NOVEL CONDITION MONITORING TECHNIQUES APPLIED TO IMPROVE THE DEPENDABILITY OF RAILWAY POINT MACHINES

by

TOMOTSUGU ASADA

A thesis submitted to the University of Birmingham for the degree of DOCTOR OF PHILOSOPHY

Department of Electronic,
Electrical and Computer
Engineering
School of Engineering
University of Birmingham

May 2013

ABSTRACT

Point machines are the key actuator used in railways to provide a means of moving a switch blade from one position to the other. Failure in the point actuator has a significant effect on train operations. Condition monitoring systems for point machines have been therefore implemented in some railways, but these condition monitoring systems have limitations for detecting incipient faults. Furthermore, the majority of condition monitoring systems which are currently in use cannot diagnose faults. The ability to diagnose faults is useful to maintenance staff who need to fix problems immediately.

This thesis proposes a methodology to detect and diagnose incipient faults using an advanced algorithm. In the main body of this thesis the author considers a new approach using Wavelet Transforms and Support vector machines for fault detection and diagnosis for railway electrical AC point machines operated in Japan.

The approach is further enhanced with more data sets collected from railway electrical DC point machines operated in Great Britain. Furthermore, a method to express the qualitative features of healthy and faulty waveforms was proposed to test the transferability of the specific algorithm parameters from one instance of a point machine to another, which is tested on railway electrical DC point machines used in Great Britain.

Finally, an approach based on Wavelet Transforms and Neural networks is used to predict the drive force when the point machine is operating. The approach was tested using electrical DC point machines operated in Great Britain.

It is shown through the use of laboratory experimentation that the proposed methods have potential to be used in a real railway system.

ACKNOWLEDGEMENTS

The author thanks the following people/parties for their support and contributions to the research:

Professor Clive Roberts, for his experienced guidance and support as the first supervisor

Associate professor Takafumi Koseki, for his advice as the local supervisor in Japan

Professor Martin Russell, for his advice as the second supervisor

Dr Edd Stewart, for his help and assistance with the data collection

Dr Hiroto Takeuchi, for his experienced advice throughout the course of the research as a senior in the company (Central Japan Railway Company)

The Central Japan Railway Company, for covering the course and maintenance fees during the course of the research and for facilitating the data collection from NTS point machine

London Underground, for facilitating the data collection from M63 and Surelock point machine

Network Rail, for facilitating the data collection from HW point machines

Finally, I would like to thank my wife, Manami, for her love, understanding, and encouragement throughout the course of PhD.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1 INTRODUCTION

1.1 Background

1.1.1 Railways and point machines

Recently, in part due to environmental issues and also because of congestion on the roads, the utilisation of railways has been increasing all around the world. The safe and reliable operation of trains is therefore becoming ever more important. To achieve this objective, most of the actuators in railways are designed to be redundant; when one of the actuators fails, the railway can still maintain its function using another actuator. Although many actuators in railways are designed in this way, there are some that cannot be designed to be redundant because of their inherent structural and mechanical nature. One of them is the point machine.

The point machine is the actuator that drives the switch blade from one position to the opposite position in order to offer different routes to trains. Failure in the point actuator has a significant effect on train operations. If this failure occurs, it leads to a less reliable service and causes discredit to the railway company. It can also lead to more disastrous consequences. In 2002 a train derailment accident caused by poor maintenance of a point machine occurred near Potters Bar railway station in the UK, killing seven people. As a result, the railway infrastructure company paid several million pounds as

compensation to victims and their relatives [1]. It is therefore important for all infrastructure companies to minimise the occurrence of point machine failure.

1.1.2 Point machines in Japan

In Japan, many types of points are used, according to company preferences and geographic areas. Table 1-1 shows the different type of point machines used by the Central Japan Railway Company.

All the power sources to the points are AC single phase (50 or 60 Hz). The NS-type model is the prevalent model and is used on local lines and also in rolling stock depots of the High Speed Train (HST), known as Shinkansen. The CS-type model is a successive model of the NS-type model developed by the Central Japan Railway Company and has a stronger force (approximately 30% greater) than the NS-type model [2]. The TS-type model is used exclusively on HST routes. The NTS-type is a successive model of the TS-type model and most of the main tracks on HST lines now use the NTS-type model [2].

Figure 1-1 shows the point mechanism of a NS-type point.

Table 1-1 Types of points in Japanese Railway

Type of point	Power source	Usage
NS type	AC 105 V Single Phase (50 or 60 Hz)	Local line or rolling stock depot for High Speed Train
CS type		Rolling stock depot for High Speed Train
TS type		Main line for High Speed Train
NTS type	AC 210 V Single Phase(50 or 60 Hz)	Main line for High Speed Train

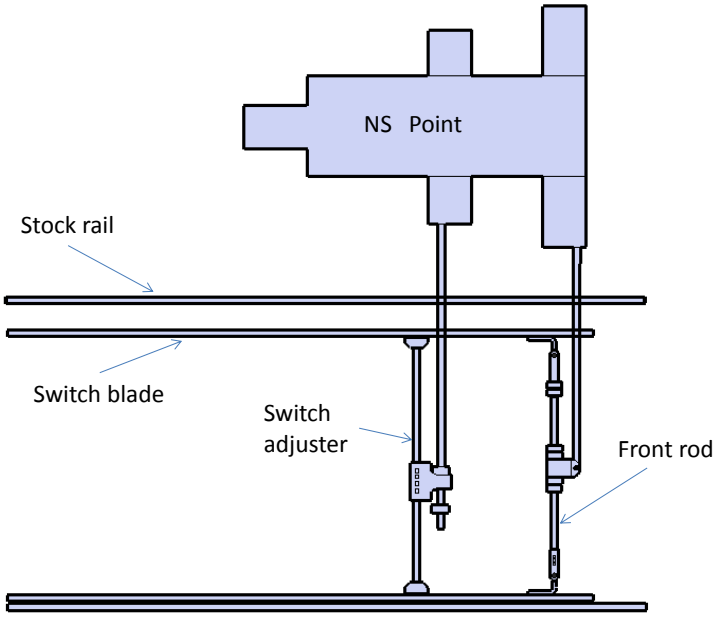

Figure 1-1 Point mechanism (NS-type model)

1.2 Current maintenance practice on the railways

1.2.1 Definition

The terminology and general concept of fault diagnosis is described in this chapter.

A “failure” is defined as ‘the inability of any asset to do what its users want it to do’ [3]. Furthermore, anything that users want it to do is defined as a “function”; there are usually several different functions in every system. A “functional failure” is defined as ‘the inability of any asset to fulfil a function to a standard of performance that is acceptable to the user’ [3]. A “fault” (potential failure) is defined as ‘an identifiable condition that indicates that a functional failure is either about to occur or is in the process of occurring’ [3]. In general, a “fault” can be identified by an experienced maintenance engineer if the engineer went to the site to carry out a full inspection.

To prevent assets from failure, there is a need to detect a “fault” (potential failure) before a “failure” occurs. Tasks designed to check and detect a “fault” to prevent the “functional failure” or to avoid the consequences of the “functional failure” are known as “on-condition tasks”. The interval between the occurrence of a “fault” (potential failure) and its degradation into a “functional failure” is defined as the “P-F interval”: “P” and “F” stand for “potential failure” and “functional failure” respectively [3]. “On-condition tasks” have to be undertaken during the “P-F interval” to avoid “functional failure”. Figure 1-2 shows the outline of a “P-F interval”.

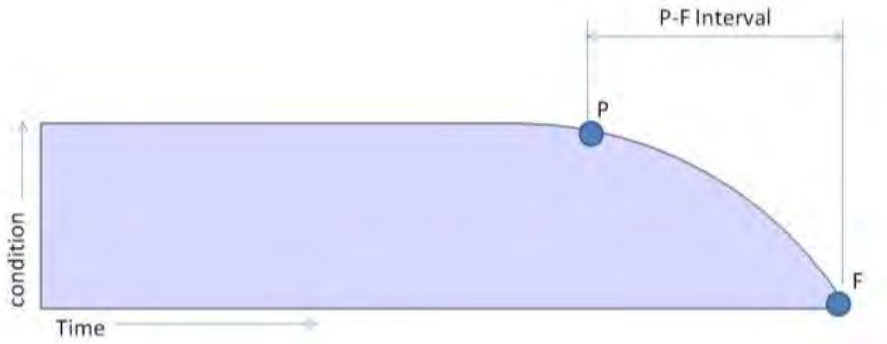

Figure 1-2 P-F interval [3]

1.2.2 Current practice and condition monitoring for point machines

Currently, maintenance tasks are undertaken at fixed intervals to minimise the likelihood of any failure of the point machines. Intervals may differ depending on the level and type of use of the machine and also company policy; maintenance tasks are undertaken every 2 weeks by the Central Japan Railway Company, whereas London Underground undertakes maintenance tasks every 6 weeks. When these maintenance tasks take place, maintenance staff use their own experience and their own senses (sight, sound, touch and smell) to detect any faults (potential failure). Condition monitoring systems have been developed to aid maintenance work. These condition monitoring systems use thresholds for the monitored parameters, creating an alarm whenever the monitored parameters exceed the predetermined thresholds. Experience in the field has shown that a problem exists with many of these systems, as an alarm is only created either after a failure has occurred or when the points are very close to failure [4]. Furthermore, the majority of condition monitoring systems which are currently in use cannot diagnose faults; the system simply creates an alarm to indicate abnormal behaviour. This can result in a delay in repairing the point machine, as the maintenance

staff do not know what the fault is until they arrive on site. A condition monitoring system for point machines which generates an alarm in the early stages of the development of the fault and which diagnoses faults correctly is therefore desired.

1.2.3 Condition monitoring systems for point machines in Japan

This sub-section provides details of the current condition monitoring systems used in Japanese point machines. Table 1-2 shows the parameters and technical methods of condition monitoring by type of point. The TS type and NTS type use four parameters for condition monitoring, whereas the NS type and CS type use only two parameters. Both systems are based on "thresholding techniques": a technique that uses *a priori* thresholds for monitoring.

Table 1-2 Condition monitoring systems in Japanese Railways

Type of point	Parameters	Method of condition monitoring
TS type NTS type	• Current • Voltage • Motor speed • Lock position	• Thresholding
NS type CS type	• Current • Voltage • (Lock error detection)	• Thresholding

All of the NTS type point machines in the field are equipped with condition monitoring systems on the Japanese HST lines. The sensors equipped inside a NTS type point machine are a current sensor, voltage sensor, rotary encoder and a lock position sensor. A lock position sensor is a sensor that sensing when the machine is in the locked

position and is used specifically to detect ‘locking failure’. The system creates an alarm if the lock position exceeds the pre-determined position. The system also utilises three sensors (a current sensor, voltage sensor and rotary encoder) to detect an abnormality of current, voltage and force in the point machine. The torque of the point machine is calculated based on a calculation table that is called the ‘Torque measurement table’. The table relates voltage and rotational speed to torque. Maintenance personnel will set an upper limit for the calculated torque, and the system generates an alarm if the torque exceeds the limit. Current and voltage are also monitored using a thresholding technique in the system. An outline of the system is depicted in Figure 1-3.

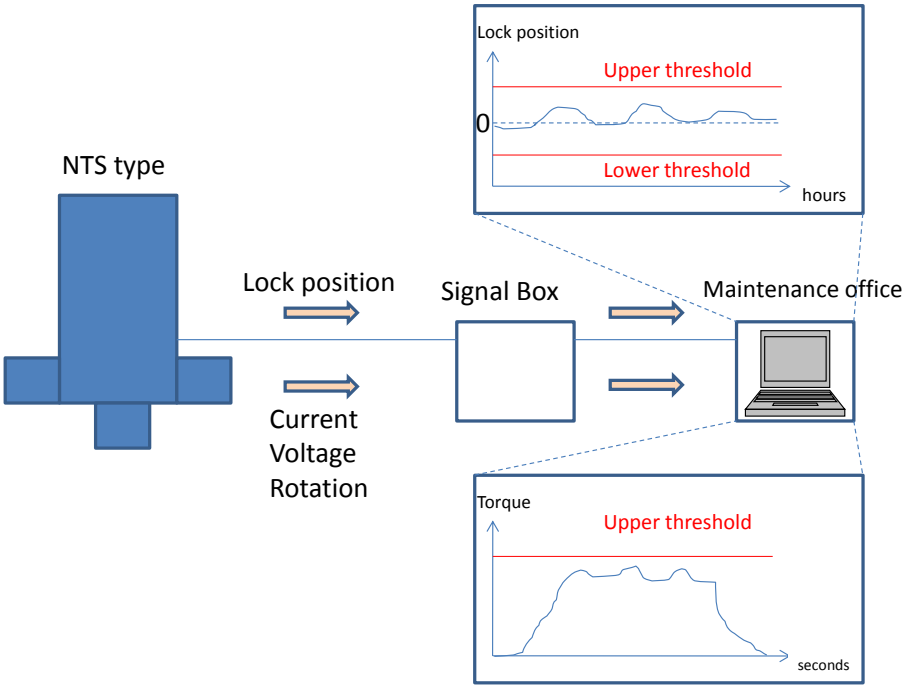

Figure 1-3 Outline of condition monitoring system (NTS type point machine)

In the NS type, on the other hand, only two sensors are equipped: a current sensor and a voltage sensor. NS type condition monitoring systems are only utilised on local lines (not on HST lines). Torque is estimated based on a lookup table that relates current and

voltage of the motor to the value of torque. The user sets an upper threshold for torque, and the system generates an alarm if the calculated torque exceeds the threshold. Although the NS type cannot acquire continuous measurement of the locking position, it does have an ability to confirm that the lock is in the correct position by means of a LED detector inside the point machine that generates an alarm if a locking failure occurs.

1.3 Proposal for improvement of the condition monitoring system

It is proposed that an improvement could be made if the condition monitoring system could detect and diagnose faults in the early stage of their development. By developing a system with this capability it would be possible to reduce the number of point machine failures, which would reduce the cost that an infrastructure company would have to pay to train operation companies and increase the reputation of the company.

It would also be possible to increase the safety of point machines by the condition monitoring system which will prevent critical accidents from happening that will potentially save a passenger's life.

It would also lead to a reduction in maintenance costs, as the frequency of maintenance tasks could be reduced. In Europe, there is a movement in the railway industry to shift from scheduled maintenance to condition based maintenance, in which the condition of the actuator determines whether maintenance is necessary (with the aid of sophisticated

condition monitoring systems). This will potentially lead to huge cost savings on maintenance tasks.

1.4 Systems engineering approach

Systems engineering is gaining popularity among many industries for developing complex and time-consuming systems. Although there is no commonly accepted and clear definition of systems engineering in the literature, the definition of systems engineering is usually based on the background and experience of the individual or the organisation [5]. The International Council on Systems Engineering defines systems engineering as "*An interdisciplinary approach and means to enable the realisation of successful systems*" [6].

Since the system which is needed to be developed in this thesis is related to a real industrial problem, it will be beneficial to utilise an approach which is used in industrial system development. It is therefore appropriate to use a systems engineering approach to develop a condition monitoring system for point machines. In this thesis, a requirements analysis is carried out in the initial stage of system development in order to avoid inefficiency. The requirements analysis was carried out using a top-down approach.

1.5 Scope of this thesis

The structure of the thesis is outlined below.

Chapter 2

Requirements for the system are presented in this chapter. Firstly, a high level mission is presented. Secondly, this is decomposed into single, testable statements so that each can be tested individually. This analysis helped to determine the direction of the solution presented in this thesis.

Then, details of the experimental setup and fault simulations are presented. Firstly, the different point machines considered in this thesis are introduced and the method of data acquisition is described. Secondly, fault simulations considered are described.

Chapter 3

A full literature review was conducted in this chapter. Firstly, a general survey of fault detection and diagnosis methodologies was undertaken and the methodologies were categorised into three categories. Secondly, a survey of condition monitoring specifically for point machines was carried out. Finally, the conclusion considered the current state-of-the art methodologies and requirements of the system set out in Chapter 2.

Chapter 4

The aim of this chapter is to develop an algorithm for fault detection and diagnosis utilising parameters collected from low-cost and practical sensors. Data was collected from an AC point machine (NTS-type point machine) used in Japan. Drive force, electrical current and electrical voltage data were collected and analysed from an AC

point machine (NTS-type point machine). Three fault conditions were simulated. Firstly, a cluster analysis was carried out to choose the best parameter for condition monitoring. Secondly, a feature extraction method was considered. Thirdly, a classification method was considered. Finally, experimental results were written. It was found that the method presented can detect and diagnose faults to a high degree of accuracy. It was also proved that the approach can provide an indication of the severity of the faults, which is important for practical implementation.

Chapter 5

The aim of this chapter is to: (1) test the method developed in Chapter 4 for other types of point machines to verify the transferability of the approach and (2) test the transferability of the specific algorithm parameters from one instance of a point machine to the next. To achieve this, data was collected from three different types of DC point machine used in Great Britain (Surelock-type point machine, M63-type point machine and HW-type point machine).

Drive force, electrical current and electrical voltage data were collected and analysed from two types of DC point machine: the Surelock-type and the M63-type. Five fault conditions were simulated. A cluster analysis was carried out to choose the best parameter for condition monitoring, as in Chapter 4. The same feature extraction and classification method presented in Chapter 4 were applied to the DC point machine data. It was found that the method can detect and diagnose faults to a high degree of accuracy. It was also proved that the approach can provide an indication of the severity of the faults, similarly to the results obtained for the Japanese point machine.

Then, the approach is used to test the transferability of the specific algorithm parameters from one instance of a point machine to the next (using HW-type point machines). To achieve this, the feature extraction method is modified to express the feature qualitatively.

Chapter 6

The aim of this chapter is to further develop the algorithm discussed in Chapter 4 and Chapter 5 so that the system can directly predict drive force which can be useful for inspection and maintenance. Data collected from British point machines (Surelock-type point machine and HW-type point machine) was used to demonstrate the method. It was found that the method can predict drive force to a high degree of accuracy.

Chapter 7

In this chapter, all the methods and results are reviewed and conclusions are made. Future work including recommendation of system architecture is discussed.

CHAPTER 2 REQUIREMENTS ANALYSIS AND EXPERIMENTAL SETUP

2.1 Requirements analysis

2.1.1 Initial requirements analysis

In this section a requirements analysis is carried out to identify a suitable condition monitoring system. A high level mission is presented, requirement decomposition is carried out and finally a conclusion is drawn. This requirements analysis, however, was carried out to help understand the general aspirations for the system and to provide the direction to the research.

The first stage of a requirements analysis is to set out a clear top-level requirement. Railway condition monitoring systems were classified by Roberts into three levels [7]:

'Level one: Data Logging and Event Recording Systems – primarily used to provide hard evidence in cases where major incidents happen.'

'Level two: Event Recording and Data Analysis Equipment – have the same functions as Level one, but are also equipped with basic data analysis options, such as

statistical/sequence analysis, (generally equipped with additional communication modules for remote access).'

'Level three: On-Line Health Monitoring Systems – defined as the highest level condition monitoring systems. These analyse data into characteristic signatures, compare these with an in-built database of healthy and simulated faulty operational modes, and flag alarms and fault diagnosis information to the operator-maintainers.' [7]

By categorising condition monitoring system into three levels, the advantages of the higher level systems are highlighted, helping to inform the direction of development. Most existing condition monitoring systems for railways are categorised as level two (or even level one in some instances). Japanese condition monitoring systems for railway points are categorised as level two, where warning alarms are generated according to the pre-determined thresholds in the system. A more advanced condition monitoring system for railway points is therefore needed to improve the safety and reliability of train operations.

The top-level requirement for this thesis is therefore to propose the highest level condition monitoring system, a 'level three' condition monitoring system for point machines. In 'Level three' condition monitoring systems, not only fault detection information but also fault diagnosis information will be informed to the user. This will be extremely useful since the maintenance staff can know before arriving on site the fault type, making it quicker to fix the fault.

2.1.2 Requirements decomposition

The top-level requirement is stipulated in Section 2.1 above. The purpose of requirements decomposition is to break down the initial requirement into a set of individual testable statements which are described in detail [5]. The following requirements were derived from the top-level requirement; these are further decomposed into testable statements.

1. Early detection and diagnosis: The system should detect and diagnose faults before failure occurs.

2. High accuracy: The system should detect and diagnose faults to a high level of accuracy so that the user will trust the system.

3. Informing fault level: The system should inform users of the fault level. When the minor fault level is alarmed, maintenance staff can wait until non-service time to repair the machine. However, if the major fault level is alarmed, it will be repaired during in-service time.

4. Not affecting operation: The system should utilise sensors which will not affect the operation of the machine. Even if the sensors break, the point machine must be able to operate normally.

5. The approach developed should be generic so that it can be applied to many different types of point machine.

These requirements can be decomposed further and written explicitly to result in a set of single, testable statements.

1. *Early detection and diagnosis: The system should detect and diagnose faults before failure occurs.*
    1.1 The system should classify the fault free condition and abnormal conditions.
    1.2 The system should diagnose faults; the system shall classify different fault conditions and inform users.
    1.3 Point machines should not fail to operate during the simulated fault conditions.

2. *High accuracy: The system should detect and diagnose faults to a high level of accuracy so that the user will trust the system.*
    2.1 The system should detect and diagnose the test data to high accuracy. (The highest accuracy achieved in the previous work was 91% [8].)
    2.2 The system should not raise false alarms (an alarm which is incorrectly triggered during "Fault free" conditions) so that users will trust the system

3. *Informing fault level: The system should inform users of the fault level. When the minor fault level is alarmed, maintenance staff can wait until non-service time to*

repair the machine. However, if the major fault level is alarmed, it will be repaired during in-service time.

3.1 (optional) The system should inform the user of two levels of severity: minor fault alarm and major fault alarm (to be later defined).

4. *Not affecting operation: The system should utilise sensors which will not affect the operation of the machine. Even if the sensors break, the point machine must be able to operate normally.*

   (No further decomposition required)

5. *The approach used should be generic so that it can be applied to many different types of point machine.*

   5.1 The approach should be demonstrated on at least two different types of point machine.

   5.2 (Optional) Once the algorithm has been trained, it should be usable on multiple instances of the same type of point machine.

2.1.3 Conclusion from requirements analysis

A requirements analysis has been carried out to make the objectives of this thesis clear. The initial requirement is to propose a method which is categorised as a 'level three' condition monitoring system for point machines. The initial requirement was decomposed to five requirements and described further in detail. The requirements of

the thesis were sufficiently decomposed and can be used to fix the direction of the research. By defining clear requirements from the beginning of the thesis, it is possible to establish the method and develop the system quickly and efficiently.

2.2 Experimental setup

2.2.1 Point machines and data acquisition box

In order to help address requirement 5.1, data was acquired from four different types of point machine. Table 2-1shows a list of the point machines considered in this thesis.

Table 2-1 A list of Point machines tested

Point machine type	Country operated	Supply-voltage to motor	Date of data collection, place
NTS-type	Japan	AC 210V	January 2011, Japan
Surelock	GB	DC 110V	September 2011, London
M63-type	GB	DC 110V	September 2011, London
HW-type	GB	DC 110V	May 2012, Derby

The NTS-type point machine is operated by a single-phase AC 210 V power supply. The single phase induction motor is located inside the point machine. M63-type, Surelock-type and HW-type point machines are operated using a DC 110V power supply. The M63-type machine is fitted with snubbing gear which brings the machine quickly to rest at the end of its stroke [9]. A brushed DC motor is located inside the M63-type point machine.

A photograph of each of the point machines taken during the data collection phase of the research is shown in Figure 2-1.

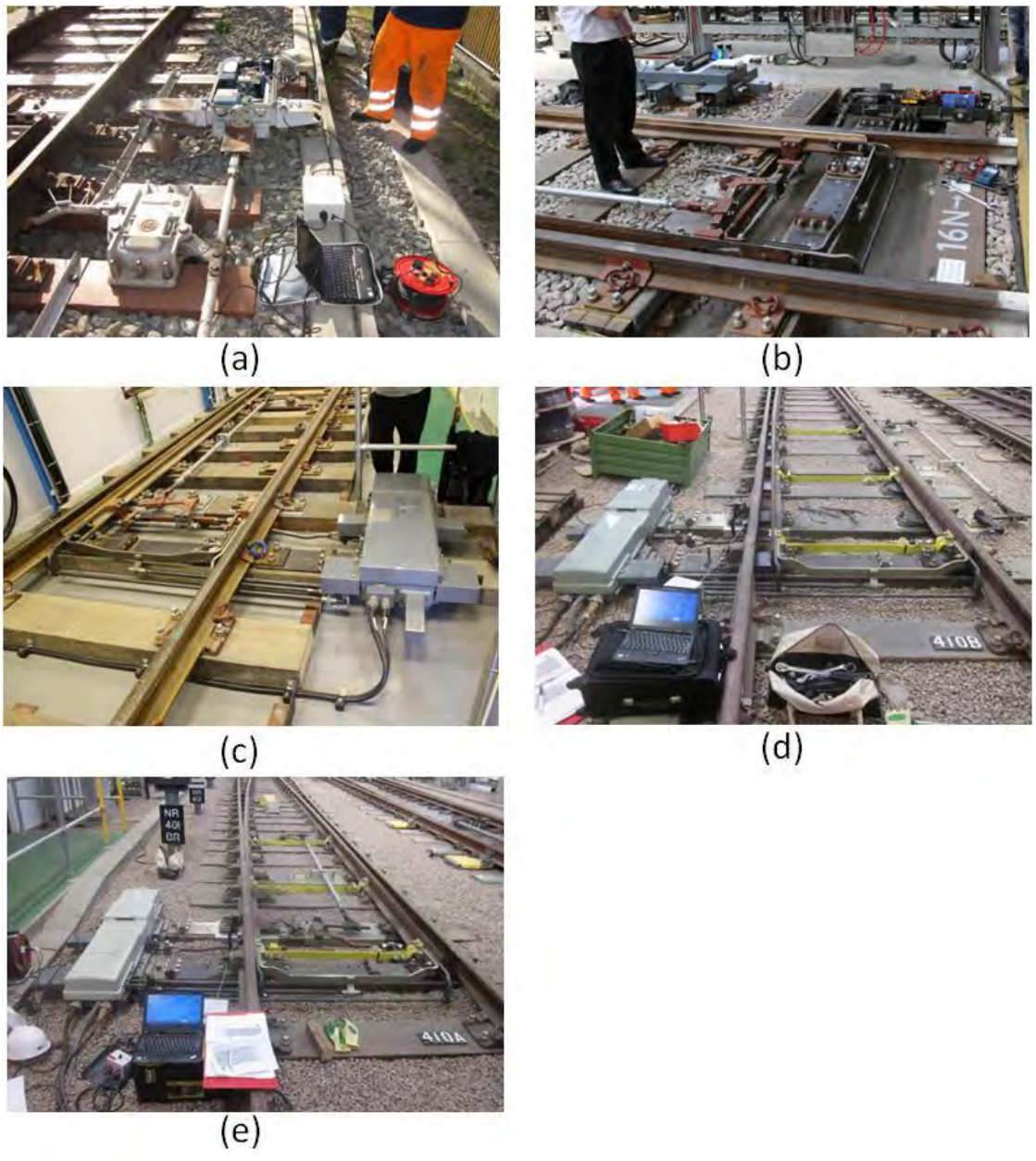

Figure 2-1 Photograph of (a) NTS-type point machine, (b) Surelock-type point machine, (c) M63-type point machine, (d) HW-type point machine 1 and (e) HW-type point machine 2

In order to acquire data from the point machines, a data acquisition box was developed. Figure 2-2 shows a circuit schematic of the data box. Five sensors can be connected to

the data box to acquire electrical current data, displacement data, force data and electrical voltage data. A National Instruments NI-6210 data acquisition unit is installed inside the data box for data collection. All the data were captured at a sampling rate of 10 kHz. (A high frequency rate generally does not harm the data whereas low frequency can harm the data considering “Nyquist” sampling theorem. 10kHz, which is close to the upper limit for the data acquisition box utilised, was therefore used.) Force in the drive rod was measured using a load pin specially designed to fit into point machine. The load pin for the Japanese point machine was manufactured by Strainstall UK Ltd, whereas the load pin for the UK point machine was manufacture by Applied Measurements Ltd. The load pin was fitted to replace a pin in the drive assembly. Electrical current data were collected inside the point machine using a LEM PCM-30 transducer capable of a range of -30 A to +30 A; electrical voltage data were collected using a LEM AV100-500 sensor, capable of a range of -750 V to +750 V.

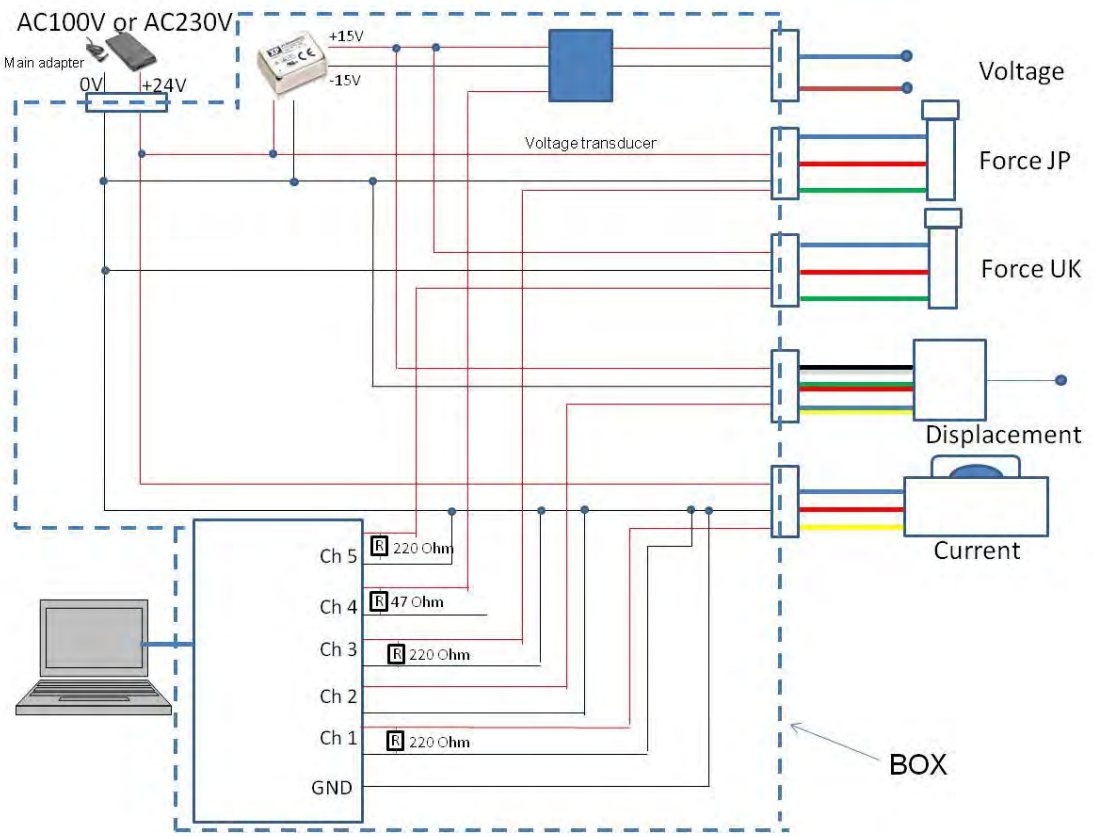

Figure 2-2 A circuit schematic of the data acquisition box

2.2.2 Fault simulations for point machines

In order to help address requirement 1.1 and 1.2, common faults experience in the field for point machines are considered. Generally, there are three types of faults which can affect the operation of point machines [4]:

Abrupt fault – a fault that appears suddenly without any prior indication

Intermittent fault – a fault that appears sporadically

Incipient fault – a fault that develops gradually over a period of time

An abrupt fault is generally difficult to detect in advance because of its inherent nature; a point machine may seem to be operating normally but then suddenly fails without any prior indication of having a fault. An intermittent fault is also difficult to detect in advance because the indication of a fault may disappear at any moment. Conversely, incipient faults can be predicted if the parameters and methodology are adequate in the system.

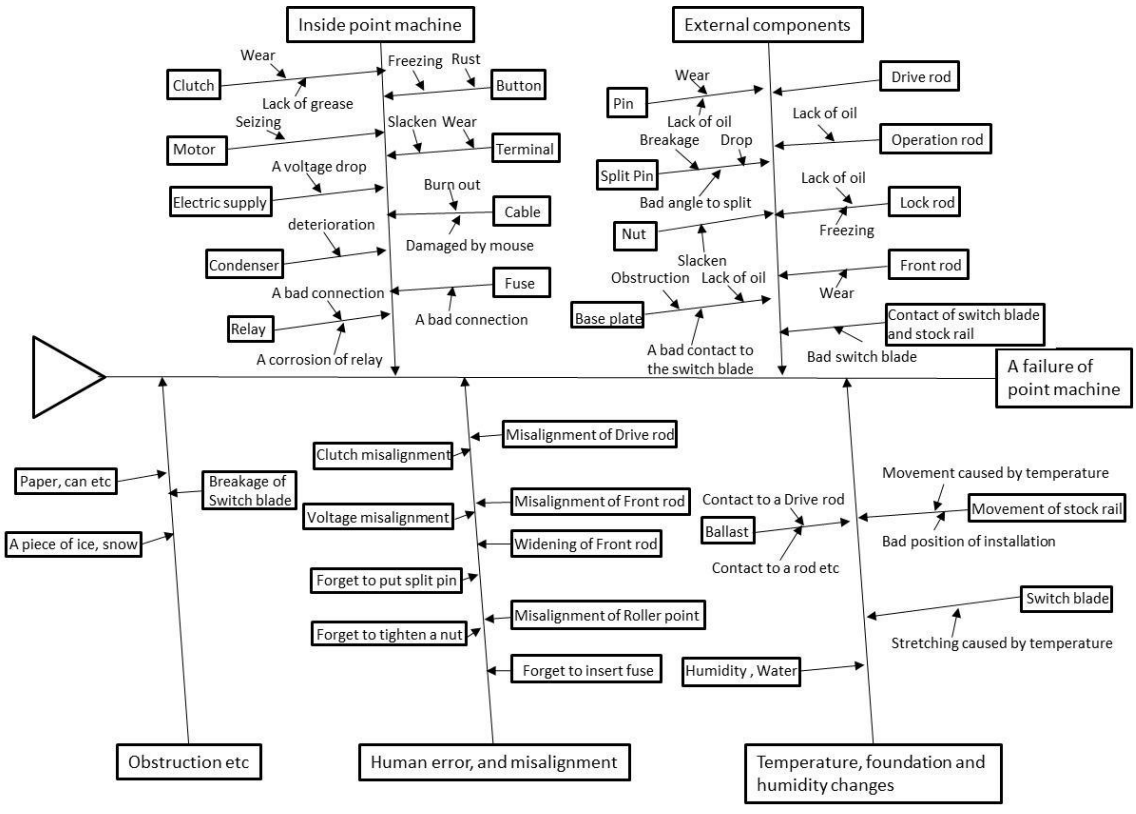

Figure 2-3 Fish bone diagram of faults for point machines [10]

There are a number of faults that will lead to the failure of a point machine. Figure 2-3 shows a fish-bone diagram of point machine failures [10]. The original diagram has been translated into English and faults that are irrelevant to the point machine currently in operation have been removed from the diagram. As shown in Figure 2-3, point machine faults can be categorised as follows.

1. Faults caused by components in the point machine
2. Faults caused by external components
3. Faults caused by human error and misalignment
4. Faults caused by temperature, track foundation movement and humidity changes
5. Faults caused by an obstruction

Faults '1' and '2' can be avoided if adequate maintenance is carried out and the components are replaced at suitable intervals. Since '4' and '5' tend to be either abrupt faults or intermittent faults, they are generally difficult to predict. Fault '3' (faults caused by misalignment) was selected as the fault type that would benefit most from automatic detection and diagnosis by a condition monitoring system.

An example of a point machine installation as used in Japan and Great Britain is depicted in Figure 2-5 and Figure 2-6 respectively.

A Point machine moves the switch blade in the following procedure:

(1) Transmit the signal to turn over the point machine from the signal box to the point machine

(2) Short-circuit the indication relay (which indicates whether the point machine is in 'Normal' position or 'Reverse' position)

(3) Complete the electrical circuit to move the motor

(4) Unlock the point machine with the locking mechanism

(5) Move the drive rod until the switch blade is fully moved

(6) Lock the point machine with the locking mechanism

(7) Break the electrical circuit to move the motor and complete indication circuit

(8) Indication relay operates and the signal indicating the position of the point machine is transmitted from the point machine to the signal box.

A mechanical locking mechanism is implemented inside the point machine, as illustrated in Figure 2-4. A lock dog will be situated inside a notch after the completion of the throw, which mechanically locks the point machine in place.

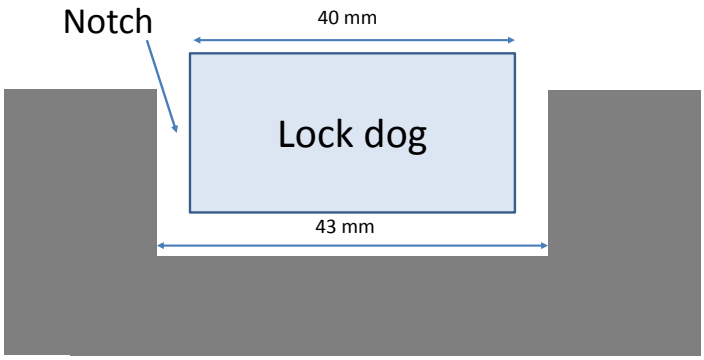

Figure 2-4 A mechanical locking mechanism implemented inside a point machine

Periodic railway maintenance is usually carried out to check the internal lock position and also the force between the stock rail and the switch blade. Since this maintenance is carried out by maintenance staff, there is always the possibility that human error may

occur, and hence misalignment. In Japan, an existing condition monitoring system is used to detect and diagnose a lock position fault, which uses a displacement sensor located inside the point machine [11]. However, there is currently no method to monitor misalignment of the driving rod, that is, the force between the stock rail and the switch blade.

Furthermore, Table 2-2 shows the 'top 5' defects recorded in Network Rail's failure management system for the point machines in use [12]. Considering the table, it has been found that the faults related to drive rod misalignment (HW3 and HW5 in the table) consisted approximately 14% of the total faults [12]. Additionally, the drive rod misalignment can potentially cause severe consequences such as train derailment (the previously mentioned accident at Potters Bar was partly caused by overdriving of the drive rod, stressing and eventually fracturing the lock stretcher bar, which led to the derailment of the train [1]), whereas other faults generally do not cause such severe consequences.

Table 2-2 Defects recorded in Network Rail's failure management system [12]

Defective Subassembly	Defect code	Defect text	Count of fail no	Percentage	Potential Risk
	HW1	T.O.K RIGHT ON ARRIVAL	666	17.02%	-
ROD DETECTOR	HW2	OUT OF ADJUSTMENT	356	9.10%	• Disrupt the operation (Failure to turnover)
ROD DRIVE	HW3	OUT OF ADJUSTMENT	348	8.89%	• Train derailment (Fracture the rod) • Disrupt the operation (Failure to turnover)
FACING POINT LOCK	HW4	OUT OF ADJUSTMENT	245	6.26%	• Disrupt the operation (Failure to turnover)
ROD DRIVE	HW5	OUT OF ADJUSTMENT /GAUGE	196	5.01%	• Disrupt the operation (Failure to turnover)

In this thesis, three fault conditions were therefore considered: (1) fault free, (2) overdriving of the drive rod, and (3) underdriving of the drive rod.

Fault free is the condition where the point machine is functioning within its normal operating conditions.

Overdriving of the drive rod was simulated by adjusting the nut of the drive rod (turning the nut clockwise and therefore increasing the force between the stock rail and the switch blade) which is indicated by the red circle in Figure 2-5 and Figure 2-6 for Japanese and UK point machines respectively. Overdriving is a condition in which the force between the switch blade and the stock rail is over the range of the ideal force (in the fault free condition). As written earlier, this condition led to the Potters Bar accident which was caused by the fracture of the lock stretcher bar [1]. The force can ideally be monitored by inserting the load pin inside the drive assembly, but usually the force will be checked using spanners or other tools, actually opening the switch blade when the switch blade was attached to the stock rail. All the conditions apart from the drive force condition were in an adequate condition when the fault was simulated.

Underdriving of the drive rod was simulated by adjusting the nut of the drive rod (turning the nut anti-clockwise and weakening the force between the stock rail and the switch blade) which is indicated by the red circle in Figure 2-5and Figure 2-6 for Japanese and UK point machines respectively. Underdriving is a condition in which the force between the switch blade and the stock rail is under the range of the ideal force (in fault free condition). The underdriving condition can affect the operation of the train because the point machines may fail to move fully if the underdrving condition is

sufficiently severe. All the conditions apart from the drive force condition were in an adequate condition when the fault was simulated.

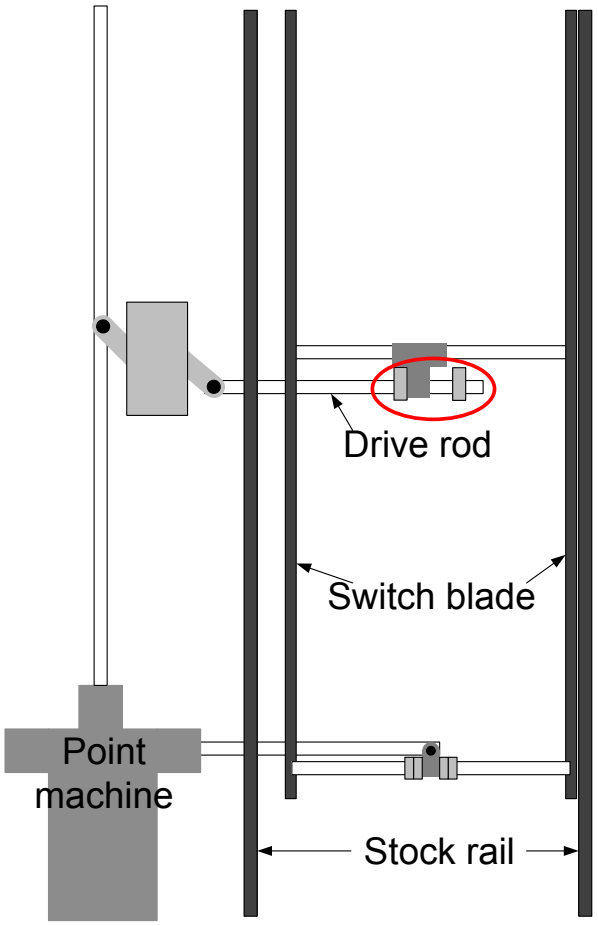

Figure 2-5 A schematic of a Japanese point machine

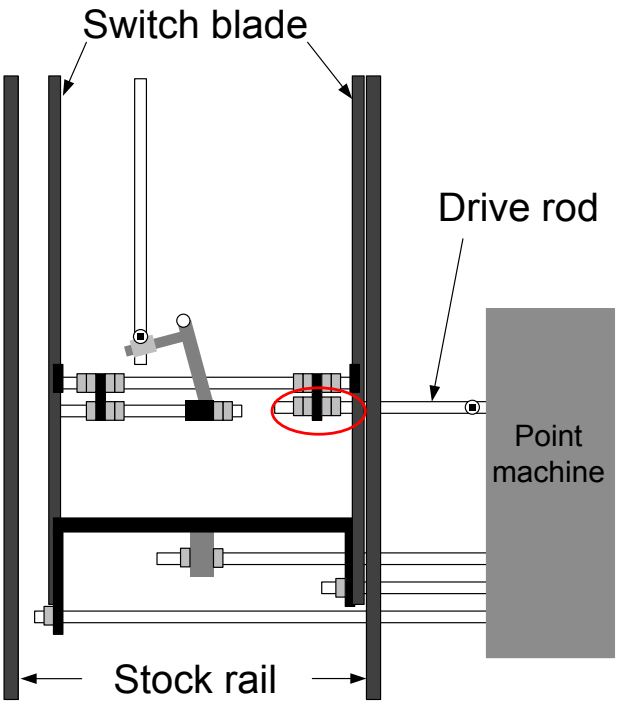

Figure 2-6 A schematic of a British point machine

CHAPTER 3 LITERATURE REVIEW

This chapter provides a review of published literature on fault detection and diagnosis for point machines. Firstly, a general categorisation of the methodology for fault detection and diagnosis is given, followed by a review and categorisation of published literature on fault detection and diagnosis for point machines. Finally, conclusions are drawn from the literature review.

3.1 A general categorisation for fault detection and diagnosis methods

There has been much research into fault detection and diagnosis, particularly in the field of chemical engineering. Venkatasubramanian *et al* categorised various research methods into three classes [13-15]: (1) quantitative model-based methods; (2) qualitative model-based methods; and (3) process history based methods. Table 3-1 shows the categorisation of the various methods for fault detection and diagnosis.

Table 3-1 Categorisation of methods on fault detection and diagnosis [13-15]

Category	Methodology
Quantitative model based	Diagnostic observer Kalman filter
Qualitative model based	Expert systems Fault trees
Process history based	Qualitative trend analysis Neural network Support vector machine

3.1.1 Quantitative model based method

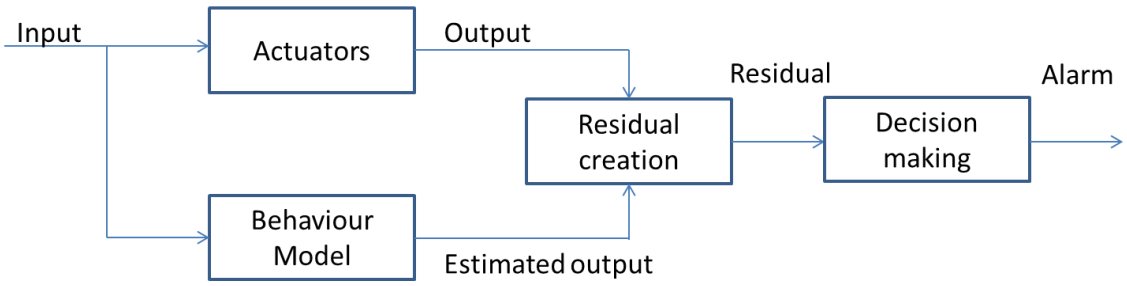

Figure 3-1 Quantitative model based method

The general idea of quantitative model-based approaches is to create a mathematical or physical model that expresses the behaviour of the monitored object based on a fundamental understanding of the process. These approaches typically entail both input and output parameters. A model generates estimated outputs from input parameters and these estimated outputs are then compared to the real outputs. Fault detection is generally carried out in the following steps:

(1) Construction of a mathematical model which estimates outputs from the gained inputs;

(2) Calculate residuals by comparing the monitored outputs and estimated outputs from (1);

(3) Make a decision from calculated residuals (a simple threshold function is used in residual evaluation in most work [13-15]).

Figure 3-1 depicts the block diagram of the quantitative model based method.

Typical methods that are used in the quantitative model based approach include:

- Diagnostic observer, which is 'expressed in state-space equations and generates a set of residuals that detect and uniquely identify different faults' [13]. Frank proposed a solution to the fundamental problem of robust fault detection, providing the maximum achievable robustness by decoupling the effects of faults from each other and from the effects of modelling errors [16].
- The Kalman filter, which is 'equivalent to an optimal predictor for a linear stochastic system in the input-output model' [14]. Chang and Hwang presented a technique using a suboptimal extended Kalman Filter to make the computations required for fault detection simpler [17].

3.1.2 Qualitative model based method

Qualitative model-based approaches are based on a fundamental understanding of the process that is expressed in terms of qualitative functions [14]. Additionally, diagnostic activity generally comprises of two important components: *a priori* domain knowledge and search strategy [14]. *A priori* domain knowledge utilises the empirical knowledge of professional engineers or experts; this knowledge is then expressed by rules. A search strategy will be carried out to search rules for a monitored system. If the rules indicating faulty conditions are found in the monitored system, the system will generate alarms. Typical methods that are used in the qualitative model based approach include:

- Expert system, which is 'generally a very specialised system which solves problems in a narrow domain of expertise' [15]. Wu *et al* proposed an expert fault diagnosis strategy for industrial chemical processing [18]. An example of an expert system for a chemical process is depicted in Figure 3-2. By utilising the knowledge of experts, the expert system can detect and diagnose faults.
- Fault tree, which is 'a logic tree that propagates primary events or faults to the top level event or hazard' [14]. Fault trees determine a fault condition by using layers of nodes which perform different logic operations such as "AND" or "OR". Ulerich and Powers present a method for fault diagnosis in a chemical process using a Digraph and Fault tree [19].

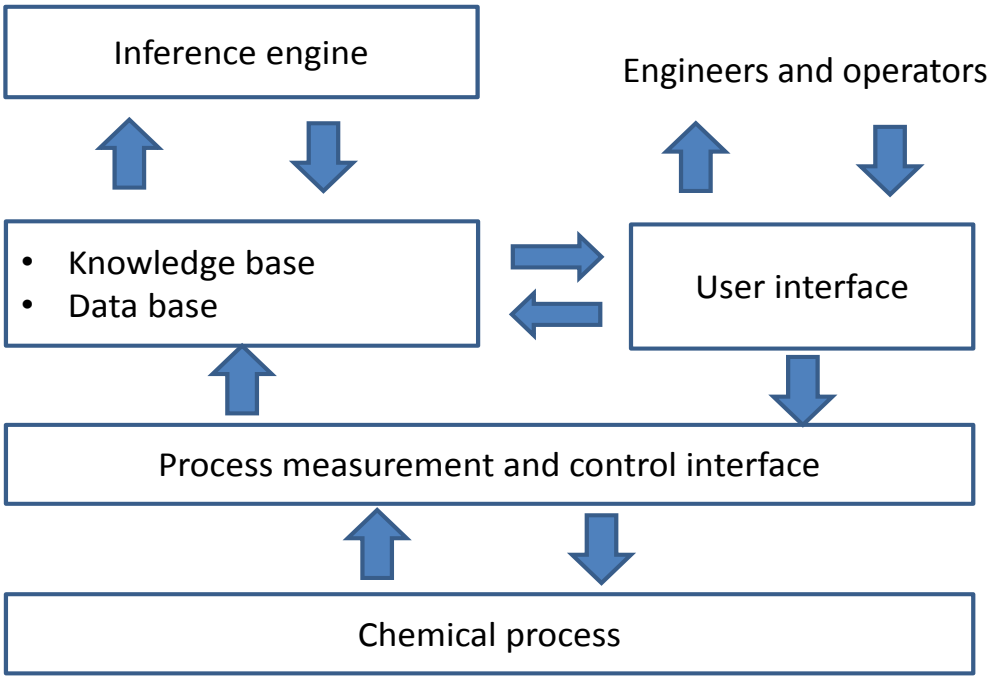

Figure 3-2 An expert system for chemical process [18]

3.1.3 Process history based method

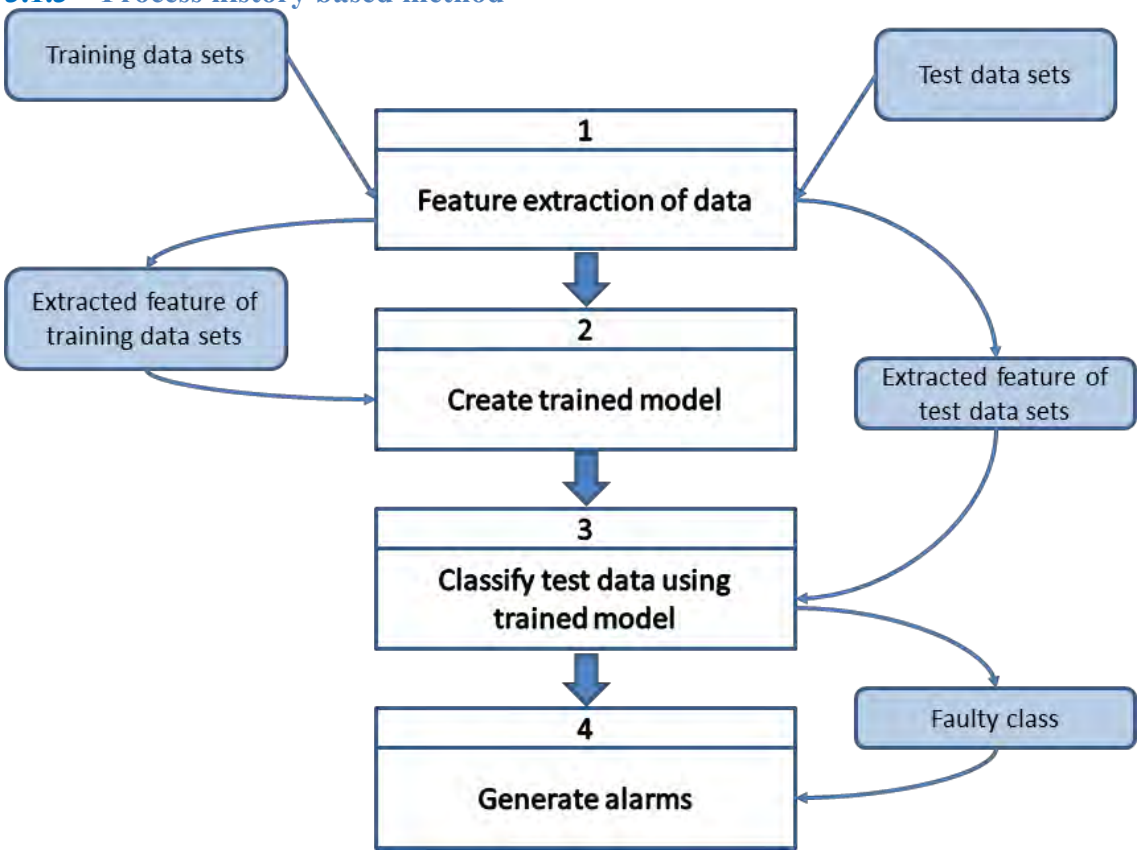

Figure 3-3 Process history-based approach

Process history-based approaches require 'the availability of a large amount of historical process data' [15]. Process history-based approaches collect data from experiments and then this data is transformed and presented as *a priori* knowledge to the system. This process is known as feature extraction. Feature extraction can be either quantitative or qualitative [15]. Figure 3-3 depicts a block diagram of the process history based approach. Typical methods that are used in the process history based approach include:

- Qualitative trend analysis, which uses the abstraction of trend information. Trend modelling can be used to explain the various important events occurring in the process, carrying out malfunction diagnosis and prediction of future states [15]. QTA is often combined with other classification methods. Rengaswamy and Venkatasubramanian proposed a method that utilises neural networks for classification of a waveform into an alphabetic expression [20]. Wong combined QTA with a Hidden Markov model for classification of fault free and faulty waveforms of process data [21].
- Neural network, which is a network of artificial neurons based on human brain. Wang compared recurrent neural networks (RNNs) with neuro-fuzzy (NF) systems, concluding that a properly trained NF system performs better than RNNs [22]. Wang also presented a neuro-fuzzy condition prognostic system for rotary machinery [22].
- Support vector machine (SVM) transfer the acquired data to a higher dimension through the use of a specific function (kernel function). An appropriate kernel function is selected in order to maximise the margin between the faulty data sets

when separated by a hyper-plane (border) in the higher dimension. Ge proposed a fault diagnosis method using the SVM for a sheet metal stamping operation [23].

3.2 Research in condition monitoring of point machines

Much research related to point machines has already been carried out. This section provides a review and classification of research related to railway point machines. Table 3-2 shows research related to point machines classified by research methods. It shows that much research is based on simple thresholding techniques: techniques setting a threshold to a monitored parameter to detect a fault or failure. Research using advanced techniques, such as neuro-fuzzy and qualitative trend analysis, have recently been presented.

Different researchers have acquired and used different parameters for condition monitoring of point machines. In Figure 3-4, the point machine is represented by a black box model; using this model it can be seen that the parameters acquired for the condition monitoring of point machines can be categorised into four classes: inputs, external influences, internal measures and outputs. Internal measures are generally only measured in specific types of point machine where it is possible to add sensors within the main actuator, for example: the pneumatic point machine and the hydraulic point machine. Table 3-3 shows research related to points classified by the class of parameters and acquired parameters of the system. It is shown that parameters such as force, current, voltage and position are commonly used for point machine research.

Table 3-2 Research of point according to method

		Point mechanism					
Method classification	Method	DC Electric-motor	AC electric-motor	Pneumatic-motor	Hydraulic-motor	Unknown	Synthetic data
Quantitative model-based	Thresholding technique	[24] [25] [26] [27]	[11] [28]	[29]			
	Polynomial curve-fitting					[30]	
	Time domain analysis					[30]	
	Regression modelling	[31]			[31]		
	Moving average filter	[32]					
	Kalman filter	[32, 33] [34, 35] [36] [31, 37]			[31]		
	Unobserved component model	[34-37]					
	H2 norm	[38]					
Qualitative model-based	Binary decision diagram					[39]	
Process history-based	Statistic parameters	[40]					
	Wavelet transform	[40]					
	Principal component analysis	[40]					
	Transient analysis	[41]					
	Spectral analysis	[41]					
	Net energy analysis	[42] [41]					
	Mixture discriminant analysis		[8] [43]				
	Expectation maximisation		[8] [43]				
	Qualitative trend analysis	[4] [44]	[45]				[46] [47]
	Neuro fuzzy (ANFIS)			[48] [49]			
	State automata	[48]			[48]		

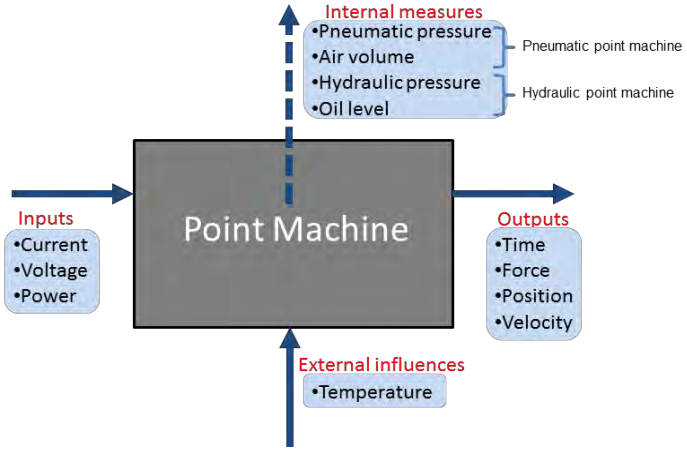

Figure 3-4 Model representation of a point machine

Table 3-3 Research of point according to parameter

Parameter Classification	Parameter	DC Electric-motor	AC Electric-motor	Pneumatic-motor	Hydraulic-motor	Unknown	Synthetic data
				point mechanism			
Inputs	Current	[32] [33] [34, 40] [41, 42] [31] [37] [4, 38, 48] [44] [25-27]	[45]		[48]	[30]	
	Voltage	[32] [41, 42] [24, 38] [25-27]					
	Power	[41, 42]	[8] [43] [28]				
External influences	Temperature	[25-27]					
Internal measures	Pneumatic pressure			[29] [48] [49]			
	Air volume			[48] [49]			
	Hydraulic pressure				[48]		
	Oil level				[31, 48]		
Outputs	Time	[31, 35, 36]		[29]			
	Force	[32, 34-36, 40] [31, 37] [48] [4, 25-27] [44]	[45]		[48]		
	Position	[42] [41] [4, 25-27] [44]	[11] [45]	[29] [48] [49]		[30]	
	Velocity	[42] [41]		[48] [49]			

3.2.1 Quantitative model-based approaches for railway point machines

Numerous point condition monitoring research projects have been carried out to date using thresholding techniques. This research can be categorised as the simplest of the quantitative model-based approaches, assuming that all of the output signals should be smaller than the pre-determined thresholds in healthy conditions. All techniques (including process history methods) eventually use a threshold of some sort in the final decision making step, but thresholding techniques defined here simply means the techniques in which a fault detection of the system was carried out by defining a threshold of the monitored parameter and creating an alarm when the parameter exceeds a predetermined threshold.

Shaw used thresholding techniques for DC electric point machines and pointed out the necessity of a universal approach to condition monitoring [24]. It was found that failure detection of the point machine can be achieved by monitoring the detection voltage.

Zhou *et al* used an array of sensors to monitor all relevant parameters of the point machine. Thresholding techniques are used to create alarms in the system [25-27]. Various parameters are monitored such as force, electrical current, displacement, voltage and temperature.

Igarashi and Shiomi proposed a magnetic sensor to detect lock warp on an electric point machine. This sensor can monitor a slight lock warp displacement according to the variation of temperature. This system also generates an alarm if the lock warp is larger than the given threshold [11]. Since a displacement sensor dedicated to detect lock warp

is implemented inside the point machine, this method can accurately detect and diagnose lock position faults.

Pabst used a thresholding technique for the AC electric point machine by considering the power over time [28]. It was found that the power changed significantly when there was an obstruction inside the stock rail and the switch blade, and also when a poor lock adjustment occurred.

Abed *et al* proposed a PC-based condition monitoring system based on a thresholding technique for a pneumatic point machine [29]. This system creates a position profile model by simply averaging ten consecutive operations of the point machine. An alarm is generated if the sum of absolute errors calculated by comparison exceeds a pre-set margin. Since a failure or fault did not actually happen during the monitoring period, the effectiveness of the system is unknown.

There is also other research using quantitative model-based approaches which do not use thresholding techniques.

Rouvray *et al* presented a polynomial curve-fitting technique and time domain analysis (ARMAX) modelling techniques using Matlab and Simulink to model signals of point machines [30]. It was found that, due to the non-linearity, the use of polynomial curve fitting did not achieve a sufficient level of accuracy to model the position signal.

Marquez *et al* presented a method which uses three steps for detecting a fault signal, comparing the reference signal data and actual data: (1) detection of irregularity in the curves, (2) checking the maximum position of the curve and (3) checking whether the curve is symmetrical with respect to the position of the maximum with a margin of a

given width [35, 36]. It was found that a good result in terms of fault detection was achieved employing a Kalman filter to the original signal with this approach.

Marquez and Pedragal used an Unobserved Components model, set up in a State Space framework. The detection of faults is carried out based on the correlation estimate between a curve free from faults with the current curve data [34, 37]. The method is reported to detect faults to a high degree of accuracy.

Pedragal *et al* also proposed a method using state space models for predicting the throw times of points, and created estimated shapes using harmonic regression based on calculated throw times. These estimated shapes are then compared with actual data and generate an alarm if the standard deviation of errors exceeds the pre-determined limit [31]. The method is reported to detect faults to a high degree of accuracy.

Later, Marquez presented a method using a Kalman filter for filtering the current curve, and similarly, a moving average filter for noisy signals [32, 33]. It was found that by using these filters, faults could be detected accurately. It was showed that a moving average filter outperformed a Kalman filter in terms of fault detection accuracy [32].

Zattoni proposed a method to detect incipient failure using residuals calculated from H2 norm. Experimental results show that this method can detect a silted bearing error [38]. Whether this method can be applied to other types of faults is yet to be examined.

3.2.2 Qualitative model-based approaches for railway point machines

There has been very little research based on purely qualitative model-based approaches to date. This is because most of the approaches which use qualitative model-based approaches in part are categorised as process history-based approaches.

Marquez proposed binary decision diagrams for remote condition monitoring and a case study for point mechanisms in railway systems has been analysed using this method [39].

3.2.3 Process history-based approaches for railway point machines

Most of the recent research is focused on the process history-based approach. Approaches range from a method using statistic parameters to a method using a neuro-fuzzy network.

McHutchon *et al* used statistic and geometric parameters, a wavelet transform and Principal Component Analysis (PCA) to classify nine different faults and concluded that applying statistical parameters to the decomposed wavelets gives a good degree of clustering [40].

Oyebande and Renfrew made analysis of six statistical parameters to discriminate between the performance of a DC electric-point machine under faulty conditions: (1) transient analysis of current waveforms, (2) spectral analysis of current waveforms, (3) transient analysis of position waveforms, (4) analysis of throw times, (5) analysis of end-of-stroke positions and (6) net energy analysis [41, 42].

Chamroukhi *et al* presented a method, based on Mixture Discriminant Analysis (MDA) and Expectation-Maximisation (EM) algorithms, which utilise seven statistic parameters acquired from the AC electric point machine, and classifies the signals into three classes: class without defect, class with minor defect and class with critical defect [43]. They also proposed a method for modelling a signal by using a regression approach, and classified signals similarly using MDA and EM algorithms utilising estimated parameters acquired from the regression approach as the feature vector [8].

Roberts *et al* proposed a method using a neuro-fuzzy network. They created mathematical models that estimate the outputs from acquired inputs, and calculated

residuals by comparing actual data and estimated outputs. These residuals were then used as inputs to a neuro-fuzzy network for classifying the process faults [48, 49]. In the research, the seven residual outputs are used as inputs for the ANFIS. After an adequate amount of training using training data sets (each data set comprising the seven residual inputs and associated fault code), the ANFIS was able to diagnose faults. An example of a two input, single output ANFIS is depicted in Figure 3-5.

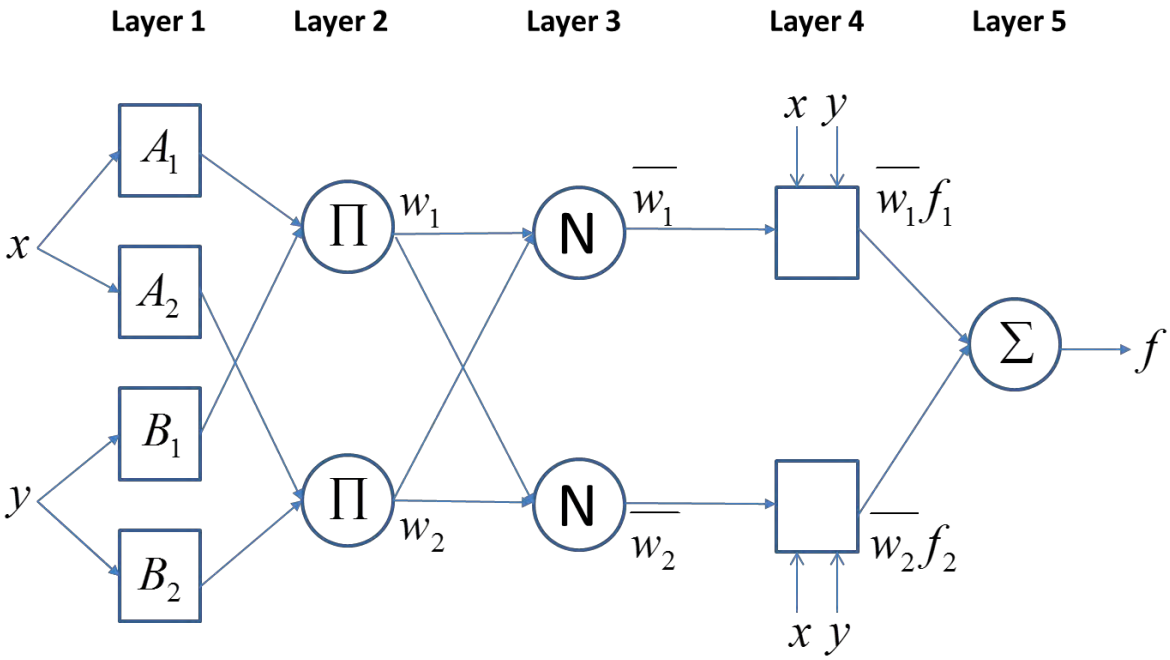

Figure 3-5 An example of ANFIS

Silmon and Roberts presented a new method to detect incipient failure based on Qualitative Trend Analysis (QTA). In this method, the shapes of waveforms that are common to all fault conditions, called "common episodes", are found based on QTA, and a signal is classified using a fuzzy rule based on the difference of the value of "common episodes" between fault free and fault signals [46, 47]. The QTA can express

the rough shape of the waveforms using alphabetic characters. An illustration of QTA used in the method is depicted in Figure 3-6.

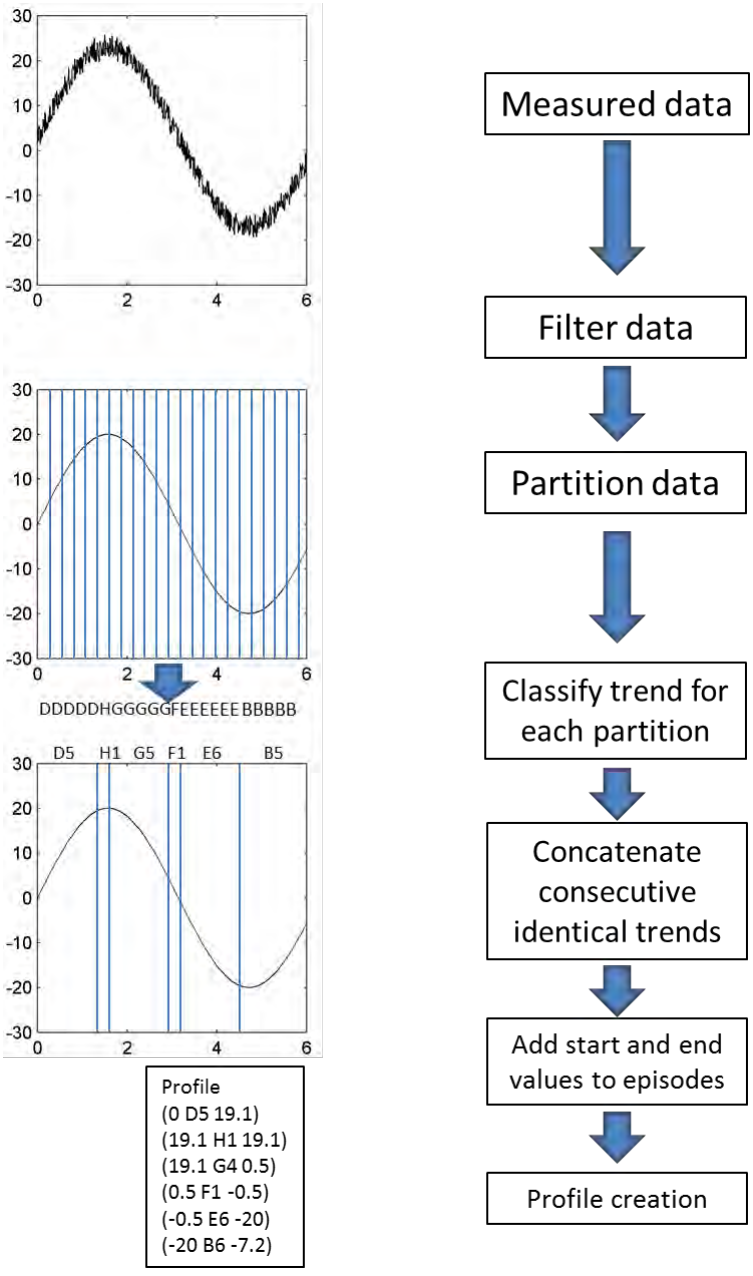

Figure 3-6 An illustration of QTA [4]

3.3 Conclusion from literature and future work

A thorough literature review has been carried out on condition monitoring research for point machines. The research methods have been categorised into three categories: Quantitative model-based methods, Qualitative model-based methods and Process history based methods. It has been found that most of the research was focused on either Quantitative model-based methods or Process history based methods. The 'categories' of fault detection schemes are not mutually exclusive. A combination of methodologies is likely to be applied in practice.

A significant amount of research on the Quantitative model-based method has been carried out using the thresholding technique. The problem with the thresholding technique is that often it cannot detect subtle changes of waveforms, therefore a system which utilises this technique can only detect failure, or faults that are very close to failure [4]. Other research in this category utilised state space modelling, such as the Kalman filter. These techniques appear to detect faults to a high degree of accuracy but do not diagnose faults, limiting their usefulness. If the system can diagnose faults, the maintenance staff can fix faults quickly because they know what is actually wrong with the point machine in advance.

A more sophisticated method that can detect and diagnose the subtle changes of waveforms is therefore required.

As for the Process history method, recent research has been carried out using techniques that are also used in the pattern recognition field, such as Mixture discriminant analysis, Neuro-fuzzy, Qualitative trend analysis and Fuzzy logic. By using these methods, it

may be possible to detect and diagnose faults. These methods would potentially meet the requirements discussed in Chapter 2 of this thesis since methods can carry out not only fault detection but also fault diagnosis. Among the previous research utilising process history, only the research carried out by Roberts and Silmon enabled both fault detection and diagnosis which was one of the key requirements discussed in Chapter 2. Consequently, the methods presented by Roberts and Silmon will be further examined in the rest of this section.

Roberts *et al* used neuro-fuzzy technology, an extension of neural-networks, for classifying process faults of the point machine [48, 49]. An Adaptive Neuro-Fuzzy Inference System (ANFIS), which is proposed by Jang [50], was used in the method.

The research acquired good results, and fuzzy rules created by the system can be transferred between the same type of point machine. The disadvantage of the research, however, is that the method entails both input and output parameters to create residuals. In many railway fields, such parameters are not readily available because the implementation of many sensors would be necessary to acquire the parameters, which would cost a lot of money and labour. The introduction of many sensors also potentially introduces additional failure modes; a breakage of a sensor can adversely affect the operation of point machine. Furthermore, a large amount of training data is needed to train the model since the method requires a significant number of parameters to be optimised. As a consequence, a research method which only uses parameters commonly available in point machines and also only requires a small amount of data for training will be necessary in this thesis.

Silmon presents research that uses parameters such as time, current, position and force [4, 45-47] for electric points. Interestingly, his research method can be used for any parameters acquired because the method is generic. This research utilises 'Qualitative Trend Analysis (QTA)' for feature extraction, and 'Fuzzy logic' for classification. In addition, since the research method is generic, it can be applied to different equipment such as train doors and level crossings.

However, this research method has two disadvantages. The first is that it utilises a number of filters so that the original waveform changes dramatically [4]. Without this filtering, the output trend from the QTA can fluctuate and be difficult to interpret afterwards. These filters may remove significant features from the monitored waveform. The second disadvantage is that through the use of QTA, which classifies a partition of the original waveforms into only nine characters by the first and second differentials, the original information of the waveforms is reduced significantly in a discrete form, as will be demonstrated in the following chapter. The classification result can be further improved if an advanced pattern recognition method is implemented instead of using simple fuzzy logic. A feature extraction method which expresses the original data more precisely and a pattern recognition method that can classify the data in high accuracy will therefore be required in this thesis.

To summarise, the method that meets the following statements will be needed for this thesis.

(1) A feature extraction method which represents the original shape information of the

waveforms in a reduced number of samples.

(2) A method that implements a state-of-the-art pattern recognition method to diagnose fault conditions.

Condition monitoring systems for railways are still at an early stage of development; there is not yet a fixed way to accomplish these requirements. There is a need for the new area of work to create effective condition monitoring systems for point machines. By using an effective process history method, there is potential to develop practical solutions for condition monitoring.

The following chapters consider railway point machine case studies. By analysing acquired data the methodology which meets these requirements will be proposed.

CHAPTER 4 DEVELOPMENT OF AN ALGORITHM FOR FAULT DETECTION AND DIAGNOSIS

4.1 Introduction

In this chapter, an algorithm for fault detection and diagnosis is developed using the data collected from an NTS-type point machine. Drive force, electrical current and electrical voltage data were acquired using a data acquisition box whilst simulating three fault conditions: 'Fault free', '(Left-hand) Overdriving' and '(Left-hand) Underdriving' (as discussed in Chapter 2).

4.2 Parameter selection

Figure 4-1 shows the drive force, electrical current and electrical voltage data (three data sets per fault condition) where each plot shows one throw of the point machine ('Right to Left' throw). Electrical current and electric voltage data were converted to root mean square (r.m.s.) format after capture, as this is a more appropriate for analysis in a fault diagnosis system; using raw AC data, it is generally difficult to interpret due to data fluctuation. The sum of the square of the current (and voltage) over one time period

(20ms) was calculated, and it was divided by one time period (window size). Then, the square root was calculated.

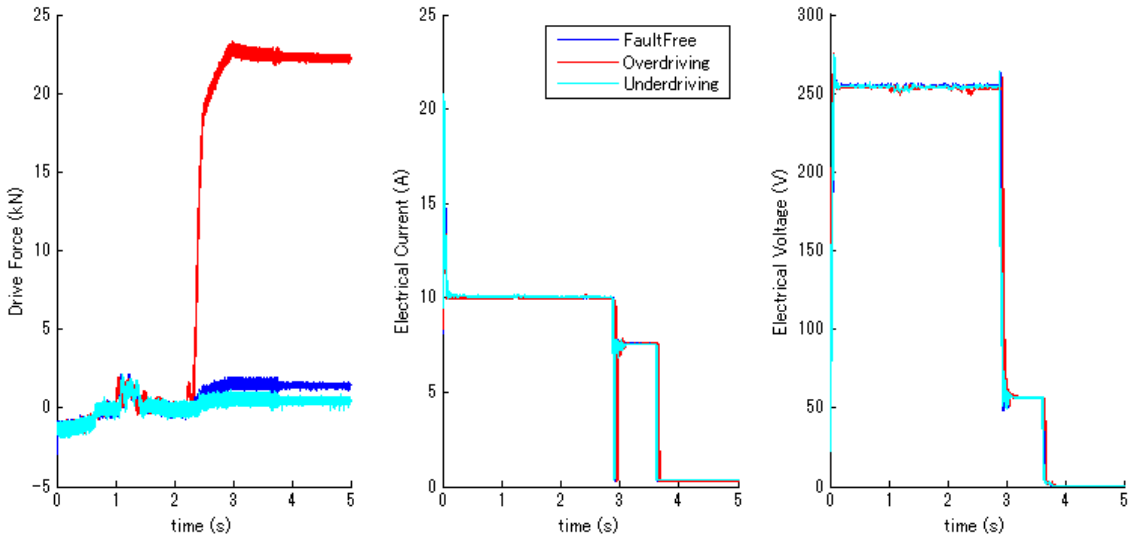

Figure 4-1 Waveforms acquired during point machine operation (Right to Left): (a) Drive Force, (b) Electrical Current, and (c) Electrical Voltage

As can be seen from Figure 4-1, the force plot clearly shows a distinction between different fault conditions, whereas the electrical current and electrical voltage data plots do not appear to show any visible distinction between different fault conditions. The drive force graph in the figure shows that the operation to turn over the point machine ends within 3 seconds, as there are no changes of the force after 3 seconds. The electrical circuit to move the motor was broken at 3 seconds, at this point the brake circuit is applied and the rotational energy of the motor is consumed by the brake circuit, producing voltage and current. After 3.6-3.7 seconds, the point machine returns to a static state.

A cluster analysis is one of the techniques used in data mining to categorise data into subsets (or clusters). Here it is used to cluster the acquired data associated with different fault conditions. The output of the cluster analysis can be used to evaluate the appropriateness of a parameter (electrical current, electrical voltage and electrical power) for condition monitoring. Appropriate parameters are those that have a clear distinction (i.e. large distance) between different fault conditions and are similar (small distance) for the same fault condition. If the parameter is appropriate for condition monitoring, the data (including various fault conditions) should divide in to clusters where each cluster contains the data with the same fault conditions after cluster analysis.

A cluster analysis using the k-means method [51] was carried out with the force, electrical current and electrical voltage data to investigate which parameter would be the best to use for a condition monitoring system. The k-means method is explained in detail by Han [52].

(1) The k-means method randomly selects k objects, and these k objects represent cluster centres.

(2) Each remaining object is assigned to the cluster which is the most similar based on the distance between the cluster centre and the object (after calculation of distance).

(3) After assigning all the objects to clusters, the k-means algorithm computes the new cluster centre (called as centroid) using the assigned objects.

(4) Each object is assigned again based on the distance between the new cluster centre (centroid) and the object.

(5) (3) and (4) continue iteratively until the new cluster centres (centroid) calculated in the current iteration are the same as those calculated in the previous iteration [52].

A subset of the total data set (10 data per each fault condition) has been used to carry out the cluster analysis. This subset is sufficient to characterise the data, while retaining unseen data in the test data set to be used later in the experiment. The centroid of the cluster was calculated by mean and the squared-Euclidean distance was used to choose the points for clusters; 10 data per fault condition were used. Figure 4-2 shows the result of the k-means clustering for force, electrical current and electrical voltage data.

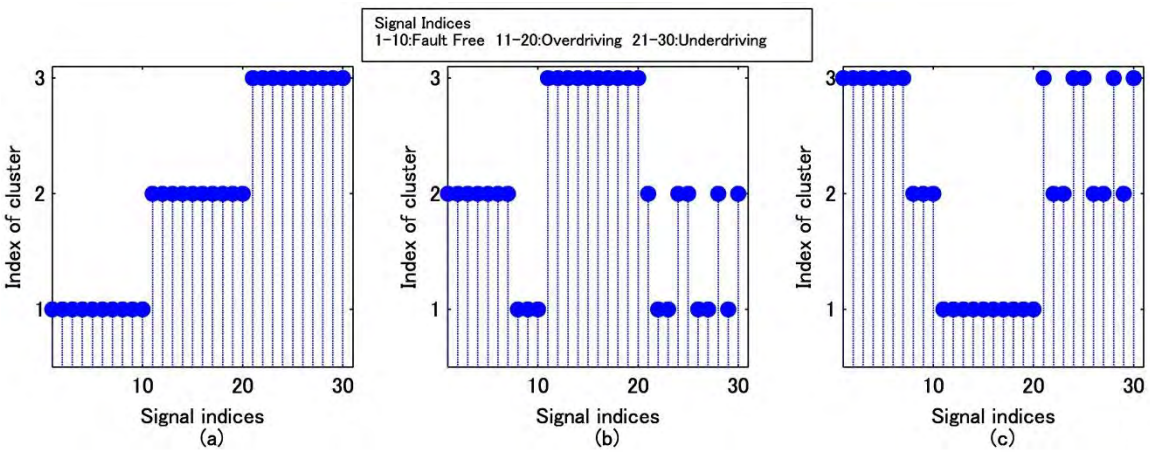

Figure 4-2 Cluster analysis for (a) Drive Force, (b) Electrical current, and (c) Electrical Voltage

It can be seen that the force data were divided clearly by fault conditions, whereas electrical voltage and electrical current were not. Although the force data would give a good result in terms of fault diagnosis, in a practical condition monitoring system it would be more difficult to acquire this data, as it requires the additional cost of the load

pin and the necessity of the pin to be fitted into the drive assembly. The introduction of the load pin also potentially introduces additional failure modes. The use of a load pin was therefore discounted for practical application as the parameters of the condition monitoring system should be acquired using sensors that will not directly affect the operation, as per requirement '5'.

Figure 4-3 shows an electrical active power waveform, calculated using the electrical current and voltage data. It can be seen that the middle part of the electrical active power shows a distinction between different fault conditions, whereas the beginning and ending part of the waveform does not show a distinction between different fault conditions.

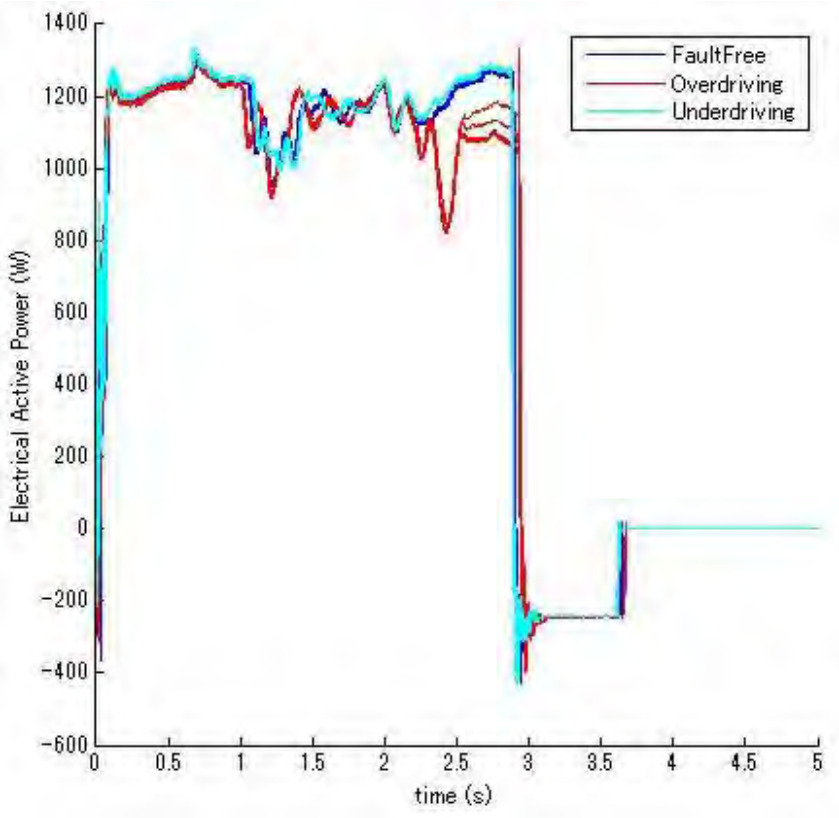

Figure 4-3 Electrical active power data for an AC point machine (Right to Left operation)

Figure 4-4 shows an electrical active power waveform removing the beginning and ending of the waveforms. Electrical reactive power and electrical apparent power were also calclulated but neither showed clearer distinction between different fault conditions compared to electrical active power.

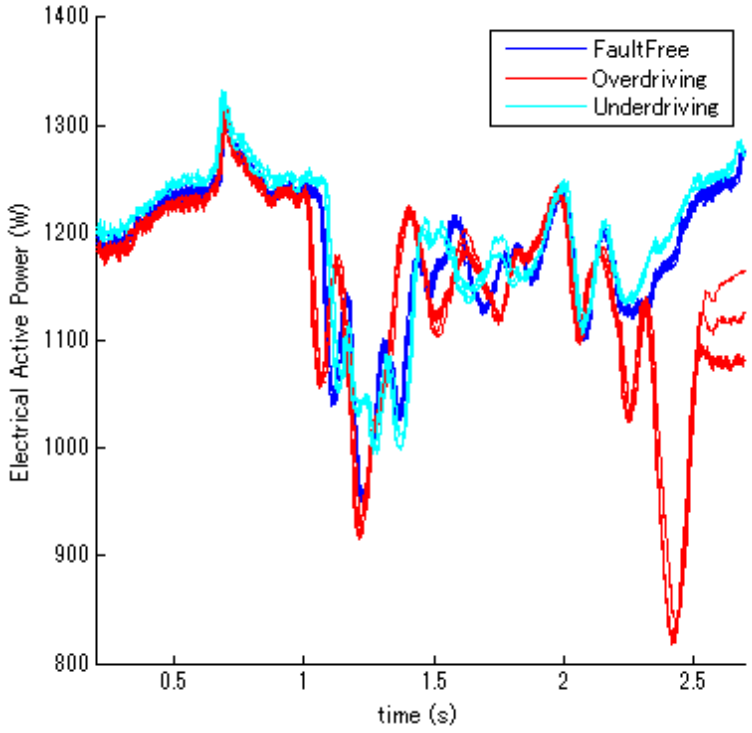

Figure 4-4 Electrical active power data for an AC point machine (Right to Left operation removing the beginning and ending)

Figure 4-5 shows the result of cluster analysis carried out on the electrical active power data.

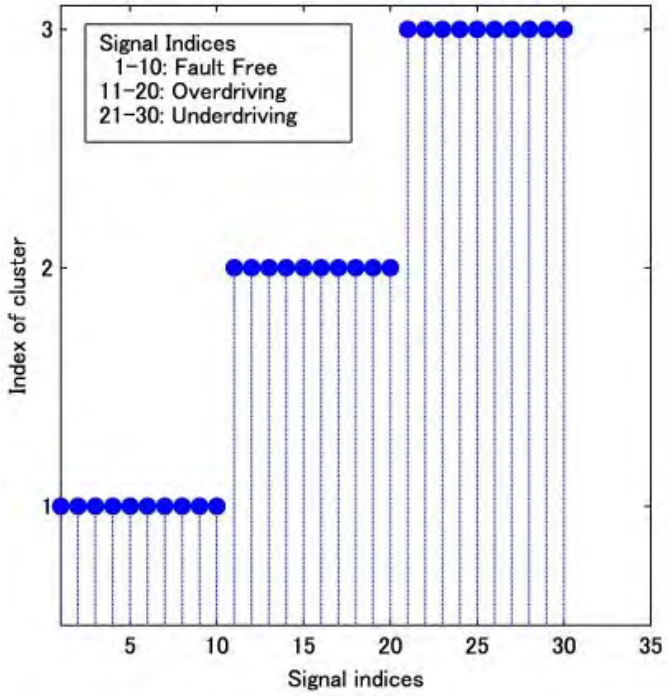

Figure 4-5 Cluster analysis for electrical active power data

It can be seen that the clusters are clearly divided by fault conditions, indicating that electrical active power can be used as a parameter for fault detection and diagnosis.

To investigate how well the cluster has been divided for force and electrical active power data after k-means clustering, the silhouette width [53] was calculated. Silhouette width is one of the methods (called intrinsic methods) to assess the clustering quality. By using silhouette width, it is possible to evaluate a clustering by examining how well the clusters are separated and how compact the clusters are [52]. Equation (4-1) shows the formula used to calculate the silhouette width, sw_i.

$$\mathrm{sw}_i = \frac{(b_i - a_i)}{\max(a_i, b_i)} \qquad (4\text{-}1)$$

where a_i is the value which reflects compactness of the cluster to which i belongs and b_i is the value which reflects the degree to which i is separated from other clusters [52].

Figure 4-6 shows the silhouette width for force and electrical active power.

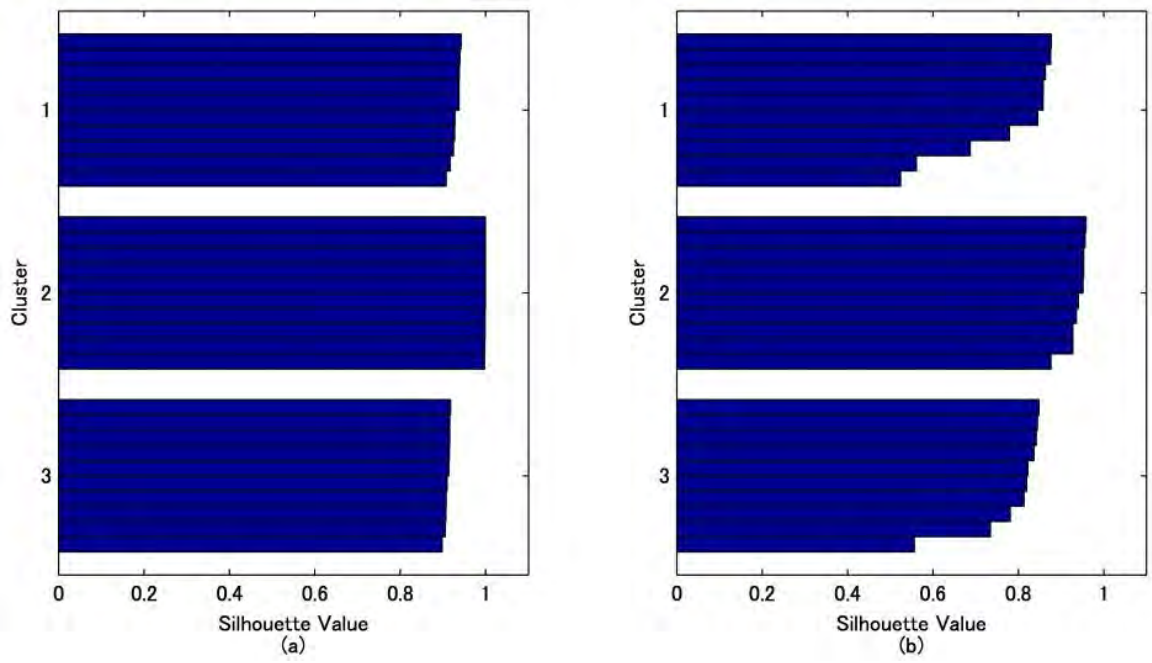

Figure 4-6 Silhouette width for (a) Drive force and (b) Electrical active power

From Figure 4-6 it can be seen that the force data were divided more clearly than the electrical active power data. The mean silhouette width gives an indication of how well the clusters divided. The mean silhouette width for the force data was found to be 0.946, whereas for the electrical active power it was 0.833. From the calculation, it can be seen that the force data has more useful information than the electrical active power data, however, as previously discussed, electrical active power is more practical to acquire.

In summary, both force data and electrical active power have potential to be used as parameters for fault detection and diagnosis. Cluster analysis has been carried out to verify this hypothesis. Although the force data contains more useful information than

the electrical active power data, using electrical power is the only feasible option because acquiring the force data is impractical.

4.3 Proposed method

4.3.1 Feature extraction

The active power data is not easy to handle by a classifier without feature extraction, as the data size is too large. A method which extracts the features of the original waveform is therefore required. Previous researchers have used Fourier analysis [54] and QTA [44] to carry out this operation. These feature extraction methods will be demonstrated and validated.

4.3.1.1 Fourier analysis

Previous work was carried out using the spectral features, derived from a discrete Fourier transform [54]. It was found that this approach is effective for certain types of train door faults, but it is not known whether this approach can be effective for incipient faults of point machines. A discrete Fourier transform of electrical active power (three fault conditions) was carried out. It has been founded that the main features were lost from the original waveforms after the discrete Fourier transform. A more effective feature extraction method that retains the original feature is therefore required for this thesis.

4.3.1.2 Qualitative trend analysis (QTA)

Qualitative trend analysis was used for incipient faults of point machines by Silmon [44]; the same feature can potentially be used in this thesis. Since the parameters used in the previous work were different from those in this thesis, it is not known whether this feature extraction method can still be applied. Figure 4-7 shows graphically how values

are chosen and stored from the k'th partition. Then, qualitative classification is carried out using the acquired values (x, y, $\dot{x}$, $\dot{y}$) [44]. Table 4-1 shows the rules determining the qualitative state of the partition.

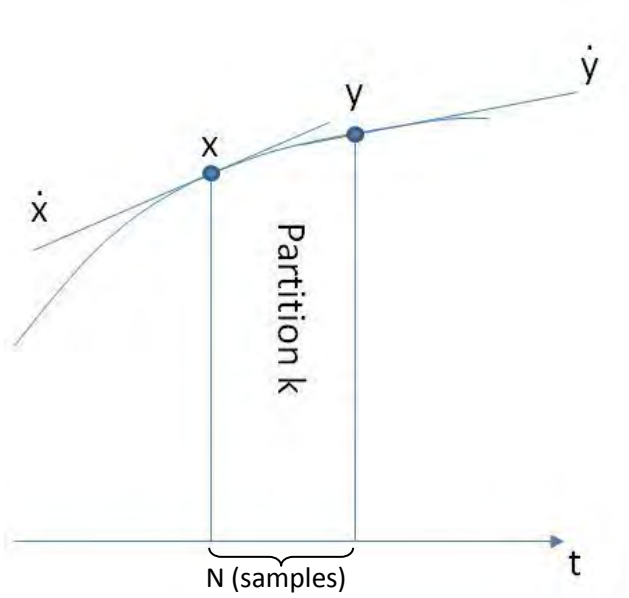

Figure 4-7 Assignment of values in the partition k [44]

Table 4-1 Table of criteria for deducing the qualitative state of a partition [44]

Qualitative state	$[y - x]$	$[\dot{y} - \dot{x}]$	$\text{sign}(\dot{x} \times \dot{y})$
1	[0] (<e1)	[0] (<e2)	any
2	[+]	[+]	[+]
3	[+]	[0] (<e2)	[+]
4	[+]	[−]	[+]
5	[−]	[+]	[+]
6	[−]	[0] (<e2)	[+]
7	[−]	[−]	[+]
8	Any	[−]	[−]
9	Any	[+]	[−]

Figure 4-8 shows the result of QTA for two data sets (each data set contains 'Fault free', 'Overdriving' and 'Underdriving' data) after filtering the electrical active power using a

moving average filter (without filtering, the result will fluctuate and be difficult to interpret). Since the original waveform is divided into only nine discrete states, it can be seen that the original information of the waveforms is significantly lost. It can be seen that the waveforms are not consistent for the same fault condition, making it difficult to use as a feature extraction method. The result may change depending on the filtering method and the parameters such as N (number of samples per partition), e1 (threshold for first differential) and e2 (threshold for second differential). There is a need to adjust these parameters depending on monitoring equipment and parameters. A feature extraction method that makes features consistent in the same fault condition and does not require any adjustment of parameters is required for this thesis.

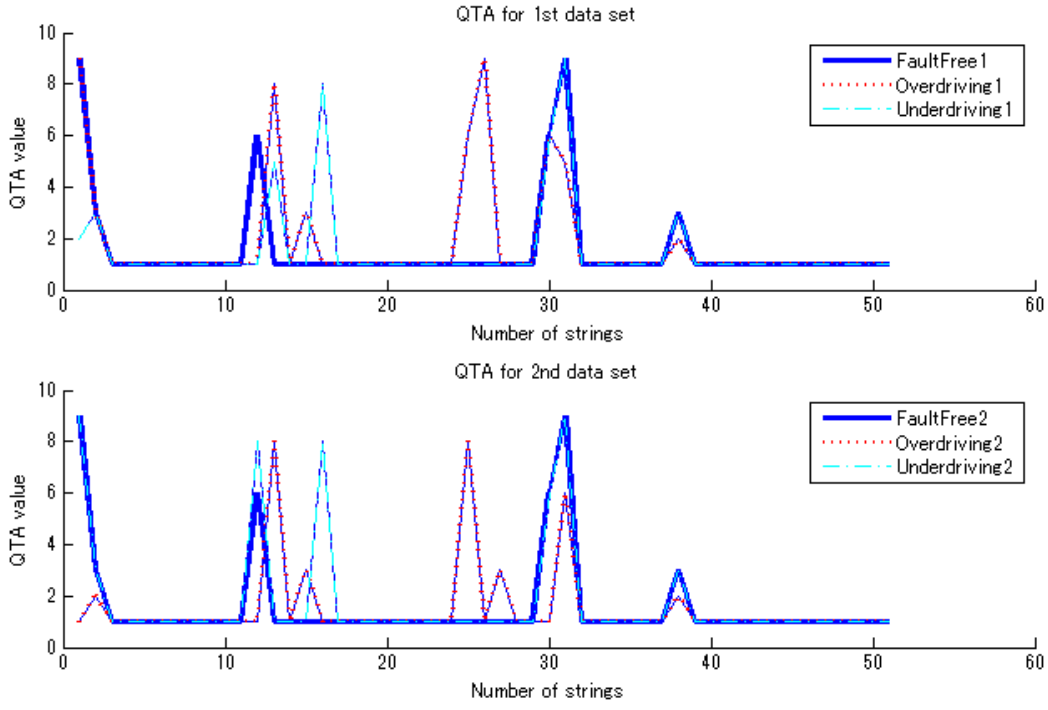

Figure 4-8 Generation of qualitative strings from electrical active power

4.3.1.3 Discrete Wavelet Transform

One method which does retain the original information is the Discrete Wavelet Transform (DWT) [55]. By using DWT, the original waveform is transformed to multiple levels of resolution, sustaining local (time) information in each level of resolution. The DWT decomposes a discrete signal into two sub-signals of (approximately) half its length. One sub-signal is a running average or trend (referred to as scaling coefficients); the other sub-signal is a running difference or fluctuation (referred to as detail coefficients) [56]. There are various types of wavelets and the calculation (for trend and fluctuation) will change depending on the type of wavelet.

The 'Haar' wavelet is the simplest type of wavelet and widely used for many applications. The theory of the discrete wavelet transform using the 'Haar' wavelet is explained in detail by Walker [56]. If a discrete signal is expressed in the form:

$$f = (f_1, f_2, \dots, f_N)$$

Scaling coefficients can be calculated as:

$$a_m = \frac{f_{2m-1} + f_{2m}}{\sqrt{2}} \qquad (4\text{-}2)$$

for m = 1,2,3,…,N/2.

Detail coefficients can be calculated as

$$d_m = \frac{f_{2m-1} - f_{2m}}{\sqrt{2}} \qquad (4\text{-}3)$$

for m = 1,2,3,…,N/2.

The scaling coefficients can be further decomposed recursively. These decomposed scaling coefficients will represent the trend of original waveforms in a much smaller number of samples using an adequate level of scaling coefficients.

Figure 4-9 shows an example of scaling coefficients using 'Haar' wavelets at level 9. After DWT waveforms have been normalised by the maximum value of the 'Fault free' waveform; only the shape information is important, not the actual value of the original waveform. It can be seen that Figure 4-9 expresses the approximate shape of the original waveform (Figure 4-4) while the original dimension, 25000, was decreased to 49 after feature extraction.

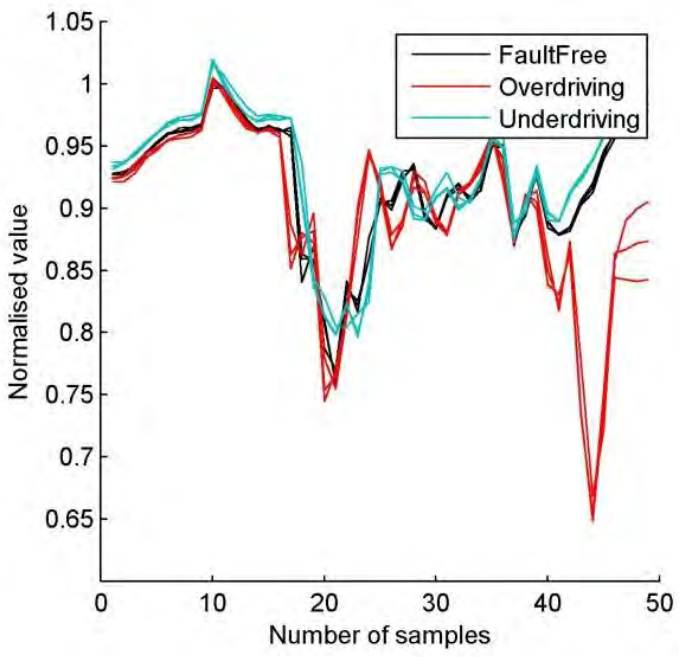

Figure 4-9 Scaling coefficients using 'Haar' wavelets at level 9

Table 4-2 shows the result of k-means clustering for scaling coefficients using different wavelets at different levels of decomposition. ○ means that the data divided by cluster correctly whereas × means that the data did not divided by cluster correctly. It can be

seen that some wavelets are able to make the right clustering even at decomposition level 10.

Table 4-2 Result of cluster analysis using different wavelets at different levels of decomposition

	Levels of decomposition by DWT			
Wavelets	Level 8	Level 9	Level 10	Level 11
Haar	◯	◯	◯	×
Daubechies 2	◯	◯	◯	×
Daubechies 3	◯	◯	◯	×
Daubechies 4	◯	◯	◯	×
Daubechies 5	◯	◯	◯	×
Daubechies 6	◯	◯	◯	×
Symlets 4	◯	◯	◯	×
Symlets 5	◯	◯	×	×
Symlets 6	◯	◯	×	×
Coiflets 1	◯	◯	×	×
Coiflets 2	◯	◯	×	×
Coiflets 3	◯	◯	×	×
Coiflets 4	◯	◯	×	×
Coiflets 5	◯	◯	×	×

Table 4-3 shows the result of mean silhouette width using different wavelets at different levels of decomposition. From Table 4-3, it can be seen that 'Haar' wavelets had the best performance. Consequently, scaling coefficients at level 9 using 'Haar' wavelets are chosen as the feature. The mean silhouette width after feature extraction was 0.844, while the mean silhouette width before the feature extraction was 0.833. A slight improvement of the silhouette width may be caused by the noise reduction of the original waveform. Finally, the features extracted from 'Right to left' and 'Left to right' operations have been concatenated (the features of the "Right to Left" and "Left to

Right" were calculated individually and then simply concatenated to make a longer feature), as shown in Figure 4-10. A concatenation of "Right to Left" and "Left to Right" features may have drawbacks; this approach may cause discontinuity to the features at the concatenation point; some point machines may not operate frequently and there might be a time lag between the data collection of the "Right to Left" waveform and the "Left to Right" waveform. It is, however, reasonable to use all the information available to carrying out fault detection and diagnosis.

Table 4-3 Mean silhouette width using different wavelets

	Different levels of decomposition by DWT		
Wavelets	Level 8	Level 9	Level 10
Haar	0.840	0.844	0.838
Daubechies 2	0.840	0.843	0.823
Daubechies 3	0.839	0.839	0.757
Daubechies 4	0.836	0.837	0.780
Daubechies 5	0.835	0.837	0.824
Daubechies 6	0.835	0.837	0.804
Symlets 4	0.836	0.834	0.753

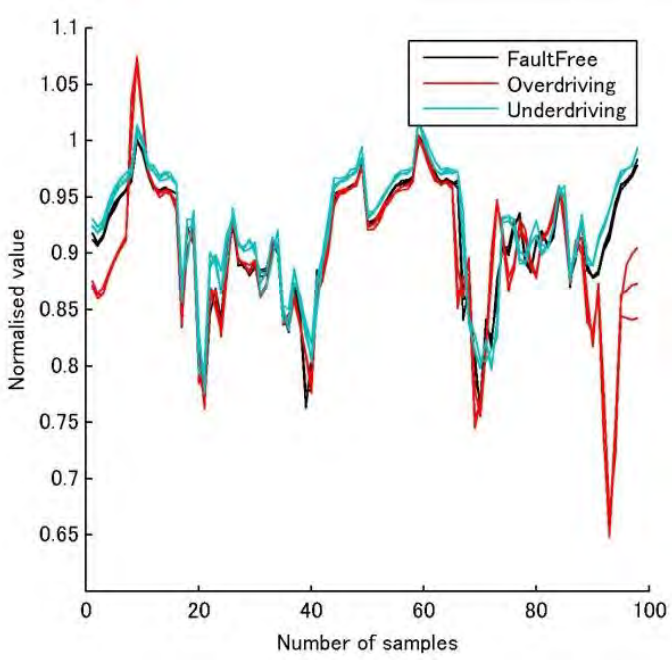

Figure 4-10 Concatenation of 'Right to left' and 'Left to right' operations

4.3.2 Fault detection and diagnosis

Nowadays, one of the most widely supported classifiers for pattern recognition is the Support Vector Machine (SVM). The basic principle of the SVM is to make a maximum margin possible between different classes after transferring to a higher dimension using the kernel function. C-SVM is an extension of SVM, allowing an overlap of data by introducing a penalty, making the boundary of the classes more general [57].

The theory of SVM is explained in detail by Bishop [58]. In the support vector machine, a decision boundary (hyper-plane: $y(x) = w^T \emptyset(x) + b$) is selected to be the one for which the margin (which is defined to be the smallest distance between the decision boundary and any of the samples) is maximised [58], as illustrated in Figure 4-11.

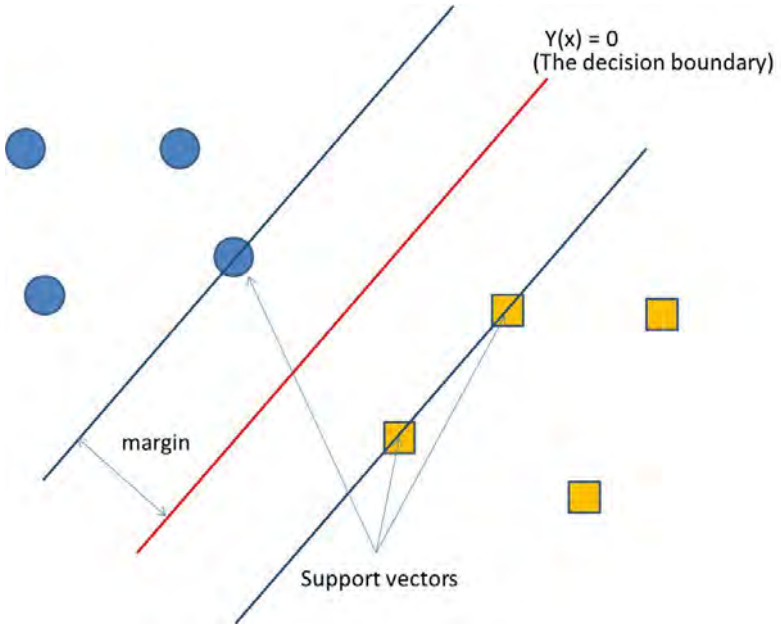

Figure 4-11 The decision boundary for support vector machine [58]

N input vectors $x_1, \ldots, x_N$ with target values $t_1, \ldots, t_N$ where $t_n \in \{+1, -1\}$, are used as the training data set [58]. New data points *x* are classified based on the sign of y(x) that is given in the form

$$y(x) = w^T \emptyset(x) + b \qquad (4\text{-}4)$$

where w denotes the normal vector to the hyper-plane, b denotes the bias parameter and $\emptyset(x)$ denotes a fixed feature-space transformation.

The distance between $y(x) = 0$ (the decision boundary) and x_i can be given as follows

$$\frac{|y(x_i)|}{\|w\|} \qquad (4\text{-}5)$$

This equation can be transformed as follows using t_n (target values)

$$\frac{t_n y(x_n)}{\|w\|} = \frac{t_n(w^T \emptyset(x_n) + b)}{\|w\|} \qquad (4\text{-}6)$$

The equation that makes the margin between two classes maximum can be written as follows

$$\arg\max_{w,b}\{\frac{1}{\|w\|}\min_n[t_n(w^T \emptyset(x_n) + b)]\} \qquad (4\text{-}7)$$

Noting that, even if we make the rescaling (*w* to *kw*) and (*b* to *kb*), the distance from any point x_n to the decision boundary remains the same, this freedom can be used to set

$$t_n(w^T \emptyset(x_n) + b) = 1 \qquad (4\text{-}8)$$

for the point that is nearest to the decision boundary. Then, all data points shall satisfy the constraints

$$t_n(w^T \emptyset(x_n) + b) \geq 1 \qquad (4\text{-}9)$$

The optimisation problem then simply requires to maximise $\|w\|^{-1}$,which is equivalent to optimising the form

$$\arg\min_{w,b} \frac{1}{2}\|w\|^2 \qquad (4\text{-}10)$$

$$\text{Subject to } t_n(w^T \emptyset(x_n) + b) \geq 1$$

To solve this constrained optimisation problem, Lagrange multipliers $a_n \geq 0$ can be introduced. Equation (4-10) can be written in the form

$$\mathrm{L(w, b, a)} = \frac{1}{2}\|w\|^2 - \sum_{n=1}^{N} a_n\{t_n(w^T\emptyset(x_n) + b) - 1\} \qquad (4\text{-}11)$$

where $\mathrm{a} = (a_1, \ldots, a_N)^T$

Optimising equation (4-11) requires minimising (4-11) with respect to *w* and *b* and maximising with respect to *a* (considering the minus sign in front of the Lagrange multiplier term).

If the derivatives of equation (4-11) are set equal to zero with respect to *w* and *b*, the following two conditions can be obtained

$$\frac{\partial L}{\partial w} = 0 \rightarrow 0 = \sum_{n=1}^{N} a_n t_n \emptyset(x_n) \qquad (4\text{-}12)$$

$$\frac{\partial L}{\partial b} = 0 \rightarrow \sum_{n=1}^{N} a_n t_n = 0 \qquad (4\text{-}13)$$

w and *b* can be eliminated from equation (4-11) using equation (4-12) and (4-13). Then, the problem to select the decision boundary which maximises two classes can be expressed in the dual representation where the following equation is maximised

$$\tilde{L}(a) = \sum_{n=1}^{N} a_n - \frac{1}{2}\sum_{n=1}^{N}\sum_{m=1}^{N} a_n a_m t_n t_m k(x_n, x_m) \qquad (4\text{-}14)$$

with respect to a subject to the constraints

$$0 \leq a_n \qquad (4\text{-}15)$$

$$\sum_{n=1}^{N} a_n t_n = 0 \qquad (4\text{-}16)$$

This problem is known as a quadratic programming problem (a common mathematical optimisation problem [58]).

After solving this problem, *w* and *b* will be decided. Then the decision function can be expressed in the form

$$\mathrm{sgn}(w^T \emptyset(x) + b) \qquad (4\text{-}17)$$

C-SVM is an extension of SVM, allowing an overlap of data by introducing a penalty [57]. The theory of SVM explained above is modified so that data points are allowed to be on the 'wrong side' of the decision boundary, but with a penalty (if the distance from the decision boundary gets larger, the penalty gets larger) [58]. By doing this, the boundary for two classes can be further generalised. The theory of C-SVM is explained in detail by Bishop [58].

For problems involving $I>2$ classes, classification can be carried out by (1) training $I(I-1)/2$ different 2-class SVMs on all possible pairs of classes and (2) classifying test data according to which class has the highest number of 'votes' [58].

If the system can accurately classify three fault conditions (including fault free) using the classifier, it can be said that the system can carry out fault detection and diagnosis successfully.

4.4 Experiments and Results

4.4.1 Experiment 1: fault detection and diagnosis for 'Overdriving' and 'Underdriving'

In this experiment, 92 data sets were used; each data set comprised of electrical active power data collected during'Right to Left' operation and'Left to Right' operation whilst simulating a single fault condition (from five fault conditions). These 92 data sets were made up as follows: 31 'Fault free' data sets, 30 'Overdriving' data sets and 31 'Underdriving' data sets.

A feature extraction was carried out using scaling coefficients (using 'Haar' wavelet at decomposition level nine) after a DWT and a classification was done by C-SVM using Linear kernel and RBF kernel. All of the calculations in the experiment were made using Libsvm software [59].

5-fold cross validation was performed to evaluate the method. [Step 1] 92 data sets were divided into five subsets. [Step 2] C-SVM was trained from 4 subsets excluding one subset. [Step 3] The performance (the percentage: the number of correctly classified data sets against the total number of data sets) was evaluated with the excluded subset. [Step 4] The same procedure was repeated for each subset (five times). [Step 5] Average performance for all the subsets (cross validation accuracy) was calculated.

Before the training using 4 subsets in Step 2, parameters of C-SVM have to be tuned properly using 4 subsets (since the performance changes depending on parameters). Parameters of C-SVM were tuned using a "grid-search" on parameters (again calculating 5-fold cross-validation accuracy). The following parameters were selected;

C = 0.0625 and gamma = 4 for the RBF kernel and C = 0.25 was selected for Linear kernel.

As a result, 100% cross validation accuracy (standard deviation = 0) was accomplished for both the Linear kernel and RBF kernel. This result demonstrates that highly accurate fault detection and diagnosis can be achieved by applying the proposed method to electrical active power data. Furthermore, a high degree of classification accuracy achieved in the Linear kernel implies that the data split linearly in multi dimension.

4.4.2 Experiment 2: fault detection and diagnosis for 'Overdriving (minor severity)', 'Overdriving', 'Underdriving (minor severity)' and 'Underdriving'

Ten data sets were added to those used in Experiment 1, making the total number of data sets 102. The ten added data sets were made up as follows: 5 data sets of'Overdriving (minor severity)', simulating an intermediate severity of faults between'Fault free' and'Overdriving'; 5 data sets of 'Underdriving (minor severity)', simulating an intermediate severity of faults between 'Fault free' and 'Underdriving'. Figure 4-12 shows the drive force data acquired in 5 fault conditions: 'Fault free', 'Overdriving (minor severity)', 'Overdriving', 'Underdriving (minor severity)' and 'Underdriving'.

A feature extraction and a classification were carried out in the same way as in Experiment 1.

5-fold cross validation was performed to evaluate the method. [Step 1] 102 data sets were divided into five subsets (each subset containing a similar number of five fault conditions). [Step 2] C-SVM was trained from 4 subsets excluding one subset. [Step 3]

The performance was evaluated with the excluded subset. [Step 4] The same procedure was repeated for each subset (five times). [Step 5] Average performance for all the subsets (cross validation accuracy) was calculated.

Before the training using 4 subsets in Step 2, parameters of C-SVM have to be tuned properly using 4 subsets (since the performance changes depending on parameters). Parameters of C-SVM were tuned using a "grid-search" on parameters (again calculating 5-fold cross-validation accuracy). The following parameters were selected; C = 1 and gamma = 32 were selected for RBF kernel and C = 32 was selected for Linear kernel.

As a result, 100% cross validation accuracy (standard deviation = 0) was accomplished for both Linear kernel and RBF kernel. This result shows that the proposed method can accurately diagnose two levels of severity.

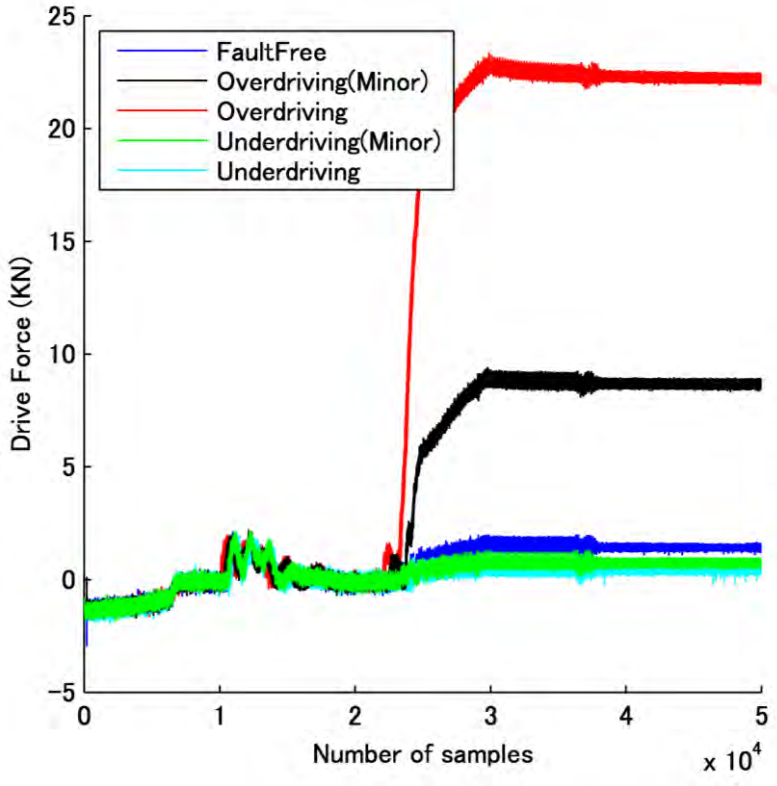

Figure 4-12 Introducing intermediate severity levels of fault between 'Fault free' and 'Overdriving' and 'Fault free' and 'Underdriving'

4.5 Conclusions

In this case study, adjustment faults were induced on a Japanese point machine. Electrical active power (which can be collected using electrical current and voltage sensors) was selected as a parameter for condition monitoring taking into account the practical requirements of the railway. A novel approach for fault detection and diagnosis using Wavelet Transforms and Support Vector Machines was presented. The new approach was shown to be accurate for classifying different fault conditions, including different levels of severity. This methodology will be further tested in different types of point machines in the next chapter.

CHAPTER 5 TRANSFERABILITY OF THE ALGORITHM TO OTHER TYPES OF POINT MACHINE AND TRANSFERABILITY OF THE SPECIFIC ALGORITHM PARAMETERS TO MULTIPLE POINT MACHINES

5.1 Introduction

In this chapter, the method developed in Chapter 4 will be further tested and developed.

First, the method is tested on other types of point machine operated in Great Britain (the Surelock-type and the M63-type) to verify the transferability of the approach to other types of point machine. If the same method developed in Chapter 4 can be applied to other types of point machine, it will make the approach more general and stronger.

Second, transferability of the specific algorithm parameters from one instance of a point machine to the next is also tested using the data collected from multiple point machines (HW-type point machines). The approach will be slightly modified to fulfil this task.

5.2 Transferability of the algorithm to other types of point machine (Surelock-type and M63-type point machine)

5.2.1 Parameter selection and feature extraction (Surelock-type point machine)

Figure 5-1 shows the drive force, electrical current, electrical voltage and electrical power data (three data sets per fault condition) acquired from the Surelock-type point machine where each plot shows one throw of the point machine (‘Left to Right’ throw on the top row, and ‘Right to Left’ throw on the bottom row).

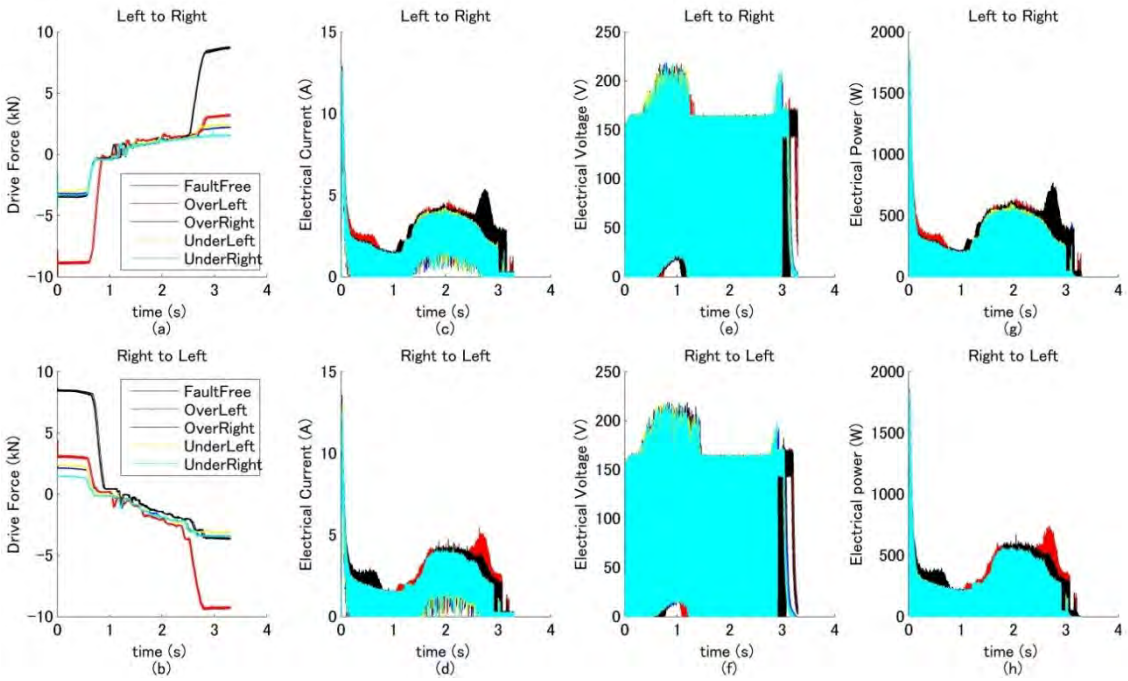

Figure 5-1 Waveforms acquired during point machine operation: (a) Drive force during left to right operation, (b) Drive force during right to left operation, (c) Electrical current during left to right operation, (d) Electrical current during right to left operation, (e) Electrical voltage during left to right operation, (f) Electrical voltage during right to left operation, (g) Electrical power during left to right operation and (h) Electrical power during right to left operation

As can be seen from Figure 5-1, only the force plot clearly shows a distinction between different fault conditions, whereas the electrical current, electrical voltage and electrical power data plots do not appear to show a clear visible distinction between different fault conditions. This result is similar to Japanese point machine data described in Chapter 4. The voltage experienced a sharp fluctuation due to the switched mode power supply. The similar sharp fluctuation of the current may be caused by the switched mode power supply and also picking up the noise from adjacent high voltage cables. Furthermore, the peak voltage varied between 150 and 200 Volts. In ideal conditions, the voltage remains constant even though the load varies as the internal resistance is close to 0 ohm.

In reality, however, there is a small amount of internal resistance in the power supply, affecting the voltage of the power supply.

Again, although the force data would give a good result in terms of fault diagnosis, in a practical condition monitoring system, it would be difficult to acquire this data, as discussed in Chapter 4. Electrical current and voltage sensors can be implemented away from the track side. They are usually less expensive than the load pin and more practical to install, as they can be implemented inside a signal box or relay room where electrical current and voltage to the point machine are supplied. The sensors will not affect the operation of the point machine as they are not implemented directly on the machine, as discussed in Chapter 4.

The electrical current, electrical voltage and electrical power parameters cannot be used directly for a condition monitoring system because the data size is too large to handle and the data acquired fluctuates (Figure 5-1) due to the switched-mode power supply. A feature extraction method is therefore needed for these parameters to be used in a condition monitoring system.

The Discrete Wavelet Transform (DWT) is used to extract features from the original waveforms, as discussed in Chapter 4.

Figure 5-2 shows scaling coefficients using the 'Haar' wavelet at decomposition level nine (three data sets per fault condition). After the DWT, waveforms have been normalised by the maximum value of the 'Fault free' waveform, as was the case in Chapter 4.

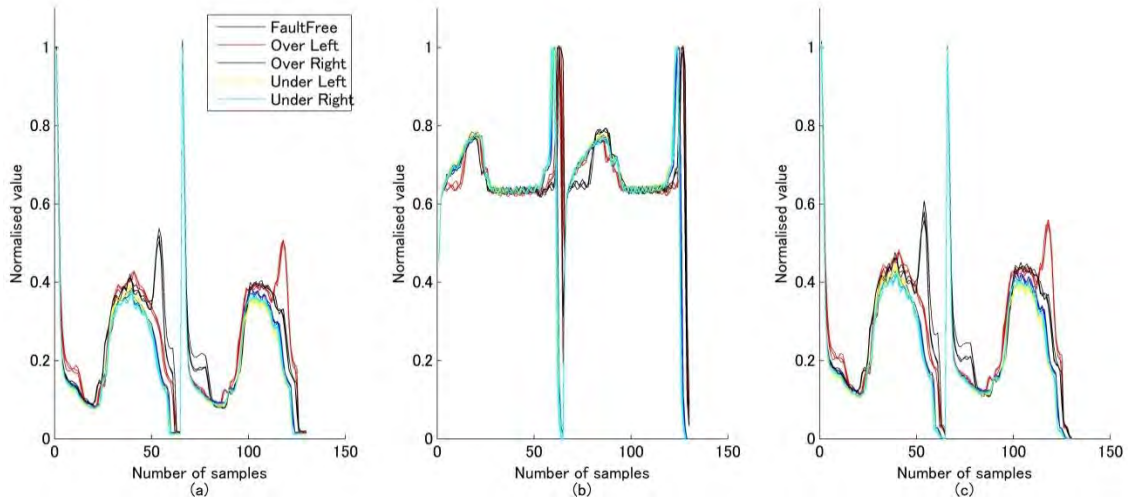

Figure 5-2 Trend features extracted from (a) Electrical Current, (b) Electrical Voltage and (c) Electrical Power

From the figure, it can be seen that electrical current and electrical power data shows a distinction between different fault conditions, whereas the electrical voltage plot does not appear to show a clear visible distinction between different fault conditions. Furthermore, it can be seen from the figure that the underdriving condition is difficult to distinguish from fault free conditions for all of the parameters.

A cluster analysis using the k-means method [51] was carried out to the electrical current, electrical voltage and electrical power data to investigate which parameter would be the best to use for a condition monitoring system, as in Chapter 4. The centroid of the cluster was calculated by finding the mean value and the squared-Euclidean distance was used to select the points for clusters; 10 data per fault condition were used. Figure 5-3 shows the result of the k-means clustering for electrical current, electrical voltage and electrical power data.

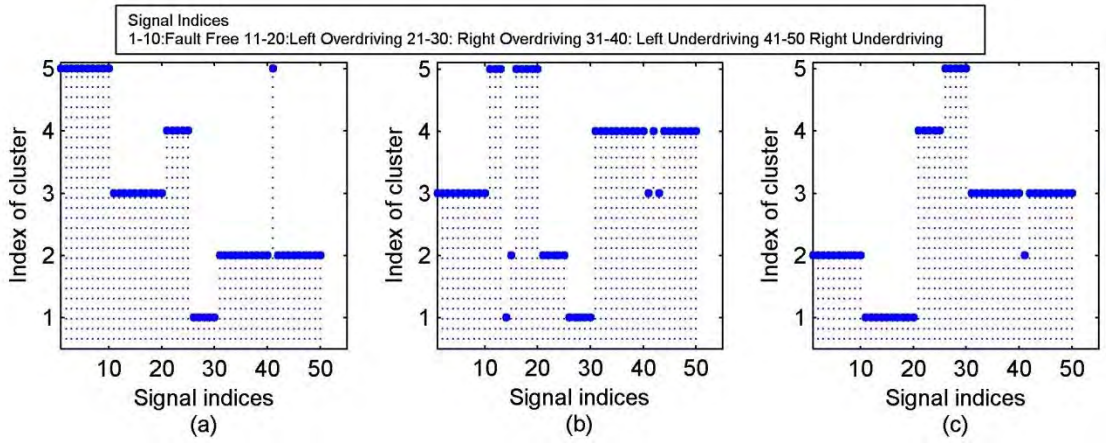

Figure 5-3 Cluster analysis using five clusters for five fault conditions of (a) Electrical Current, (b) Electrical Voltage and (c) Electrical Power

Unfortunately, clusters did not clearly separate by fault conditions in any of the three parameters. This result may be caused by the subtle difference between the 'Fault Free' condition and the 'Underdriving' condition. A further cluster analysis was therefore carried out using three fault conditions: 'Fault Free', 'Left hand Overdriving' and 'Right hand Overdriving'. Figure 5-4 shows the result of cluster analysis for these three classes.

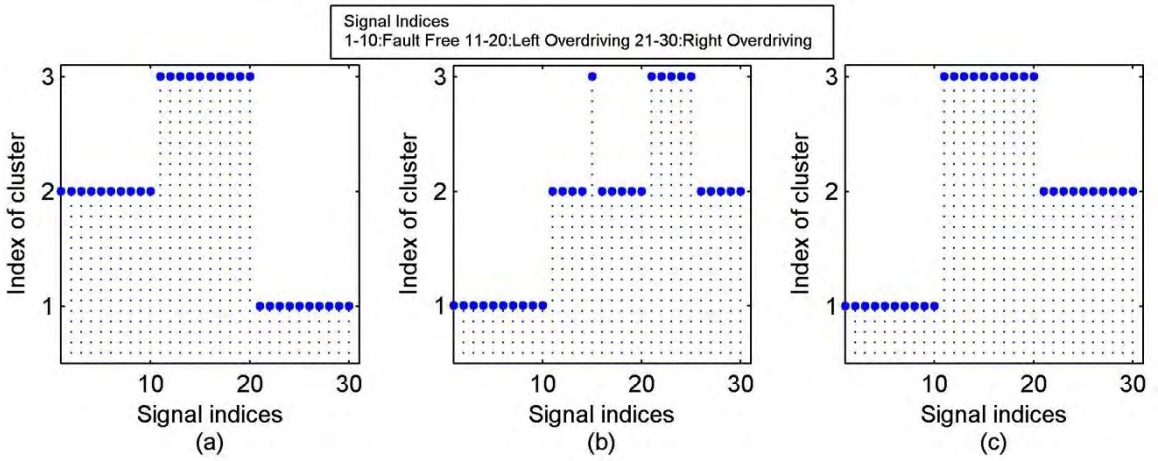

Figure 5-4 Cluster analysis using three clusters for three fault conditions of (a) Electrical Current, (b) Electrical Voltage and (c) Electrical Power

From the figure, it can be seen that the electrical current and electrical power data divided clearly by different fault conditions. Both electrical current and electrical power therefore have the potential to be used as the parameter for a condition monitoring

system. To investigate how well the cluster was divided for these two parameters after k-means clustering, the silhouette width [53] was calculated, as in Chapter 4 (Japanese point machine case).

Figure 5-5 shows the silhouette width for electrical current and electrical power.

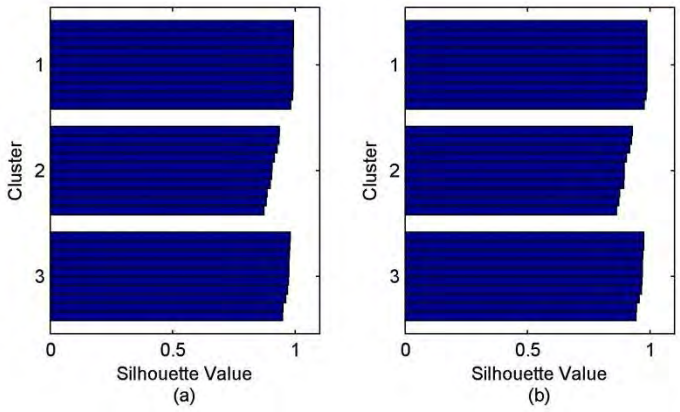

Figure 5-5 Silhouette width for: (a) Electrical current and (b) Electrical power

From Figure 5-5 it can be seen that electrical current and electrical power showed a similar result. Calculating the mean value of the silhouette width, electrical current showed a slightly better result than electrical power; 0.955 for electrical current and 0.949 for electrical power. This result implies that the electrical current has more useful information in terms of condition monitoring than the electrical power. Consequently, the electrical current data has been selected as the parameter for the condition monitoring system.

The calculations above were carried out using scaling coefficients acquired from the 'Haar' wavelet at the decomposition level of nine, as it was previous used in the Japanese case. It is, however, important to verify that the use of the 'Haar' wavelet at a decomposition level of nine is a good choice. The same methodology (calculating the mean value of the silhouette width) was followed to perform the analysis of different

wavelets and decomposition levels, as was the case in Chapter 4. Electrical current data (with three fault conditions: 'Fault free', 'Left hand Overdriving' and 'Right hand Overdriving') were considered for this analysis. The result of the calculation is shown in Table 5-1.

Table 5-1 Mean silhouette width using different wavelets at different levels of decomposition

	Levels of decomposition by DWT					
wavelets	Level 8	Level 9	Level 10	Level 11	Level 12	Level 13
Haar	0.938	0.955	0.961	0.968	0.969	0.966
Daubechies2	0.943	0.955	0.959	0.962	0.964	0.935
Daubechies3	0.944	0.953	0.958	0.959	0.950	0.912
Daubechies4	0.943	0.953	0.958	0.960	0.944	0.925
Daubechies5	0.942	0.954	0.958	0.961	0.953	0.933
Daubechies6	0.942	0.954	0.957	0.960	0.957	0.924
Symlets4	0.944	0.952	0.957	0.958	0.941	0.900
Symlets5	0.941	0.952	0.955	0.955	0.946	0.881
Symlets6	0.943	0.952	0.955	0.954	0.936	0.867
Coiflets1	0.945	0.953	0.959	0.959	0.950	0.908
Coiflets2	0.943	0.952	0.955	0.954	0.938	0.863
Coiflets3	0.943	0.952	0.956	0.956	0.943	0.881
Coiflets4	0.943	0.953	0.958	0.961	0.954	0.911
Coiflets5	0.943	0.953	0.959	0.962	0.955	0.916
Average	0.943	0.953	0.958	0.959	0.950	0.909

It can be seen from the table that the levels of decomposition at nine, ten, eleven and twelve showed relatively good results. At these levels of decomposition, the 'Haar' wavelet showed the highest value.

Although the 'Haar' wavelet at decomposition level twelve showed the best result in the table, it is important to consider that this result was calculated from only three classes ('Fault Free', 'Left hand overdriving' and 'Right hand overdriving') since it was not possible to make the right clustering using five classes ('Fault Free', 'Left hand

Overdriving', 'Right hand Overdriving', 'Left hand Underdriving' and 'Right hand Underdriving'). It is therefore reasonable to say that retaining information using a lower level of decomposition would be necessary for more complex classification: i.e. five class classification. Consequently, it can be considered that using the 'Haar' wavelet at a decomposition level nine is good choice.

Finally, a cluster analysis using three clusters for five classes ('Fault Free', 'Left hand Overdriving', 'Right hand Overdriving', 'Left hand Underdriving' and 'Right hand Underdriving') of electrical current was carried out, as shown in Figure 5-6.

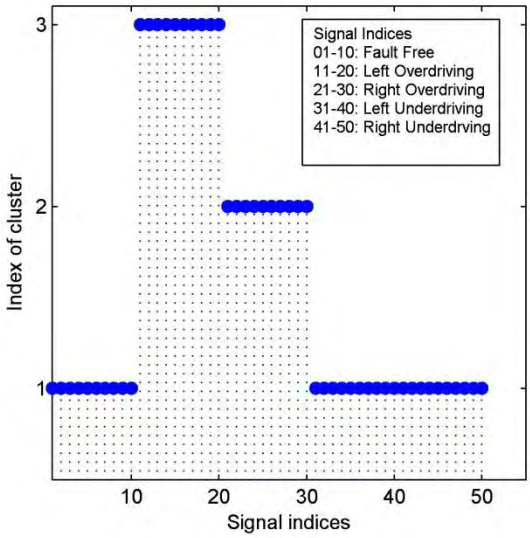

Figure 5-6 A cluster analysis using three clusters for five classes of electrical current

The figure shows that the 'Fault Free (indices 1-10)', 'Left hand Underdriving (indices 31-40)' and 'Right hand Underdriving (indices 41-50)' were grouped in the same cluster, implying that these three classes have similar data. It was not possible to make the correct clustering using five clusters for five fault conditions, however, it may still be possible to make the right classification using a more sophisticated classifying algorithm.

To summarise, four potential parameters were examined: Drive force, Electrical current, Electrical voltage and Electrical power. For practical reasons, drive force was excluded. Trend information was then extracted using the DWT (as in Chapter 4) for electrical current, electrical voltage and electrical power. After a cluster analysis, it was discovered that electrical current is the best of the three parameters for condition monitoring.

5.2.2 Fault detection and diagnosis (Surelock-type point machine)

If the system can accurately classify five fault conditions (including fault free) using the classifier, it can be said that the system can carry out fault detection and diagnosis successfully. C-SVM is used as a classifier, as in Chapter 4.

An experiment was carried out to verify the approach, as in Chapter 4. In this experiment, 142 data sets were used; each data set comprised of electrical current data collected during 'Right to Left' operation and 'Left to Right' operation whilst simulating single fault condition (from five fault conditions). These 142 data sets were made up as follows: 30 'Fault free', 26 'Left hand Overdriving', 26 'Right hand Overdriving', 30 'Left hand Underdriving' and 30 'Right hand Underdriving' data sets. A feature extraction was carried out using scaling coefficients (using the 'Haar' wavelet at decomposition level nine) after DWT and a classification was done by C-SVM. A five-fold cross validation was performed to evaluate the method, as in Chapter 4. First, 142 data sets were divided into five subsets. Second, C-SVM was trained from 4 subsets, excluding one subset. Third, the performance (the percentage: the number of correctly classified data sets against the total number of data sets) was evaluated with the

excluded subset. Fourth, the same procedure was repeated for each subset (five times). Finally, average performance for all the subsets (cross validation accuracy) was calculated.

Table 5-2 shows the result of the classification using C-SVM. Two kernel functions were tested: 'Linear kernel' $k\left(x, x^{'}\right) = x^{T}x^{'}$ and 'RBF kernel' $k\left(x, x^{'}\right) = \exp(-\gamma \left\| x - x^{'} \right\|)$. All the classifications in the experiment were made using the Libsvm software [59], as was the case in Chapter 4.

Table 5-2 Classification accuracy for five fault conditions (Linear kernel and RBF kernel)

	Fault Free	LH Over	RH Over	LH Under	RH Under	(Accuracy)
Subset 1	6/6	6/6	6/6	6/6	5/6	96.66% (29/30)
Subset 2	6/6	5/5	5/5	6/6	6/6	100% (28/28)
Subset 3	6/6	5/5	5/5	6/6	6/6	100% (28/28)
Subset 4	6/6	5/5	5/5	6/6	6/6	100% (28/28)
Subset 5	6/6	5/5	5/5	6/6	6/6	100% (28/28)
CV accuracy						99.33% ±1.49

As can be seen from Table 5-2, a high rate of accuracy was achieved using the C-SVM approach, similarly to the results obtained for the Japanese point machine. The result of the classification for Linear kernel and RBF kernel was the same (Table 5-2).

Further work was carried out to consider the intermediate severity of faults (as in Chapter 4) between (1) 'Fault free' and 'Left hand Overdriving', (2) 'Fault free' and 'Right hand Overdriving' and (3) 'Fault free' and 'Right hand Underdriving', as shown

in Figure 5-7. It was not possible to introduce the intermediate severity of faults between 'Fault free' and 'Left hand Underdriving' since a failure occurred to the point machine after underdriving the drive rod nut for only 2/6th of a turn; underdriving the drive rod nut for 1/6th of a turn was the maximum severity for 'Left hand Underdriving' and there was no intermediate severity for this fault condition.

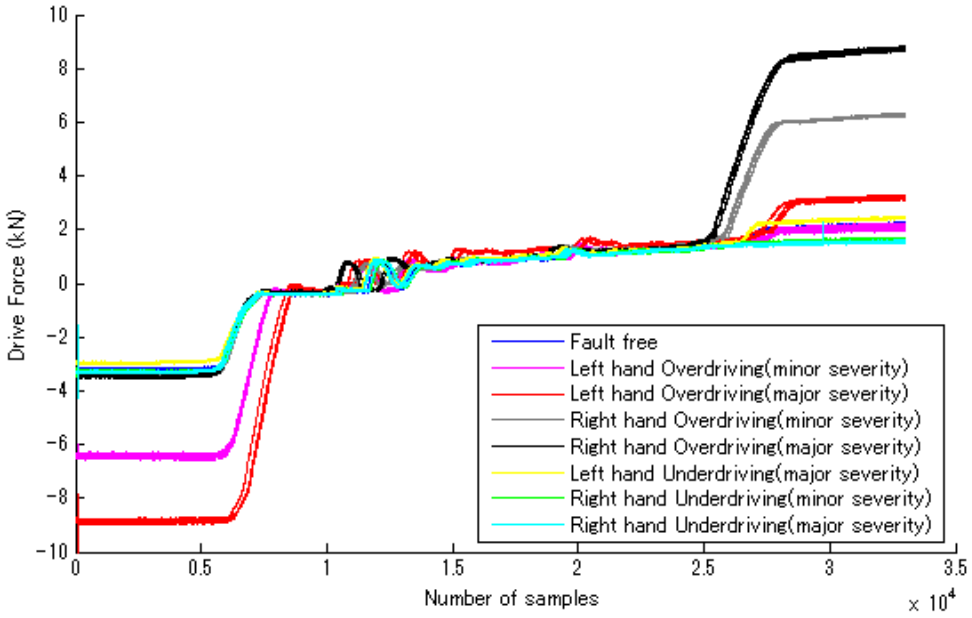

Figure 5-7 Drive force during Left to Right operation introducing intermediate severity of faults between (1) 'Fault free' and 'Left hand Overdriving', (2) 'Fault free' and 'Right hand Overdriving' and (3) 'Fault free' and 'Right hand Underdriving'

The intermediate severity of faults and the original fault conditions were re-named according to Table 5-3. If the system is able to classify these eight fault conditions accurately, it can then be said that the system can diagnose the severity of faults, making it more practical for real use, as discussed in Chapter 4.

Table 5-3 Re-naming of fault conditions including intermediate severity of faults

Fault condition (Original name)	New name
Fault free	Fault free
(Left hand Overdriving)	Left hand Overdriving (major severity)
(Right hand Overdriving)	Right hand Overdriving (major severity)
(Left hand Underdriving)	Left hand Underdriving (major severity)
(Right hand Underdriving)	Right hand Underdriving (major severity)
Intermediate condition between 'Fault free' and 'Left hand Overdriving'	Left hand Overdriving (minor severity)
Intermediate condition between 'Fault free' and 'Right hand Overdriving'	Right hand Overdriving (minor severity)
Intermediate condition between 'Fault free' and 'Right hand Underdriving'	Right hand Underdriving (minor severity)

In this experiment, 227 data sets were used; each data set comprised of electrical current data collected during 'Right to Left' operation and 'Left to Right' operation. These 227 data sets were made up as follows: 30 'Fault free', 26 'Left hand Overdriving (major)', 25 'Left hand Overdriving (minor)', 26 'Right hand Overdriving (major)', 30 'Right hand Overdriving (minor)', 30 'Left hand Underdriving (major)', 30 'Right hand Underdriving (major)' and 30 'Right hand Underdriving (minor)'. A feature extraction was carried out using scaling coefficients (using the 'Haar' wavelet at decomposition level nine) after DWT and a classification was done by C-SVM. Five-fold cross validation was performed to evaluate the method, as was the case in Chapter 4.

Table 5-4 and Table 5-5 show the result of classification. As can be seen from Table 5-4 and Table 5-5 , a high accuracy rate has been achieved using the C-SVM approach,

similarly to the results obtained for the Japanese point machine. The system can now diagnose the faults at two levels of severity: minor severity and major severity (for Surelock-type point machine).

Table 5-4 Classification accuracy for eight fault conditions (Linear kernel)

	Fault Free	LH Over	RH Over	LH Under	RH Under	LH Over (Minor)	RH Over (Minor)	RH Under (Minor)	(Accuracy)
Subset 1	6/6	6/6	6/6	6/6	5/6	5/5	6/6	6/6	97.87% (46/47)
Subset 2	6/6	5/5	5/5	6/6	6/6	5/5	6/6	6/6	100% (45/45)
Subset 3	6/6	5/5	5/5	6/6	6/6	5/5	6/6	6/6	100% (45/45)
Subset 4	6/6	5/5	5/5	6/6	6/6	5/5	6/6	6/6	100% (45/45)
Subset 5	6/6	5/5	5/5	6/6	6/6	5/5	6/6	6/6	100% (45/45)
CV accuracy									99.57% ±0.95

Table 5-5 Classification accuracy for eight fault conditions (RBF kernel)

	Fault Free	LH Over	RH Over	LH Under	RH Under	LH Over (Minor)	RH Over (Minor)	RH Under (Minor)	(Accuracy)
Subset 1	6/6	6/6	6/6	6/6	5/6	5/5	6/6	6/6	97.87% (46/47)
Subset 2	6/6	5/5	5/5	6/6	6/6	5/5	6/6	6/6	100% (45/45)
Subset 3	6/6	5/5	5/5	6/6	5/6	5/5	6/6	6/6	98.68% (44/45)
Subset 4	6/6	5/5	5/5	6/6	6/6	5/5	6/6	6/6	100% (45/45)
Subset 5	6/6	5/5	5/5	6/6	6/6	5/5	6/6	5/6	97.78% (44/45)
CV accuracy									98.68% ±1.20

5.2.3 Parameter selection and feature extraction (M63-type point machine)

Figure 5-8 shows the drive force, electrical current, electrical voltage and electrical power data (three data sets per fault condition) acquired from the M63-type point machine.

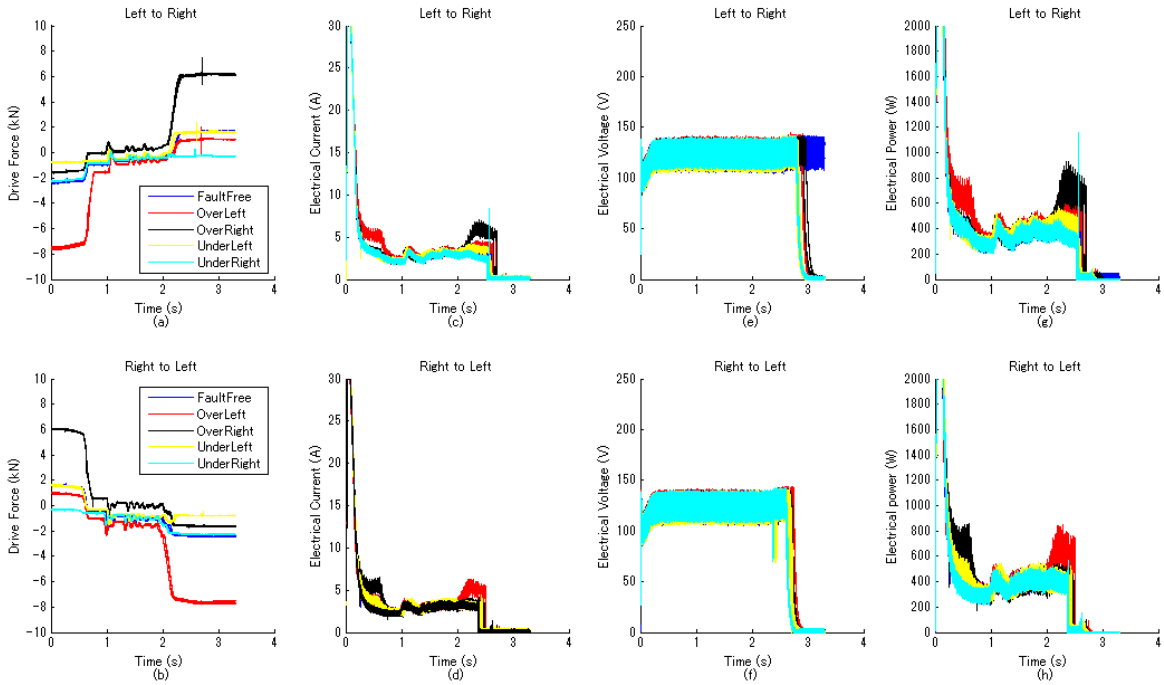

Figure 5-8 Waveforms acquired during point machine operation: (a) Drive force during left to right operation, (b) Drive force during right to left operation, (c) Electrical current during left to right operation, (d) Electrical current during right to left operation, (e) Electrical voltage during left to right operation, (f) Electrical voltage during right to left operation, (g) Electrical power during left to right operation and (h) Electrical power during right to left operation

From the discussion and analysis carried out in the previous section (which implies that electrical current is the best parameter), the electrical current was selected as a parameter for the M63-type point machine. Although the current data shows a lot of noise, it is still possible to notice a distinction between different fault conditions (Figure 5-8).

The feature extraction method (using Discrete Wavelet Transform) was also applied to the electrical current data of the M63-type point machine, as was in Chapter 4 (Japanese point machine).

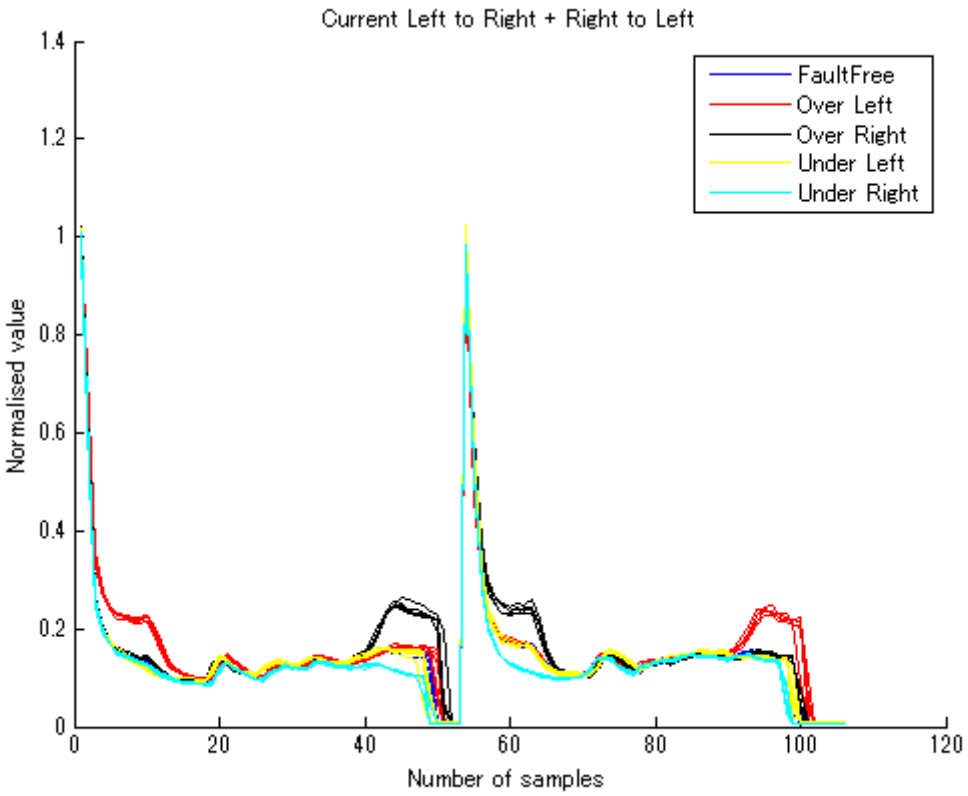

Figure 5-9 Trend features extracted from electrical current

Scaling coefficients using the 'Haar' wavelet at decomposition level nine were used as a feature, following the discussion and analysis carried out in 5.2.1 (Surelock-type point machine). Figure 5-9 shows scaling coefficients using the 'Haar' wavelet at decomposition level nine (three data sets per fault condition). After the DWT, waveforms have been normalised by the maximum value of the 'Fault free' waveform, as was the case in Chapter 4. The distinction between different fault conditions can be seen from the figure.

5.2.4 Fault detection and diagnosis (M63-type point machine)

C-SVM was also used for making a fault detection and diagnosis from the extracted feature, as in Chapter 4.

An experiment was carried out to verify the approach, as in Chapter 4. In this experiment, 140 data sets were used; each data set comprised of electrical current data collected during 'Right to Left' operation and 'Left to Right' operation whilst simulating single fault condition (from five fault conditions). These 140 data sets were made up as follows: 30 'Fault free', 25 'Left hand Overdriving', 25 'Right hand Overdriving', 30 'Left hand Underdriving' and 30 'Right hand Underdriving' data sets. A feature extraction was carried out using scaling coefficients (using the 'Haar' wavelet at decomposition level nine) after DWT and a classification was done by C-SVM. A five-fold cross validation was performed to evaluate the method, as in Chapter 4. First, 140 data sets were divided into five subsets. Second, C-SVM was trained from 4 subsets, excluding one subset. Third, the performance (the percentage: the number of correctly classified data sets against the total number of data sets) was evaluated with the excluded subset. Fourth, the same procedure was repeated for each subset (five times). Finally, average performance for all the subsets (cross validation accuracy) was calculated.

Table 5-6 and Table 5-7 show the result of classification using C-SVM. Two kernel functions were tested: Linear kernel and RBF kernel, as was the case in Chapter 4.

Table 5-6 Classification accuracy for five fault conditions (Linear kernel)

	Fault Free	LH Over	RH Over	LH Under	RH Under	(Accuracy)
Subset 1	6/6	5/5	5/5	5/6	6/6	96.43% (27/28)
Subset 2	6/6	5/5	5/5	6/6	6/6	100% (28/28)
Subset 3	6/6	5/5	5/5	6/6	6/6	100% (28/28)
Subset 4	6/6	5/5	5/5	6/6	6/6	100% (28/28)
Subset 5	6/6	5/5	5/5	6/6	6/6	100% (28/28)
CV accuracy						99.33% ±1.49

Table 5-7 Classification accuracy for five fault conditions (RBF kernel)

	Fault Free	LH Over	RH Over	LH Under	RH Under	(Accuracy)
Subset 1	6/6	5/5	5/5	4/6	6/6	92.86% (26/28)
Subset 2	6/6	5/5	5/5	6/6	6/6	100% (28/28)
Subset 3	6/6	5/5	5/5	6/6	6/6	100% (28/28)
Subset 4	6/6	5/5	5/5	6/6	6/6	100% (28/28)
Subset 5	6/6	5/5	5/5	6/6	6/6	100% (28/28)
CV accuracy						98.57% ±3.19

As can be seen from Table 5-6 and Table 5-7, a high rate of accuracy was achieved using the C-SVM approach, similarly to the result obtained for the Japanese point machine.

Further work was carried out to consider the intermediate severity of faults (as in Chapter 4) between (1) 'Fault free' and 'Left hand Overdriving', (2) 'Fault free' and 'Right hand Overdriving', (3) 'Fault free' and 'Left hand Underdriving' and (4) 'Fault free' and 'Right hand Underdriving'.

The intermediate severity of faults and the original fault conditions were re-named according to Table 5-8.

Table 5-8 Re-naming of fault conditions including intermediate severity of faults

Fault condition (Original name)	New name
Fault free	Fault free
(Left hand Overdriving)	Left hand Overdriving (major severity)
(Right hand Overdriving)	Right hand Overdriving (major severity)
(Left hand Underdriving)	Left hand Underdriving (major severity)
(Right hand Underdriving)	Right hand Underdriving (major severity)
Intermediate condition between 'Fault free' and 'Left hand Overdriving'	Left hand Overdriving (minor severity)
Intermediate condition between 'Fault free' and 'Right hand Overdriving'	Right hand Overdriving (minor severity)
Intermediate condition between 'Fault free' and 'Left hand Underdriving'	Left hand Underdriving (minor severity)
Intermediate condition between 'Fault free' and 'Right hand Underdriving'	Right hand Underdriving (minor severity)

In this experiment, 256 data sets were used; each data set comprised of electrical current data collected during 'Right to Left' operation and 'Left to Right' operation. These 256 data sets were made up as follows: 30 'Fault free', 25 'Left hand Overdriving (major)', 30 'Left hand Overdriving (minor)', 25 'Right hand Overdriving (major)', 25 'Right hand Overdriving (minor)', 30 'Left hand Underdriving (major)', 30 'Left hand Underdriving (minor)', 30 'Right hand Underdriving (major)' and 31 'Right hand Underdriving (minor)'. A feature extraction was carried out using scaling coefficients

(using a 'Haar' wavelet at decomposition level nine) after DWT and a classification was done by C-SVM. A five-fold cross validation was performed to evaluate the method, as was the case in Chapter 4. Table 5-9 and Table 5-10 show the result of classification. As can be seen from the table, a high accuracy rate has been achieved using the C-SVM approach, similarly to the result obtained for the Japanese point machine. The system can now diagnose the faults at two levels of severity: minor severity and major severity (for M63 point machine).

Table 5-9 Classification accuracy for nine fault conditions (Linear kernel)

	Fault Free	LH Over	RH Over	LH Under	RH Under	LH Over (Minor)	RH Over (Minor)	LH Under (Minor)	RH Under (Minor)	(Accuracy)
Subset 1	6/6	5/5	5/5	6/6	6/6	6/6	5/5	6/6	7/7	100% (52/52)
Subset 2	6/6	5/5	5/5	6/6	6/6	6/6	5/5	6/6	6/6	100% (45/45)
Subset 3	6/6	5/5	5/5	6/6	6/6	6/6	5/5	6/6	6/6	100% (45/45)
Subset 4	6/6	5/5	5/5	6/6	6/6	6/6	5/5	6/6	6/6	100% (45/45)
Subset 5	6/6	5/5	5/5	6/6	6/6	6/6	5/5	6/6	6/6	100% (45/45)
CV accuracy										100% ±0

Table 5-10 Classification accuracy for nine fault conditions (RBF kernel)

	Fault Free	LH Over	RH Over	LH Under	RH Under	LH Over (Minor)	RH Over (Minor)	LH Under (Minor)	RH Under (Minor)	(Accuracy)
Subset 1	6/6	5/5	5/5	5/6	6/6	6/6	5/5	6/6	7/7	98.08% (51/52)
Subset 2	6/6	5/5	5/5	6/6	6/6	6/6	5/5	6/6	6/6	100% (45/45)
Subset 3	6/6	5/5	5/5	6/6	6/6	6/6	5/5	6/6	6/6	100% (45/45)
Subset 4	6/6	5/5	5/5	6/6	6/6	6/6	5/5	6/6	6/6	100% (45/45)
Subset 5	6/6	5/5	5/5	6/6	6/6	6/6	5/5	6/6	6/6	100% (45/45)
CV accuracy										99.62% ±0.86

5.2.5 Conclusions

A cluster analysis was carried out to the data collected from the DC-type point machine which is operated in Great Britain (Surelock-type point machine) and it was concluded that electrical current is the best parameter for condition monitoring of the DC-type point machine. Then, an approach using Wavelet Transforms and Support Vector Machines was applied to both the Surelock-type point machine and the M63-type point machine, as was in Chapter 4 (Japanse point machine). Good classification results were achieved using the C-SVM for both types of point machine, similarly to the result for Japanese point machine. It was proved that this method can also make an accurate classification for two levels of severity, making the system more practical for real use.

5.3 Transferability of the specific algorithm parameters to multiple point machines (HW-type point machine)

5.3.1 Introduction and motivation

The aim of this section is to test the transferability of the specific algorithm parameters from one instance of a point machine to the next.

So far, training data sets and test data sets have been collected from the same point machine for fault detection and diagnosis. Since condition and characteristics of a point machine (including conditions and characteristics of external components such as drive rod) vary from one to another, achieving a high classification accuracy essentially requires the training data sets collected from the point machine to be monitored in the system. The problem for doing this is that it may cause a lot of time and labour to collect training data sets from all the point machines to be monitored (for example, train depot). Collecting data sets from a single point machine may require only 1-2 hours if

the maintenance engineers get used to the data collection procedure, but since there are a lot of point machines operated in the field, collecting the training data sets from all the point machines to be monitored requires a tremendous effort. Furthermore, it is necessary to collect a new training data set each time a new point machine (or an external component) is installed in the field, or even if significant maintenance and/or adjustment is carried out on a machine.

It is, therefore, more practical and useful if the training data sets collected from one point machine can be used to monitor other point machines. If this is the case, massive training data sets can be collected from training or test facilities and these data sets can be used to monitor all the point machines operated in the field.

In this chapter, a method using qualitative features to make the features transferable from one point machine to another is presented.

5.3.2 Data analysis and qualitative features

Data were collected from two HW-type point machines, as shown in Figure 2-1 in Chapter 2: point machine 1 and point machine 2. The main difference between the two point machines is that the backdrive of one point machine is located inside the track, whereas the backdrive of the other point machine is located outside the track.

Figure 5-10 shows the electrical current collected from two HW-type point machines: point machine 1 and point machine 2.

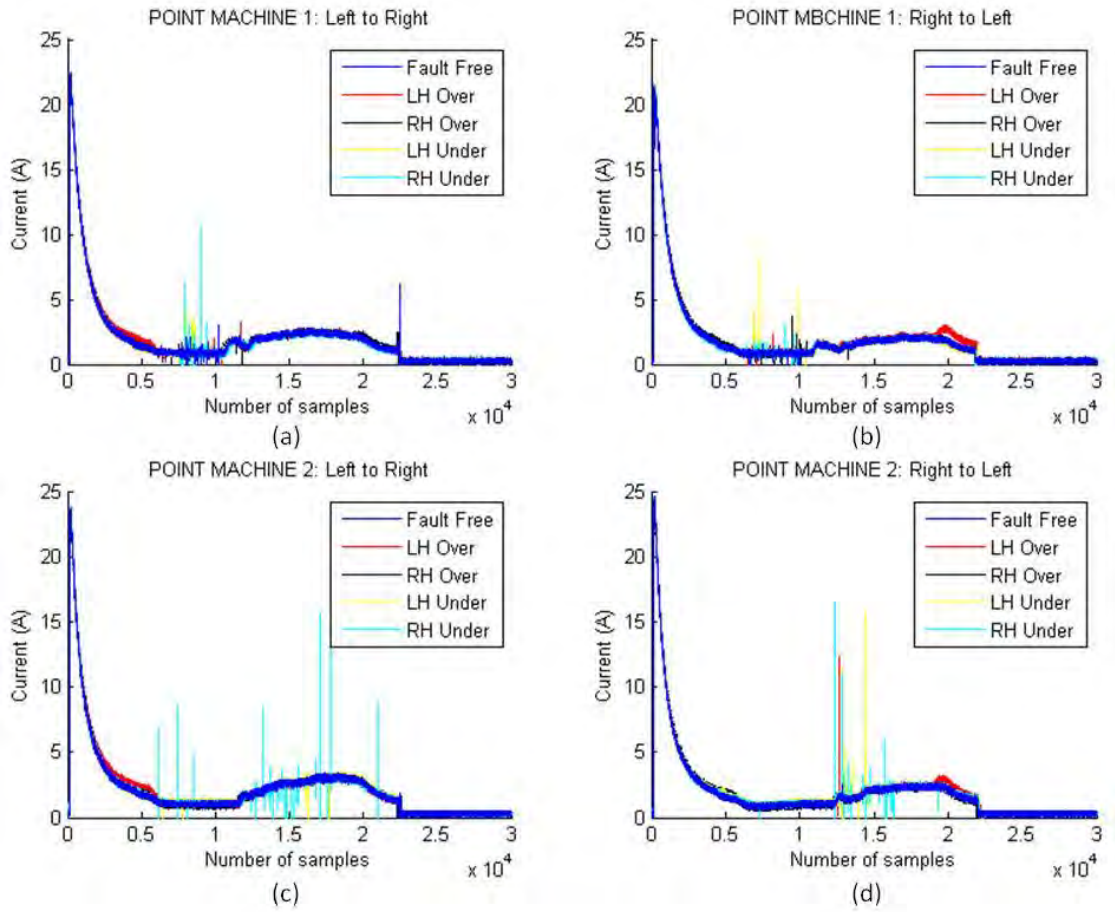

Figure 5-10 Electrical current for (a) point machine 1 during Left to Right operation, (b) point machine 1 during Right to Left operation, (c) point machine 2 during Left to Right operation and (d) point machine 2 during Right to Left operation

As can be seen from the figure, the shape of the waveforms is not identical, even in the "Fault free" condition, although data were collected from the same type of point machine. This is caused by the different settings of switch components: backdrive, slide chairs, rails, nuts, etc.

So far, feature extraction has been carried out using DWT, which extracts the approximate shape of the original waveforms. However, it was discovered that even the

shape of ‘Fault free’ waveforms can be different because of the different settings of the switch components. The feature extraction method has to be changed in order to make the specific algorithm parameters transferable to the other point machine.

Figure 5-11 shows the feature extracted from two point machines using DWT: point machine 1 and point machine 2.

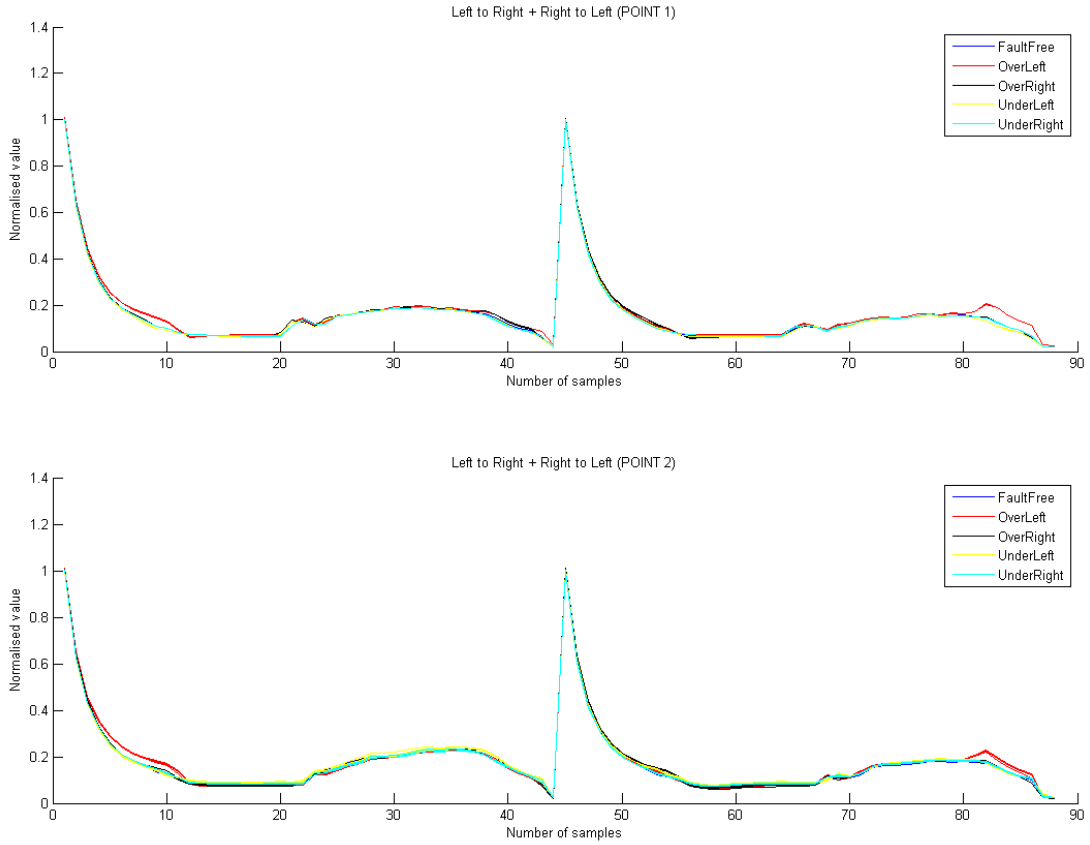

Figure 5-11 A feature extracted from (a) point machine 1 and (b) point machine 2

Again, it can be seen from the figure that the shape of the waveforms is not same, even for the same fault condition. It seems, however, that the change of tendency (from

"Fault Free" condition) remains similar to some extent. In order to qualitatively express the change from a "Fault Free" condition to a faulty condition, the features of "Fault Free" were subtracted from all of the features. Figure 5-12 depicts the qualitative features subtracting the "Fault Free" waveform from the original features.

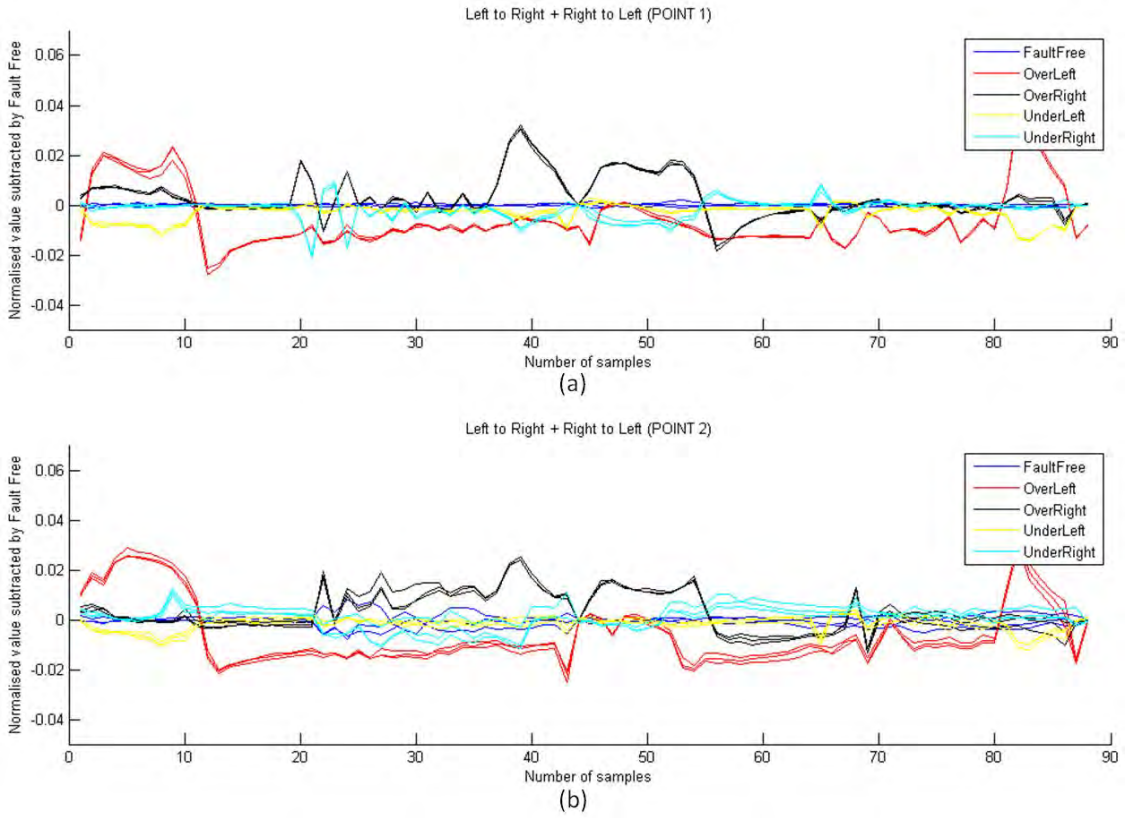

Figure 5-12 A qualitative feature extracted from (a) point machine 1 and (b) point machine 2

From the figure, it can be seen that the qualitative features remain similar (if not identical) between point machine 1 and point machine 2. These qualitative features can therefore be used to make the specific algorithm parameters transferable to the other point machine.

5.3.3 Fault detection and diagnosis

5.3.3.1 Experiment 1

The aim of this experiment is to prove that the qualitative features proposed in Section 5.3.2 will not decrease the high classification accuracy achieved in Section 5.2 for one instance of the point machine.

In this experiment, 61 data sets were used; each data set comprised of electrical current data collected during 'Right to Left' operation and 'Left to Right' operation. These 61 data sets were made up as follows: 10 'Fault free', 10 'Left hand Overdriving', 11 'Right hand Overdriving', 15 'Left hand Underdriving' and 15 'Right hand Underdriving'. A feature extraction was carried out using scaling coefficients (using a 'Haar' wavelet at decomposition level nine) after DWT subtracting the feature of 'Fault free' from all the features (to make the features qualitative) and a classification was carried out by C-SVM. A five-fold cross validation was performed to evaluate.

Table 5-11 shows the result of classification using C-SVM.

Both the Linear and RBF-kernel were used as kernel functions. All the classifications in the experiment were made using the Libsvm software [59].

Table 5-11 Classification accuracy for Experiment 1

Kernel function	Cross validation accuracy (Standard deviation)
Linear kernel	100% (± 0)
RBF kernel	100% (± 0)

As can be seen from the table, a high rate of accuracy was achieved using the C-SVM approach. From this experiment, it was proved that accurate classification can be done with qualitative features for one instance of the point machine.

5.3.3.2 Experiment 2

The aim of this experiment is to prove that the qualitative features can be transferable from one point machine to others (from point machine 1 to point machine 2).

All the data sets collected from point machine 1 (61 data sets in total) were used as training data sets; each data-set comprised of qualitative features extracted from electrical current data collected during the 'Left to Right' operation and 'Right to Left' operation (subtracting the 'Fault Free' feature from original features).

Test data sets were collected from the different point machine (point machine 2). 56 data sets were used as test data sets: 6'Fault Free', 15 'Left hand Overdriving', 15 'Right hand Overdriving', 15 'Left hand Underdriving' and 5 'Right hand Underdriving' data sets. Table 5-12 shows the result of classification using C-SVM.

Both Linear and RBF kernel were used as kernel functions. All the classifications in the experiment were made using the Libsvm software [14].

Table 5-12 Classification accuracy for Experiment 2

	Fault Free	LH Over	RH Over	LH Under	RH Under	(Accuracy)
Linear kernel	6/6	15/15	15/15	9/15	0/5	80.4% (45/56)
RBF kernel	6/6	15/15	11/15	9/15	0/5	73.2% (41/56)

As can be seen from the table, classification accuracy decreased from the previous experiment. It can be seen from the table that especially 'Right hand Underdriving (class 5)' had a low classification accuracy (the data sets for 'Left hand Underdriving' and 'Right hand Underfdriving' were classified as 'Fault free' class).

5.3.3.3 Experiment 3

The aim of this experiment is to prove that the qualitative features can be transferable from one point machine to others (from point machine 2 to point machine 1).

All the data sets collected from point machine 2 (56 data sets in total) were used as training data sets; each data-set comprised of qualitative features extracted from electrical current data collected during 'Left to Right' operation and 'Right to Left' operation (subtracting the 'Fault Free' feature from original features).

Test data sets were collected from the different point machine (point machine 1). 61 data sets were used as test data sets: 10'Fault Free', 10 'Left hand Overdriving', 11 'Right hand Overdriving', 15 'Left hand Underdriving' and 15 'Right hand Underdriving' data sets. Table 5-13 shows the result of classification using C-SVM.

Both Linear and RBF kernel were used as kernel functions. All the classifications in the experiment were made using the Libsvm software [14].

Table 5-13 Classification accuracy for Experiment 3

	Fault Free	LH Over	RH Over	LH Under	RH Under	(Accuracy)
Linear kernel	10/10	10/10	5/11	15/15	0/15	65.57% (40/61)
RBF kernel	10/10	10/10	11/11	15/15	0/15	75.41% (46/61)

As can be seen from the table, classification accuracy was similar to Experiment 2.

As was seen in the previous experiment, it can be seen from the table that especially 'Right hand Overdriving (class 5)' had a low classification accuracy (the data sets for 'Right hand Overdriving' and 'Right hand Underdriving' were classified as 'Fault free' class).

5.3.4 Conclusions

The transferability of the specific algorithm parameters from one instance of a point machine to the next was tested in this section. To make the features transferable, qualitative features were proposed, subtracting the 'Fault Free' feature from all the features. Classification accuracy was not as high as the results of previous experiments (section 5.1 and 5.2); however, 65%-80% classification accuracy was still achieved using the proposed method. A classification for 'Right hand Underdriving' had a particularly low classification accuracy. This result may be partly caused by different setting of the backdrive: the backdrive of point machine 1 was located outside the track (right of the right stock rail) whereas the backdrive of point machine 2 was located inside the track (between the left stock rail and the right stock rail). This change of the

setting may have influenced the change of behaviour between point machine 1 and point machine 2, eventually leading to poor accuracy for this particular fault condition.

5.4 Conclusions

In this chapter, the method proposed in Chapter 4 was further developed and applied into DC-type point machines operated in Great Britain. It was found that accurate fault detection and diagnosis can also be carried out for DC-type point machines using the approach, including different levels of severity.

Furthermore, in order to make the training data sets collected from one point machine transferable to others, a method to express the qualitative features was proposed. If this is successful, data sets can be collected from training or test facilities and these data sets can be used to monitor all the point machines in the field, which can save a lot of time and labour. The classification accuracy was eventually lower than that of single instance (particularly, the classification of 'Right hand Underdriving' was low), although the result shows that fault detection and diagnosis can be carried out with 65%-80% classification accuracy. This result may be improved in the future by collecting more data and carrying out further data analysis.

As a conclusion, it has been shown that using training data sets collected from the point machine to be monitored provides much better performance from the condition monitoring system. This means that each point machine requires specific preliminary measurement (collecting training data) before monitoring. Although this preliminary

measurement is time consuming, this may only require 1-2 hours for maintenance staff to carry out and will significantly improve the performance.

CHAPTER 6 DRIVE FORCE PREDICTION

6.1 Introduction

In this chapter, the approach will be further developed so that the system can directly predict drive force; this can be useful for inspection and maintenance purposes.

So far in this thesis, a classification using the SVM algorithm was demonstrated for fault detection and diagnosis from waveforms of point machines. As a result, it was found that accurate fault detection and diagnosis can be carried out using the approach (as described in Chapter 4 and 5), including different levels of fault severity. In the experiments in Chapter 4 and 5, two levels of severity were tested: major severity and minor severity. The natural progression of these experiments is to develop an accurate prediction for the force between stock rail and switch blade. Since drive force is a direct and understandable parameter for maintenance staff, predicting drive force can be beneficial in real operational use.

Drive force and electrical current data were acquired using a data acquisition box whilst simulating different strengths of drive force from Surelock-type and HW-type point machines.

6.2 Further developing the algorithm to predict drive force: Surelock-type point machine

6.2.1 Neural network for drive force prediction

In Chapter 5, drive force faults were detected and diagnosed by utilising electrical current as an input parameter. Feature extraction was carried out using the DWT, and the classification was done using a SVM. For drive force prediction, it is necessary to modify the approach. The basic idea is to predict the drive force (either at the beginning or end of the throw) from the extracted feature, so the DWT feature extracted from the electrical current will be used as an input to the neural network with the drive force (either at the beginning or end) being the output of the neural network.

As for the parameter and feature extraction method, a similar approach to the method used in Chapter 5 can be used for drive force prediction, considering the successful results of classification in Chapter 5. Unlike the case in Chapter 5, there are no 'Fault free' waveforms that can be used to normalise the extracted features, since the aim of the method is to directly predict the drive force. Scaling coefficients (without normalisation) are therefore used for drive force prediction. Additionally, there could be slight changes in the value of drive force for each operation, so the prediction of drive force is carried out separately for each operation ('Left to Right' or 'Right to Left' operation).

Figure 6-1 shows the extracted features (for drive force prediction) from 'Left to Right' for electrical current.

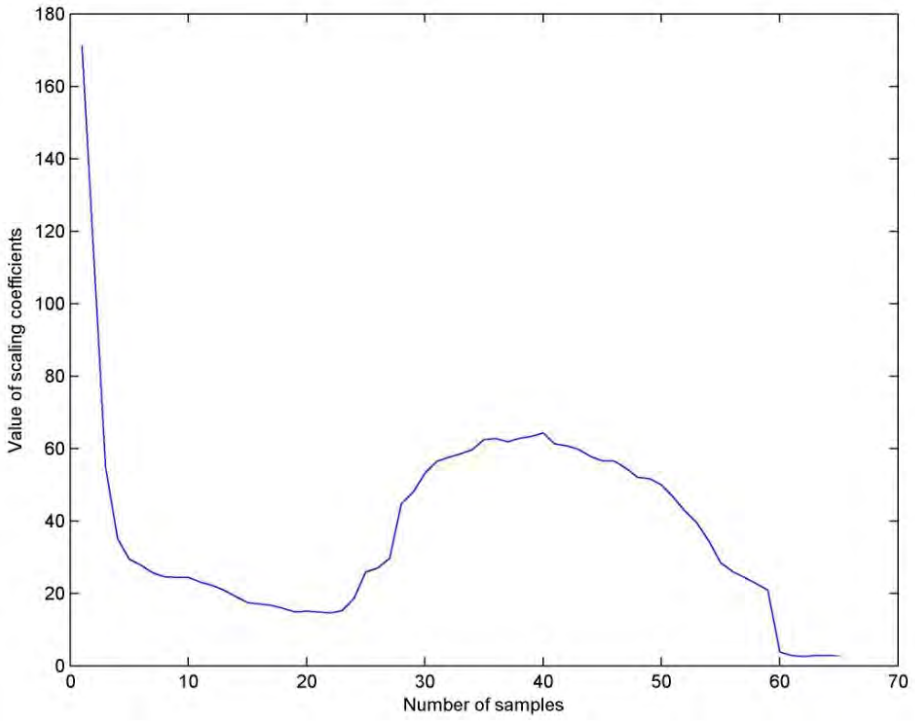

Figure 6-1 Extracted feature from the 'Left to Right' electrical current

The classification algorithm also has to be changed so that the system can predict drive force from extracted features.

Predicting drive force from inputs (extracted features from electrical current) can be conducted by non-linear function approximation. Neural networks are well known to have the ability to address non-linear function approximations [60]. An appropriate neural network can fit any input and output data relations into a non-linear function after training (*e.g.* house price prediction [61]). As for predicting drive force, the neural network shall take the extracted features (which are the scaling coefficients after the discrete wavelet transform) as inputs and predict drive force as the output. A typical three-layer neural network architecture, which consists of 'input layer', 'hidden layer' and 'output layer', can be used to address the problem.

Figure 6-2 shows a neural network designed to predict drive force.

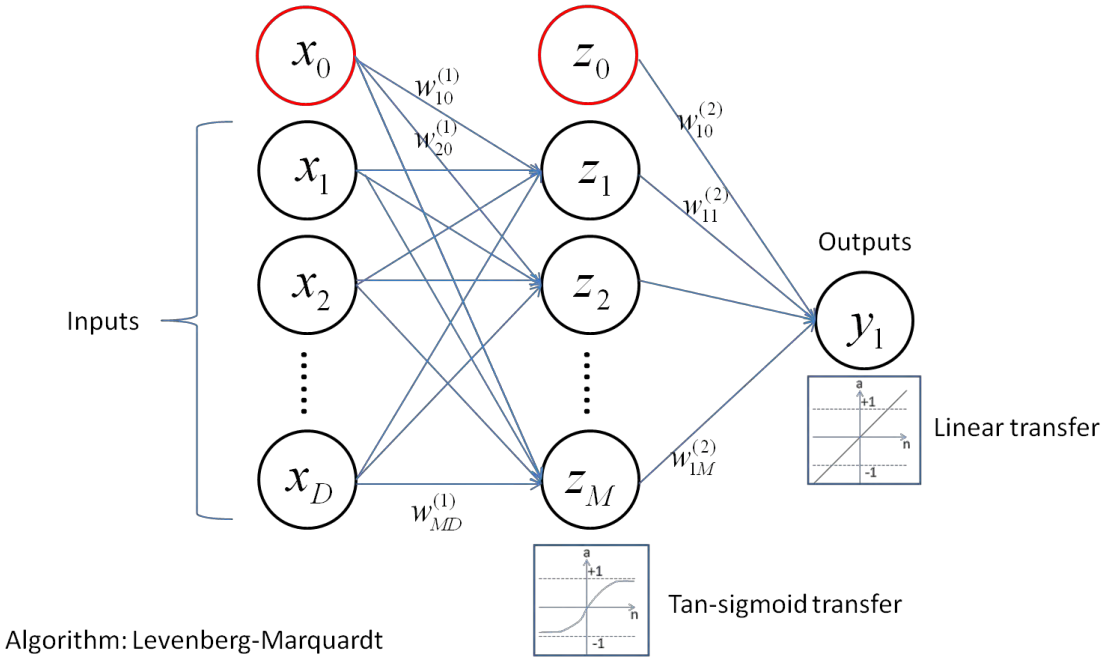

Figure 6-2 Neural network architecture

The theory of neural networks is provided in detail by Bishop [58].

Input variables $\boldsymbol{x_1, \ldots, x_D}$ will be transferred to the j'th hidden layer $\mathbf{a_j}$ in the form

$$\mathbf{a_j} = \sum_{\mathbf{i=1}}^{\mathbf{D}} \mathbf{w_{ji}^{(1)} x_i} + \mathbf{w_{j0}^{(1)}} \qquad (6\text{-}1)$$

Each $\mathbf{a_j}$ is then transformed using an activation function h(.) to give

$$\boldsymbol{z_j} = \boldsymbol{h}(\mathbf{a_j}) \qquad (6\text{-}2)$$

where an activation function is given by

$$\boldsymbol{h}(\boldsymbol{a}) = \boldsymbol{tanh}(\boldsymbol{a}) = \frac{\mathbf{e^a} - \mathbf{e^{-a}}}{\mathbf{e^a} + \mathbf{e^{-a}}} \qquad (6\text{-}3)$$

These values are again combined to give single output unit activation

$$\mathbf{y_1} = \sum_{\mathbf{j=1}}^{\mathbf{M}} \mathbf{w_{1j}^{(2)} z_j} + \mathbf{w_{10}^{(2)}} \qquad (6\text{-}4)$$

Since the output activation function is a linear function, this $\mathbf{y_1}$ will directly become the output of the network.

Parameters (**w**) have to be tuned to minimise the performance function (defined later) in order to make accurate drive force prediction.

A training algorithm, which is called the 'Levenberg-Marquardt' algorithm [62], was applied to tune the parameter, since this algorithm is known to perform efficiently [60].

The performance function is defined as:

$$\boldsymbol{F = mse = \frac{1}{N}\sum_{i=1}^{N}(t_i - y_i)^2} \qquad (6\text{-}5)$$

which is simply a mean squared error between targets and network outputs (*i.e.* neural network predictions).

Generally, by using a training algorithm, parameters of neural networks will be tuned to fit the relations between inputs and outputs of the training data sets. If the training is overdone, however, the network will only perform well for training data sets and will perform poorly for other data sets which are not used as training data sets [58]. This phenomenon is called 'Overfitting' and it is therefore necessary to avoid this occurring.

One of the methods to avoid 'Overfitting' is (1) to provide other data sets apart from training data sets (which are called '*validation data sets*') and (2) to stop training (adjusting weights) at the early stage (at the point of smallest error with respect to the 'validation data sets') [58]. This approach is called 'Early stopping' and is widely used

to avoid ‘Overfitting’ of neural networks [58]. This ‘Early stopping’ approach is used to prevent the neural network from ‘Overfitting’ in this method.

6.2.2 Deciding the number of hidden neurons

The number of hidden neurons in the hidden layer (Layer 2 in Figure 6-2) is important to make an accurate prediction of the drive force. The optimum number of hidden neurons may change depending on the problem, so it is needed to evaluate the performance changing the number of hidden neurons.

An experiment to optimise the number of hidden neurons was therefore conducted.

Twenty-one drive force conditions were simulated and data were collected from the point machine, as shown in Figure 6-3.

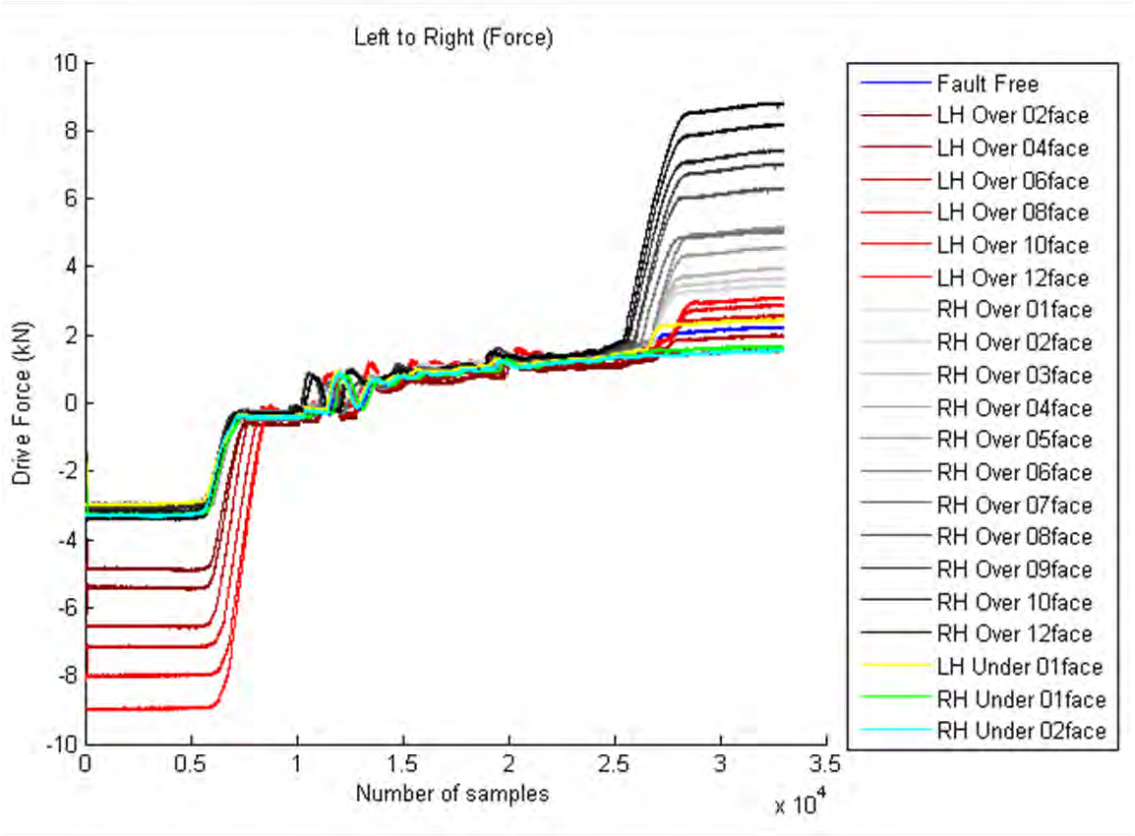

Figure 6-3 Drive force condition simulated

In the experiment, 63 data sets (3 data sets per drive force condition) were used as training data sets; each data set comprised of electrical current data collected during 'Left to Right' operation associated with 'Left-hand force'. Twenty-one data sets (with a single data set per drive force condition) were used as validation data sets. Validation data sets were used for 'Early stopping' and also measuring performance.

A feature extraction was carried out using scaling coefficients (using the 'Haar' wavelet at decomposition level nine) after DWT and the neural network was used for drive force prediction.

Performance (which is the mean squared error between targets and predictions) of validation data sets is examined changing the number of hidden neurons. The average performance over thirty iterations was calculated (changing initial weights each time), since the performance results fluctuated depending on the initial value (of the weight).

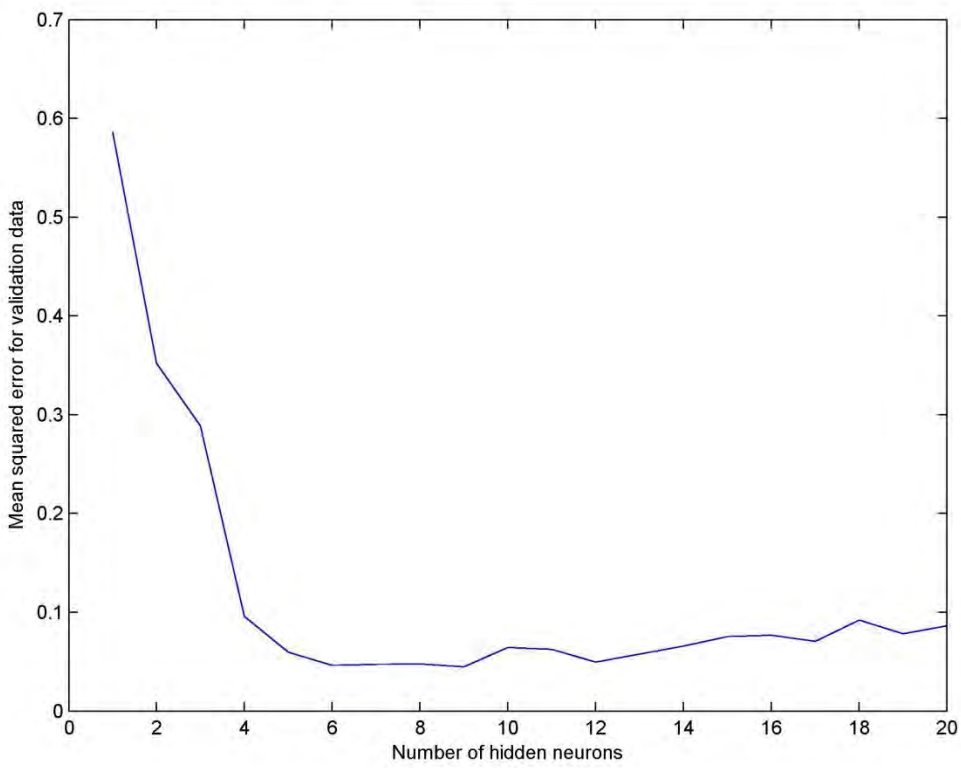

Figure 6-4 Performance for validation data changing number of hidden neurons

Figure 6-4 shows the result of the experiment. It was found that using a neural network with nine hidden neurons had the best performance for the particular problem. Nine hidden neurons are therefore used for the rest of the experiments detailed in this chapter.

6.2.3 Experiments for predicting drive force

The aim of the first experiment is to make a 'Left hand' drive force prediction from 'Left to Right' electrical current data.

Sixty-three training data sets (3 data sets per drive force condition), 21 validation data sets (single data set per drive force condition) and 206 test data sets (remaining data sets) were used in this experiment; each data set comprised of electrical current data collected during 'Left to Right' operation associated with 'Left-hand force'.

Validation data sets were used for 'Early stopping' to avoid 'Overfitting'. By using the 'Early stopping' method, the performance of the validation data is monitored and the iteration (to adjust parameters) is stopped when the magnitude of the gradient of the performance is small or the number of successive iterations that the performance fails to decrease reaches a predetermined number.

Nine hidden neurons were utilised, given the result of the previous experiment. Since there is a risk that the neural network would tune the weights into local optima (and therefore perform poorly) because of bad initial weights, the neural network which had the best performance for validation data over fifty calculations (changing initial weights each time) was used for drive force prediction. Figure 6-5 shows the result of performance for 'Training data sets', 'Validation data sets' and 'Test data sets. Figure 6-6 shows the regression plot for test data sets.

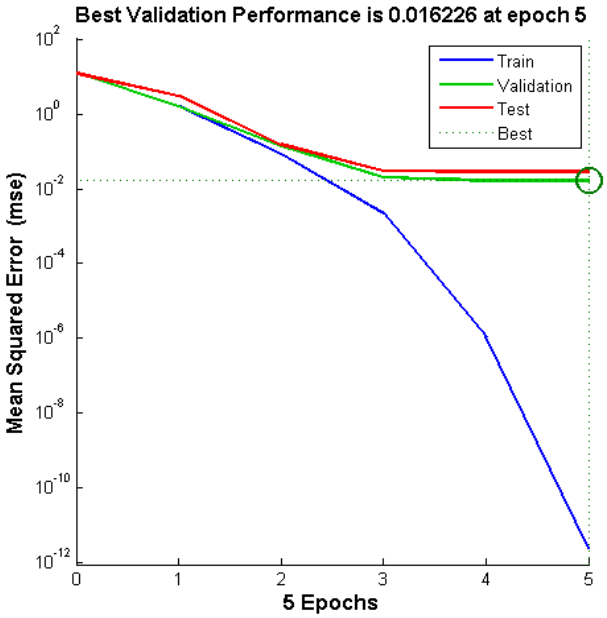

Figure 6-5 Performance of the neural network for Training, Validation and Test data sets

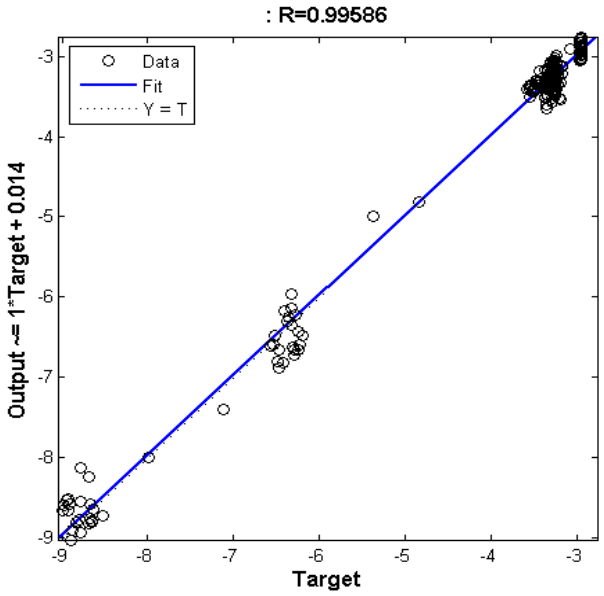

Figure 6-6 Regression plot for test data sets

From Figure 6-5, it can be seen that the performance of the validation data sets and the test data sets was similar and the mean squared error for the test data sets was small, which implies that the drive force prediction was accurate.

From Figure 6-6 it can be seen that prediction of the drive force is highly correlated with the real measurement (target). The value "R" in the figure shows the correlation coefficient which means 'how well the variation in the output is explained by the targets' [60].

From the experiment carried out so far, it was found that an accurate 'Left hand' drive force prediction can be carried out using electrical current waveform acquired during 'Left to Right' turn. Drive force predictions for other combinations are yet to be examined: Right hand drive force prediction from 'Left to Right' waveform, Right hand drive force prediction from 'Right to Left' waveform and Left hand drive force prediction from 'Right to Left' waveform.

Further experiments were carried out to predict 'Right hand' force from 'Left to Right' electrical current data, to predict 'Right hand' force from 'Right to Left' electrical current data and to predict 'Left hand' force from 'Right to Left' electrical current data, using the same method. Figure 6-7, Figure 6-8 and Figure 6-9 show the results of the experiments. The number of training data sets, validation data sets and test data sets remains the same as the first experiment in this section for all the experiments: 63 training data sets (3 data sets per drive force condition), 21 validation data sets (single data set per drive force condition) and 206 test data sets (remaining data sets).

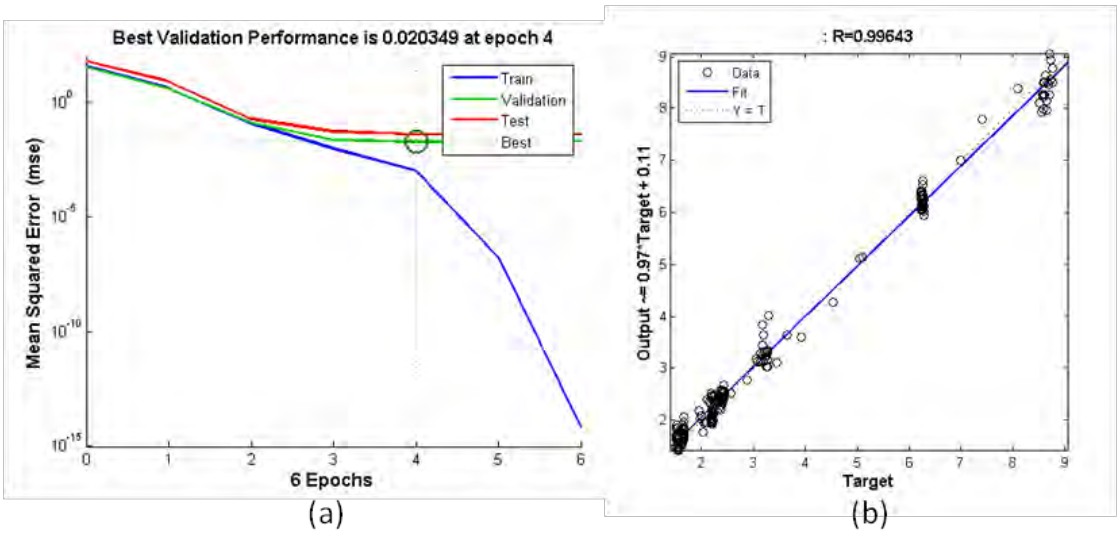

Figure 6-7 (a) Performance plot and (b) regression plot for predicting 'Right hand' force from 'Left to Right' electrical current data

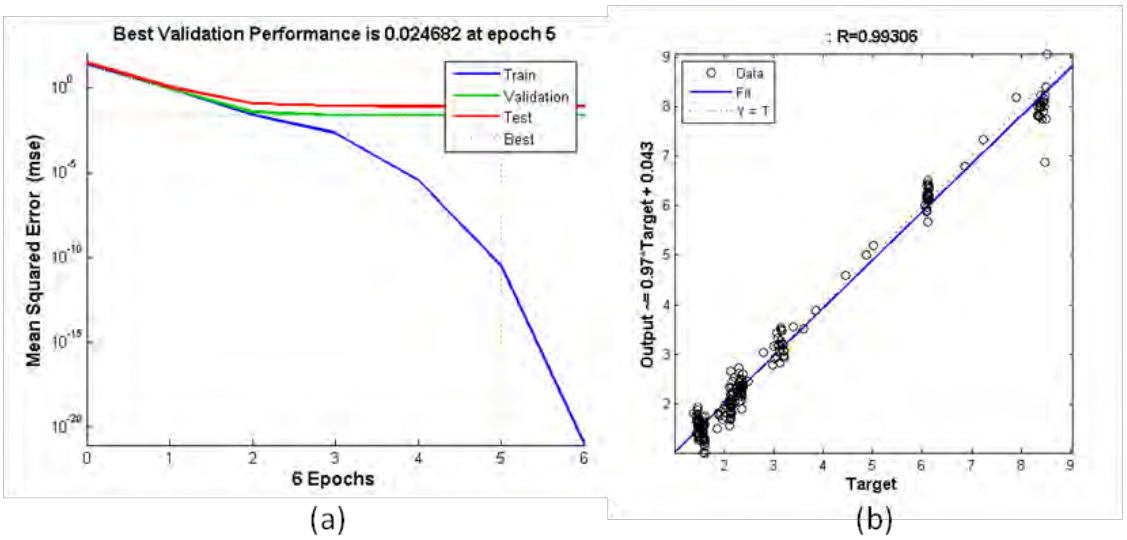

Figure 6-8 (a) Performance plot and (b) regression plot for predicting 'Right hand' force from 'Right to Left' electrical current data

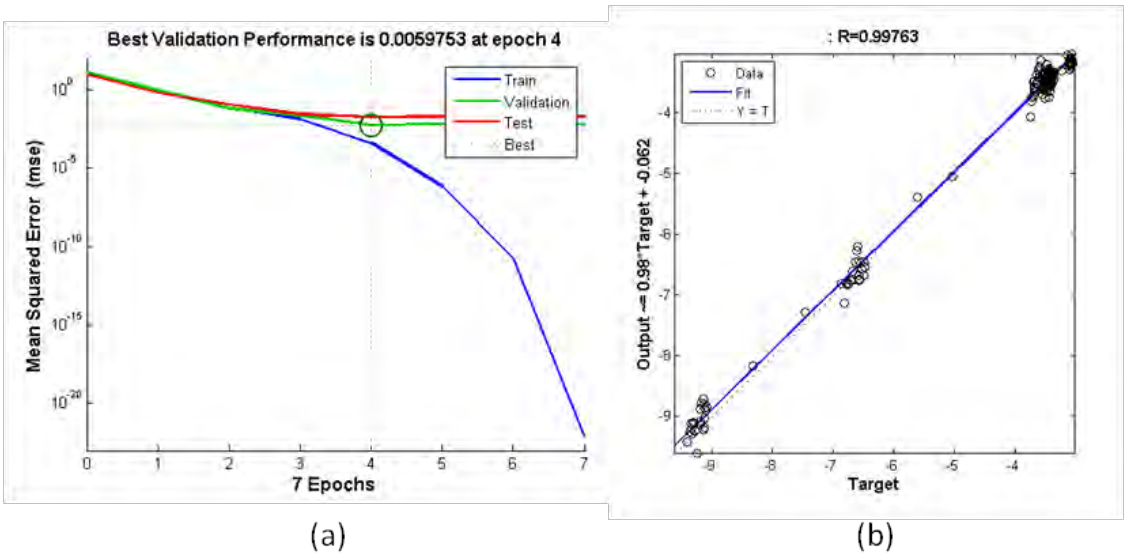

Figure 6-9 (a) Performance plot and (b) regression plot for predicting 'Left hand' force from 'Right to Left' electrical current data

It was found that accurate drive force prediction can be carried out in all cases (Right hand drive force prediction from 'Left to Right' waveform, Right hand drive force prediction from 'Right to Left' waveform and Left hand drive force prediction from 'Right to Left' waveform), showing similar results to the first experiment in 6.2.3. From this result, it can be said that each waveform contains information for both direction of drive force; i.e. the 'Left to Right' waveform contains both Left hand drive force and Right hand drive force information, and the 'Right to Left' waveform contains both Left hand drive force and Right hand drive force information.

6.3 Transferability of the algorithm to other types of point machine and further testing the ability of the algorithm: HW-type point machine

The aims of this section are: (1) to examine whether the method proposed in the previous section can be applied to other types of point machine (HW-type point machine), and; (2) to further test the ability of the method using further data sets.

Feature extraction and drive force prediction were carried out in the same way as described in the previous section; scaling coefficients after DWT were used as extracted features and drive force prediction was carried out using a Neural Network (the same number of hidden neurons as set out in the previous section were utilised, since the data and difficulty of the prediction remain the same for the HW-type point machine). Figure 6-10 shows the extracted feature (for drive force prediction) from the 'Left to Right' electrical current.

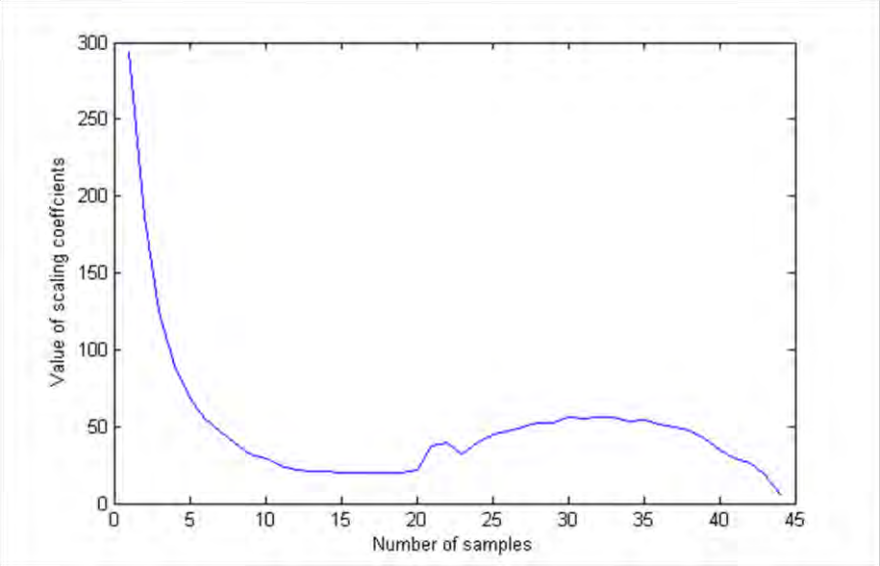

Figure 6-10 Extracted feature from the 'Left to Right' electrical current

6.3.1 Experiment 1: Testing the transferability of the algorithm to other types of point machine (HW-type point machine)

The aim of this experiment is to test the transferability of the algorithm to other types of point machine: the HW-type point machine.

The first experiment is carried out to predict 'Left hand' drive force from 'Left to Right' electrical current data, as in the previous section.

Nine drive force conditions were simulated and data were collected from the point machine, as shown in Figure 6-11, as in the previous section.

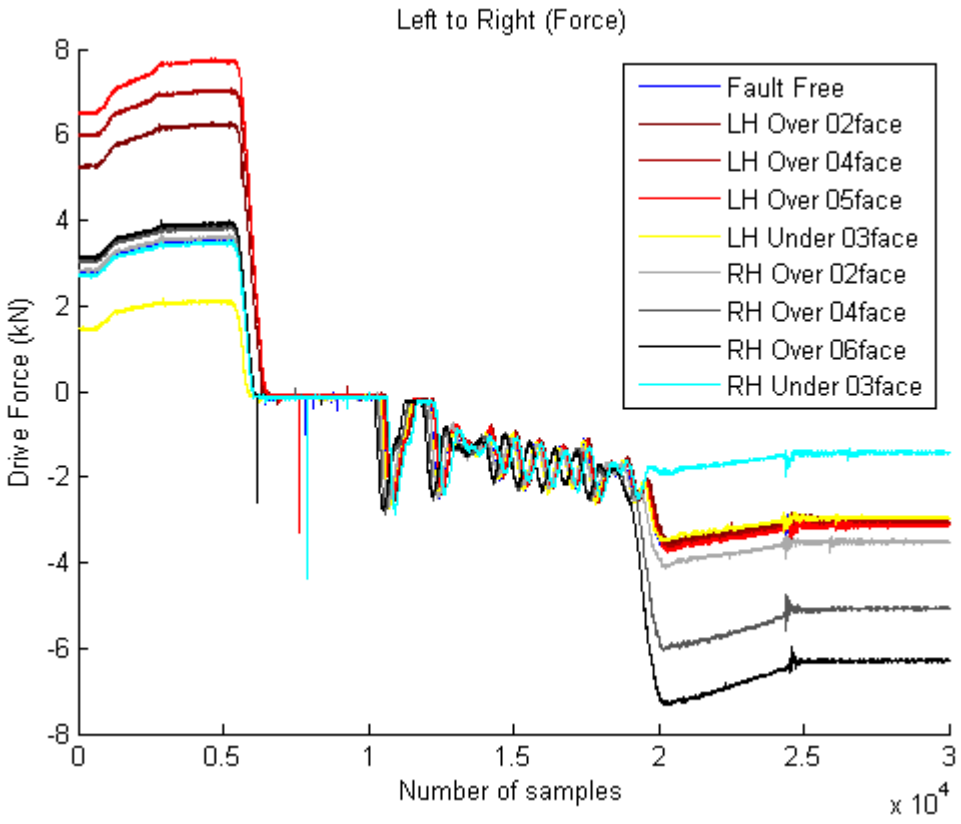

Figure 6-11 drive force conditions simulated

Twenty-seven training data sets (3 data sets per drive force condition), 9 validation data sets (single data set per drive force condition) and 78 test data sets (remaining data sets) were used in this experiment; each data set comprised of electrical current data collected during 'Left to Right' operation associated with 'Left-hand force'.

Validation data sets were used for 'Early stopping' to avoid 'Overfitting'. Nine hidden neurons were utilised, as in the previous section. The neural network which had the best performance for validation data over fifty calculations (changing initial weights each time) was used for drive force prediction. Figure 6-12 shows the performance results for 'Training data sets', 'Validation data sets' and 'Test data sets. Figure 6-13 shows the regression plot for test data sets.

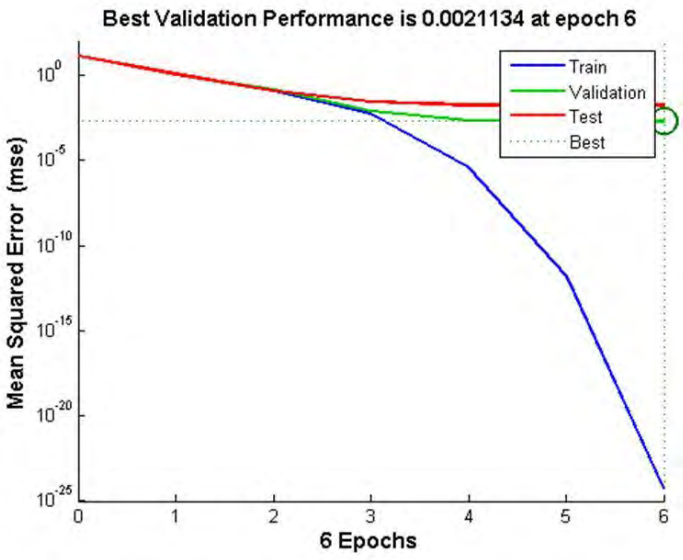

Figure 6-12 Performance of the neural network for Training, Validation and Test data sets

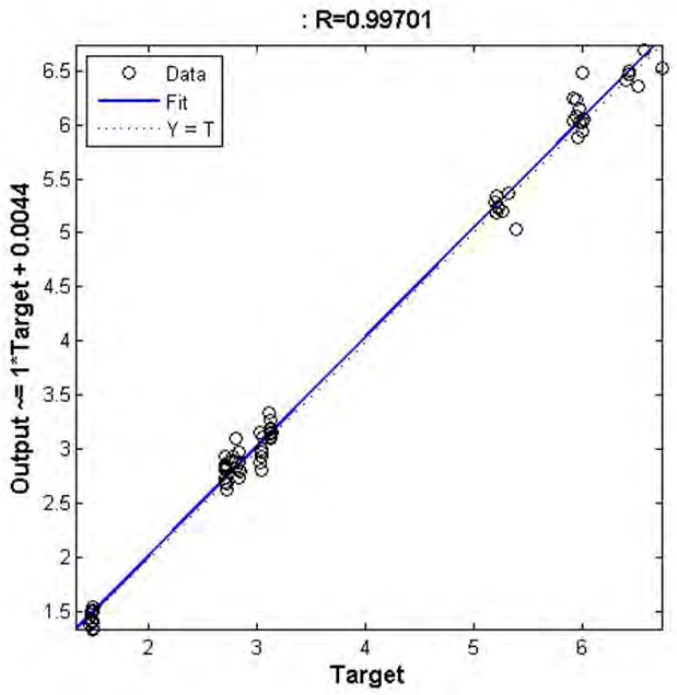

Figure 6-13 Regression plot for test data sets

From Figure 6-12, it can be seen that the performance of the validation data sets and test data sets was similar and the mean squared error for the test data sets was small, which implies that the drive force prediction was accurate.

From Figure 6-13 it can be seen that prediction of drive force is highly correlated with real measurement (target), similarly to the results obtained for the Surelock-type point machine.

It was found that the method used to predict drive force can be used for HW-type point machines. From the experiment carried out so far, it was found that an accurate 'Left hand' drive force prediction can be carried out using electrical current waveform acquired during 'Left to Right' turn. Drive force predictions for other combinations are

yet to be examined: Right hand drive force prediction from 'Left to Right' waveform, Right hand drive force prediction from 'Right to Left' waveform and Left hand drive force prediction from 'Right to Left' waveform.

Further experiments were carried out to predict 'Right hand' force from 'Left to Right' electrical current data, to predict 'Right hand' force from 'Right to Left' electrical current data and to predict 'Left hand' force from 'Right to Left' electrical current data using the same method, as was in 6.2.3 (Surelock-type point machine). Figure 6-14, Figure 6-15 and Figure 6-16 show the results of the experiments. The number of training data sets, validation data sets and test data sets remain the same as for the first experiment in this section for all the experiments: 27 training data sets (3 data sets per drive force condition), 9 validation data sets (single data set per drive force condition) and 78 test data sets (remaining data sets).

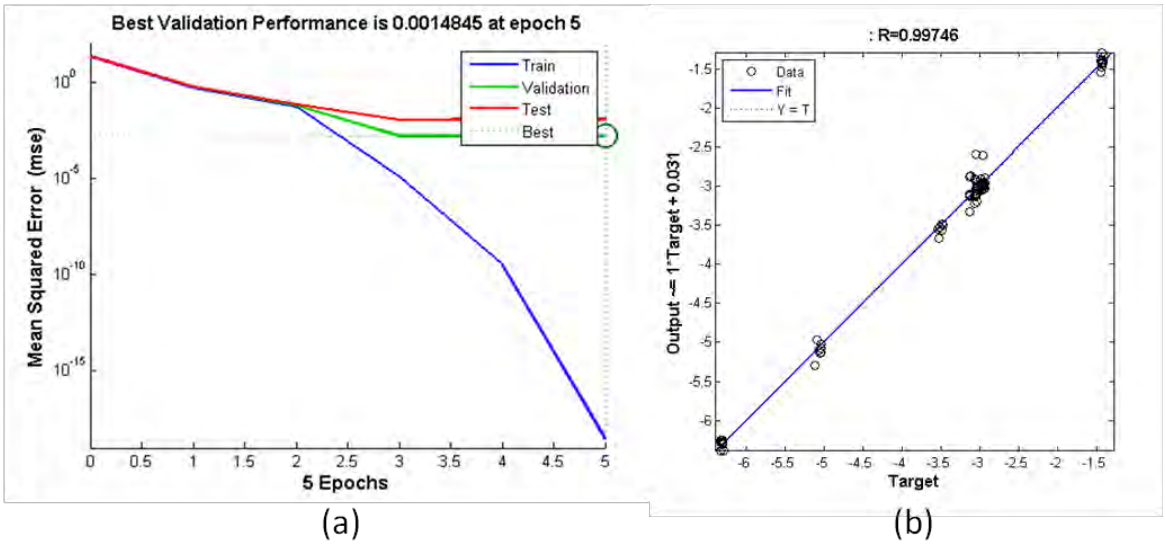

Figure 6-14 (a) Performance plot and (b) regression plot for predicting 'Right hand' force from 'Left to Right' electrical current data

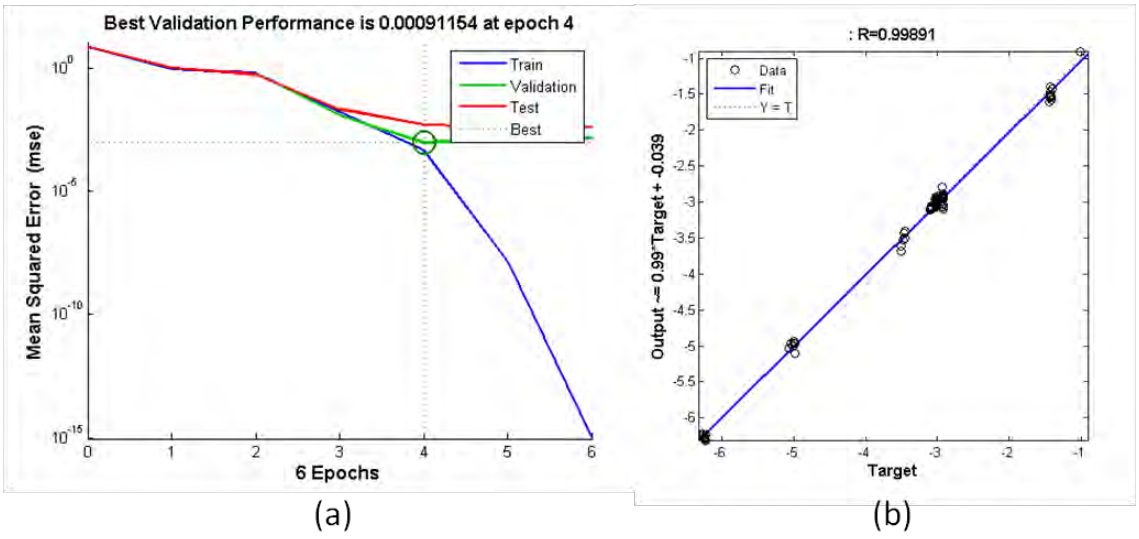

Figure 6-15 (a) Performance plot and (b) regression plot for predicting 'Right hand' force from 'Right to Left' electrical current data

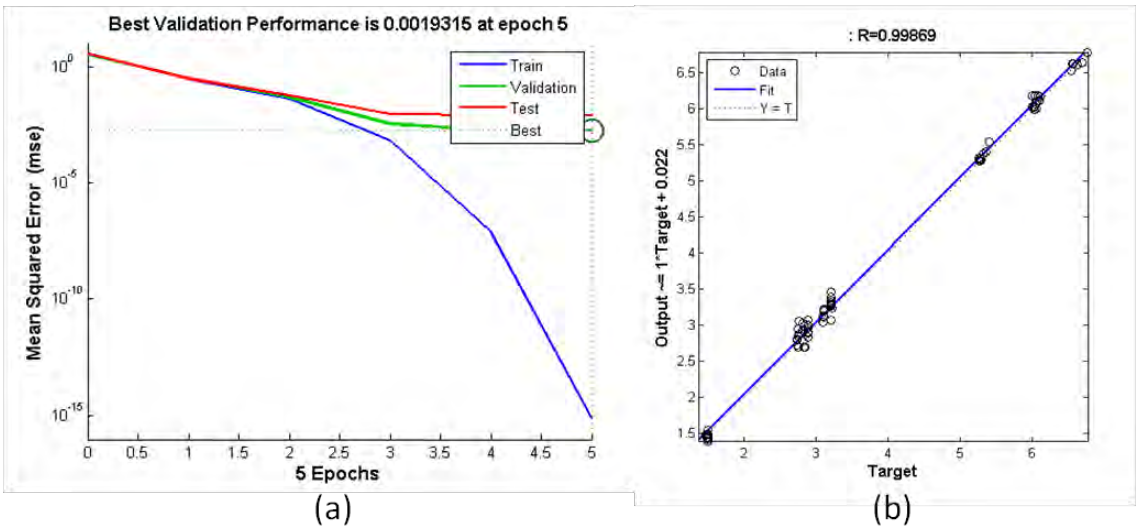

Figure 6-16 (a) Performance plot and (b) regression plot for predicting 'Left hand' force from 'Right to Left' electrical current data

It was found that accurate drive force prediction can be carried out for all the cases showing similar results to the first experiment carried out in 6.3.1. From this result, it

can be said that each waveform contains information for both side of drive force, as was the case for the experiment carried out in 6.2.3 (Surelock-type point machine).

6.3.2 Experiment 2: Further testing the ability of the algorithm (increasing data sets)

The aim of this experiment is to test the algorithm for further data sets. So far, the algorithm has been tested for the data in which one side of the drive force was fixed to the normal drive force. It is, however, possible that both sides of drive force will be misaligned. Figure 6-17 shows 16 drive force conditions added to the drive force conditions used in Experiment 1 (Figure 6-11).

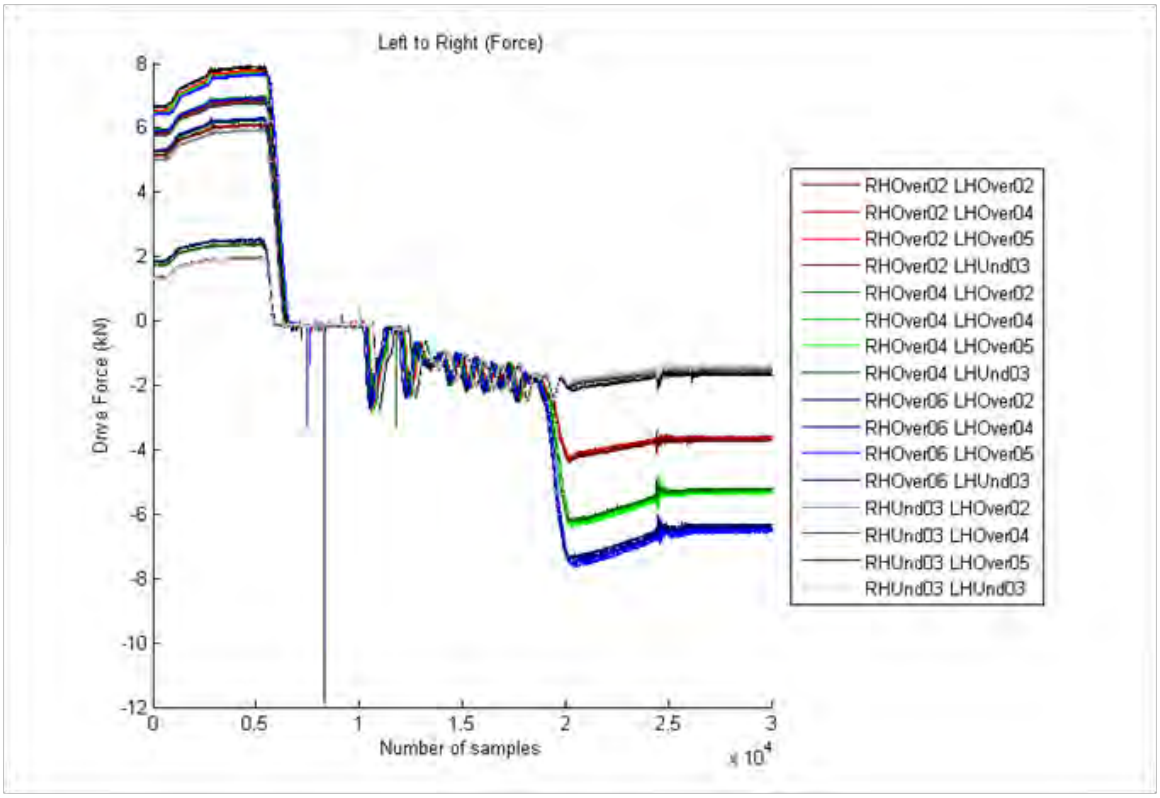

Figure 6-17 Drive force conditions simulated

Together with the drive force conditions used in Experiment 1, there are 25 drive force conditions.

75 training data sets (3 data sets per drive force condition), 25 validation data sets (single data set per drive force condition) and 178 test data sets (remaining data sets) were used in this experiment; each data set comprised of electrical current data collected during 'Left to Right' operation associated with 'Left-hand force'.

Validation data sets were used for 'Early stopping' to avoid 'Overfitting'. Nine hidden neurons were utilised, as in 6.2.3 (Surelock-type point machine). The neural network which had the best performance for validation data over fifty calculations (changing initial weights each time) was used for drive force prediction. Figure 6-18 shows the result of performance for 'Training data sets', 'Validation data sets' and 'Test data sets'. Figure 6-19 shows the regression plot for test data sets.

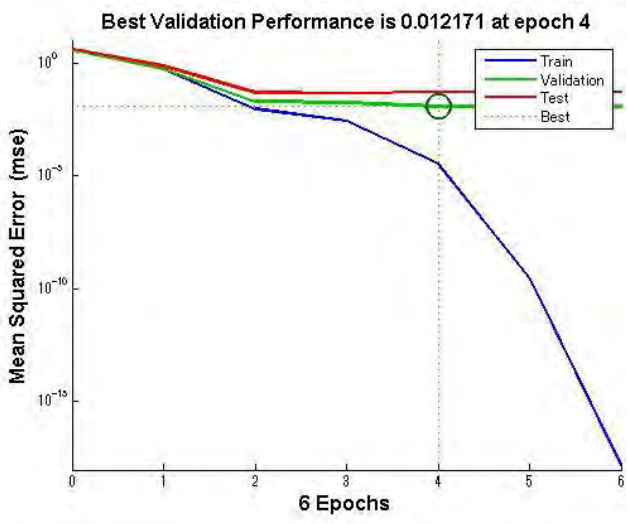

Figure 6-18 Performance of the neural network for Training, Validation and Test data sets

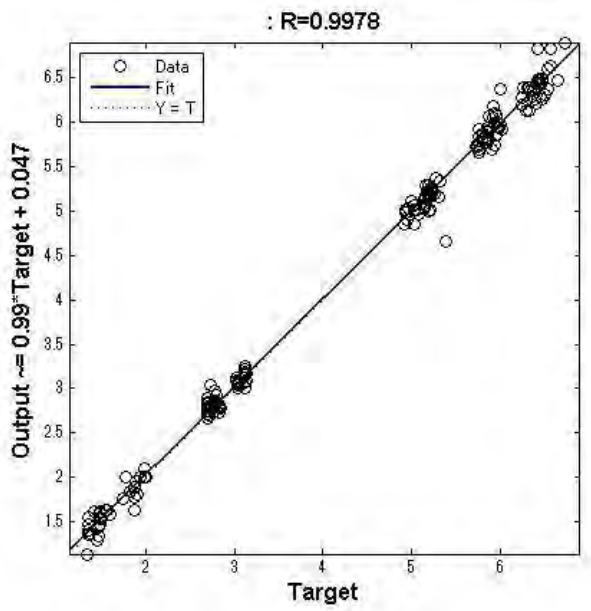

Figure 6-19 Regression plot for test data sets

From Figure 6-18, it can be seen that the performance of the validation data sets and the test data sets was similar and the mean squared error for test data sets was small, which implies that the drive force prediction was accurate.

From Figure 6-19 it can be seen that prediction of drive force is highly correlated with real measurement (target), similarly to the results obtained for the Surelock-type point machine.

It was found that the method used to predict drive force can be used for drive force conditions in which one side is not fixed to normal drive force. From the experiment carried out so far, it was found that an accurate 'Left hand' drive force prediction can be carried out using electrical current waveform acquired during 'Left to Right' turn. Drive force predictions for other combinations are yet to be examined: Right hand drive force prediction from 'Left to Right' waveform, Right hand drive force prediction from

'Right to Left' waveform and Left hand drive force prediction from 'Right to Left' waveform.

Further experiments were carried out to predict 'Right hand' force from 'Left to Right' electrical current data, to predict 'Right hand' force from 'Right to Left' electrical current data and to predict 'Left hand' force from 'Right to Left' electrical current data using the same method, as was in 6.2.3 (Surelock-type point machine). Figure 6-20, Figure 6-21 and Figure 6-22 show the results of the experiments. The number of training data sets, validation data sets and test data sets remain the same as the first experiment in this section for all the experiments: 27 training data sets (3 data sets per drive force condition), 9 validation data sets (single data set per drive force condition) and 78 test data sets (remaining data sets).

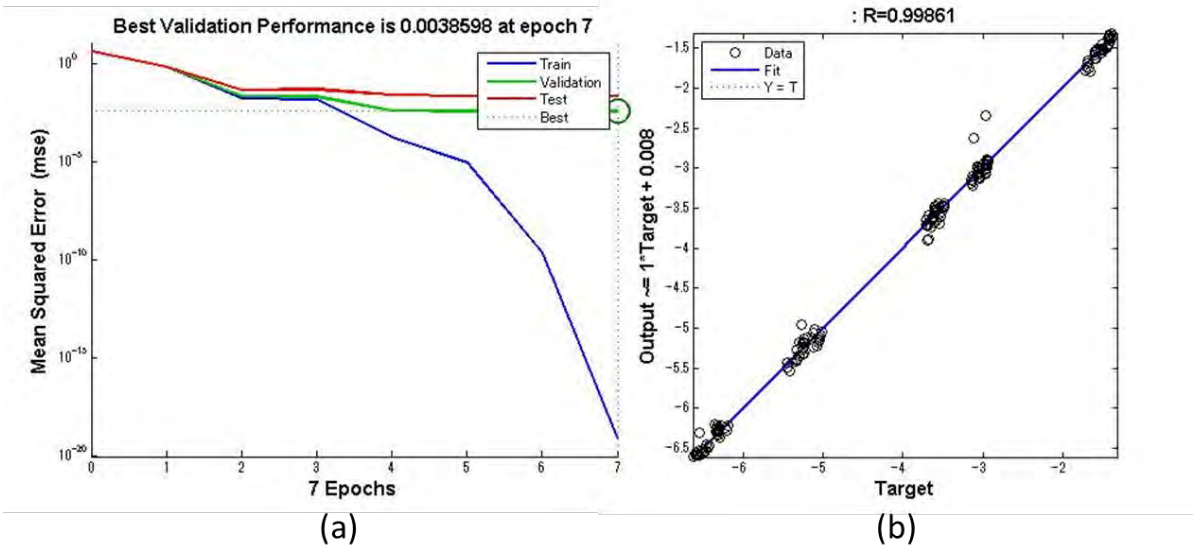

Figure 6-20 (a) Performance plot and (b) regression plot for predicting 'Right hand' force from 'Left to Right' electrical current data

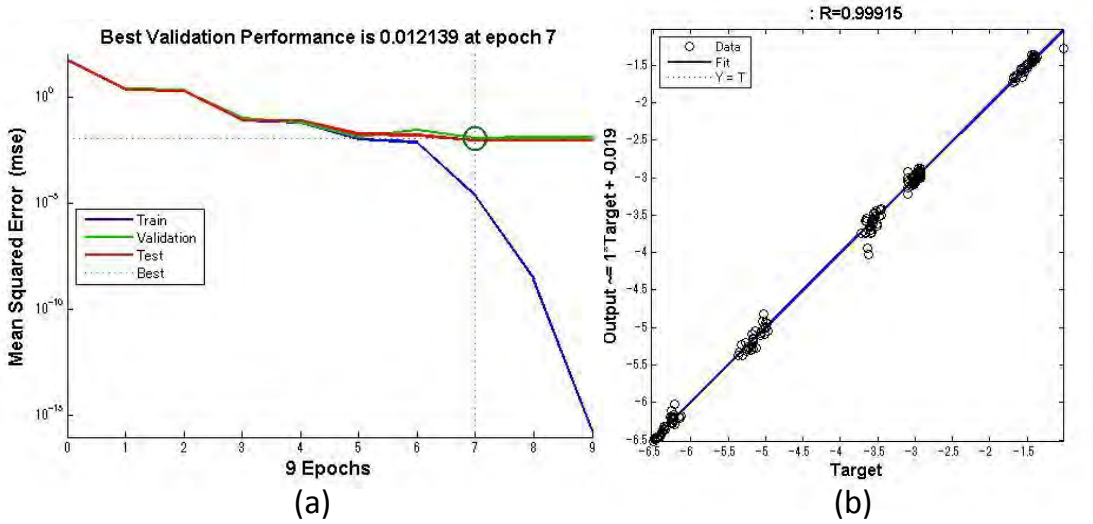

Figure 6-21 (a) Performance plot and (b) regression plot for predicting 'Right hand' force from 'Right to Left' electrical current data

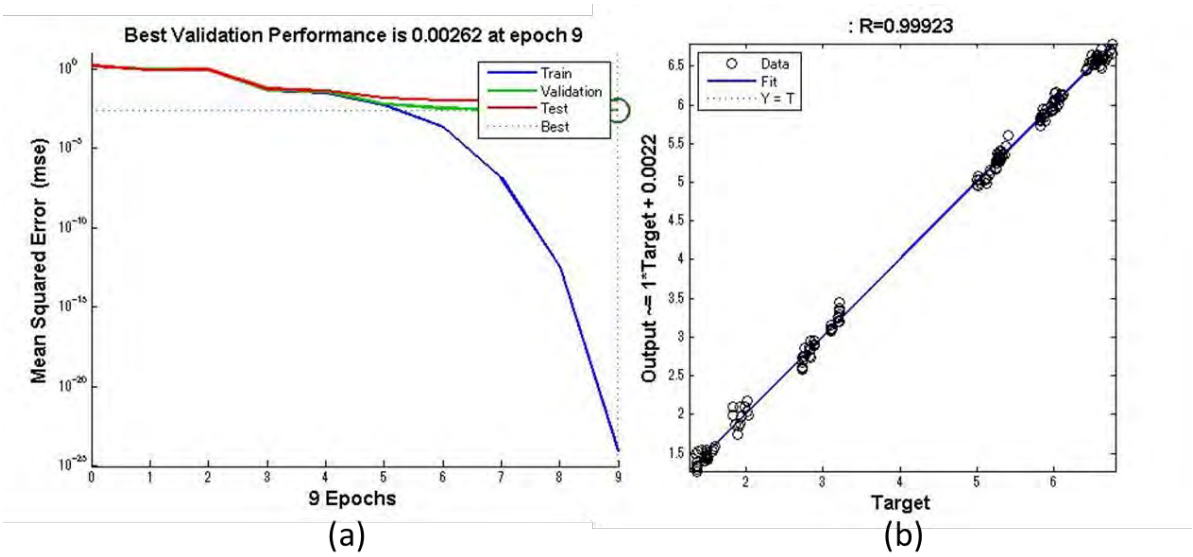

Figure 6-22 (a) Performance plot and (b) regression plot for predicting 'Left hand' force from 'Right to Left' electrical current data

It was found that accurate drive force prediction can be carried out for all cases, showing similar results to the first experiment carried out in 6.3.2.

6.3.3 Experiment 3: Testing the data which is not from the drive force conditions as in training data sets

The aim of this sub-section is to test the data which is not from the drive force conditions as in the training data sets (to check the generalisation ability of the algorithm).

The Neural network model trained in Experiment 2 was used for predicting drive forces. Figure 6-23 shows the drive force conditions simulated for the test data sets (the drive force simulated was in-between the conditions of the drive force condition used in Experiment 1 in 6.3.1).

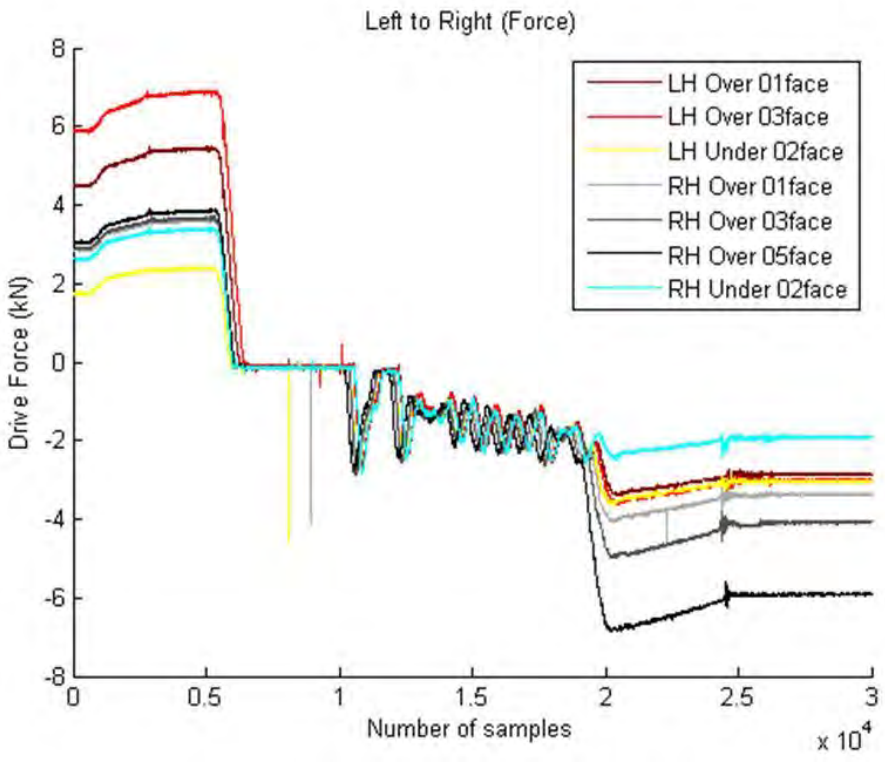

Figure 6-23 Drive force conditions simulated

Twenty-one data sets (3 data sets per drive force condition) were used as test data sets. Figure 6-24 and Figure 6-25 shows the regression plot for test data sets predicting 'Left hand' force from 'Left to Right' electrical current data.

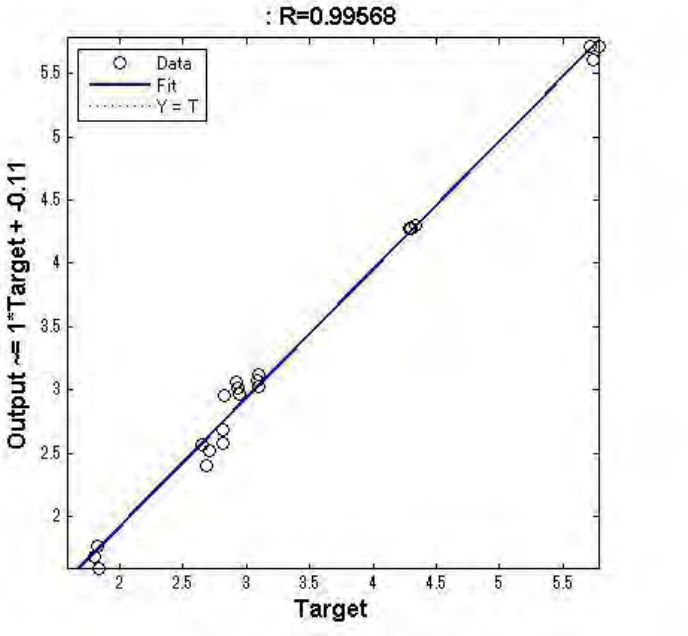

Figure 6-24 Regression plot for test data sets predicting 'Left hand' force from 'Left to Right' electrical current data

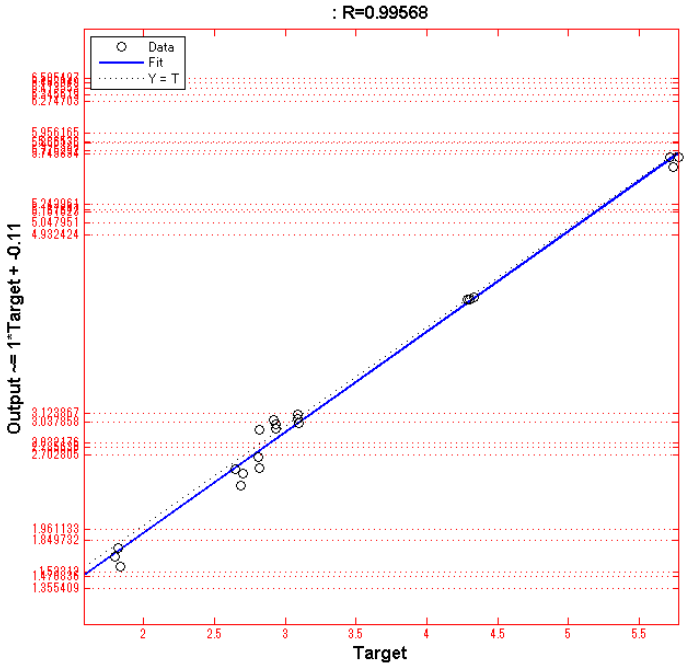

Figure 6-25 Regression plot for test data sets predicting 'Left hand' force from 'Left to Right' electrical current data (drive force for training data is plotted in dotted lines)

From Figure 6-24, it can be seen that prediction of drive force is highly correlated with real measurement (target). It was found that the method can achieve a good result for the data from in-between conditions of training data sets. Furthermore, Figure 6-25 shows regression plot in which drive force for training data is also depicted (an average drive force for each drive force condition as in training data is calculated and depicted in dotted lines in the figure). From Figure 6-25, it can be seen that the outputs (drive force prediction) were accurate even in the case where no similar drive force condition was contained in the training data. This result shows the generalisation ability of the trained network.

From the experiment carried out so far, it was found that an accurate 'Left hand' drive force prediction can be carried out using electrical current waveform acquired during

'Left to Right' turn. Drive force predictions for other combinations are yet to be examined: Right hand drive force prediction from 'Left to Right' waveform, Right hand drive force prediction from 'Right to Left' waveform and Left hand drive force prediction from 'Right to Left' waveform.

Further experiments were carried out to predict 'Right hand' force from 'Left to Right' electrical current data, to predict 'Right hand' force from 'Right to Left' electrical current data and to predict 'Left hand' force from 'Right to Left' electrical current data using the same method, as was in 6.2.3 (Surelock-type point machine). Figure 6-26, Figure 6-27, Figure 6-28, Figure 6-29, Figure 6-30 and Figure 6-31 show the results of the experiments. The number of test data sets remains the same as the first experiment in this sub-section for all the experiments: 21 test data sets (3 data sets per drive force condition).

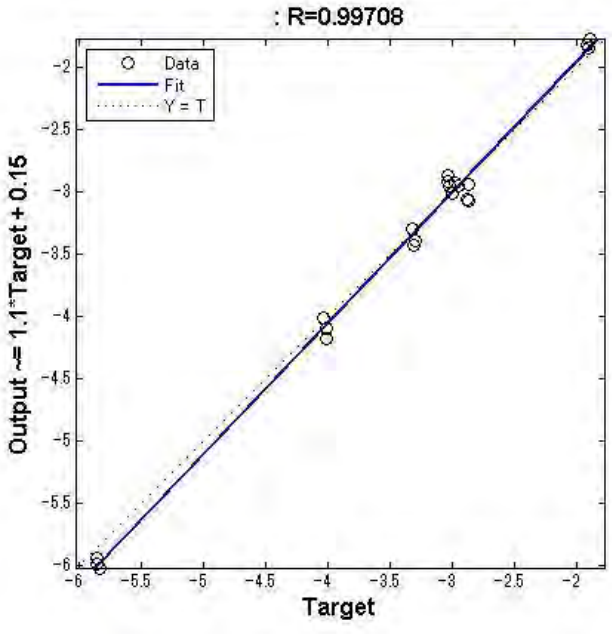

Figure 6-26 Regression plot for test data sets predicting 'Right hand' force from 'Left to Right' electrical current data

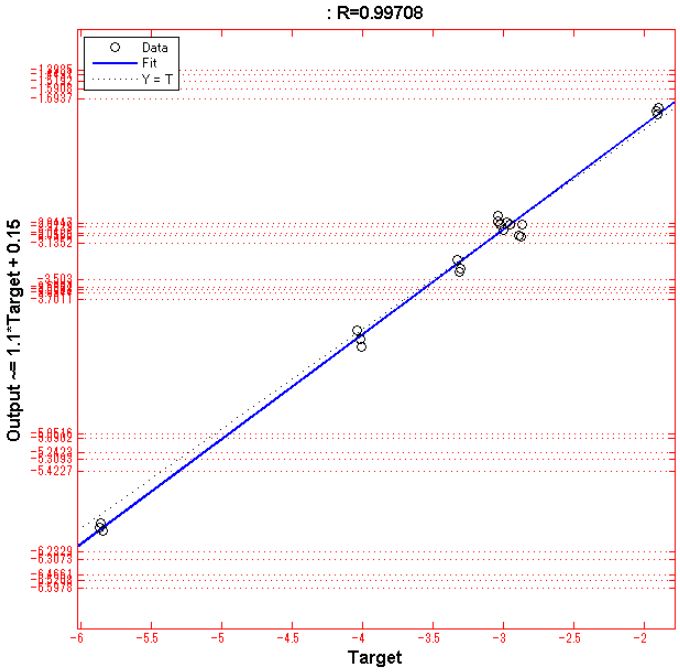

Figure 6-27 Regression plot for test data sets predicting 'Right hand' force from 'Left to Right' electrical current data (drive force for training data is plotted in dotted lines)

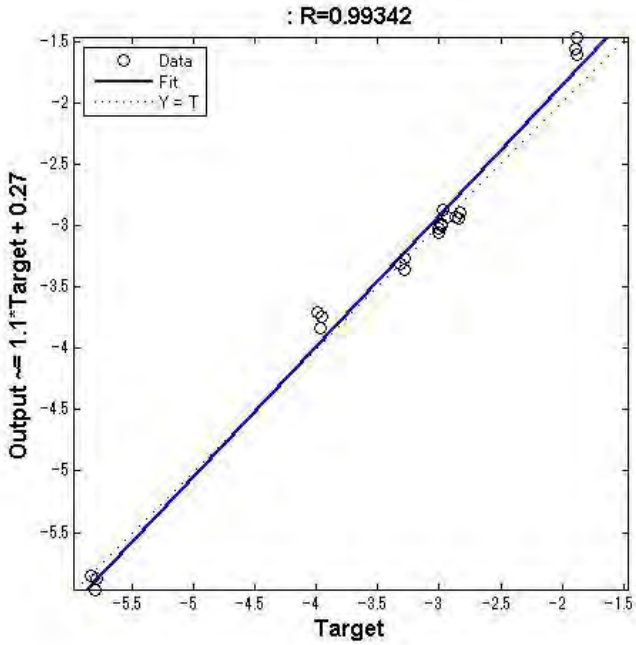

Figure 6-28 Regression plot for test data sets predicting 'Right hand' force from 'Right to Left' electrical current data

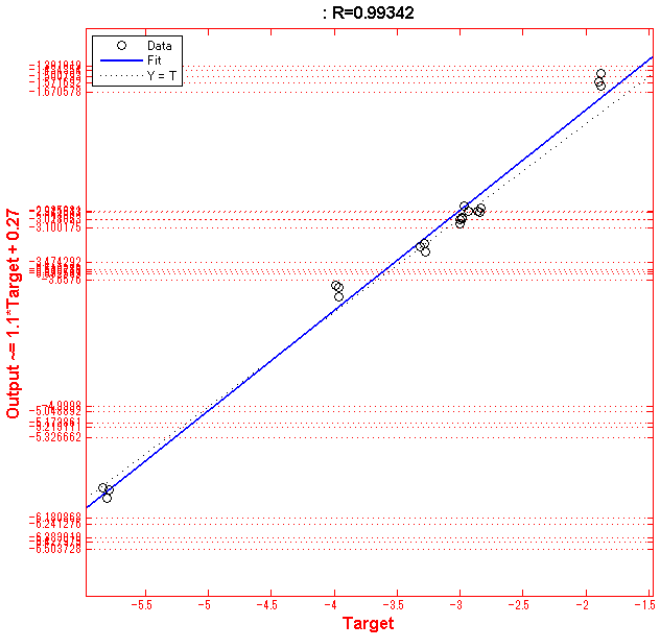

Figure 6-29 Regression plot for test data sets predicting 'Right hand' force from 'Right to Left' electrical current data (drive force for training data is plotted in dotted line)

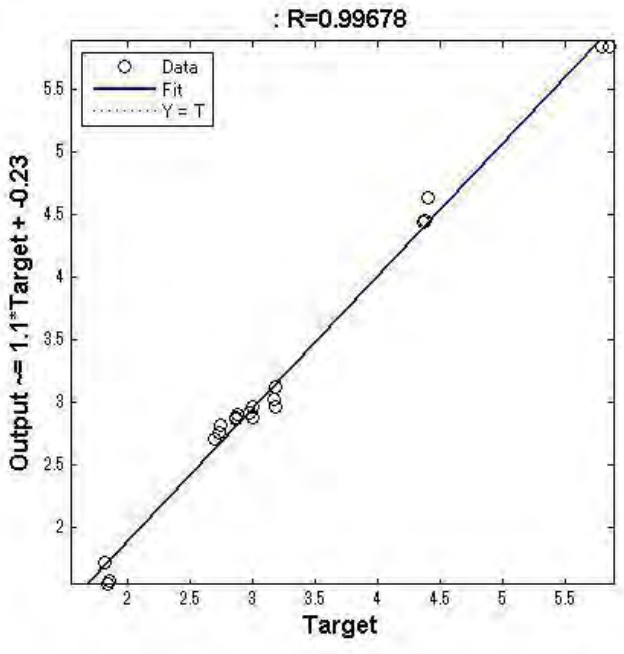

Figure 6-30 Regression plot for test data sets predicting 'Left hand' force from 'Right to Left' electrical current data

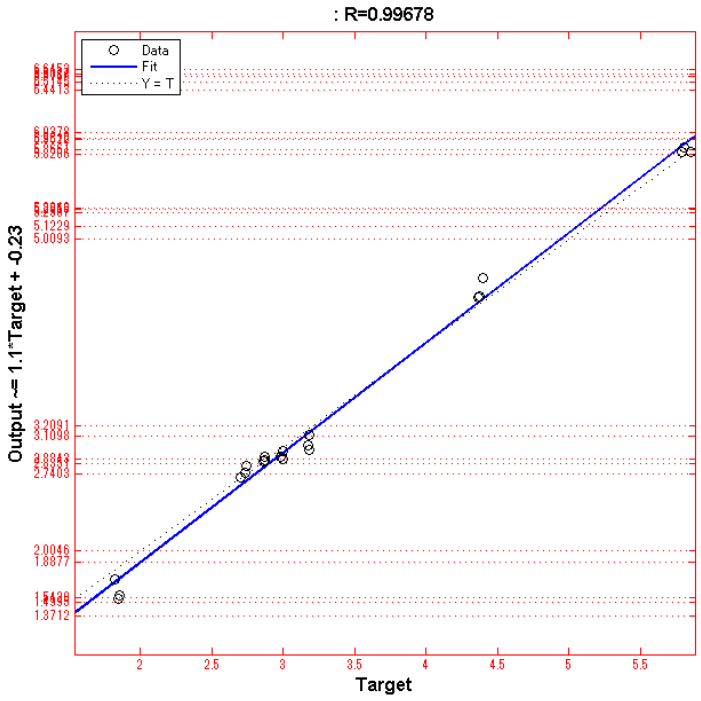

Figure 6-31 Regression plot for test data sets predicting 'Left hand' force from 'Right to Left' electrical current data (drive force for training data is plotted in dotted lines)

It was found that accurate drive force prediction (for test data where no similar drive force condition was obtained in the training data) can be carried out for all the cases, showing the similar results to the first experiment carried out in 6.3.3. It can be said that the trained neural network had generalisation ability for all the cases.

6.4 Conclusions

In this chapter, a method to predict drive force was developed based on the method written in Chapter 4 and 5. The method utilises the DWT for feature extraction and a Neural network for drive force prediction.

Two point machines were tested: the Surelock-type point machine and the HW-type point machine. It was found that an accurate drive force prediction can be carried out using the proposed method for both point machines. As for HW-type point machine, it was found that an accurate drive force prediction can be done even if both sides of drive force are misaligned simultaneously. Finally, the data which is not from the drive force conditions as in the training data sets were tested to check the generalisation ability of the algorithm, and it was found that the system can accurately predict the drive force. Eventually, the error between the target (actual drive force) and the output (drive force prediction) was approximately 20% for the worst case.

An advantage of using this method is that it can directly predict drive force. Since drive force is a direct and understandable parameter for maintenance staff, predicting drive force can be beneficial in real operational use.

A disadvantage, however, might be that it is necessary to collect more data (including different drive force condition) than the classification method to make an accurate prediction. Additionally, since the neural network is susceptible to bad initial weights, it was necessary to carry out the calculation a number of times to find the initial weights of the neural network which has good performance (otherwise, the neural network could possibly tune the weights into local optima and therefore perform poorly).

CHAPTER 7 CONCLUSIONS AND FURTHER WORK

7.1 Introduction

In this chapter, conclusions are presented that have been drawn during the course of the research. Details of suggested future work including recommendations for an appropriate system architecture for practical implementation are also presented in the chapter.

Many years have passed since the first point machine condition monitoring systems were installed in the field. Currently condition monitoring systems are becoming widely accepted by many railway companies around the world, however, the fault detection of the point machine is still carried out by simple threshold algorithms. During this period of time, computers (both hardware and software) have developed significantly. Machine learning algorithms have also developed and they are used in many applications around the world (e.g. speech recognition and character recognition [61]). In this thesis, the fault detection and diagnosis of point machines are carried out using machine learning algorithms (k-means, silhouette width, support vector machine, and neural networks).

During the course of the research a total of five data sets from both Japanese and British point machines were collected to ensure that the results of the research were generic.

Research regarding point machine condition monitoring has been carried out previously (for more than 10 years) at the Birmingham Centre for Railway Research and Education at the University of Birmingham; the research detailed in this thesis was developed from this existing knowledge.

7.2 Conclusions

The current methodology for condition monitoring of point machines has limitation in terms of the ability to detect incipient faults prior to failure (so when the system detects faults the point machine has already failed or nearly failed). Significant improvements of the railway service (in terms of dependability) can be made if the system can detect these incipient faults in the early stage of their development and inform users of faults before these faults get worse.

To accomplish this task, a detailed analysis has been carried out to the waveforms and advanced machine learning algorithms have been utilised in this thesis.

The following is a short summary of the achievements presented in this thesis:

- Collected data from Japanese point machine (NTS-type point machine used in Central Japan Railway Company, Japan) simulating multiple fault conditions (data acquired: January 2011);

- Developed an algorithm for fault detection and diagnosis for Japanese AC point machine, as presented in Chapter 4. 100% cross validation accuracy was achieved for data relating to three fault conditions (experiment 1 in Chapter 4) and five fault conditions (experiment 2 in Chapter 4);
- Collected data from a type of point machine operated in Great Britain (Surelock-type point machine used in London Underground, London) where multiple fault conditions were simulated (data acquired: September 2011)
- Collected data from a type of point machine operated in Great Britain (M63-type point machine used in London Underground, London) where multiple fault conditions were simulated (data acquired September 2011)
- Collected data from a type of point machine (two data sets) operated in Great Britain (HW-type point machine used in Network Rail, Derby) where multiple fault conditions were simulated (data acquired May 2012)
- Applied the algorithm of fault detection and diagnosis (developed for Japanese point machine) to DC point machines operated in Great Britain to check the transferability of the algorithm to other types of point machine, as written in Chapter 5.

  For the Surelock-type point machine, a 99.33% cross validation accuracy was achieved for five fault conditions using both Linear and RBF kernel, whereas a 99.57% and a 98.68% cross validation accuracy were achieved for eight fault conditions using Linear and RBF kernel respectively.

  For M63-type point machine, a 99.33% and a 98.57% cross validation accuracy were achieved for five fault conditions using Linear and RBF kernel respectively,

whereas a 100% and a 99.62% cross validation accuracy were achieved for nine fault conditions using Linear and RBF kernel respectively;

- A method to express the qualitative features was proposed to test the transferability of the specific algorithm parameters from one instance of a point machine to the next, which was tested on point machines operated in Great Britain (two HW-type point machines), as presented in Chapter 5. An 80.4% and a 73.2% classification accuracy were achieved using Linear kernel and RBF kernel respectively (data collected from point machine 1 were used as training data; data collected from point machine 2 were used as test data). A 65.57% and a 75.41% classification accuracy were achieved using Linear Kernel and RBF kernel respectively (data collected from point machine 2 were used as training data; data collected from point machine 1 were used as test data).
- A method to directly predict the drive force (from current data) was proposed and tested, as written in Chapter 6. An accurate drive force prediction was achieved (an average of the correlation coefficients over all the combinations was 0.996) for the Surelock-type point machine.

  An accurate prediction was also achieved (an average of the correlation coefficients over all the combinations was 0.998) for the HW-type point machine. More data were added, but still an accurate prediction was achieved (an average of the correlation coefficients over all the combination was 0.999).

As listed above, there are a lot of benefits using the fault detection and diagnosis methods proposed in this thesis. Both fault free data and faulty data of the point machine have been collected from various type of point machines and it was found that

the proposed method can accurately detect and diagnose the faults (when the training data sets and test data sets have been collected from the same point machine). It was not possible to accomplish these results using conventional methods such as thresholding methods since the changes of the waveforms are subtle. Furthermore, a method to predict drive force was proposed. A good result (accurate prediction) was also acquired using the proposed method, as written in the list above.

The method proposed in the thesis requires only current and voltage transducers for condition monitoring (a single current transducer for DC point machines and both current and voltage transducers for AC point machines). This is important in real implementation since it is possible to acquire the data without the possibility of affecting the operation of the point machine. (Conversely, as described in Chapter 4, a load pin inserted in the drive assembly of the point machine can directly affect the operation of the point machine if the sensor goes wrong and this can be a significant risk for the railway) Therefore, the fact that a good result was accomplished only using accessible sensors (current transducer and voltage transducer) has a significant advantage for practical condition monitoring of point machines.

The work presented in this thesis has been published in an academic conference paper and two journals, which can be found in Appendix A.

7.3 Further work

It has been demonstrated in the thesis that the proposed method using the developed machine learning algorithm is valid for detecting and diagnosing incipient faults. It is, however, important to mention the further work required before fully implementing the proposed algorithm.

Firstly, a system architecture should be provided to achieve a practical condition monitoring system.

A recommendation towards a condition monitoring system architecture is provided. It is important to propose a system architecture as well as an algorithmic approach to support the practical realisation of a system.

Condition monitoring systems are already used in Japanese railway and it is practical to make use of the existing system architecture and make necessary changes in order to realise the proposed condition monitoring method. A system architecture for the existing Japanese point machine condition monitoring system in Central Japan Railway Company is depicted in Figure 7-1.

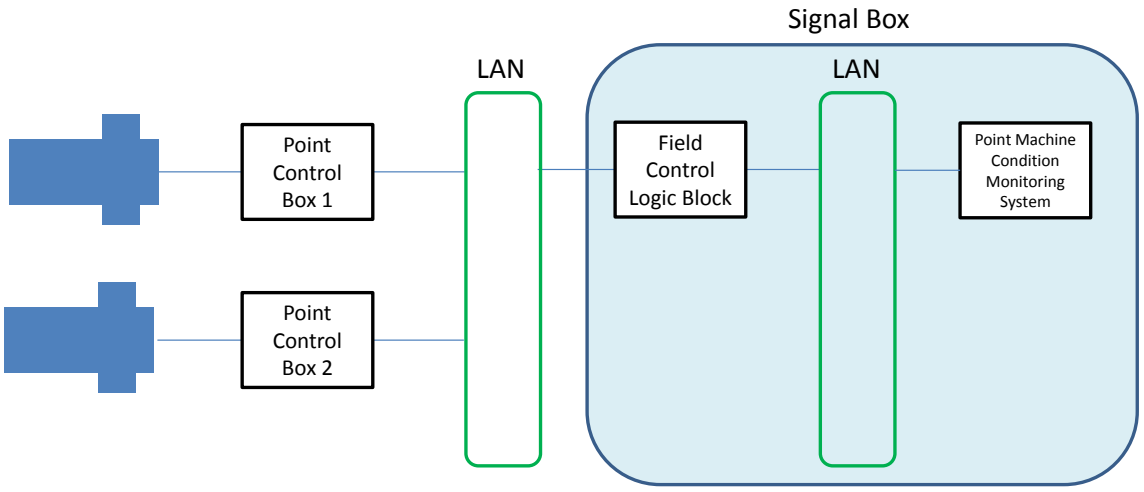

Figure 7-1 A system architecture for Japanese point machine condition monitoring system

A Point control box is installed near the point machine and a computer (for processing control and status signals) and a transmission device (for transmitting data) is installed in the box. Electrical current and voltage are provided from a point control box to a point machine, and electrical current and voltage transducers are already implemented in the point control box in the Japanese railway (Central Japan Railway Company).

A point control box is connected to the field control logic block (which processes all the control and status signals from all the point machines and signals in the field) via a LAN. Finally, a point machine condition monitoring system (a computer) is connected to the field control box via the LAN.

An approach proposed for Japanese point machines (in Chapter 4) is using (1) wavelet transforms for feature extraction and (2) a Support Vector Machine for classifying. Changes to the existing software are enough to implement the proposed method since

the proposed method only requires current and voltage transducers which are already implemented in the point control box.

Feature extraction should be carried out locally to the point machine. Feature extraction can be carried out through the upgrading of software on the existing embedded computer located in the point control box. Once a point machine is operated, the embedded computer inside the point control box can calculate the active power of the point machine and then run the feature extraction routine. The extracted features will be passed to the point condition monitoring system via the LAN using the existing transmission device in the box. Using this approach the amount of data transmitted across the LAN can be contained, thus ensuring a minimal load on the network. The Support Vector Machine classifier can be installed in the point machine condition monitoring system inside the signal box. It is important to note that these changes can be carried out by making only software changes (therefore not changing the system architecture of the existing system).

Secondly, further data collection and analysis should be carried out. The data collection and analysis needed to be carried out in the future is listed below:

- More data should be collected from point machines actually operating in the field (outside). All of the data of the point machines (throughout the thesis) was collected from the training facilities maintained by railway companies. These point machines contains all the switch components (switch blade, drive rod, etc.), so it can be said that the data collected are highly realistic data. But considering

the fact that some of the point machines (Surelock, HW and M63 point machines) are installed inside buildings (therefore, maintained in a good condition) and that there are no trains run over these switches, it may be necessary to collect and validate the approach using the data from point machines installed in the railway field in the future;

- Data including temperature changes of the switch blade should be collected. All of the data of point machines (throughout the thesis) was collected within 2-3 days (during the daytime) because of the limitation of the usage of the facility (the data was collected from training facilities in railway companies). It is important to carry out an analysis of how temperature changes (i.e. seasonal changes) affect the performance of the fault detection and diagnosis approach;
- Data including various lengths and various components of switches should be collected. It is needed to analyse how the length of switches (and components) can affect the data. This may help develop a more sophisticated algorithm that can make the training data sets collected from one point machine transferable to others and thus reliably carrying out fault detection and diagnosis;
- Long-term data should be collected. Due to the limitation in time for the research, it was not possible to collect long-term data (for example, one or two years of operational data of point machines in the field). Data from the point machine may change because of the changes of condition and maintenance actions. It may be necessary to collect and analyse the long-term data before actually implementing the proposed method.

APPENDIX A PUBLISHED PAPERS

The following shows the papers that have been published during the course of the PhD study.

1. Asada T, Roberts C. (2011). Development of an effective condition monitoring system for AC point machines, 5th International Conference on Railway Condition Monitoring and Non-destructive Testing (RCM 2011), UK.
2. Asada T, Roberts C, Koseki T. An algorithm for improved performance of railway condition monitoring equipment: Alternating-current point machine case study (published in Transportation Research Part C – Emerging Technologies)
3. Asada T, Roberts C. Improving the dependability of DC point machines with a novel condition monitoring system, (published in the Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit).

REFERENCES

1. Potters bar derailment: report and recommendations. Rail Safety & Standards Board; 2010 [cited 2011 19/10]; Available from: http://www.rssb.co.uk.
2. Singou system no sinpo to hatten [Development of Japanese signaling system]: Japan Railway Electrical Engineering Association; 2009.
3. Moubray J. Reliability Centred Maintenance 2nd Ed.: Industrial Press Inc; 1997.
4. Silmon J. Operational industrial fault detection and diagnosis: Railway actuator case studies. PhD Thesis, University of Birmingham2009.
5. Blanchard BS, Fabrycky WJ. Systems engineering and analysis. 5th ed. Boston, Mass. ; London: Pearson; 2011.
6. Engineering ICoS. What is Systems Engineering? [cited 2012 25/01]; Available from: http://www.incose.org/practice/whatissystemseng.aspx.
7. Roberts C, Goodall R. Strategies and techniques for safety and performance monitoring on railways. 7th IFAC Symposium on Fault detection, Supervision and Safety of Technical Processes; Spain 2009. p. 746-55.
8. Chamroukhi F, Same A, Govaert G, Aknin P. Time series modeling by a regression approach based on a latent process. Neural Netw. [Proceedings Paper]. 2009 Jul-Aug;22(5-6):593-60.
9. Style 63 point machine installation and maintenance: Westinghouse Brake and Signal Co. Ltd.
10. Nakamura K. Tentetsu Souchi [The point machine]: Japan Railway Electrical Engineering Association; 1992.
11. Igarashi.Y, Shiomi.S. Development of monitoring system for electric switch machine. QR of RTRI 2006;47(2).
12. Silmon JA. Understanding the potential for points condition monitoring as part of an intelligent infrastructure initiative: Network Rail internal report 2008.
13. Venkatsubramanian V, Rengaswamy R, Yin K, Kavuri SN. A review of process fault detection and diagnosis Part I: Quantitative model-based methods. Computers & Chemical Engineering 2003 Mar 15;27(3):293-311.
14. Venkatasubramanian V, Rengaswamy R, Kavuri SN. A review of process fault detection and diagnosis Part II: Quantitative model and search strategies. Computers & Chemical Engineering 2003 Mar 15;27(3):313-26.
15. Venkatasubramanian V, Rengaswamy R, Kavuri SN, Yin K. A review of process fault detection and diagnosis Part III: Process history based methods. Computers & Chemical Engineering 2003 Mar 15;27(3):327-46.

16. Frank PM. Fault-Diagnosis in Dynamic-Systems Using Analytical and Knowledge-Based Redundancy - a Survey and Some New Results. Automatica 1990 May;26(3):459-74.
17. Chang CT, Hwang JI. Simplification techniques for EKF computations in fault diagnosis - suboptimal gains. Chemical Engineering Science 1998 Nov;53(22):3853-62.
18. Wu M, Gui WH, Shen DY, Wang YL. Expert fault diagnosis using rule models with certainty factors far the leaching process. Proceedings of the 3rd World Congress on Intelligent Control and Automation, Vols 1-52000:238-41, 3766.
19. Ulerich NH, Powers GJ. Online Hazard Aversion and Fault-Diagnosis in Chemical Processes - the Digraph + Fault-Tree Method. Ieee Transactions on Reliability 1988 Jun;37(2):171-7.
20. Rengaswamy R, Venkatasubramanian V. A Syntactic Pattern-Recognition Approach for Process Monitoring and Fault-Diagnosis. Engineering Applications of Artificial Intelligence 1995 Feb;8(1):35-51.
21. Wong JC, McDonald KA, Palazoglu A. Classification of process trends based on fuzzified symbolic representation and hidden Markov models. Journal of Process Control 1998 Oct-Dec;8(5-6):395-408.
22. Wang WQ, Golnaraghi MF, Ismail F. Prognosis of machine health condition using neuro-fuzzy systems. Mechanical Systems and Signal Processing 2004 Jul;18(4):813-31.
23. Ge M, Du R, Zhang GC, Xu YS. Fault diagnosis using support vector machine with an application in sheet metal stamping operations. Mechanical Systems and Signal Processing 2004 Jan;18(1):143-59.
24. Shaw DC. A universal approach to points condition monitoring. IET International Conference on Railway Condition Monitoring 2008.
25. Zhou F, Archer N, Bowles J, Duta M, Henry M, Tombs M, Zamora M, Baker S, Burton C. Remote condition monitoring and validation of railway points. Computing & Control Engineering Journal 2002 Oct;13(5):221-30.
26. Zhou FB, Duta MD, Henry MP, Baker S, Burton C, editors. Remote condition monitoring for railway point machine. Railroad Conference, 2002 ASME/IEEE Joint; 2002 23-25 April 2002.
27. Zhou F, Duta M, Henry M, Baker S, Burton C, editors. Condition monitoring and validation of railway point machines. Intelligent and Self-Validating Instruments -- Sensors and Actuators (Ref No 2001/179), IEE Seminar on; 2001 14 Dec. 2001.
28. Pabst M, editor. Remote monitoring of points based on effective power dissipation. Condition Monitoring for Rail Transport Systems (Ref No 1998/501), IEE Seminar on; 1998 10 Nov 1998.
29. Abed SK, Fararooy S, Allan J. Pc-based condition monitoring system to verify reliability-centred maintenance of electro-pneumatic point machines on London underground. International Conference on Developments in Mass Transit Systems 1998:222-7.
30. Rouvray P, Hallam P, Danaher S, Thorpe MG, editors. The application of Matlab to railway signalling system fault modelling. The Use of Systems Analysis and Modelling Tools: Experiences and Applications (Ref No 1998/413), IEE Colloquium on; 1998 20 Mar 1998.

31. Pedregal DJ, Garcia FP, Roberts C. An algorithmic approach for maintenance management based on advanced state space systems and harmonic regressions. Annals of Operations Research 2009 Feb;166(1):109-24.
32. Marquez FPG, Paul W, Roberts C. Failure analysis and diagnostics for railway trackside equipment. Engineering Failure Analysis 2007 Dec;14(8):1411-26.
33. Marquez FPG, Schmid F. A digital filter-based approach to the remote condition monitoring of railway turnouts. Reliability Engineering & System Safety 2007 Jun;92(6):830-40.
34. Marquez FPG, Tercero DJP, Schmid F. Unobserved component models applied to the assessment of wear in railway points: A case study. European Journal of Operational Research 2007 Feb 1;176(3):1703-12.
35. Marquez FPG, Schmid F, Collado JC. Wear assessment employing remote condition monitoring: a case study. Wear 2003 Aug-Sep;255:1209-20.
36. Marquez FPG, Schmid F, Collado JC. A reliability centered approach to remote condition monitoring. A railway points case study. Reliability Engineering & System Safety 2003 Apr;80(1):33-40.
37. Pedregal DJ, Garcia FP, Schmid F. RCM2 predictive maintenance of railway systems based on unobserved components models. Reliability Engineering & System Safety 2004 Jan;83(1):103-10.
38. Zattoni E. Detection of incipient failures by using an H-2-norm criterion: Application to railway switching points. Control Engineering Practice 2006 Aug;14(8):885-95.
39. Marquez FPG. Binary decision diagrams applied to fault tree analysis. IET International Conference on Railway Condition Monitoring 2008.
40. McHutchon MA, Staszewski WJ, Schmid F. Signal processing for remote condition monitoring of railway points. Strain 2005 May;41(2):71-85.
41. Oyebande BO, Renfrew AC. Condition monitoring of railway electric point machines. Iee Proceedings-Electric Power Applications 2002 Nov;149(6):465-73.
42. Oyebande BO, Renfrew AC. The design and implementation of a simulated trackside environment test system for an electric point machine. Proceedings of the Institution of Mechanical Engineers Part F-Journal of Rail and Rapid Transit 2007 Dec;221(4):477-85.
43. Chamroukhi F, Same A, Aknin P, Antoni M. Switch mechanism diagnosis using a pattern recognition approach. IET International Conference on Railway Condition Monitoring 2008.
44. Silmon JA, Roberts C. Improving railway switch system reliability with innovative condition monitoring algorithms. Proceedings of the Institution of Mechanical Engineers Part F-Journal of Rail and Rapid Transit 2010;224(F4):293-302.
45. Silmon J. D3.3.4 Algorithms for detection and diagnosis of faults on S&C. INNOTRACK 2009.
46. Silmon J, Roberts C. A systems-engineered intuitive adaptive failure prediction system. IET International Conference on Railway Condition Monitoring 2008.
47. Silmon J. A systems approach to fault detection and diagnosis for condition-based maintenance. IET International Conference on Railway Condition Monitoring 2006.

48. Roberts C. Methods for fault detection and diagnosis of railway actuators. PhD Thesis, University of Birmingham 2007.
49. Roberts C, Dassanayake HPB, Lehrasab N, Goodman CJ. Distributed quantitative and qualitative fault diagnosis: Railway junction case study. Control Engineering Practice 2002 Apr;10(4):419-29.
50. Jang JSR. Anfis - Adaptive-Network-Based Fuzzy Inference System. Ieee Transactions on Systems Man and Cybernetics 1993 May-Jun;23(3):665-85.
51. Lloyd SP. Least-Squares Quantization in PCM. IEEE Transactions on Information Theory 1982;28(2):129-37.
52. Han J, Kamber M, Pei J. DATA MINING concepts and techniques. USA: Elsevier; 2012.
53. Rousseeuw PJ. Silhouettes - a graphical aid to the interpretation and validation of cluster-analysis. Journal of Computational and Applied Mathematics 1987 Nov;20:53-65.
54. Goodman CJ, Lehrasab N, Dassanayake HPB, Roberts C, Fararooy S. Industrial fault diagnosis: pneumatic train door case study. Proceedings of the Institution of Mechanical Engineers Part F-Journal of Rail and Rapid Transit 2002;216(3):175-83.
55. Mallat SG. A theory for multiresolution signal decomposition - the wavelet representation. IEEE Transactions on Pattern Analysis and Machine Intelligence 1989 Jul;11(7):674-93.
56. Walker JS. A Primer on WAVELETS and Their Scientific applications second edition. US: Chapman & Hall/CRC; 2008.
57. Cortes C, Vapnik V. Support-Vector Networks. Machine Learning 1995 Sep;20(3):273-97.
58. Bishop C. Pattern Recognition and Machine Learning. Jordan M, Kleinberg J, Scholkopf B, editors: Springer Science + Business Media LLC; 2006.
59. Chang C, Lin C. LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology 2011;2(3):27:1-:.
60. Hagan MT, Demuth HB, Beale MH. Neural network design. [United States Boulder, Colo.: : s.n.] ; Campus Pub. Services; 1995.
61. Beale M, Hagan MT, Demuth HB. Neural network toolbox 7 User's guide: Mathworks; 2010.
62. Hagan MT, Menhaj MB. Training Feedforward Networks with the Marquardt Algorithm. Ieee Transactions on Neural Networks 1994 Nov;5(6):989-93.