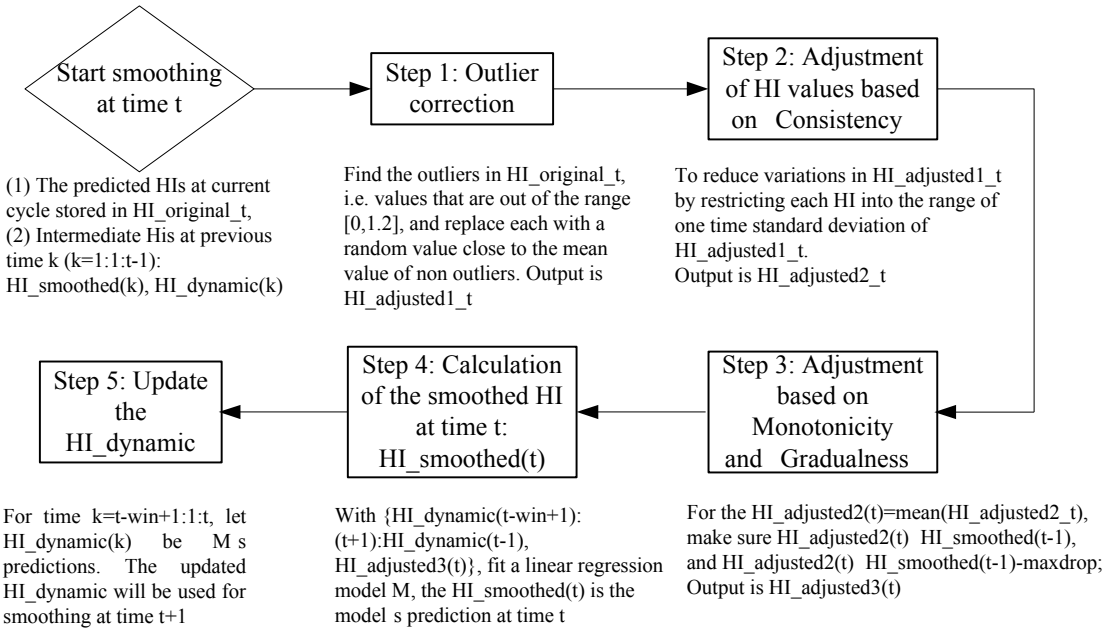

Figure 5.7 HI dynamic smoothing procedure [13]

5.3.1 Using Yang's Method in HI Smoothing for Planetary Gearboxes

Based on the originally modeled HI in Section 5.2.3, we implement the reported HI dynamic smoothing procedure with the fixed window size of 5 and the fixed maxdrop of 0.1 [13]. We have done data analysis using the reported HI dynamic smoothing procedure with the fixed window size of 5 and the fixed maxdrop of 0.1. It results in an RMSE=0.0203, which is lower than the RMSE=0.0219 evaluated by the originally modeled HI without smoothing. It is found that the reported HI dynamic smoothing procedure smooths the modeled HI in terms of the RMSE evaluation. Figure 5.8 shows that the fluctuation of the smoothed HI using the reported method is lower than the originally modeled HI without smoothing. In addition, we believe that the reported HI dynamic smoothing procedure may perform even better by changing the fixed window size of 5 and the fixed maxdrop of 0.1.

Thus, we propose to improve the reported HI dynamic smoothing procedure [13] in the following section.

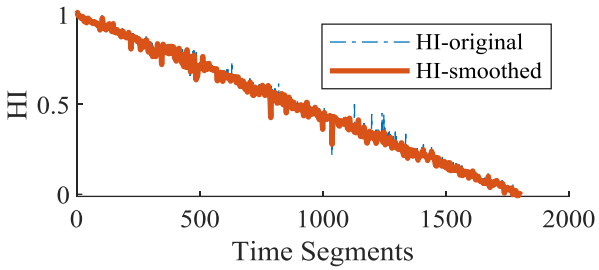

Figure 5.8 Smoothed HI vs. time (using reported method in [13])

5.3.2 Using Improved HI Dynamic Smoothing Procedure

Based on the originally modeled HI in Section 5.2.3, we could implement HI dynamic smoothing procedure with modifications. We treat this problem as an optimization problem, the objective function is to minimize the RMSE between the assumed true HI and the smoothed HI, the variables to be adjusted are the window size and the maxdrop in the reported HI dynamic smoothing procedure. We use a For-loop in Matlab to consider different combinations of the window size and the maxdrop in order to find the best combination of the window size and the maxdrop which results in an HI with the highest accuracy. We will introduce how we implement the For-loop in the following paragraph.

We propose to use a window size in a range [2,100] instead of the fixed window size of 5. In addition, we propose to use a maxdrop in a range [0,0.5] instead of the fixed maxdrop of 0.1. Two variables are used in this optimization problem, we assume that a trade-off between the window size and the maxdrop can be achieved. For the For-loop as mentioned in the previous paragraph, we use a parameter i that equal to the integer value from 2 to

100 (a range of [2,100] of the window size) with a stepsize of 1 in pairing with another parameter j that equal to the value from 0 to 0.5 (a range of [0,0.5] of the maxdrop) with a stepsize of 0.0001. Then a large number of combinations of the window size and the maxdrop will be produced. The smoothed HIs using the produced combinations of the window size and the maxdrop will be used in comparisons regarding their corresponding RMSE values. The HI dynamic smoothing procedure using the best combination of the window size and the maxdrop with the lowest RMSE value will be used as the combination of these two parameters in the HI smoothing for planetary gearboxes.

Figure 5.9 a) shows the RMSEs with respect to their corresponding window sizes at the optimal maxdrop value. Figure 5.9 b) shows the window sizes with respect to their corresponding optimal maxdrops. Figure 5.10 provides a 3-d plot showing the objective function as a function of the window size and the maxdrop. It is found that when the window size equal to 91 and the maxdrop equal to 0.4439, the lowest RMSE of 0.0087 can be obtained. The RMSE of 0.0087 using our improved HI dynamic smoothing procedure is much lower than the RMSE of 0.0203 using the reported HI dynamic smoothing procedure with the fixed window size of 5 and the fixed maxdrop of 0.1.

For different planetary gearboxes, the best value of the window size and the maxdrop may be different. With our proposed improved HI dynamic smoothing procedure, we can find the best combinations of these two parameters for other planetary gearboxes. The idea to propose the improved HI dynamic smoothing procedure is to find the best combination of the parameters that best fit in smoothing the modeled HI for different planetary gearboxes. In addition, the improved HI dynamic smoothing procedure may be applied in other systems beyond the planetary gearboxes.

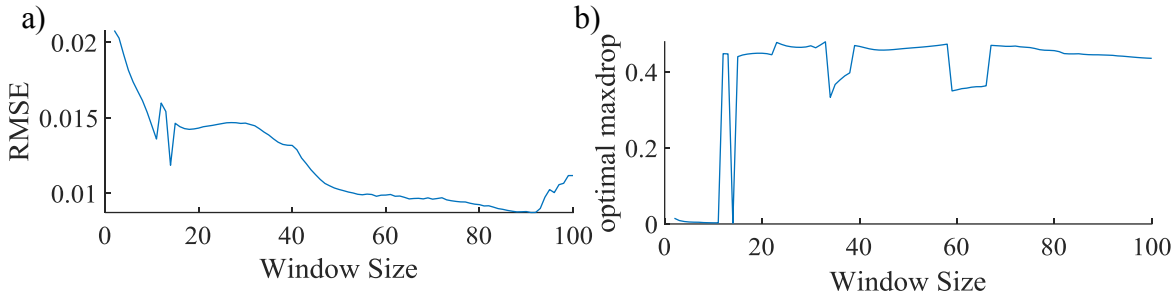

Figure 5.9 a) RMSE vs. window size; b) Window sizes with respect to their corresponding optimal maxdrops

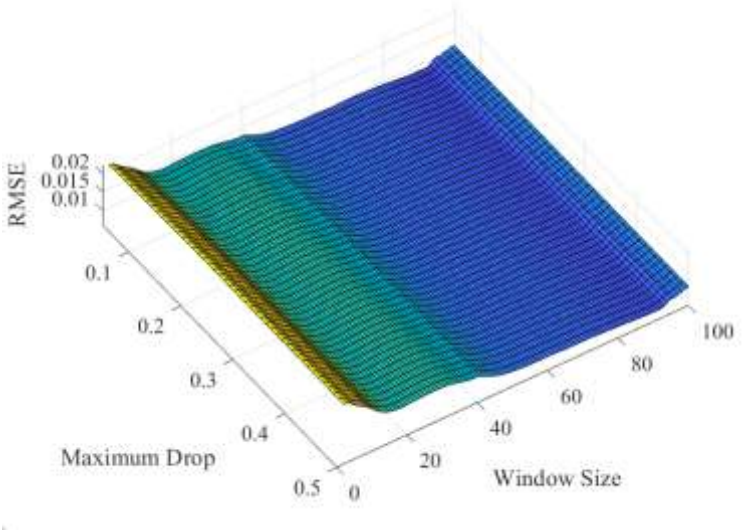

Figure 5.10 3-d plot for RMSE in HI smoothing

Figure 5.11 also shows that the smoothed HI using our improved HI dynamic smoothing procedure is much smoother than the smoothed HI using the reported method with the fixed window size and the fixed maxdrop [13].

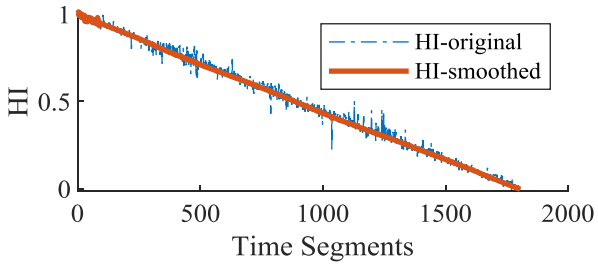

Figure 5.11 HI vs. time

5.3.3 Summary

In Section 5.3, we have investigated the impact of the combinations of the window size parameter and the maxdrop parameter in the reported HI dynamic smoothing procedure. Two variations including the reported method using the fixed window size of 5 and the fixed maxdrop of 0.1 and the improved method using the changeable window size and the changeable maxdrop are employed to test the performance of HI smoothing. The results of data analyses show that the improved method using the changeable window size and the changeable maxdrop outperformed the reported method. Based on the data analyses using the RTF experiment data, the improved method which smoothed the HI with the lower RMSE is recommended for the HI smoothing for the planetary gearboxes.

5.4 Comparisons with Yang's Method

This section compares the performance using the reported Yang's method and our improved method in HI development for planetary gearboxes with the observations obtained in the earlier sections.

For the HI modeling part, three feature selection methods are compared, Table 5.7 lists the detailed comparisons. The accuracy of the modeled HI using the fixed stepsize of 1

improves 5.36% compared with the modeled HI using the fixed stepsize of 5 [13]. In addition, the accuracy of the modeled HI using the GA-based method improves 16.09% compared with the modeled HI using the fixed stepsize of 5 [13]. Therefore, the GA-based feature selection method is recommended in the HI modeling for the planetary gearboxes.

Table 5.7 Comparisons among these 3 different feature selection methods

Methods	RMSE values	Relative improvement over [13]
Stepsize of 5 [13]	0.0261	-
Stepsize of 1	0.0247	5.36%
GA-based method	0.0219	16.09%

For the HI smoothing part, two HI dynamic smoothing procedures are compared, Table 5.8 lists the detailed comparisons. The accuracy of the smoothed HI using the improved HI dynamic smoothing procedure improves 57.14% compared with the smoothed HI using the reported HI dynamic smoothing procedure with the fixed window size of 5 and the fixed maxdrop of 0.1 [13]. Therefore, the improved HI smoothing is recommended in the HI smoothing for the planetary gearboxes.

Table 5.8 Comparisons between these 2 different HI dynamic smoothing procedures

Procedures	RMSE values	Relative improvement over [13]
Reported fixed window size and fixed maxdrop [13]	0.0203	-
Improved HI smoothing	0.0087	57.14%

5.5 Summary

This whole chapter investigates two aspects of HI development using the FFNN and HI dynamic smoothing procedure based on vibration signals in a lab planetary gearbox. The first aspect is the impact of the combinations of input features in HI modeling. A smaller fixed stepsize of 1 is found for selecting better input features in HI modeling. In addition, a GA-based method in input feature selection is proposed and performs better compared with the reported fixed stepsize of 5 and the improved fixed stepsize of 1. The second aspect is the impact of the different combinations of parameters in the reported HI dynamic smoothing procedure. An applicable range of [2,100] for window size and an applicable range of [0,0.5] for maxdrop are found able to provide reasonable results of smoothed HI using the improved HI dynamic smoothing procedure. In addition, the best combination of window size and maxdrop that results in the highest accuracy of the HI is found using optimization.

Chapter 6

Summary and Future Work

6.1 Summary

HI development for planetary gearboxes using vibration signals includes two crucial parts: HI modeling and HI smoothing. This thesis investigated a reported method which uses an FFNN in HI modeling, and the reported method also uses an HI dynamic smoothing procedure in HI smoothing. Two aspects were investigated and the work of this thesis study is summarized as follows:

1) The impact of the combinations of the input features in the FFNN-based HI model was investigated. A finer feature selection method using the fixed stepsize of 1 was investigated based on the reported feature selection method using the fixed stepsize of 5. The results of data analyses showed that the finer feature selection method using the fixed stepsize of 1 performed better than the reported method. Both the finer feature selection method using the fixed stepsize of 1 and the reported method using the fixed stepsize of 5 used the candidate features following the top-down MR ranking. Furthermore, a GA-based feature selection method was investigated to find a subset of features not necessarily following the top-down MR ranking. This method outperforms the improved feature selection method using the fixed stepsize of 1 and the reported feature selection method using the fixed stepsize of 5 according to the results.

2) The impact of both the window size parameter and the maxdrop parameter in the reported HI dynamic smoothing procedure was investigated. An improved HI dynamic smoothing procedure using the optimized window size parameter and the optimized maxdrop parameter was proposed. An optimized combination of the window size and the maxdrop was found in the HI dynamic smoothing procedure. The results showed that compared to the reported HI dynamic smoothing procedure, the improved HI dynamic smoothing can implement a more effective HI smoothing. The improved HI dynamic smoothing obviously decreased the fluctuations of the originally modeled HI. The improved HI dynamic smoothing procedure can help to improve the accuracy of the HI development.

6.2 Future Work

This thesis investigated two aspects of a reported method in order to improve the performance of the reported method. However, in this study several other issues are found and need to be addressed in the future:

1) Other condition monitoring data types such as thermographic data, acoustic data, and oil debris data may be useful in HI development. Using more data types may help to develop a more accurate HI. Further studies need to be conducted on figuring out the impact of using such data types.

2) Further investigations on other modeling algorithms such as Recurrent Neural Networks that may perform better on time-series data analysis need to be conducted in order to better fit the HI development. This will be a study topic in the future.

Bibliography

[1] H. C. Andrew, K.S., Albert, "Maintenance, Replacement, and Reliability Theory and Applications," vol. 3, no. September, 1981.

[2] A. Khadersab and S. Shivakumar, "Vibration Analysis Techniques for Rotating Machinery and its effect on Bearing Faults," *Procedia Manufacturing*, vol. 20, pp. 247–252, 2018.

[3] B. H. Kapil, P. N. Pundlik, P. Y. Raghunath, K. H. Bagul, P. N. Patil, and R. Y. Patil, "Thermal Analysis of Journal Bearing Using CFD Software for Performance Enhancement," *International Journal of Advance Research*, vol. 3, pp. 287–293.

[4] W. M. Alobaidi, E. A. Alkuam, H. M. Al-Rizzo, and E. Sandgren, "Applications of Ultrasonic Techniques in Oil and Gas Pipeline Industries: A Review," *American Journal of Operation Research*, vol. 05, no. 04, pp. 274–287, 2015.

[5] Plant Wellness Way Eam, *web page* . https://www.lifetime-reliability.com/cms/machinery-health-measurement/.

[6] Y. Lei, J. Lin, M. J. Zuo, and Z. He, "Condition monitoring and fault diagnosis of planetary gearboxes: A review," *Measurement*, vol. 48, no. 1, pp. 292–305, 2014.

[7] S. Sheng, "Gearbox Reliability Collaborative Update," *National Renewable Energy Laboratory*, 2013. https://www.nrel.gov/docs/fy14osti/60141.pdf.

[8] L. Liu, "Vibration Signal Analysis for Planetary Gearbox Fault Diagnosis," PhD Dissertation, University of Alberta, 2018.

[9] L. Liu, X. Liang, and M. J. Zuo, "Vibration signal modeling of a planetary gear set with transmission path effect analysis," *19th World Conference on Non-Destructive Testing*, vol. 85, pp. 20–31, 2016.

[10] M. Hoseini, D. Wolfe, Y. Lei, and M. J. Zuo, "Planetary Gearbox Test Rig," Technical Report, University of Alberta, 2008.

[11] M. Hoseini, Z. Liu, T. Tao, and M. J. Zuo, "Original Features Used for Condition Monitoring of Planetary Gearboxes," Technical Report, University of Alberta, 2011.

[12] J. Qu, Z. Liu, M. J. Zuo, and H. Z. Huang, "Feature selection for damage degree classification of planetary gearboxes using support vector machine," *Proceeding of the Institution of Mechanical Engineering, Part C: Journal of Mechanical Engineering Science*, vol. 225, no. 9, pp. 2250–2264, 2011.

[13] F. Yang, M. S. Habibullah, T. Zhang, Z. Xu, P. Lim, and S. Nadarajan, "Health Index-Based Prognostics for Remaining Useful Life Predictions in Electrical Machines," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 4, pp. 2633–2644, 2016.

[14] R. Heywood and T. Mcgrail, "Generating Asset Health Indices Which Are Useful and Auditable," *Doble Engineering*, 2015.

[15] I.T. Yu and C.D. Fuh, "Estimation of Time to Hard Failure Distributions Using a Three-Stage Method," *IEEE Transactions on Reliability*, vol. 59, pp. 405–412, 2010.

[16] A. N. Jahromi, R. Piercy, S. Cress, J. R. R. Service, and W. Fan, "An approach to power transformer asset management using health index," *IEEE Electrical Insulation Magazine*, vol. 25, no. 2, pp. 20–34, 2009.

[17] Y. J. Chen, W. X. Tang, X. H. Liang, and M. J. Zuo, "Degradation Assessment for Critical Assets in Power Generation," Technical Report, University of Alberta, 2017.

[18] R. K. Youree, J. S. Yalowitz, A. Corder, and T. K. Ooi, "A Multivariate Statistical Analysis Technique for On-Line Fault Prediction," *International Conference on Prognostics and Health Management*, 2008.

[19] A. M. Riad, H. K. Elminir, and H. M. Elattar, "Evaluation of neural networks in the subject of prognostics as compared to linear regression model," *International Journal of Engineering and Technology*, vol. 10, no. 06, pp. 52–58, 2010.

[20] X. L. and L. Z. Cheng, N. Hu, "Health indicator extraction based on sparse representation of vibration signal for planetary gearbox," *Prognostics and System Health Management Conference (PHM-Chengdu)*, pp. 1–4, 2016.

[21] E. S. K. J. F. Kenney, "Mathematics of statistics, 3rd Edition Princeton," *Van Nostrand Company*, 1962.

[22] P. D. Samuel and D. J. Pines, "A review of vibration-based techniques for helicopter transmission diagnostics," Journal of Sound and Vibration, vol. 282, no. 1–2. 2005.

[23] "Harmonic," *Wiki Encyclopedia*. https://en.wikipedia.org/wiki/Harmonic.

[24] "Sideband," *Wiki Encyclopedia*. https://en.wikipedia.org/wiki/Sideband.

[25] "Modulation," *Wiki Encyclopedia*. https://en.wikipedia.org/wiki/Modulation.

[26] P. J. Dempsey and S. Sheng, "Investigation of data fusion applied to health monitoring of wind turbine drivetrain components," *Wind Energy*, vol. 16, pp. 479–489, 2013.

[27] R. Wald, T. Khoshgoftaar, and J. Sloan, "Fourier transforms for vibration analysis: A review and case study," IEEE International Conference on Information Reuse and Integration, 2011.

[28] "Frequency domain," *Wiki Encyclopedia*. https://en.wikipedia.org/wiki/Frequency_domain.

[29] E. K. P. Chong and S. H. Zak, "An Introduction to Optimization: 4th Edition," *Wiley*, 2011.

[30] "Genetic algorithm," *Wiki Encyclopedia*. https://en.wikipedia.org/wiki/Genetic_algorithm.

[31] "Regression analysis," *Wiki Encyclopedia*. https://en.wikipedia.org/wiki/Regression_analysis.

[32] V. Cherkassky, "Neural networks and nonparametric regression," *Neural Networks for Signal Processing II Proceedings of the 1992 IEEE Workshop*, 1992.

[33] "Backpropagation," *Wiki Encyclopedia*. https://en.wikipedia.org/wiki/Backpropagation.

[34] "Integer programming," *Wiki Encyclopedia*. https://en.wikipedia.org/wiki/Integer_programming.

[35] M. Pandey, T. Patel, X. Liang, T. Tian, and M. J. Zuo, "Descriptions of Pitting Experiments, Run-to-Failure Experiments, Various Load and Speed Experiments, and Crack Experiments Carried out on the Planetary Gearbox Test Rig," Technical Report, University of Alberta, 2011.

[36] X. Zhao, R. Moghaddass, T. Patel, and M. J. Zuo, "Feature Evaluation and Feature Ranking," Technical Report, University of Alberta, 2011.

[37] V. T. Do, M. R. Hoseini, A. Sahoo, M. Pandey, and M. J. Zuo, "Metal Scan Data , Gear Weight Loss Data , and Gear Damage Images Recorded during the Run-to-Failure Experiments," Technical Report, University of Alberta, 2010.