Machine Learning Approaches to Support Life Extension of Nuclear Power Reactors

Craig Berry

A thesis submitted for the degree of Doctor of Philosophy to the Department of Electronic and Electrical Engineering University of Strathclyde

March 2020

Declaration

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Signed: Craig Berry
Date: 04/03/2020

Acknowledgements

I would like to acknowledge a few people for their support through my PhD, without whom this work would not have been possible.

Firstly, I would like to thank my supervisors Dr Graeme West and Professor Stephen McArthur who have provided me with support and guidance throughout my PhD.

I would like to thank Duncan Hawthorne for providing the funding for my PhD through his scholarship. Without his generosity vital research, such as my own would not be completed.

I would like to thank Anna Rudge and EDF Energy for providing me with a real engineering problem to tackle. I hope that the outputs of my research will be helpful to ensure the safe operation of the UK's nuclear fleet as they approach the end of their operational lives.

Lastly, on a personal note, I would like to thank my family and friends for always being there for me when I need them, allowing me to take my mind of my work when I needed it most. I would especially like to thank my wife Georgina for her love, support and patience.

Abstract

Condition monitoring is the process of observing a parameter, or multiple parameters, extracted from industrial assets which provide an indication of their overall health or presence of faults. In many instances the data produced is numerous, difficult to understand and time consuming to analyse. In the nuclear power industry large volumes of data are captured through inspections and monitoring but only a relatively small amount of the data has associated health information. Machine learning has been identified as a solution to these issues allowing data to be analysed quicker and in greater volumes, alleviating some of the demands on the relevant engineers.

Within this thesis two key contributions are made. The first is the development of a model of fuel channel bore estimation and subsequent deployment to make estimations of operational data. Fuel Grab Load Trace (FGLT) data is operational data which was never intended to be used in condition monitoring. Greater understanding of FGLT data was obtained by targeted analysis of specific regions of the data based on physical understanding of nuclear assets. The model was produced by combining techniques from machine learning supplemented with formalised engineering knowledge to provide decision support in a safety-related environment. The model was found to produce estimations of fuel bore within 1mm root mean square error on real historical FGLT data.

The second is the exploration and implementation of semi-supervised machine learning techniques applied to the detection of cracks within nuclear reactor monitoring data. Semi-supervised machine learning is attractive as it provides comparable results to supervised machine learning techniques without the costly burden of labelling large volumes of data, a commonly occurring challenge in nuclear reactor core condition monitoring. Three different semi-supervised machine learning approaches have been evaluated and for the detection of cracks using FGLT data it has been found that Transductive Support Vector Machines have the best performance.

In this thesis I demonstrate the benefits of machine learning and semi-supervised machine learning to decision support of nuclear reactors through two health monitoring related problems of nuclear reactor cores. Challenges still remain with the application of machine learning in the nuclear condition monitoring domain including

the high cost of obtaining labelled data, lack of historical operate to failure data and incorporating future data in the training of algorithms as it is made available.

TABLE OF CONTENTS

1 Introduction

1.1 Introduction to Research

Commercial nuclear power stations have been in use throughout the world since the 1950s. In the following years there has been a steady increase in new technologies which facilitated nuclear power stations as a commercial enterprise around the world, leading to the development of multiple different designs of nuclear power reactors including Pressurised Water Reactors (PWR), Boiling Water Reactors (BWR), High Power Channel Type Reactors (RMBK), Canada Deuterium Uranium (CANDU) and Advanced Gas-cooled Reactors (AGR), to provide some examples. Nuclear power is now integral to providing a base load generation of electricity to people across the world, in 2018 roughly 10% of the global electricity generation was produced through nuclear power.

Monitoring and inspection are important for the safe operation of all nuclear power stations in the world. Many of the reactors around the world are nearing or have exceeded their initial design lives, which has led to an increase in the use of both inspection and monitoring equipment to ensure safe operation of life-limiting components within reactors (Coble, et al., 2015). Additionally, there has been a change in asset management since the design and construction of older nuclear stations such as extracting additional health information from data that is recorded through analysis and increasing the number of sensors being utilised for condition monitoring.

In the United Kingdom the fleet of Advanced Gas-Cooled Reactors (AGR) are either approaching the end of their initial design lives or have had their operational lifetime extended. The current generation of nuclear reactors in the UK began their construction in 1965, with the first beginning generation in 1976 and the most recent beginning its commercial operation in 1989. The reactors had initial design lives of 35 years, depending on the individual station. Since 2005 the UK's AGR stations have been periodically inspected and have received lifetime extensions based on current understanding of their operational health, allowing for their continued operation beyond their initial closure dates. There are two reactors operating at each of the seven

AGR power stations in the UK and at the time of writing, the closure dates for these stations range from 2023 to 2030.

When the reactors initially start generating electricity, the standard time between inspections of the graphite cores of AGRs was every 3 years. However, some of the older reactors in this generation are now being inspected yearly to gather more information about their current operational health. The increased number of inspections coincides with the recent discovery of a new type of crack, known as a keyway root crack, being found in some of the older reactors which had only previously been predicted (Robinson & Maul, 2011). These cracks have been expected to occur and are an indication of the reactor approaching the later stages of its operational life and therefore these new cracks must be considered when defining safety cases for the operation of these reactors (McLachlan, et al., 1995).

Inspecting the core is a time-consuming task which requires the reactor to be offline. This results in economic losses due to extended periods where electricity is not being generated. While an individual reactor is taken offline it is estimated that there is a loss in revenue on average of roughly £0.5 million each day. Therefore, it is desirable to minimise the duration of statutory outages or safely reduce the frequency requirements of these inspections. Monitoring data, known as Fuel Grab Load Trace (FGLT) data, is obtained from onload and offload refuelling of the reactors has been demonstrated in recent years to be suitable for providing additional information about the health of the core (West, et al., 2006), despite previously only being used as a measure of safety. Manual analysis of this data is a lengthy process which requires specialised domain expertise. Condition monitoring through machine learning has been identified as a method to provide decision support tools to the manual analysis of monitoring data.

Condition monitoring is the process of observing a parameter, or multiple parameters, extracted from industrial assets which provide an indication to its overall health or the presence of faults. In many applications the information produced is used to inform the operation of assets, as well as the planning of maintenance regimes. The most basic condition monitoring involves using sensors on structures, systems and components to measure characteristics, such as vibration or temperature, and checking that they are within a predefined range. If the values are outside of the defined range, it is a possible

indication of the presence of a fault. Applications of condition monitoring have been demonstrated in a vast range of fields including aerospace (Wechsler, et al., 2012) (Mazzoleni, et al., 2019), rail (Hodge, et al., 2015), manufacturing (Cholette, et al., 2013) (Shi, et al., 2018), rotating electrical machines (Tavner, 2008) (Vishwakarma, et al., 2017), power distribution (Livshitz, et al., 2004) (Reddy, et al., 2017), wind turbines (Qiao & Lu, 2015) and in robot-assisted surgery (Dai, et al., 2015). The primary focus of this work is the application and subsequent analysis of data generated within condition monitoring of nuclear power stations through case studies of AGR stations. Machine learning has been identified (Ma & Jiang, 2011) as an approach, which when utilised can provide decision support to engineers by automating the process of data analysis.

Machine learning allows automated learning of a desired subject without being explicitly programmed by a human. An example of this, is the exploration of the feature space to provide a means for automatically discovering and assessing patterns contained within the data. Machine learning has seen an increase in use recently due to the volume and frequency of collected data becoming larger as well as the improvement in computing performance (Wu, et al., 2014) (O'Leary, 2013) (Loyd's Register Foundation, 2014) (Dean, et al., 2018). Machine learning can perform analysis upon large volumes of data faster and more consistently than humans but only if it has been implemented and trained correctly.

In the nuclear power condition monitoring domain, obtaining associated health information from monitoring data can be difficult. There are two major challenges, the first is the limited number of domain experts to analyse the data. This is primarily due to the aging and mature nature of the nuclear power industry. In the UK no new nuclear power station has been built since 1988 resulting in a gap in skill within the industry. Additionally, many of the people with the domain expertise and involvement in the design of the nuclear reactors have retired.

The second major challenge is the lack of available training data. The worldwide trend has been an increase in the amount of data that is obtained but in many instances this data has no associated information about them. In the context of AGRs, there are limited examples of labelled data to train machine learning algorithms. Ground truth data can only be obtained when the reactor is inspected at shutdown and on only ~10%

of the core per inspection. A challenging problem is to accurately train machine learning algorithms which generalise well on future data when there is insufficient training data, it is also difficult to evaluate their performance when there are limited instances of ground truth training data. One method to address this problem within the nuclear domain has been to use unsupervised machine learning techniques (Baraldi, et al., 2013). Unsupervised machine learning does not require labelled ground truth training data and instead explores unknown relationships in the data such as sorting instances into similar clusters or detecting anomalous data. This will not be able to assign a classification to new data. A different identified solution to situations with a lack of labelled data is to use semi-supervised learning (Gouriveau, et al., 2013). However, as of writing this there has only been a limited number of applications within nuclear condition monitoring (Ma & Jiang, 2015) (Moshkbar-Bakhshayesh & Ghofrani, 2016).

The key focus of the work in this thesis is to utilise data analytical techniques and machine learning to derive additional important information about the health of operating reactors. Challenges overcome in this domain include lack of labelled data, need for transparency of decisions and time pressures which leads to the need to explore semi-supervised techniques. They are demonstrated on the monitoring and inspection health data that is obtained from the graphite core, however the processes and approaches demonstrated in this thesis can be more broadly applied to other nuclear condition monitoring problems. The graphite core is one of the life limiting factors of the reactors since it cannot be repaired or replaced. It is therefore required to periodically assess the health of the core of the reactor through inspections. These inspections target the graphite bricks that make up the structure of the core and look for cracks or other signs of ageing that may affect the reactors ability to operate safely. There are two challenges that are addressed in this research. Firstly, the ability to provide an estimate of the dimensional properties of the graphite bricks within the core using monitoring data, which was never originally designed to provide any health information about the core. This is achieved through a physical understanding the underlying of the contributing components of the data. The second challenge addressed determines the presence of cracked bricks within the core and does so by using both supervised and semi-supervised machine learning techniques. The results are

compared to demonstrate the applicability of using semi-supervised machine learning on nuclear condition monitoring data. Both tasks are currently undertaken at inspections with dedicated measurement and visual equipment. Therefore, the aim is to allow the data collected at monitoring events to produce the same high-quality measurement information that is obtained at inspections but more frequently.

1.2 Novel Contribution

This thesis makes the following novel contributions to the application of machine learning techniques for intelligent decision support in nuclear applications:

The development of an approach for extracting reactor health information from data not originally designed for condition monitoring purposes. This approach augments existing reactor inspection regimes by providing a supporting health indicator. This is demonstrated through an application of fuel channel bore estimation.

- An in-depth analysis of the fuel grab load trace by using knowledge of the physical properties of the reactor core and a supply of historical data
- Development of a framework that removes unwanted components present in the fuel grab load trace, on a trace by trace basis
- Improving the understanding of the different contributions of the fuel grab load trace by using historical data to explore its relationship to channel bore measurements and allows it to be generalised
- Demonstration and evaluation of the performance of the algorithm on real nuclear power station fuel grab load trace data
- Recommendation for the application of the algorithm on operational data for monitoring and inspection regimes

This is demonstrated in chapter 4, where the algorithm is subjected to historical data which has been recorded from refuelling events at AGR stations. This data is analysed using knowledge of the physical design of the reactors and the mechanisms present during refuelling events to deconstruct the Fuel Grab Load Trace (FGLT) signals and extract the desired individual components. Through doing this a frictional component can be extracted which is related to the dimensions from the associated fuel channel.

Using inspection data as ground truth information multiple regression machine learning algorithms were trained and compared. The best implementation of the algorithm is then suggested for application on real world refuelling events.

An evaluation and demonstration of semi-supervised learning approaches with respect to nuclear condition monitoring data, indicating improvements over traditional supervised machine learning classification techniques.

- A critical review of three semi-supervised techniques exploring their suitability to nuclear condition monitoring data where obtaining labelled operational data of faults is a costly, time-consuming process
- The first demonstration of semi-supervised machine learning techniques being applied to the UK's fleet of nuclear power stations for condition monitoring purposes
- Recommendation of the semi-supervised machine learning algorithm for the application of crack detection and its implementation

This is demonstrated in chapter 5; the majority of historical FGLT data has been recorded and stored but never provided with an associated health information label. This has resulted in a large amount of unlabelled data compared to the limited amount of labelled health data available. Manually analysing the unlabelled data is a difficult and time-consuming process which only a small number of people would be able to extract useful health information from. In this scenario the health information is whether individual graphite bricks are cracked. The semi-supervised machine learning techniques that are being investigated are co-training, label propagation and transductive support vector machines. They are compared based on their performance of correctly classifying bricks; due to the imbalanced nature of the problem additional metrics other than the accuracy of the algorithms are assessed. These algorithms are then compared to the performance of their typical supervised counterparts. Finally, a recommendation of the technique to use on real world data is provided based on the evaluation of results.

1.3 Thesis Outline

This thesis continues in chapter 2 by providing an overview of nuclear power generation. An objective of this chapter is detailing AGRs design and the methods of inspecting and monitoring them, with a primary focus on the graphite reactor core. Additionally, a subsection of the chapter delves into the monitoring data that is obtained at refuelling events of AGRs. This data is known as the FGLT and throughout this chapter the various physical components that contribute to its value are discussed and highlighted on an annotated diagram of a full trace.

Chapter 3 discusses the approaches and methodologies of machine learning, as well as the applications within both the nuclear and wider condition monitoring domain. It focuses particularly on identifying what supervised, unsupervised and semi-supervised machine learning are and the situations where they are used.

Chapter 4 details the novel contribution that has been made to the channel bore prediction model. It details the process by which the four components of the FGLT have been analysed and generalised. Through analysing years of historical data, it is shown how the extraction of the component, which is directly related to the estimation of fuel channel bore, is achieved. The chapter then goes on to describe the machine learning training of a linear regression model used to estimate fuel channel bore from FGLT data.

Chapter 5 details the novel contributions that have been made to the evaluation of different semi-supervised machine learning algorithms for crack detection of graphite bricks in AGRs. This chapter includes the process of identifying and extracting important features from the data, incorporating a method of ensuring that there is class balance and testing the semi-supervised algorithms in an accurate way. The three semi-supervised learning methods used for crack detection are the co-training algorithm, the label propagation algorithm and semi-supervised support vector machines.

Chapter 6 summarises the conclusions contained in this body of work and discusses the further work that can be undertaken to progress this research further. A notable focus is the ambition to create a machine learning framework that combines the research already undertaken on semi-supervised machine learning with incremental machine learning.

1.4 Associated Publications

Journal Papers

Young, A, Berry, C, West, G.M., McArthur, S.D.J. (2019) 'A generalized model for fuel channel bore estimation in AGR cores', *Nuclear Engineering and Design*, 352(1)

Conference Papers

Berry, C, Pattison, D, West, GM, McArthur, SDJ & Rudge, A 2017, Estimating fuel channel bore from fuel grab load trace data. in *The 5th EDF Energy Nuclear Graphite Symposium.* Warrington, 5th EDF Energy Generation Ltd Nuclear Graphite Conference, Southampton, United Kingdom, 10/05/16.

Berry, C, West, G, McArthur, S & Rudge, A 2017, 'Semi-supervised learning approach for crack detection and identification in advanced gas-cooled reactor graphite bricks' Paper presented at 10th International Topical Meeting on Nuclear Plant Instrumentation, Control, and Human-Machine Interface Technologies, NPIC and HMIT 2017, San Francisco, United States, 11/06/17 - 15/06/17

2 Nuclear Power Background

2.1 Introduction

Nuclear power is important in the United Kingdom and around the world as a reliable method to generate electricity. As of 2020, there are currently 442 nuclear power reactors in operation worldwide that have an installed net generating capacity of 389,791 MWe. In addition to this, there are roughly another 50 nuclear power reactors that are under construction worldwide. In the UK, nuclear power contributes to roughly 20% of the energy generation portfolio, providing stable and reliable base-load generation capabilities. A major challenge facing the nuclear industry, both in the UK and abroad, is the age of the stations. Many are in the latter stages of their operating lives and there is the desire to continue and extend their operation, provided it is safe to do so. The key to making the case of lifetime extension is to understand both current and future health of the major reactor components, which is done by interpreting data gathered both during routine operations and through inspection during planned statutory outages. A brief introduction into the principles of nuclear power generation will be provided, before focussing on the design and operation of the UK's AGR stations and the associated data they generate, which are used as a case study for this thesis.

2.2 Nuclear Power Generation

There are five key components to generating electricity from nuclear power stations, these are:

- A fuel that is used to generate heat
- A moderator that allows for a sustainable fission reaction
- Control rods which control the rate of the reaction within the core
- A coolant which transfers the heat to boil water and generate steam
- A steam turbine generator which transfers the energy from the steam to electricity

Nuclear fission occurs when a free neutron bombards a fissionable atom, typically a heavy element such as uranium, and becomes absorbed into the nucleus. The nucleus splits and in the case of uranium it splits into barium and krypton. In addition to the energy this process releases, it is possible that free neutrons are released in this process. A chain reaction process can result if there are further fissionable nuclei to absorb the free neutrons. Controlling the chain reaction process is of great importance to ensure that the reaction is sustainable. The chain reaction can be controlled by changing the fuel mix and raising and lowering control rods. Choosing a suitable moderator is the main method for ensuring the reaction is sustainable.

The primary function of a moderator is to slow down the speed of the fast neutrons produced in nuclear fission to change them into thermal neutrons, allowing the chain reaction to occur. Common types of moderators used in commercial nuclear are water (both light and heavy) and graphite.

The purpose of control rods is to manage the rate of reaction by absorbing free neutrons within the core. They are inserted into the core by varying amounts as required to limit the reaction and control the amount of electricity that is generated. Additionally, they can be fully inserted under certain circumstances to completely shut down the reactors. The most common occurrence of this is when the reactor is being taken offline for inspection, but the control rods are used in the event of an emergency shutdown of the reactor.

Once the reactor is self-sustaining the generated heat energy is required to be changed into electrical energy. Under most circumstances this is achieved by using a substance, known as a coolant, to transfer heat away from the core to the water in the steam generator. Lastly the coolant is used to boil water to convert it into steam. The steam is then used to power turbines which produce electricity from the generator. Some reactors differ when it comes to the process, such as Boiling Water Reactors (BWR), which use water as both a coolant and to produce steam to generate electricity.

In the United Kingdom the main type of reactor design that is present is the AGR and as such further information will be provided about its design.

2.3 Advanced Gas-Cooled Reactor

There are 15 nuclear reactors in the United Kingdom, 14 of these reactors are AGRs and the remaining type is a Pressurised Water Reactor (PWR). For the purpose of this thesis further detail is placed on AGRs. In the UK the AGRs are located at six different sites: Hunterston, Hinkley Point, Dungeness, Hartlepool, Torness and Heysham (with two stations).

Dungeness B was the first station to begin construction in 1965 with an initial stated completion date of 1970. The design of the first generation AGR was based on a small-scale experimental prototype which demonstrated the design of a gas-cooled nuclear reactor. However, throughout construction the scaling up of the prototype resulted in some problems resulting in a delayed completion date of 1983, 13 years after its initial date. In 2005 the operational lifetime of Dungeness B was extended by 10 years from its initial design life of 2008 (British Energy, 2005). This life extension was repeated in 2015 with upgrades and Dungeness B is scheduled to be closed in 2028 (EDF Energy, 2015).

Hinkley Point B and Hunterston B were the same design as the Dungeness B reactors and began construction in 1967. The stations were completed and generating electricity in 1976. Since their construction they have had their operational lifetimes extended twice. The first, in 2007, extended the initial 35-year design life of the stations by an additional five years from 2011 to 2016 (British Energy, 2007). The second occurred in 2013 when the stations' operational lives were extended by 7 years with decommissioning of the station planned in 2023 (EDF Energy, 2012). These stations however have been operating at lower power than they were stated. In 2006 they reduced power generation to roughly 70% of nominal operation (British Energy, 2007); this was then increased to 80% in 2013 following modifications to the boilers.

Hartlepool and Heysham I were then designed and built based on a slightly modified version of the first generation but with a multi-cavity pressure vessel. Hartlepool began construction in 1968 and was followed by Heysham I in 1970 but both were connected to the grid to begin generating in 1983. Both stations had an initial closure date of 2009. Similarly to the other stations which have been mentioned, both Hartlepool and

Heysham I have had two individual 5 year life extensions resulting in the current closure dates of these stations being 2024 (EDF Energy, 2012) (EDF Energy, 2016).

Lastly, Torness and Heysham II were designed and constructed last as a second generation AGR design that incorporated further safety features, primarily the use of more complex computerised control and instrumentation systems. These began construction in 1980 and were connected to the grid to begin generating in 1988. As these are the newest stations it is no surprise that they are expected to be the last of the current fleet to be in operation. At the time of writing they are expected to be closed in 2030 having already been given an additional 7 years operational life extension (EDF Energy, 2016).

2.3.1 Design

At each AGR station in the UK there are two individual reactors. In an AGR the core, boilers and gas circulators are all contained within a single pressure vessel. The core of an AGR is composed of thousands of graphite bricks, interlocked over multiple layers. These bricks are not only important to form the structure of the core but also act as a moderator for the fission reaction. There are multiple different types of graphite bricks in the core including fuel channel, control channel, gas cooling channel and structural bricks which are all interconnected by graphite keys. An example of the graphite bricks within the core can be seen in Figure 2.1 which clearly shows the interlocking nature of the core. A typical AGR in the UK will have over 300 graphite fuel channels with 10 or 11 brick layers within the core.

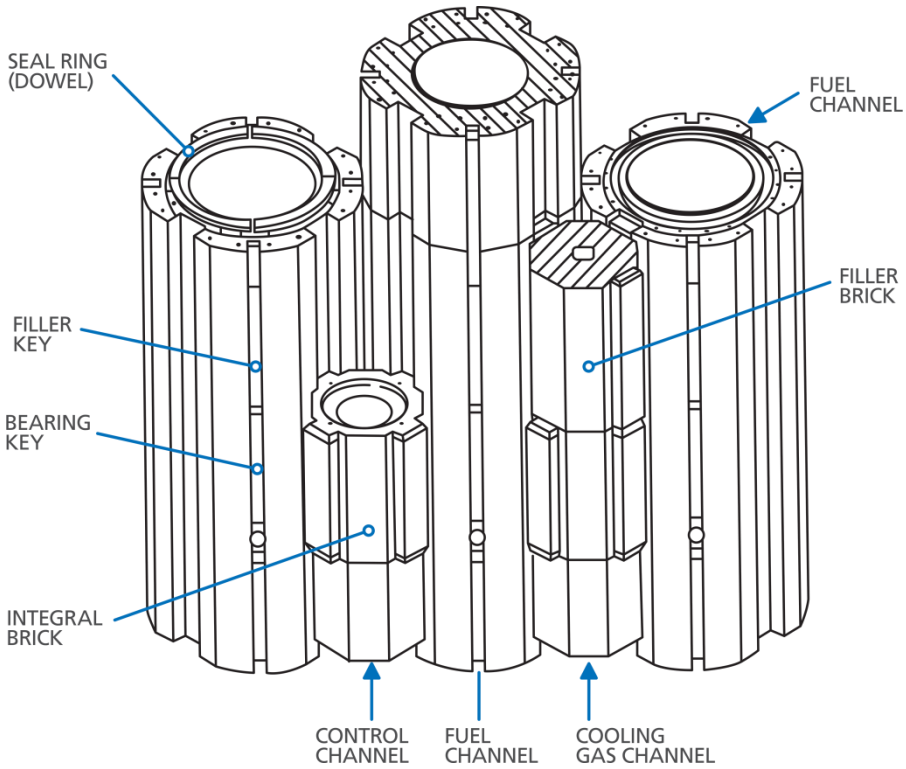


Figure 2.1 - Diagram of the keying structure of the graphite bricks within the core of an AGR (EDF Energy, 2018).

Enriched uranium oxide (UO_2) is used as the fuel and is present as small hollow cylindrical which are located inside the fuel elements. In each fuel stringer there are eight fuel elements which are linked together by a tie bar, permitting for some flexibility of the fuel stringer and allowing it to conform to slight changes in the fuel channel shape. A diagram of a fuel assembly can be seen in Figure 2.2, containing a fuel stringer and a plug unit. The entire fuel assembly is roughly 26m in length,

At the bottom of the fuel stringer, below the eight fuel elements, there is the bottom support and reflector, at the top of the fuel stringer there is the top reflector, also known as the central inertial collector. An important aspect of the design of the reflectors is the steel stabilising brushes which are in contact with the fuel channel walls. Attached to the bottom support and reflector is a set of three steel brushes which is known as the lower stabilising brush (LSB) and attached to the top reflector is a single stabilising brush known as the upper stabilising brush (USB).

The plug unit contains a closure unit which allows for the pressure vessel boundary to be sealed. In addition, there is also a heat shield to reduce the flow of heat to the concrete and stand pipe, a biological shield is present to reduce radiation on the pile cap which is located on top of the pressure vessel. The purpose of the gag actuator located above the shields is to adjust the gas flow through the fuel channel to match the fuel's heat production. Thermocouples are present to measure the temperature of the outlet gas which is used for control, indication and fault detection purposes. Inside the core the number of fuel stringers varies at the different stations but there are approximately 300 individual fuel stringers.

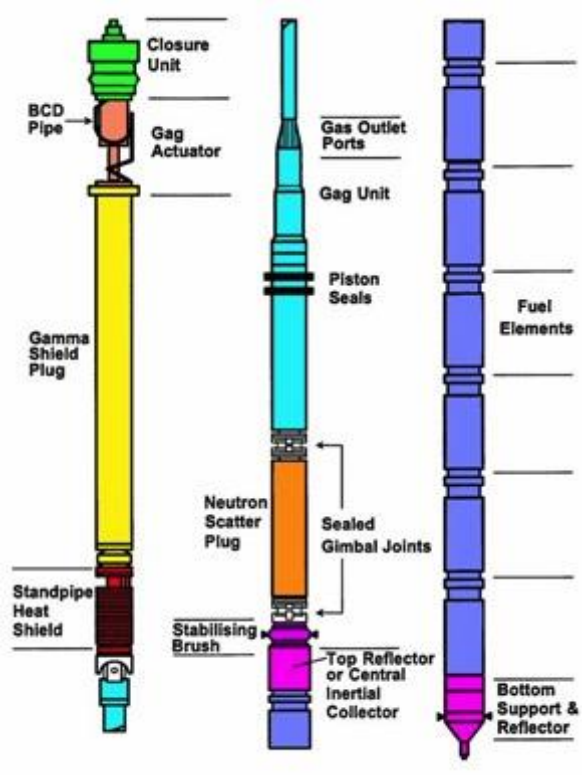


Figure 2.2 - Diagram of a fuel assembly identifying its key points (Steer, 2005).

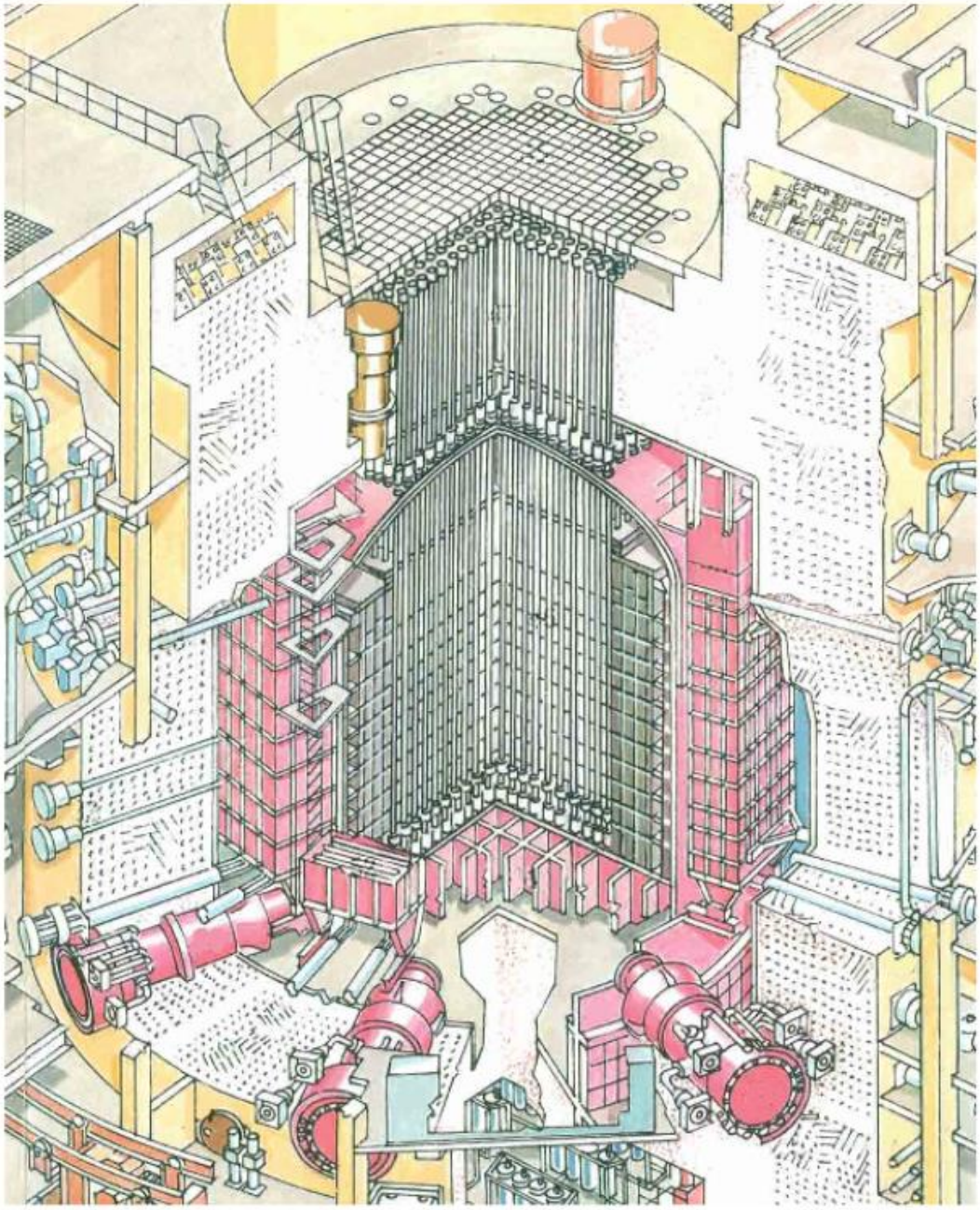


Figure 2.3 – A cutaway drawing of an AGR core enclosed within its pressure vessel.

The coolant used within AGRs is carbon dioxide (CO_2) and it is inserted into the bottom of the core to circulate throughout. Methane gas in present in the gas mixture of coolant, this is as a measure to reduce the amount of degradation that occurs within the graphite bricks of the core. The process of graphite brick degradation is discussed in a later subsection of this chapter.

There are roughly 80 boron control rods inside AGR reactors which can be used to control the rate of the reaction and are used to shut down a reactor completely. Additionally, in the event of a reactor trip they are inserted fully in to the core by gravity via an electromagnetic clutch being de-energised. The AGR design also has a backup if control rods are unable to shut down the reaction. In this scenario nitrogen gas can be injected into the core which absorbs free neutrons, stopping the reaction.

Steer (Steer, 2005) provides a more detailed overview of the design of the AGR and a detailed and comprehensive overview of all aspects of the operation and maintenance of the AGR can be found in *Modern Power Station Practice – Nuclear Power Generation* (British Electricity International, 1992).

2.3.2 Ageing Mechanisms

Industrial assets degrade over time, in the case of the graphite moderator, the prolonged exposure to the harsh environment leads to changes in the physical structure of the graphite bricks. The stress caused by the graphite weight loss results in internal strain on the bricks which can cause their shape to distort and in some cases the bricks can crack due to this process.

There are two key mechanisms that contribute to the degradation and aging of the graphite bricks inside the reactor core: radiolytic oxidation and irradiation. Radiolytic oxidation leads to significant graphite weight loss through the following process. Carbon dioxide coolant gas breaks down into carbon monoxide oxidising ions under the effects of irradiation as shown in Equation 1.

$$CO_2 \rightarrow CO + O \qquad \textbf{Equation 1}$$

Under most circumstances the oxidising ions will recombine with the carbon monoxide to produce carbon dioxide as shown in Equation 2.

$$O + CO \rightarrow CO_2 \qquad \textbf{Equation 2}$$

In addition to this, the oxidising ions can combine with the carbon atoms that are present in the graphite structure. When this occurs carbon atoms are stripped from the

graphite, resulting in an overall weight loss that occurs as readily on the surface of the graphite, as it does in the internal pores, as shown in Equation 3.

$$O + C \rightarrow C(O)$$ **Equation 3**

To slow down this process of radiolytic oxidation, carbon monoxide or methane gas can be added to help soak up the oxidising ions. (Brocklehurst & Kelly, 1993) (Neighbour, 2000)

The second process of degradation is the irradiation of the graphite material. The graphite bricks are polygranular graphite, which is defined as containing the following as described by Neighbour (Neighbour, 2000):

- *Two or more carbonaceous species originating as filler binder or impregnant*
- *A wide range of crystalline perfection and crystallite sizes in different parts of the microstructure*
- *Complex networks of pores of different types that originate at different stages in the manufacture of the graphite and the filler*

The graphite in the reactor becomes irradiated by both neutrons and ionising gamma radiation. The graphite has a lattice structure of atoms and when they undergo irradiation they expand into the porous regions of the graphite structure, resulting in an increase in their dimensions. Figure 2.4 displays the different profiles for the aging of the graphite reactor bricks, including undelayed pore generation, unlimited pore closure and delayed steady state pore generation.

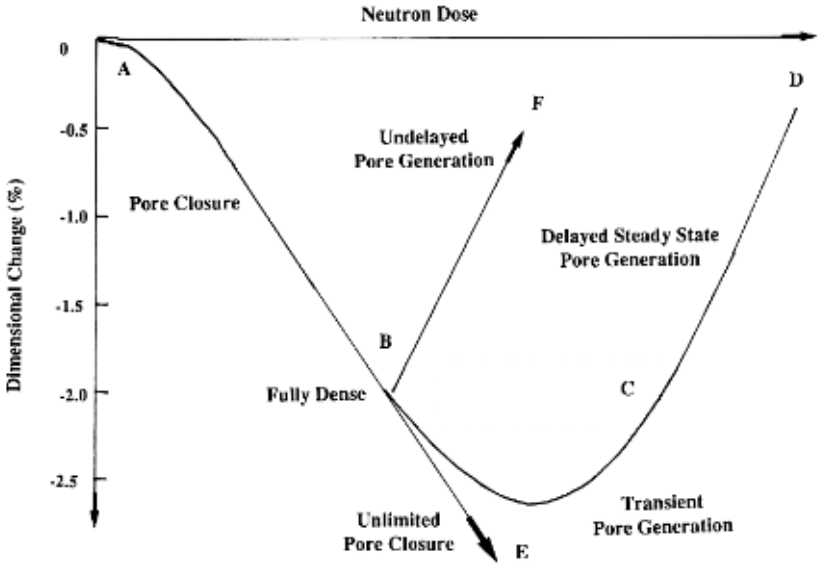


Figure 2.4 – The predicted aging profile of nuclear irradiated graphite bricks (Neighbour, 2000).

At the beginning of the life of a reactor the dimensions of the graphite bricks shrink due to the weight loss caused by radiolytic oxidation and corrosion. This shrinkage occurs for a large portion of the life of the graphite bricks until turnaround is reached. Turnaround occurs when the expansions of the graphite lattices have expanded into all available pores and subsequent expansion increases the dimensions of the graphite. Originally this process was theorised to occur beyond the design lives of AGRs in the UK. However due to the life extension, it is believed that there is a small number of graphite bricks that are approaching or have surpassed the turnaround point. The primary reason for graphite bricks within the core cracking is due to dimensional change within the bricks.

2.3.2.1 Graphite Brick Cracking

In the context of AGR graphite core brick health, there are two main categories of cracks; they are described as axial and circumferential cracks. Circumferential cracks in cylindrical graphite bricks occur horizontally around the circumference of the fuel channel. The presence of this type of crack results in the narrowing of the fuel channel

as the edges of the crack protrude into the fuel channel. These cracks have typically occurred at the beginning of the reactor's life.

Bore initiated axially cracked bricks occur vertically and propagate from the inside of the fuel channel to the outside surface of the brick. These cracks can vary from a small crack in a section of the brick to one which spans the entire vertical length. Axial cracks typically result in the fuel channels being widened, as the cracked bricks open. These types of cracks are caused by the forces acting on the graphite bricks during their operational life. Bore initiated axial cracks are caused by the stresses and strain applied to graphite bricks in the lattice structure of the core. The main cause of these cracks is due to the weight and dimensional changes over the lifetime of the graphite brick.

The last type of crack that is currently known about but is very rarely present in the graphite bricks are keyway root cracks. These cracks are very similar to bore initiated axially cracked bricks where they manifest as vertical cracks. They differ from bore initiated axial cracks by the way they are formed. They are typically formed after the graphite bricks have reached the turnaround point from radiolytic oxidation, where the graphite bricks stop shrinking and instead begin to swell. This alteration in structure results in a change in stresses and strain on the brick. This can cause keyway root cracks to form from the outsides of the bricks, in the keyway, and propagate inwards to the fuel channel.

An example of a cracked brick can be seen in Figure 2.5. Figure 2.6 shows the differences between bore initiated axial cracks and keyway initiated axial cracks. It shows how in the early life of the reactors the tensile forces affect the inner bore and at the later stages in their life the tensile forces affect the outer area of the bricks at the keyways.

Figure 2.5 - An example of a graphite brick which has an axial crack.

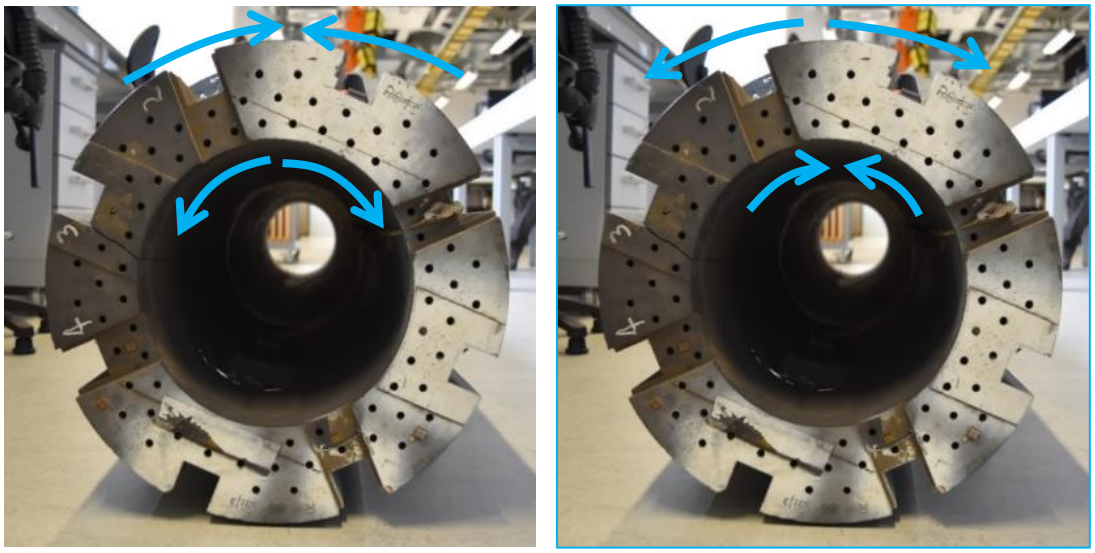


Figure 2.6 – Comparison of the different stresses acting on the graphite brick. At the start of life the brick experiences compressive forces on its outside and tensile forces on its inside. Later in life the brick experiences tensile forces on its outside and compressive forces on its inside.

2.3.2.2 Implication of Aging

There are two main implications of the aging process of the graphite bricks within the core of the reactor. The first is that the aging process causes dimensional change and cracking. A hypothetical situation where structural change of the graphite brick lattice

resulted in the control rod or fuel channels being blocked or misaligned would mean that reactor would not be able to sufficiently control of the rate of reaction.

There are safety margins built in to the design of the reactor and conservative inspection regimes that ensure that the reactor can operate safely (Office for Nuclear Regulation, 2017). The ageing process has been studied (Brocklehurst & Kelly, 1993) and the effect of cracks physically modelled (Kralj, et al., 2005) but operational data can provide insight into the current state of aging within the core.

The second implication of graphite brick aging is their reduced ability to function as a moderator. This occurs due to the weight loss associated with the oxidisation of the graphite. Losing the effectiveness of graphite as a moderator is potentially a life limiting factor of the reactor. However, it is much more likely that the loss of structural integrity through cracking and weight loss would limit the life of the reactor before this occurs.

Monitoring and inspections of the graphite core are currently used in the normal operation of AGRs to obtain fuel channel health information. Monitoring events occur much more frequently but only provide complicated data that only infer the health of the core after analysis by domain experts. Inspections provide ground truth data of the current dimensions of fuel channels but are obtained much more infrequently during schedule outages.

2.3.3 Monitoring – Fuel Grab Load Trace

The FGLT is a measurement of the mechanical load that is exerted on the refuelling machine as the fuel stringer is removed from and inserted into the core during refuelling events. This recording is split into two distinct sections, the discharge trace is recorded when an old fuel stringer is removed from the core and the charge trace is recorded when a new fuel stringer is inserted into the core. Originally the load value was recorded as a safety measure to ensure that the fuel stringer safely travels through the fuel channel. An example of the use of this data could be in determining if the fuel stringer has become snagged during the refuel process, this will be indicated in a larger than expected load value present. A situation like this would then prompt graphite core operators to respond to the problem. In the last 10-15 years FGLT data has been seen to have other uses by being able to relate the load value to the friction experienced when the fuel stringer's stabilisation brushes as it is being inserted or removed from

the core. Originally, the FGLT was recorded and stored on paper however a project was undertaken to create a database which is used to store digital copies of the traces (West, et al., 2008). Additionally, there has been a process undertaken to digitise the older paper traces and many of them have been stored in the database as well.

An example of FGLT traces of the core region can be seen in Figure 2.7 with an example of a single brick layer being identified. Brick interfaces can be identified by peaks in the discharge trace and by troughs in the charge trace.

Currently graphite core engineers manually analyse FGLTs after refuelling events to help provide an insight into the health of the graphite bricks within the core. They look for differences in the traces, such as abnormally high or low load values, and note their location in the trace. This knowledge is used to provide information and influence the operation of the station such as targeting specific fuel channels for inspection.

Refuelling is scheduled to occur in AGRs roughly every 6-8 weeks and can occur either when the reactor is at a statutory outage or when the reactor is at low power. The fuel stringers typically reside within the core for 5-7 years before being refuelled (Steer, 2005). However, fuel stringers can occasionally be removed prematurely and shuffled into other fuel channels as required to aid in the operation of the reactor. Therefore, over the course of 8 years there should be relevant FGLT information about every fuel channel in the core. It is important to note that without analysis this obtained data does not currently provide health information.

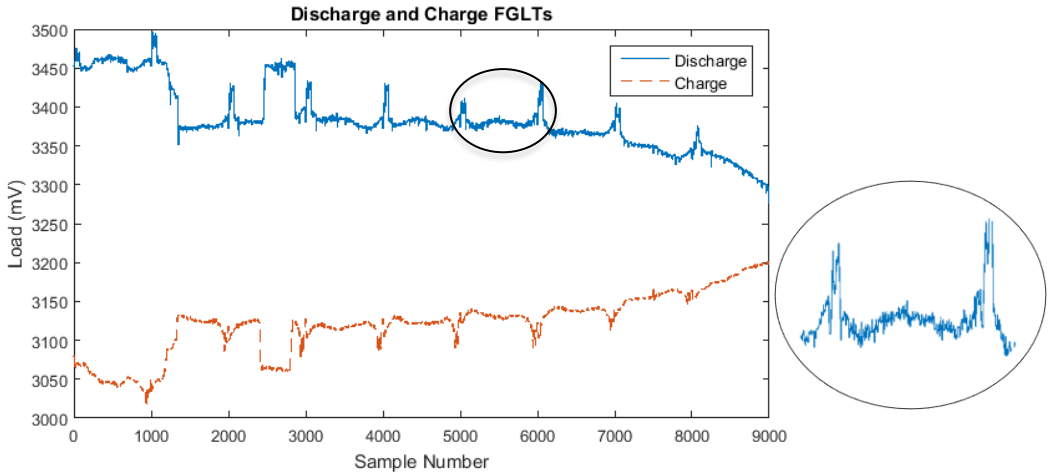


Figure 2.7 Example of a discharge (blue) and charge (red) trace obtained from the same refuelling event, with a magnified example of a single brick layer.

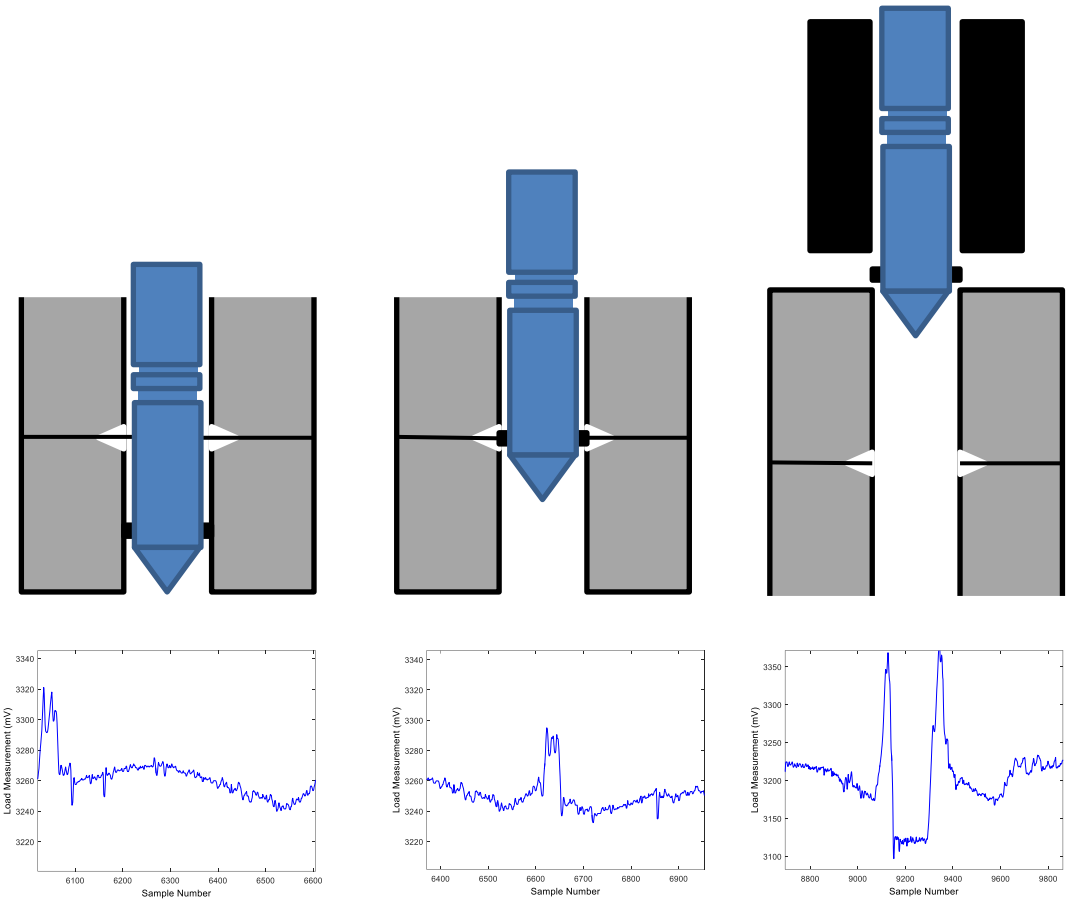


Figure 2.8 – Example of three different locations within the refuelling process and the FGLT showing: middle of the brick, interface between two bricks and leaving graphite core region

2.3.3.1 Walkthrough of a FGLT

In this section a description of the underlying physical components of the load values produced at refuelling events is provided. Each refuelling event starts by producing a discharge trace as the fuel stringer is removed and ends with a charge trace as a fuel stringer is inserted into the core, as such this order will be retained in the walkthrough as seen in Figure 2.9.

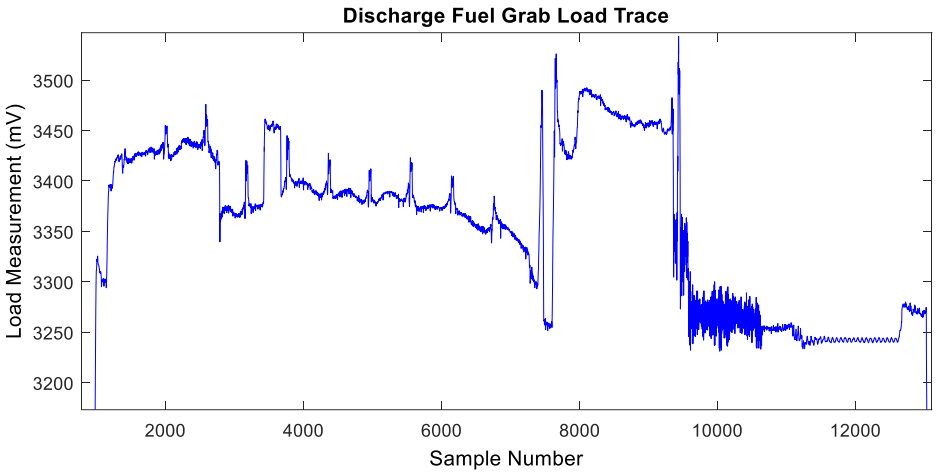


Figure 2.9 – An example of a discharge fuel grab load trace.

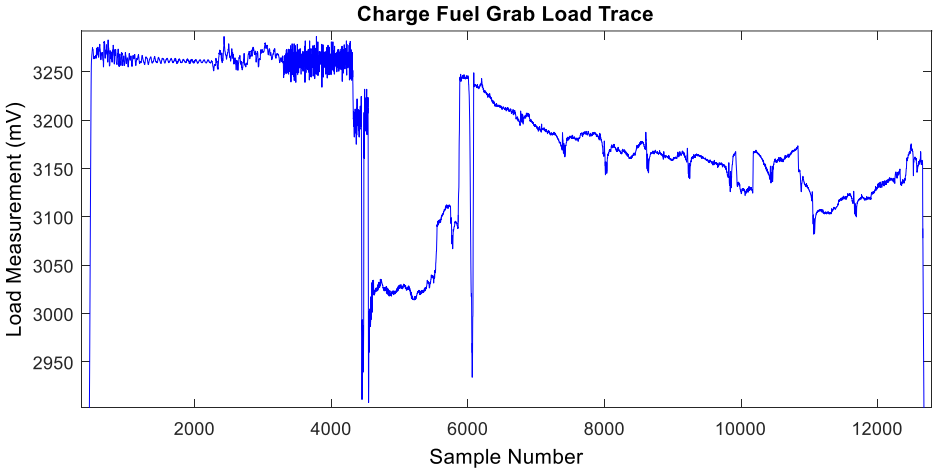


Figure 2.10 - An example of a charge fuel grab load trace

Initially the fuel stringer is located within the fuel channel and all locations are described with respect to the LSB. As it begins to be removed there is a transient region before the load value stabilises, in this region the LSB is passing through the graphite brick layers 1-4 while the USB is in the guide tube region. There is a drop off in load at the start of brick layer 4 which is due to the USB losing contact with the guide tube

as it becomes wider. Then as the LSB passes through brick layer 5 the gas ports of the fuel stringer join with the refuelling machine resulting in a change in the aerodynamic effects present in the trace. Additionally, at this time there is a further contribution from the USB when it passes through the narrower piston seal bore. After this from brick layer 6-11 there are no further USB contributions, except for some cases in brick layer 9 where slight contact can be made by the USB as it hunts for the entrance to the standpipe. Once the LSB exits the core but before it passes into the guide tube there is a region where it becomes wider leading to load values of the fuel stringer in suspension. After this region the LSB enters the narrower region of the steel guide tube and friction is once again present. The LSB then begins repeating the initial steps undertaken by USB at this point, but now when the guide tube widens, the brushes from the LSB are still in contact with the wall of the guide tube. Similarly to when the USB passes the piston seal bore there is an increase in friction when the LSB passes it. After this however the LSB enters the hotbox, a region where there should be no interaction with the brushes. Finally, the LSB passes into the stand pipe which should not cause any friction as it has a larger diameter than the brushes however some contact resulting in friction can still be present in this region.

2.3.3.1.1 FGLT Sub-Components

It is understood that there are four main components that contribute to the load value in the FGLT (West, et al., 2014). This can be seen in Equation 4.

$$L(h) = K_L\Phi(h) + K_U\Phi(h+d) + F_G(h) + m \qquad \textbf{Equation 4}$$

Where L is the measured load at height h, K_L is a parameter which relates the friction caused by the lower stabilising brushes due to the channel bore (Φ) at height h to load. K_U is a parameter which relates the friction caused by the upper stabilising brush due to the channel bore (Φ) at height h offset by d, the difference in height values between the LSB and USB. F_G is the summation of any aerodynamic effects to the load value at height h and m is the mass of the fuel stringer, also known as the deadweight, which is constant.

Each of the four contributions is discussed in the following sub chapters.

2.3.3.1.2 Deadweight

The deadweight of the fuel stringer is defined as the contribution that the mass of the fuel stringer exerts on the load mechanism. This contribution should be a constant value throughout the entire trace as the fuel stringer is not expected to change mass during the charge or discharge process. It is also believed that the weight contribution of the fuel pellets in the fuel stringer is negligible and therefore the change in their weight over the duration that they are in the core does not make a considerable difference to the deadweight contribution.

In chapter 4 the deadweight is explored further using real monitoring data from refuelling events at an AGR station over a time period of 10-15 years.

2.3.3.1.3 Aerodynamic Effects

As stated in section 2.3.1, in AGRs the coolant which transfers heat away from the core is carbon dioxide (CO_2). It is believed that there are three different effect profiles caused by coolant gas that depend on the mode of operation of the reactor during refuelling (West, et al., 2014). These are low power refuelling, offload pressurised refuelling and offload depressurised refuelling. The CO_2 supports the weight of the fuel stringer during both the discharge and charge processes of refuelling so will always result in a lower load value. Low power refuelling has the largest reduction in load compared to the offload refuelling but all three follow the same general shape of aerodynamic contribution. Initially there is believed to be a constant contribution when the fuel stringer is within the core, following this there is a step change in the load with an overshoot that occurs as coolant gas flows into the refuelling machine resulting in a second steady state value. With reference to the location of this within the core the first steady state contribution occurs from the beginning of the discharge trace until the lower stabilising brush reaches brick layer 5. The step change with overshoot occurs within brick layer 5 and finally the second steady state region exists from brick layer 6 until the fuel stringer is outside the fuel channel.

In chapter 4 the aerodynamic effects of many refuelling events are explored to obtain a better understanding of their contributing. The different modes of operation of reactors are also investigated.

2.3.3.1.4 Upper Stabilising Brush Friction

The upper stabilising brush (USB) is a single steel brush which is in the upper region of the fuel stringer. The brush, including the fuel stringer, has a diameter of 250.4mm. This is a smaller diameter than many of the regions within the core. Therefore, it is not always in contact with the fuel channel and refuelling machine.

In comparison to the lower stabilising brush (LSB), which will be discussed in the following subchapter, the USB has a smaller diameter and is at the upper end of the fuel stringer. When analysing the effects of the USB it is important to note that that the load value is referenced with respect to the location of the LSB. The USB is roughly 6 metres further up the channel; this means that any contribution from it will not have any relevance to the current brick. Another important consideration about the USB is that the maximum size of bore that is possible to be detected is roughly 250mm, as this is the diameter that the brush extends. The first region that it is present in is while the USB is in the guide tube. The guide tube has a diameter of 248.2 mm and is present from the start of the trace until near the start of brick layer 4 where it widens. The second region is in brick layer 5 where the USB enters an area of restriction from the piston seal bore with a diameter of 247.65 mm. There is sometimes one other region with a contribution from the USB which can occur midway through layer 9. This is due to the USB "hunting" for the entrance to the stand pipe which has a diameter of 250.4mm. This last contribution from the USB is not always present as the brushes may not be physically in contact with it.
The contribution of the USB is investigated using historical FGLT data later in chapter 4 to help create a generalised understanding of its contribution during operation.

2.3.3.1.5 Lower Stabilising Brush Friction

The LSB with the fuel stringer included, has a larger diameter than the USB at 266.2mm. Additionally the lower stabilising brush is not just a single brush like the USB but instead is three sets of stabilising brushes separated by steel chamfers. At the

beginning of their life the graphite bricks are machined to have an inner diameter of 263mm resulting in constant contact with the LSB at all regions within the core. The LSB is the contribution which is directly related to the fuel channel bore dimensions. Changes in the channel bore can be observed in the load value of the FGLT. This relationship can be seen by observing the different brick layers and identifying the brick interfaces (the regions between two bricks that are narrower) as peaks in the discharge trace or troughs in the charge trace.

It is also worth noting that the current belief is that there are two types of friction associated with the LSB, namely static and kinetic friction. Static friction occurs at the narrower brick interface when there is a quick change in the bore diameter resulting in a larger amount of force required to move the fuel stringer. Kinetic friction occurs during the brick layer when there are gradual changes in bore resulting in a smoother response in load value. This means that over the course of the trace there are different contributions to friction that need to be considered.

In chapter 4 the LSB is visited in greater detail by analysing 10-15 years of refuelling data and establishing its relationship to channel bore measurements.

2.3.4 Inspections

Inspections are performed at statutory outages after fuel stringers have been vacated from their respective fuel channels. Statutory outages generally occur every 2-3 years; however, many of the older reactors are now being inspected every year. At outages roughly 20-30 of the fuel channels (~10% of the core) are subjected to inspections. CHANSELA (CHANnel SELection Assistant) (Watson, et al., 2011) is a channel inspection tool which combines multiple different sources of information to provide support for channel selection, as it is not feasible to inspect every channel at outages. Inspecting every channel would result in outages that would be economically costly due to the length of time it would take and unnecessary movement of the fuel stringers is not desired. Part of this tool also involves data visualisation of the graphite core simplified to an overview of the fuel channels to provide supporting knowledge of all the previous inspections that have been undertaken. It achieves this by having a backend database which contains all the data that has been recorded at previous inspections including channel diameters and crack descriptions. The main benefit

however is a set of rules that it operates on to provide recommendations about which channels should be inspected based on several factors. These rules exist like standard database queries but when combined with the data visualisation it is a very useful tool to understand the health of the core. In addition to the standard ruleset it is possible to create custom rules for the selection of channels. This is crucial as it allows this system to be able to be utilised later in life where the previous criteria for channels to be inspected are no longer applicable. Ensuring that the inspections are performed quickly and safely is important to allow the reactor to be brought back to operation as it is economically costly to have the reactors shutdown longer than is required. It is estimated that every day that a reactor is offline costs roughly £0.5 million in lost earnings.

Inspections are important for the safe operation of AGRs as they are the only time it is possible to obtain ground truth information about the health of the core, allowing the graphite bricks to be assigned labels about their health. At monitoring, graphite core engineers can make assumptions about the current health of the graphite bricks in the core, such as widening/narrowing of fuel channels or the presence of cracks, but without inspecting the fuel channel it cannot be confirmed.

2.3.4.1 Fuel Channel Bore Measurement

Core Bore Inspection Unit/ Core Bore Measurement Unit (CBIU/CBMU) or New In-Core Inspection Unit 2 (NICIE 2) are devices which are used during offline inspections to accurately measure the diameter of inspected channels (Cole-Baker & Reed, 2005) (Steer, 2005). The CBIU/CBMU device consists of 4 feelers which are spaced 90 degrees apart. The feelers measure the distance from the centre of the channel to the channel walls as the device is lifted out of the channel. In addition to the feelers a set of four wheels (two above and two below the device) are used to ensure that the device is centred in the middle of the channel. Additionally, two tilt transducers determine the tilt experienced by the bricks on individual brick layers.

An example of the data produced can be seen in Figure 2.11 where a single fuel channel's diameter is recorded from brick layer 2 to brick layer 11.

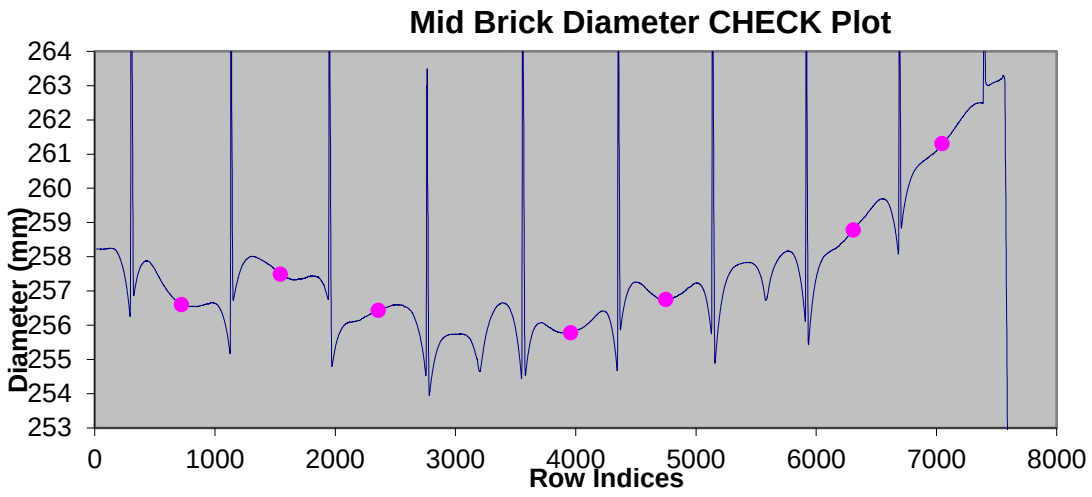


Figure 2.11 – An example of the recorded output from bore measurement instrumentation highlighting the diameter of the middle point of each brick.

The benefit of using bore measurement equipment allows core engineers ascertain core health information about the dimensions of the bricks. This includes being able to track the shrinkage rates when multiple inspections have been performed on the same channel at previous inspections. Cracks that have a significant impact on the dimensions of a fuel channel can be detected from bore measurement information however this information alone is often not enough for a confirmation of a crack.
Other useful information that is obtained from dimensional bore measurements is the measure of whether the fuel channel is elliptical or remains circular. Additionally, as the bore dimensions are measured throughout the entire channel the bow of the individual bricks can be measured which shows the difference in diameter of the brick over its vertical length. All this information is valuable to graphite core engineers to accurately understand the current health of the core and its future process of aging.

2.3.4.2 Visual Inspection

In addition to measuring the dimensional properties of fuel channels, visual inspections are performed on the fuel channels. The inspections devices have a camera attached

that records video footage of the inside of fuel channels. This process records videos of the device travelling from the bottom of the core to the top at six different orientations each rotated by 60 degrees. The field of view of the camera is greater than 60 degrees and therefore six traversals of the channel provide a complete view of it. While this is happening graphite core engineers watch the footage and identify any signs of degradation such as cracks. If any are found the camera is inserted into the channel again to capture detailed footage of the specific defect. The camera at this point is free to follow the required shape of the defect to capture a clear representation of it. Once this is complete individual frames are extracted from the video footage and are manually stitched together to create a complete 2D image of the area of interest. An example of a manually stitched defect can be seen in Figure 2.12.

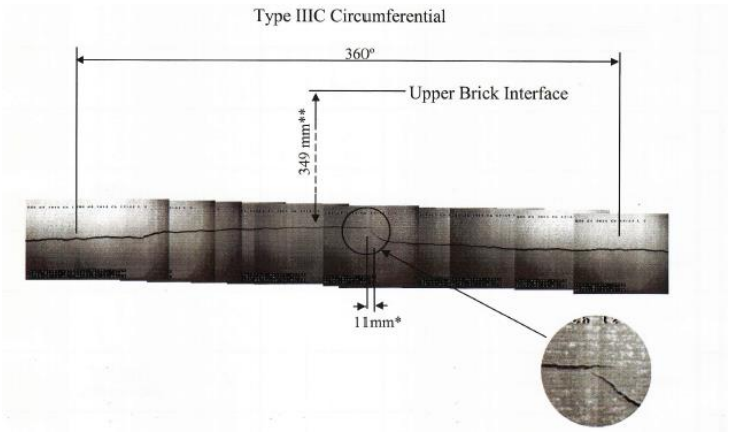


Figure 2.12 – An example of the recorded manual stitching process that occurs at scheduled offline inspections.

In addition to the recorded video footage obtained at inspections there has been recent work that aims to create a complete image of every fuel channel inspected and is being called a “chanorama” (channel panorama) (West, et al., 2015). This is performed by extracting horizontal strips from individual frames of the video footage and stitching these strips vertically to produce a complete vertical section of the channel. Once this has been performed for each of the traversals of the channel the six strips are then automatically aligned and stitched together horizontally. This process can be

completed automatically from the video footage and is quicker than the manual process of stitching the video footage together. Currently this automated image stitching is performed in addition to the manual image stitching but will replace it eventually.

The most recent state-of-the-art technique that is being researched for visual inspection of graphite cores is through a process called structure from motion (SfM). SfM uses multiple 2D image viewpoints of the same structure to recreate a 3D object. This process is completed from extracting individual frames from the video footage to recreate the entire channel; however, this is currently in the conceptual stages of research (Law, et al., 2017).

The process of visual inspection has been greatly developed over the life of the reactors; this has been highlighted in Figure 2.13. Initially any defects were hand drawn to store as a record, following this the images of defects were manually stitched together in a time-consuming process and now this image process is able to be automated in a much quicker time by software (Bradford, 2015). In the future SfM is anticipated to be undertaken to obtain 3D information about defects.

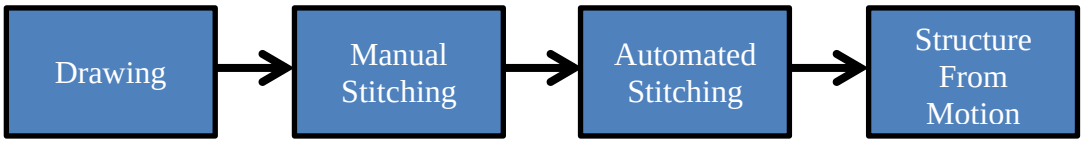


Figure 2.13 – The progression of the visual recording of graphite bricks at offline inspection events.

2.3.4.3 Trepanning

Another technique to understand the health of the graphite is an invasive inspection technique, called trepanning, which involves removing a small cylinder of graphite from the bricks in the core. From the extracted graphite, the Young's modulus and strength can be examined and tested (Tzelepi, et al., 2005). The amount of graphite removed is a diameter of 19mm with a depth of 6mm. By analysing small samples of the graphite, it is possible to obtain information about the overall health of the brick.

The monitoring and inspection data provide information about the core but to obtain further health information, additional data analytics and machine learning techniques can be used.

2.3.4.4 Eddy Current Inspection

Eddy current is the last method of inspection the graphite bricks within the core. It is a non-destructive testing method, the purpose of which is to be able to detect the presence of flaws in a conductible material. Eddy current testing can be used to monitor the health of reactor core graphite (Dekdouk, et al., 2012). This is due to there being a correlation with the electrical conductivity and the weight loss of graphite (Hacker, et al., 2000).

The first prototype device for Eddy current testing was the Proof-of-Principle Eddy Current Tool (PoPECT) which allows for in core inspections of the graphite fuel bricks. The probe includes a rotating head which allows for both circumferential and vertical scan operation modes.

Following on from PoPECT, the Prototype Eddy Current Inspection Tool (PECIT) was developed. The improvements from the PoPECT include three individual probes (a single probe coil, a differential probe and a radio gradiometer) and stabilising wheels at the upper and lower ends of the device. PECIT is also able to scan much larger areas of the bricks due to its ability to perform helical scans with a nominal pitch of 18mm.

2.3.5 Graphite Core Health Assessment Research

There have been many instances of research being undertaken to provide decision support for graphite core health assessment. The following section summarises the different approaches that have been taken the better understand the graphite core of AGRs.

A method of tracking cracks is for an operator that manually extracts frames that contain cracks from a video recording of a fuel channel inspection during a scheduled outage. A method to speed up this process has been implemented by West (West, et al., 2015) where the frames from the video footage are automatically stitched together to create a 360-degree image of the channel. A crack detection algorithm can then be applied to the image and the locations and severity of the cracks can be stored.

Wallace (Wallace, et al., 2010) has shown that the movement patterns of control rods are able to provide information, that can be used to understand the health of the graphite core. The control rods are used to control the rate of the reaction inside the

core. They can ascend from and descend into the control channels to do this and their height is recorded. As groups of control rods are provided with the same instruction, the differences in which they change in height can be used to infer possible faults in the control channels.

The British Energy Trace Analysis (BETA) system (West, et al., 2008) has been deployed and its purpose is to provide decision support for domain experts during the scheduled refuelling events. BETA automatically identifies and extracts features form the refuelling traces using a rule-based algorithm. Once these features are extracted, they can be compared to benchmarks to detect anomalies. These benchmarks are obtained by analysis of normal condition refuelling events and will evolve and change over time.

Stephen has shown how hidden Markov models can be utilised to detect deviations from expected values within FGLT data (Stephen, et al., 2009). Part of this work involved modelling Markov dynamics to learn an envelope of expected FGLT data behaviour and from this can then calculate the perplexity of this expected value to the recorded value. This work was undertaken to provide a method to interpret FGLT data without obtaining models of the underlying unknown relationships in the data. This work provided a demonstration of how a decision support tool could be implemented to allow graphite core engineers to monitor the ageing process of AGRs and draw more robust conclusions about the state of the core's health.

A substantial volume of other work has been undertaken on the health management of the UK's AGR cores. The University of Bristol has constructed both a $1/8^{th}$ and $1/4^{th}$ scale model of the core of an AGR on a shaker table to help understand the effect that seismic activity has on the core (Shaw, et al., 2005) (Castro, 2005) (Brasier & Rogers, 2011) (Dihoru, et al., 2011) (Roscow, et al., 2008) (Horgan & Webster, 2014). It was not possible to construct a 1:1 scale model of the core as this would result in a model which is over 1000 tonnes in weight, which would be too heavy for any of the earthquake simulators in Europe (Dihoru, et al., 2016).

Dihoru describes work which focusses on physically modelling the seismic effects on AGRs with degraded components (Dihoru, et al., 2016). Doubly axially cracked bricks were the primary focus of the seismic simulations consisting of a random configuration of 30% of brick layers 4-7 being axially cracked. The tests allow the core to experience

different types of vibration inputs including white noise, sinusoidal dwell and seismic. The displacements of the bricks are recorded and are used to understand the core health under different conditions to provide evidence for safe operating limits.

Modelling has also been undertaken using Finite Element (FE) solid model analysis to simulate the state of the core components under different characteristics such as cracks in graphite bricks. One application of FE modelling is SALCOR (Solid Analysis of Loaded Core) (Salih & McLachlan, 2016). This application can represent the geometry of a loaded core and allows for simulating the opening, shear and separation of singly and doubly cracked fuel bricks as well as core behaviour resulting from keyway disengagement (Salih & McLachlan, 2016). Another FE model approach is GCORE which is specifically used for seismic assessments of the graphite core (Riley, et al., 2016) (Cowell & Steer, 2016).

Finally FE modelling has been undertaken to aid in the prediction of keyway initiated axial brick cracking (Tan, et al., 2016) and the propagation of such cracks (Crump, et al., 2016) (Pearce & Kaczmarczyk, 2016). Similar to the propagation of cracks, work has been undertaken to investigate secondary cracking, these are cracks that could have been caused by the release of energy of another crack (Booler, et al., 2016) and the interaction that cracked bricks have in their neighbouring regions (Teng, et al., 2016). Kyaw has modelled the cracking of graphite bricks that have undergone radiolytic oxidation and irradiation (Kyaw, et al., 2014).

As mentioned previously and it being a large component of the work in this thesis, crack detection is of major importance to the safe operation of AGRs. A statistical method of modelling the evolution of bore initiated cracks using inspections has been shown by Robinson (Robinson, et al., 2016). Additionally, code has been developed by Robinson, called CrackSim, which enables judgements to be made about planned future inspections based on both cracking rates and uncertainties.

Understanding the process of graphite brick cracking has been an important priority of AGR health research. A statistical approach has been considered to help anticipate the number of cracks that are expected to be found at inspections as well as predicting the types of cracks that are likely to be present (Maul, et al., 2008) (Tan, et al., 2011). Hodgkins has studied crack propagation using X-ray tomography to understand the behaviour of how the energy is dissipated in nuclear graphite (Hodgkins, et al., 2005).

Additionally, Hodgkins has also studied the effect that radiolytic oxidation has on the fractures within graphite bricks (Hodgkims, et al., 2008).

Keyway initiated brick cracking research recently has been prominent in the literature in recent years due to the detection of the cracks in 2014. Bradford provides an overview of the process of reducing the uncertainties that have been present in the previous efforts to predict cracking (Bradford, 2016). Additionally, Bradford identifies areas of the uncertainty that have not been considered previously to allow for greater accuracy in the forecasts of cracks in future inspections. Leguillon has developed a tool for predicting the initiation of these cracks by accounting for both stress and energy conditions (Leguillon, et al., 2016).

Investigating the microstructure of the gilsocarbon graphite is important to understand the changes that it is undergoing due to irradiation within the core of the reactor. An overview of the micromechanical testing of graphite is provided by Liu (Liu, et al., 2016). The ACCENT program has been undertaken to help understand the impact that dimensional changes to the material and microstructure properties of graphite at various stages of irradiation (Taylor, et al., 2016). National Nuclear Laboratory has developed a micromechanistic model of polycrystalline graphite irradiation behaviour which models individual crystals (Tzelepi, et al., 2016).

There has also been research work undertaken on the material property changes of graphite due to the fast neutron irradiation and radiolytic oxidation. This work is known as project Blackstone and allows an understanding of graphite at the end of life conditions for the UK's nuclear reactors using a materials test reactor (Davies, et al., 2016). Project Blackstone has currently undergone two phases of research and phase three is currently being undertaken and is expected to be complete in 2021. Phase one and two were focussed on property changes to the predicted end of life condition of the graphite bricks at the older reactors (Dungeness, Hunterston, Hinkley Point, Hartlepool and Heysham 1). Phase three is focused on the remaining stations (Heysham 2 and Torness).

2.4 Challenges around Nuclear Monitoring

Currently the worldwide trend for nuclear power stations is to extend their operating life beyond their initial design life. This can be seen in the United States where in 2012,

71 out of the 104 nuclear power plants that were in operation have had their life extended beyond their initial design life from 40 to 60 years (Meyer, et al., 2012), with many looking to extend that further to 80 years (Bond, et al., 2011). In the United Kingdom the nuclear generation fleet is also undergoing life extension. The primary reason for this is that the nuclear reactors are still safe to operate and able to generate electricity beyond their initial designed lives. Life extension of AGRs is more economically beneficial than decommissioning existing reactors and building new reactors. Another benefit of this is that their operation allows for the maintenance of a reliable baseload for the country as old coal stations are expected to close. The current generating fleet in the United Kingdom has been generating since the 1970's (World Nuclear Assocation, 2015) and many of the stations were originally expected to be shut down in the 2000's, but now have had their operational lifetimes extended. In regard to the life extension of AGRs in the United Kingdom, McLachlan (McLachlan, et al., 1995) stated that the presence of a keyway root crack would mean that the station had reached the later stages of its operation life and that these cracks would need to be incorporated in to future safety cases. These cracks have now been identified in both Hunterston reactors and they are still planned to remain in operation until 2023 as the safety cases for operation have be redefined that show that the reactor can operate safely with the presence of these cracks. These cracks are no longer considered as life ending and require large amounts present to result in loss of structural integrity of the core.

With this increased plant life, monitoring has become more prevalent as many of the components were not expected to be in use for double their initial life. This has also shifted the type of maintenance that has been traditionally performed in nuclear power stations, from planned preventative maintenance to be more proactive in repair and maintenance of components (Agarwal, et al., 2013). This is where maintenance can be planned by using the condition monitoring data to not just detect faults but to predict when the next fault will occur or the remaining useful life (RUL) of an asset. The detection of faults is known as diagnostics and predictions of future faults or RUL are known as prognostics. This type of maintenance, if it is successful, is very economically beneficial to the nuclear operators as extended unnecessary outages are very costly in terms of lost possible revenue. Therefore, being able to plan efficient

maintenance ahead of time will reduce the money lost from outages. With suitable monitoring it will also reduce staff exposure to radiation either by having on-board sensors or by shortening the time that staff members are exposed through more efficient maintenance. This is not possible within the context of graphite reactor core bricks as they cannot be repaired or replaced. Therefore, they are inspected to ensure that they are safe to operate until the next statutory inspection. It is becoming more desirable for monitoring to provide more frequent health information. Data analysis and machine learning are integral in obtaining valuable information to influence the health management of assets.

2.4.1 Prognostic and Diagnostic Techniques

The Department of Electrical and Computer Engineering at the University of Western Ontario have produced a review of fault detection and diagnosis methods focusing on nuclear power plants (Ma & Jiang, 2011). They identify that there are two key fault identification and detection methods; these are model-based and data-driven methods. Model-based methods require an understanding of the physical properties of the structure, system or component that is being monitored that can then be translated into a mathematical model. This model can then be used to predict what the normal degradation of a piece of equipment should be at a certain time in its lifecycle. Then the difference of the measurements to this value are known as residuals and these indicate any degradation that has occurred. Data-driven models are more prominent in a lot of condition monitoring approaches due to the complexity of trying to represent some structures, systems or components mathematically. Data-driven models primarily use machine learning techniques to learn about the application through previous historical operational data. They can be inadequate for some problems as they simply look for relationships in the data or its statistical properties rather than basing their decisions on the known physical properties of the equipment. In their review of prognostics for nuclear power, Coble (Coble, et al., 2015) states that data-driven methods are effective at predicting run to failure times of equipment but that understanding the physical mechanics of the system that leads to these is also invaluable. This is due to the nature of many faults in safety critical systems where it would be impractical and unsafe to let them run to failure such as in the nuclear power industry.

The Nuclear Engineering Department at the University of Tennessee have been focussing their research on prognostics, plant life extension and management as well as providing prognostic support for new builds, especially in advanced small modular reactors. Coble (Coble & Hines, 2011) shows how a general path model is used to extrapolate a parameter in condition monitoring to be able to predict its time of failure through dynamic Bayesian updating techniques. This method won the 2008 Prognostic and Health Management Data Challenge where it could predict the degradation of an aircraft engine from multivariate simulator data when a fault condition has occurred at an unknown time. Welz (Welz, et al., 2015) has shown a lifestyle prognostic modelling method from beginning of life, through operational life and to end of life using Bayesian and bootstrap aggregation for the lifetime of a heat exchanger in a light water reactor. The modelling was then tested on a heat exchanger which was under accelerated aging. To construct the model, feature extraction was performed on the many sensor outputs to reduce the signals that would be input to the model. Barbieri (Barbieri, et al., 2015) has shown how trends and advances in prognostics has led to better and more efficient maintenance where remaining useful life estimations have been performed on experimental data from electric motors. Barbieri used feature extraction on a large amount of sensor variables to obtain the features that can be used to create a prognostic parameter that will allow the remaining useful life to be calculated by extrapolating the parameter to a failure threshold using a linear general path model. The method of generating these prognostic parameters was compared between using a genetic algorithm approach and an ordinary least squares estimation.

2.4.2 Applications

2.4.2.1 Cable Monitoring

There has been a considerable amount of research into cable aging and management in nuclear power stations. This is due to many cases of plant life extension where it was never considered that these cables would be in operation as long as they are. Shumaker (Shumaker, et al., 2012) with the Analysis and Measurement Services (AMS) Corporation have described methods taken to identify, locate and repair/replace damaged cables in nuclear power stations. They do this by utilising many non-destructive testing methods as well as human inspection to predict the RUL.

AMS in cooperation with Polymer Aging Concepts, Inc (PAC) have developed micro-sensors that can be installed into new and existing cables that can be used to indicate polymer material deterioration which can indicate the RUL. Further work has been completed by McCarter (McCarter, et al., 2015) at AMS showing how using indenter modulus measurements it is possible to estimate the RUL of instrumentation and control cables. Indenter modulus is performed by a probe gently pressing down on a cable and recording the change in displacement of the cable with respect to the force applied. This method is beneficial as it is a non-destructive type of testing compared to some methods of cable testing that can damage the cables such as elongation-at-break. The indenter modulus test can also be performed while the cable is in operation. The RUL prediction was tested using the general path model (GPM), an appended general path model and a third order polynomial fit on thermally accelerated aged cables. It was found that both GPM estimations had high accuracy even with few data points and when the third order polynomial fit obtained sufficient data points it produced similarly accurate results.

Bowler performed a review of many different cable monitoring techniques (Bowler & Liu, 2015) and the factors which affect the aging of the cables. They state that when cables are irradiated it can cause them to produce gas that can be trapped within the polymer and can cause series degradation of the electrical and mechanical properties of the cable and that under some situations the presence of oxygen can cause the cables to age at an accelerated rate even if the cables are at a lower temperature. Some methods of testing cables are through capacitive testing, permittivity of the polymer material, dielectric and mechanical testing.

2.4.2.2 Online Condition Monitoring

As stated previously more nuclear power stations are moving to using online condition monitoring and shifting from diagnostics to prognostics. There are many benefits of this such as being able to anticipate a failure when the signs first appear. This change from time scheduled maintenance to asset health scheduled maintenance may lead to overall shorter outage durations and less of them (International Atomic Energy Agency, 2008). Online condition monitoring can be seen with a collaboration of the U.S. Department of Energy's Light Water Reactor Sustainability program and the Electric Power Research Institute's (EPRI) Long-Term Operations program to create

the Fleet-Wide Prognostic and Health Management (FW-PHM) software suite (Agarwal, et al., 2013). The FW-PHM suite consists of databases for fault analysis and RUL predictions and equivalent advisers which interpret the input data to provide centre and plant staff with decision support relating to faults and maintenance required. The fault database contains a comprehensive list of faults for all the utilities which are associated with the EPRI. With this comprehensive list there is also information on what equipment will the specific fault manifest in and what characteristics are used to identify the fault. It is then the fault diagnostic advisor which uses online monitoring data to determine if there are any faults and what those faults could be. This information is then provided to a technician who will now have much more specialised knowledge of the problem to decide about a specific piece of equipment. The RUL advisor will also get the fault prediction from the diagnostic advisor and will use that and the sensor information to obtain a prediction of the RUL. These RUL predictions are then stored in the RUL database which can then be consulted and used to predict when maintenance should occur.

2.5 Conclusions

Nuclear power electricity generation is an essential part of many countries generating capacity in the UK it makes up approximately 20% and is integral to provide a base load. This chapter has provided an overview of nuclear power generation with a more focussed look at the design and data obtained from the UK's AGRs. The current fleet of AGRs that are operating are entering the later stages of their operational lives, the oldest of which have been connected to the grid since 1976. Each reactor has been individually granted life extensions based on safety cases of operation. In the context of this work the most important feature to consider about AGRs are their graphite cores. The graphite core within an AGR is made of graphite bricks which are degrading over the course of the reactor's operational lifetime. The graphite bricks cannot be repaired or replaced and are therefore considered a life-limiting factor for the reactor. The main concern of graphite brick degradation is dimensional changes which are caused by irradiation and radiolytic oxidation. Initially at the start of their operational lives the graphite bricks begin shrinking, resulting in narrowing of fuel channels within the core. After many years inside the reactor the bricks reach turnaround point where

they begin expanding, resulting in the fuel channels becoming wider. The stresses applied to the bricks are different before and after the turnaround point which can result in different cracks developing at different points in the life of the reactor.

Cracks in graphite fuel bricks are currently detected during scheduled inspections that occur when the reactor is offline. Dedicated measurement equipment is inserted into the core to measure the dimensions of the fuel channels and determine dimensional change. Additionally, this equipment has a camera to record footage of the inside of the fuel channel. Inspections are supplemented by FGLT monitoring data that is obtained more frequently at refuelling events. The load value of the FGLT is obtained from the combination of four different components: fuel stringer deadweight, aerodynamic effects from CO_2, lower stabilising brush friction and upper stabilising brush friction. It is possible to know the expected contributions from each component at each point on a trace based on physical knowledge of the core. The lower stabilising brush friction is an important contribution to understand as it is directly related to the dimensions of the fuel channel. This relationship is investigated in the future chapters of this thesis.

Ensuring that nuclear power stations are operating safely while their operational lives are being extended is of the utmost importance. In the later chapters of this thesis two different condition monitoring approaches are proposed, which provide decision support to graphite core engineers by extracting health information from FGLT data.

3 Data Analytics and Machine Learning

3.1 Introduction

Machine learning and data analytics are valuable tools for extracting useful information from operational data of assets. Machines learn by being provided with a large number of examples known as training data. Features of the examples are then analysed, such as their similarity to known historical examples, known concepts/classifications or how they are different from previous examples. This allows models to be developed that allow future examples to be sorted intelligently and correctly. Data analytics involves examining a set of data to be able to extract more information or meaning from the data than it currently provides. Data analytics can often be applied to guide the training methods of machine learning.

Kevin P. Murphy defines the three primary types of machine learning that can be undertaken in his book *Machine Learning A Probabilistic Perspective* (Murphy, 2012). These are:

- Supervised machine learning,
- Unsupervised machine learning
- Reinforcement machine learning

Common problems within machine learning often involve the distinction of whether the input data is labelled or unlabelled. Labelled data is input or training data which has meaningful information associated with them, these are typically either numerical or categorical values. Unlabelled data however, does not have associated data with them. Typically, in situations with labelled training data supervised machine learning is used and in situations with unlabelled data unsupervised machine learning is used.

Semi-supervised machine learning, which utilises both labelled and unlabelled data, has not been included in the three main types of machine learning techniques. However, it is proposed here in this work that it should be considered. Semi-supervised

learning has a large amount of overlap between supervised and semi-supervised machine learning. Before defining clearly what semi-supervised learning is, it is important to clarify the differences and training methods of both supervised and unsupervised machine learning and additionally a short overview of reinforcement machine learning is provided.

3.2 Supervised Machine Learning

Supervised machine learning is the standard approach for most classification and regression problems. The main requirement of supervised machine learning is a supply of ground truth training data that is an accurate representation of the problem. This training data includes the input data which is accompanied by an associated label. The relationship between the input data and their labels are explored to produce a generalised function that represents the relationship between them. This generalised function can produce a solution to future, previously unseen, data. A benefit of supervised machine learning is, that provided there is sufficient labelled training data that is representative of the feature space, the problem should be solvable. The two most common objectives of a supervised machine learning problem are either classification or regression.

3.2.1 Classification

Classification is a supervised machine learning problem in which an algorithm is trained to predict a discrete class when provided with a new observation of data. Training a classifier requires a source of data with associated labels indicating which class they belong to. The training data is also required to be representative of the future data that is going to be supplied for classification, allowing the classifier to generalise well with respect to the problem. The properties of the inputs and outputs are explored to be related to each other. The classifier can then use this training information to create rules or functions which allow the various classes of data to be distinctly identified. This provides the classifier the ability to learn about the feature space of the problem. Once the classifier is trained it can then use its learned knowledge to provide classifications to data which it has never seen before.

An example of a classification problem when related to condition monitoring could be being able to identify the specific type of rolling element bearings fault from vibration

signals (Abbasion, et al., 2007). As part of the classification process Abbasion was able to correctly classify the vibration input data into six different categories of faults in addition to a health classification. A benefit of many classification algorithms, in terms of fault detection, is that they not only identify that there is a fault in the data but provide an indication of what type of fault is present.

3.2.2 Regression

The main difference between classification and regression is that classification has a finite discrete amount of outputs available to the algorithm, whereas regression produces outputs that are continuous in nature. Supervised regression algorithms are trained similarly to classification algorithms; they require a training set of input data with associated output values. The algorithms explore the feature space of the data to construct a relationship between the inputs and the output that is sufficiently generalised to predict the correct output of future unseen input data. Calculating the RUL of assets, such as rolling bearing elements is an example of the use of regression (Loutas, et al., 2013). This can be achieved by recording health information paired with the degradation status of components to learn the relationship between them. This allows for predictions to be made when future data is obtained.

3.3 Unsupervised Machine Learning

Unsupervised machine learning is involved with learning the statistical properties of the input data without any knowledge of labels associated with that data. Traditional tasks that are performed by unsupervised learning are clustering, novelty detection and dimensionality reduction. It is performed by using an agnostic view of the problem and exploring the underlying density that is likely to have generated the input data's distribution.

3.3.1 Clustering

Clustering is a machine learning approach and it is the unsupervised counterpart of classification. It involves grouping instances of data together that have similar properties. Clustering can be used as a method of data exploration during data mining and visualisation of data. An advantage of data clustering is that understanding can be obtained about the groupings of sets of data without the time-consuming process of

manual analysis and labelling. It can also be very useful for finding trends in the data and for understanding the distribution.

Unsupervised clustering methods have been applied in nuclear condition monitoring applications by Baraldi to identify similar transients based on their characteristics (Baraldi, et al., 2013). This was performed firstly on an artificial case study to be verified prior to being applied to a simulation of a pressuriser in a PWR nuclear power station. The use of unsupervised clustering was very beneficial for future fault understanding and interpretation as it allowed for similar transients to be grouped together allowing for prototypical trajectories to be created per cluster.

3.3.2 Anomaly Detection

Anomaly detection is most often associated with unsupervised learning, however there are implementations of supervised anomaly detection algorithms. Anomaly detection is primarily used within data mining/exploration circumstances when all the data is from an unlabelled source. The most common application of novelty detection is when a large volume of "normal" data is accessible but there are very few instances of "abnormal" data (Pimentel, et al., 2015). There are numerous approaches to perform novelty detection; one such method is through a probabilistic approach. In this approach the training data is analysed to generate a probability density function of the data. The algorithm can then decide on a threshold to associate future data with "normal" conditions. Under many circumstances anomaly detection algorithms are viewed as a one-class classifier, where the trained data is normal operation. Typically, these algorithms are effective in being able to identify outliers in the data.

Anomaly detection of intrusive malicious attacks of simulated supervisory control and data acquisition (SCADA) systems has been shown by Yang (Yang, et al., 2006). Jin has also shown how data driven anomaly detection can be performed on nuclear power station data by using symbolic dynamic filtering to mask the sensor degradation signatures (Jin, et al., 2011). In this case anomaly detection is achieved in real time on simulated data which used Gaussian noise to represent sensor uncertainties. Jin proposes that degraded sensors can be detected in real time for components within nuclear power stations which have low false alarm and missed detection rates.

3.3.3 Dimensionality Reduction

Dimensionality reduction in the context of data analysis is involved in reducing the number of features that are used to represent the data. It is often beneficial to lower the number of dimensions in the input features of many algorithms to reduce their computational complexity in their training or their execution. Additionally, in many cases of high dimensionality data there are redundancies between features that can often be consolidated to reduced overall complexity.

In many cases the features being used to describe the problem are very numerous which often makes it impossible for humans to understand the data visually. Therefore, by reducing the dimensionality of the data down to dimensions that are visually comprehensible, it can possible for humans to understand relationships in the data.

3.4 Reinforcement Learning

The last type of machine learning is reinforcement machine learning, this type of learning involves an agent having the ability to interact with its environment, typically through trial and error. The requirements for reinforcement learning are an agent and an environment that it can interact with. The agent obtains a description of the current state of the environment and from this state performs an action that will change the environment. This results in a reinforcement reward that indicates whether the action performed was beneficial or harmful to the desired response. This process then repeats with the new state of the environment. With the reward score appropriately tuned for the specific domain this can create a very knowledgeable agent. Additionally, for situations where the environment can be simulated, powerful computers can be used to simulate large quantities of outcomes that would be infeasible to have that amount of labelled output data for supervised learning.

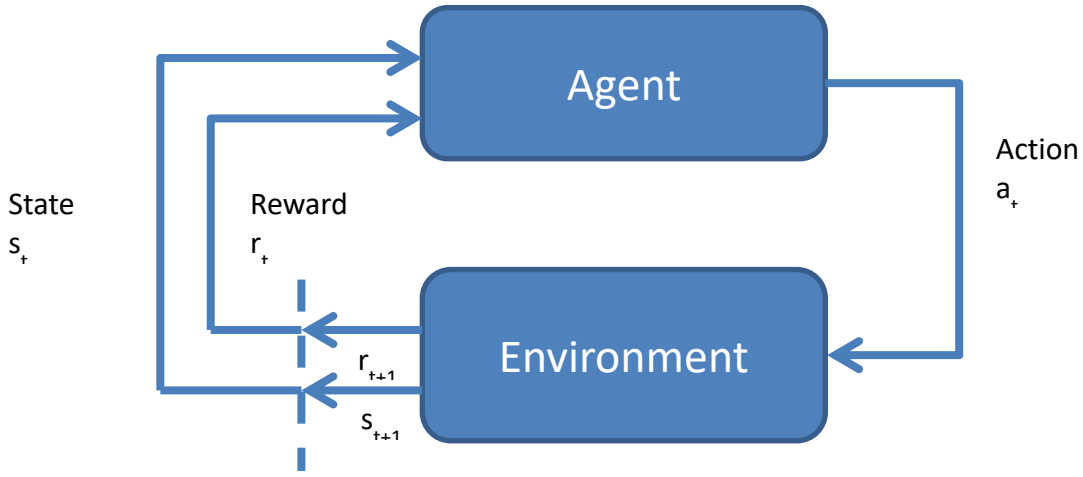


Figure 3.1 – Example of reinforcement learning where an agent can interact with the environment.

Reinforcement machine learning has been shown on simulated experiments to have use in scheduling condition-based maintenance by Knowles to move away from scheduled maintenance while maintaining high reliability of assets (Knowles, et al., 2011). The benefits of reinforcement machine learning within this context are that it is possible to optimise the performance of the algorithm with respects to long term rewards. If implemented well on a system this would save time and costs on maintenance over a long term.

3.5 Semi-Supervised Machine Learning

Semi-supervised learning combines aspects of both types of learning to produce new machine learning algorithms (Chapelle, et al., 2006) (Zhu & Goldberg, 2009). Semi-supervised learning aims to improve the capability of learning algorithms when there is an insufficient amount of data that has been labelled. It does this by considering information from the unlabelled data that would otherwise have been ignored. This can involve looking at the statistical properties of the data such as distances between samples, distribution or density of samples and using what is known about the labelled data to train the algorithm with confidence.

Semi-supervised learning so far has seen limited use in nuclear condition monitoring but there are two successfully applications in the literature (Ma & Jiang, 2015) (Moshkbar-Bakhshayesh & Ghofrani, 2016). Ma has shown how a combinational algorithm of spectral clustering and label propagation has managed to classify faults in a Canada deuterium uranium (CANDU) desktop simulator and a physical NPP simulator. Moshkbar-Bakhshayesh has shown a combinational supervised and semi-supervised method using transductive support vector machines (TSVMs) for the detection and classification of NPP transients.

3.5.1 Transductive or Inductive

Within machine learning, especially within semi-supervised learning, there is a distinction on how the specific algorithms are applied. Transductive machine learning problems involve generating a solution to the problem that is confined to the labelled and unlabelled data that has been subjected to training. Additionally, when new data becomes available the algorithm would then have to be retrained to produce predictions for that data. Inductive machine learning instead, learns a function, using labelled and unlabelled data, which produces the output when provided with input data. Therefore, when new unlabelled data is obtained after the initial training, it can be supplied to the learned function to obtain an output prediction. This difference between these two learning paradigms leads to inductive algorithms requiring less computation time when being applied to new data whereas transductive algorithms will have larger computational complexity with every new sample of data.

3.6 Algorithms

The following subsections of this chapter will provide an overview of some algorithms which are employed within the wider condition monitoring domain found when investigating the literature. Included within this subsection are the three semi-supervised algorithms which are evaluated in chapter 5, these are discussed after describing their traditional supervised counterparts.

3.6.1 Artificial Neural Networks

Artificial neural networks (ANN) are a prominent type of machine learning. They involve replicating the biological neural networks that are present within the human

brain. They can be simplified down to a series of inputs, which are the features of the data, followed by a hidden layer or multiple hidden layers of differently weighted nodes that capture the knowledge of the problem. Finally, the last part of the structure of an ANN is the outputs which provide the classification for the input or in some cases the regression. The hidden layer nodes each contain a weight value and an activation function that consolidates the outputs from the previous layer and passes an output to the next layer. The number of hidden layers and the number of nodes in each layer can vary depending on the implementation and there is a specific field of artificial neural network implementation, which involves multiple hidden layers, known as deep learning.

Neural networks are mainly trained by a process called backpropagation. It calculates the error at the output of the network using a validation set of data. This error value can then be traced backwards throughout the network to find the error value contribution from each of the nodes. From these values it is optimisation is performed on the value of the node weights in an attempt to reduce the overall error value. This is all performed under the assumption that the hidden layer nodes are initialised with random small weight values.

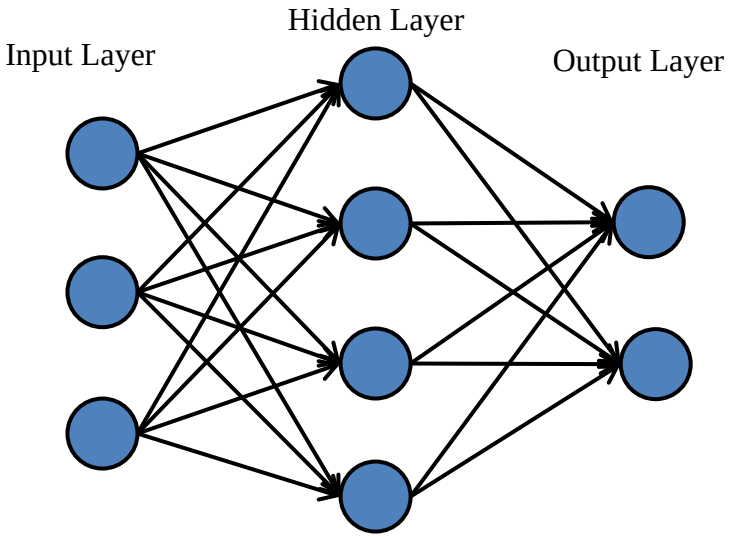


Figure 3.2 – Example of the architecture of an artificial neural network.

Within the domain of nuclear power station condition monitoring Şeker has shown the use of recurrent neural networks (a type of ANN which incorporates memory) to detect anomalies in simulated operational data (Şeker, et al., 2003). Additionally, they have

been shown to detect motor bearing damage and detect anomalies from the simulated operation of a high-temperature gas cooled reactor.

Convolutional neural networks, a deep learning technique, has been proposed to be used within nuclear power plant conditional monitoring for the detection of cracks within visual monitoring data by Chen (Chen & Jahanshahi, 2018). Chen has demonstrated the ability to detect cracks on underwater metallic surfaces from nuclear inspection videos at very high accuracy, including detecting small cracks with low contrast and variant brightness. The downsides with this method however are the long computational time associated with deep learning and the large amount of training data required, which is not always possible to obtain in the nuclear power industry.

3.6.2 K-Nearest Neighbours

K-nearest neighbours (kNN) is a very simple to implement supervised algorithm which can be used for classification and regression. When a new instance of data is required to be assigned either a classification or regression value the nearest k data instances in the feature space are used to decide that value. The distance metrics for kNN can vary depending on the application, including Euclidean distance, Hamming distance, Pearson's coefficient or Spearman's coefficient as some examples. Once the nearest neighbours have been identified, depending on the implementation, they are given a weighting based on their relevance to the instance to be labelled. In some applications the weights will be uniform and a have a value of 1/k, whereas the weights of others will be weighted based on similarity to the current instance. With weights applied a voting system will be used to assign the current instance a value.

Zio has shown the use of a fuzzy kNN algorithm in combination with genetic algorithm optimisation for diagnostic classification of nuclear transients by comparing this approach to others (Zio, et al., 2006).

3.6.3 Support Vector Machines

Support vector machines (SVM) and support vector regression (SVR) are types of supervised machine learning. In terms of classification SVMs produce a binary classification algorithm that distinguishes between two classes of data. Multi-class SVMs do exist but they are multiple SVMs stacked and working together. The basic principal of SVMs is to create a linearly separating hyperplane in the feature space,

between the two classes of data, as well as providing margins between the two classes. The bounds of this margin are represented by two additional hyperplanes which are parallel to the separating hyperplane. For the construction of a strong classifier it is desirable to maximise the distance between the parallel hyperplanes, while maintaining the separation of classes. Hard margins are used when no data is located within the margin and the data is linearly separable. However, in in situations where the data is not linearly separated soft margins are used, allowing instances of data to be present within the margin. Loss functions are used to ensure that the best margin is chosen for the overall performance, at the cost of some instances of data. The generation of the margin is achieved by choosing the data instances which are closest to the boundary between the two classes, these data instances are known as the support vectors.

One of the key techniques for the training and implementation of SVM are the use of kernels. Kernel mapping is a process where the input data is mapped onto a higher dimensional space. This is particularly useful for classifying data which is not linearly separable but by transforming it into higher dimensional space it can become linearly separable. There are many different types of kernel to use to for SVM but the most commonly used is the radial basis function (rbf) as shown in Equation 5.

$$K(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right) \qquad \textbf{Equation 5}$$

SVMs have been shown to have use in condition monitoring of nuclear power plant components by Liu, where a probabilistic SVR is used to provide short-term forecasts of out-of-control parameters (Liu, et al., 2013). This work was tested on a real data case study of a leak flow in the first seal of a reactor coolant pump where it achieved a high prediction distribution and confidence score.

3.6.4 Co-Training

Co-training is a semi-supervised learning technique that uses different views of the data to allow two classifiers to train each other using labelled and unlabelled data (Blum & Mitchell, 1998). It is important to note that there are assumptions with the

views taken of the data. Firstly, each individual view of the data must be able to generate a suitable classifier provided that the problem was fully supervised and secondly the views of the data must be conditionally independent of each other. There are variations of co-training that use different base classifiers or use more than two classifiers such as tri-training (Zhou & Li, 2005). For the purposes of the AGR monitoring co-training is undertaken with two artificial neural networks being the base classifiers.

Initially base classifiers are chosen and two classifiers are trained on their independent views of the labelled data. Each view of the unlabelled data is then subjected to their respective trained classifiers. A confidence score is obtained about the label assigned from the output of the classifier. The most confident instances of each view are then permanently given those labels and are then provided to the alternative view of the data classifier now as labelled data. This process is then iteratively repeated decreasing the instances of unlabelled data and increasing the instances of labelled data with each iteration. Once a suitable criterion is met, such as a certain number of iterations or a percentage of unlabelled data being given labels then the individual classifiers can be extracted to be used for future data based on a voting scheme. The pseudocode for the algorithm is as follows:

Split Data into Labelled (L) and Unlabelled (U) instances

Two base classifiers (C_1 and C_2)

Produce two different view of L (L_1, L_2) and U (U_1, U_2)

***WHILE** condition not met*

Train C_1 using L_1

Train C_2 using L_2

Subject U_1 to C_1

Subject U_2 to C_2

Remove most confident U_1 from U_2 and give to L_2

Remove most confident U_2 from U_1 and give to L_1

END

Co-training was initially proposed for text classification by Blum and Mitchel (Blum & Mitchell, 1998) but more recently has seen some uptake in the fields of condition monitoring, including bridge structural health monitoring (Yu, et al., 2012) and prognostic based applications (Hu, et al., 2015).

Hu (Hu, et al., 2011) (Hu, et al., 2012) has proposed semi-supervised data driven algorithm for prognostics called COPROG. The algorithm is based on the co-training algorithm but extends its usefulness into the domain of condition monitoring prognostics and remaining useful life (RUL) calculation. Hu has been able to show the benefits of using both labelled and unlabelled data in examples of a simulation of a rolling-element bearing problem and an implementation of vibration monitoring of a faulty fan. COPROG much like co-training is a wrapper algorithm to train the underlying machine learning algorithms better. In the cases explored by Hu feedforward neural networks and radial basis networks were used as the learning algorithms within COPROG. It was found that the COPROG learning produced more

accurate predictions of the RUL of the bearings and fan than traditionalised supervised approaches when there is suitable unlabelled data to aid in the training process.

3.6.5 Label Propagation

Label propagation (Zhu & Ghahramani, 2002) is a transductive machine learning algorithm based on graph theory. Label propagation requires that an assumption is made that in the feature space that the closeness of data indicates that they are likely to have a similar class label. This allows labels to propagate from data instances to other instances through dense feature spaces. To perform label propagation initially a metric for similarity is chosen in this application the Euclidean distance (d) has be chosen and is calculated for all data instances of *x* as shown in Equation 6.

$$d_{ij}^2 = \sum_{d=1}^{D} (x_i^d - x_j^d)^2 \qquad \textbf{Equation 6}$$

This creates a graph of the similarity matrix where every point of data is related to every other point of data. The Euclidian distance is scaled by the σ hyperparemeter to generate the weight matrix (w). The σ hyperparemeter controls the influence of the neighbouring data instances where tending to ∞ results in more data having an influence on labelling and a value tending to 0 results in only the closest neighbours influencing the label. A probability matrix (T_{ij}) is then constructed based on the weight matrix as shown in Equation 8. The probability matrix is then transposed into a row matrix $\bar{T}_{ij}$, Equation 9, which is then used to iteratively update the label matrix (Y_U) as show in Equation 10. As the number of iterations tends to ∞ the value of Y_U tends to the specific positive or negative class providing classification to the unlabelled data. However, it is possible to not require this iterative process by using Equation 11 and additionally, does not require a prior estimation of the labels of the unlabelled data.

$$w_{ij} = \exp\left(-\frac{d_{ij}^2}{\sigma^2}\right) \qquad \textbf{Equation 7}$$

$$T_{ij} = \frac{w_{ij}}{\sum_{k=1}^{l+u} w_{kj}}$$ **Equation 8**

$$\bar{T}_{ij} = \frac{T_{ij}}{\sum_k T_{ik}}$$ **Equation 9**

$$Y_U = \bar{T}_{uu} Y_U + \bar{T}_U Y_L$$ **Equation 10**

$$Y_U = (I - \bar{T}_{uu})^{-1} \bar{T}_{ul} Y_L$$ **Equation 11**

Label propagation is a transductive algorithm in such that it can only be used to predict the classes that make up the training set of the data. Therefore, without invoking a further learning algorithm upon the now fully labelled data; new data will require the algorithm to be fully repeated to predict their classes. Alternatively, for future classifications, supervised machine learning algorithms can be trained using all the data as it now has labels.

Ma has shown through using a combination of spectral clustering and label propagation it is possible to classify faults in a Canada deuterium uranium (CANDU) desktop simulator and a physical nuclear power plant simulator (Ma & Jiang, 2015). This approach was pursued due to the lack of training data associated to specific faults in nuclear power plant systems. When a fault is detected the data is stored as unlabelled data, where it is used with simulated fault data to be classified as a specific type of fault.

Outside of nuclear condition monitoring an adapted label propagation algorithm has been shown to have applications for structural bridge health monitoring (Chen, et al., 2014). In this work acceleration signals are collected from sensors on the bridge, however the challenging aspect of analysing these signals is the presence of many instances of unlabelled or unseen data.

3.6.6 Transductive Support Vector Machine

Transductive Support Vector Machines (TSVM) (Vapnik, 1998) are an evolution of the basic concepts of Support Vector Machines (SVM). Their goal is to generate a hyperplane which separates two classes from the available training data. The optimisation of the location of this hyperplane is of great importance. This is achieved by maximising the margin between each class and the separating hyperplane, ensuring that, provided the training data is representative of operational data, the classifier is well generalised. Therefore, the optimisation problems involved with generating the optimal hyperplane, involves minimising the error rate of training data while maximising the distance of the nearest data instances to the separating hyperplane.

TSVMs differ from SVMs by incorporating unlabelled data into the optimisation problem. An assumption of the data is that similar classifications of data should be close to each other in the feature space and that low-density regions in the feature space a naturally more likely to better classification boundaries.

TSVMs have been utilised by Moshkbar-Bakhshayesh (Moshkbar-Bakhshayesh & Ghofrani, 2016) by combining it with a supervised machine learning method to identify transients in nuclear power plants. The two-part process consists of using a previously created algorithm containing autoregressive integrated moving average (ARIMA) model which was combined with an error back propagation (EBP) algorithm (Moshkbar-Bakhshayesh & Ghofrani, 2014). The downside of this first step method is that it is unable to cluster the unknown and unlabelled transients which are rectified by the inclusion of TSVM.

The combinational proposed identifier was applied to a water-moderated reactor for the detection and classification of NPP transients. Through combining a previous supervised method for detecting transients in nuclear power stations with a semi-supervised algorithm Moshkbar-Bakhshayesh has shown that semi-supervised machine learning can be used for classification purposes. Outside of nuclear condition monitoring, TSVMs have been shown to have applications for intelligent gear fault diagnosis by Shen (Shen, et al., 2012). In this situation the number of unlabelled data instances greatly outnumbered the number of labelled data and there was a lack of fault samples within the labelled data. TSVMs in this case proved to generalise better than

traditional SVM algorithms when additional unlabelled data was provided. Under circumstances such as this it can be very beneficial to make use of unlabelled data for improved performance.

3.6.7 Decision Trees

Decision trees (DT) are a method of supervised machine learning, one of their key advantages is that they are transparent in their process. This means that the specific reasons for their decisions are defined and ensures that there is traceability, allowing users to view all decisions made by the algorithm. This action can be very useful in determining if the algorithm has been trained effectively and not overtrained to the specific training data.

Within the architecture of a DT the input data is referred to as the root of the tree. When this root is split by the decisions made during the learning process, these are referred to as branches. Finally, when the algorithm has stopped branching and each of the branches reaches a value that is referred to as the leaf nodes.

A common method of training DTs is through information theory and considering the entropy of the information gained at each split in the tree. This allows the best feature of the data to be identified, permitting a decision about that feature to split the data into branches. This process is repeated on each branch until the tree has been fully grown. The DT can be pruned by removing the lowest nodes and decisions. Pruning is normally undertaken if the DT is overtrained and does not generalise well to unseen data.

Fuzzy decision trees have been shown to have use in the diagnostics of the steam generators of pressurised water reactors by Zio (Zio, et al., 2009). This implementation built on a previous decision tree algorithm to better cope with overlapping data in the feature space a problem that can occur often in nuclear power diagnostics.

Mu has shown the use of decision trees in conjunction with fuzzy rough sets for nuclear power station diagnostics (Mu, et al., 2014). In this work a decision tree was applied to a nuclear power station simulator to confirm that faults are still able to be detected and diagnosed after the parameter reduction of the fuzzy rough sets method. This application was able to detect many faults in the data including loss of coolant

accidents, feed water pipe ruptures, steam generator tube ruptures and main steam pipe ruptures.

3.6.8 K-means

The k-means algorithm is a common type of clustering algorithm which is used to cluster unlabelled data into k (a number defined by a user's hyperparemeter) different groups. K-means shares similarities with the supervised K-nearest neighbour classifier algorithm. K-means clusters the data instances while minimising the within cluster variance. Upon the completion of the algorithm there will be "k" defined mean value locations and the data samples which are closest to these means are assigned the same grouping.

The k-means algorithm in combination with principal component auto-associative support vector regression (PCSVR) has been implemented by Seo for on-line calibration monitoring of nuclear power station data (Seo, et al., 2011). In this context k-means has been chosen as it is a quick clustering tool to train. This allowed the data to be subdivided into groups that are similar to allow for the better implementation of the PCSVR algorithm.

3.6.9 PCA

Principal component analysis (PCA) is a common method of dimensionality reduction, particularly for data visualisation. PCA transforms data into a different coordinate system to maximise the variance in the first coordinate of the data and with each successive feature having less variance. Once this is performed the features which capture the smaller amounts of variance in the data can be omitted in favour of the largest impact features. Many features within data can have high correlation between them indicating strong relationships which might not be easily recognisable by human analysis.

This can allow humans to visualise the data in comprehensible dimensional space with only a small loss of information within the data. Other uses are to reduce the computational complexity of algorithms by reducing the number of features allowing for faster training or allowing the use of the algorithms in real-time.

In the domain of nuclear condition monitoring PCA has been shown to be used in combination with artificial neural networks for fault detection by Hadad (Hadad, et al., 2008). In this case PCA was used to reduce the complexity of the inputs required for an ANN. Through doing this, more accurate networks were able to be developed to provide better fault detection for operators. PCA has also shown to be used to reduce the dimensionality of the pressure difference in nuclear steam generators as a method to create a simpler formulation of the effects (Girard, et al., 2013). The reduction in dimensionality provides simplified input and outputs, which has allowed for a much more convenient framework for future development and other diagnostic contexts.

3.7 Pre-Processing Steps

In machine learning an important process before the algorithms can be trained is pre-processing. This main reason for doing this is that the training data is not in an appropriate format for the specific machine learning algorithm. In this subsection two different pre-processing approaches are discussed: imbalanced data and feature extraction.

3.7.1 Imbalanced Data

In many machine learning problems, a big issue being faced is working with imbalanced data (He & Garxia, 2009). Imbalanced data describes the situation where one class of data is substantially numerous or lacking in data instances compared to the other classes of data in a learning problem. This is an issue as many machine learning algorithms have an inherent underlying assumption that they are operating on a balanced dataset. In many of these algorithms, minimising the absolute error is important and thus they favour the majority class. A well-known example of an imbalanced data set within the machine learning community is the mammography data set. This two-class dataset contains mammography data to determine if a patient has breast cancer, there are 10,923 instances of the negative class but there are only 260 instances of the positive class. Using this dataset as an example, if the learned algorithm simply classifies every output of as being negative the performance of the algorithm would be 97.7% accurate. This would be a very good score under normal circumstances, however with the knowledge that none of the patients who have cancer are getting correctly diagnosed, it demonstrates that this is solution is pointless.

This deceptive performance of accuracy can be avoided by considering the use of other metrics including recall and precision, which can provide context on the performance of specific classifications of data.

The most basic metric of performance of the algorithms on the test data is their average performance on the test set; this can be seen in Equation 12. True positives (TP) and true negatives (TN) are instances of correctly classified the presence or absence of the condition being tested for. False positives (FP) and false negatives are the opposite; they are instances which have misclassified the presence or absence of the condition being tested for. However, using accuracy alone is not a suitable measure of the performance of the algorithms due to there being a class imbalance problem. Therefore, it is necessary to consider both the recall, precision and the resulting F1 score to determine the most suitable classifier. In binary classification the recall value is important as it displays the mean percentage of the positive case in the testing set that is correctly detected and can be calculated as shown in Equation 13. Additionally, another important metric of performance is precision and can be calculated as shown in Equation 14. Precision is the mean percentage of the number of correctly identified positive cases compared to the total number of positive cases identified. This allows for a confidence in classification to be established and provides an insight into the likelihood into the probability of false positives.

Lastly, F1 Score can be considered to bring both precision and recall together and can be calculated from Equation 15. F1 score can be more generally described as the F-score when precision and recall are not weighted similarly to each other, this can be seen in Equation 16 where β is the weighting factor (β values greater than 1 indicate greater priority on recall with β values less than 1 indicate greater priority on precision).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad \textbf{Equation 12}$$

$$Recall = \frac{TP}{TP + FN}$$ **Equation 13**

$$Precsion = \frac{TP}{TP + FP}$$ **Equation 14**

$$F1\ Score = 2.\left(\frac{Precision.Recall}{Precision + Recall}\right)$$ **Equation 15**

$$F\ Score = (1 + \beta^2).\left(\frac{Precision.Recall}{\beta^2.Precision + Recall}\right)$$ **Equation 16**

There are many methods to solve the class imbalance problem. Firstly, in many circumstances there are opportunities to resample the data to provide a more balanced dataset. Random oversampling and random undersampling are relatively simple techniques that are used to balance the distribution of classes within asset of data. Random oversampling is a process where additional data is obtained for the minority class through randomly duplicating some of the data until the distribution of the classes is equal. The main drawback of random oversampling is that instances of data become overrepresented in the training data and as such this can cause the resultant algorithms to exhibit overfitting as opposed to algorithms which generalise well to future data. Random undersampling on the other hand randomly removes instances of the majority class until the distribution of the classes is equal. A better method of obtaining a balanced distribution of classes in a dataset through undersampling is through a method known as informed undersampling. Informed undersampling's objective is to reduce the number of samples from the majority class while retaining as much of the class information as possible.

Synthetic oversampling follows the same principal as random oversampling by providing more examples of the minority class. It differs however by creating new data that intends to be representative of the data. A popular example of this type of oversampling that has been chosen to be applied to this class imbalance problem is the

Synthetic Majority Oversampling Technique (SMOTE) (Chawla , et al., 2002). SMOTE synthetically generates the new minority data samples by considering the current samples in the feature space and randomly generates the data on the lines between two data samples.

Many learning algorithms can be altered under certain circumstances to perform better in situations of extreme class imbalance. One method of doing this is by applying different weight to the cost of classifying data instances, such as making the error contribution of misclassifying a minority class being much greater than the majority classes. Lastly another way to handle imbalanced data is to change the feature space that the algorithm is taught within to increase the size of the interclass margin between the classes.

Measures taken to ensure that data is balanced are essential for the graphite brick crack detection that is discussed further in chapter 5.

3.7.2 Feature Extraction

Feature extraction is an important part of data mining and machine learning, it is a method of dimensionality reduction. Features are often chosen to try and capture the important parts of the data and reducing the amount of redundancy that is seen within instances of data. Feature selection can often be performed from a knowledge-based approach where knowledge elicitation is undertaken by domain experts to characterise the features that they would use in manual analysis.

3.8 Evaluation of Algorithms

Within machine learning algorithms it is important that overtraining does not occur. Overtraining is a situation where the output of the algorithm does not generalise well to unseen data and can only perform correctly on the data it has seen before in training. It is for this reason that many machine learning algorithms separate the training data into training, testing and sometimes validation sets, allowing the performance of the algorithms to be assessed on data which should be representative of future unseen operating data. This is an example of using cross validation for training. K fold cross validation is the most common type of cross validation, which is performed by randomly splitting the training data into k different groups. The chosen algorithm is

then trained k number of times where Each iteration trains on all groups of data except one which is used to test the performance of the algorithm. The test set is then cycled through all the groups of data throughout the iterations. The performance of the algorithm is then evaluated on all the iterations and it is possible to obtain metrics such as the mean performance, the confidence bounds of the algorithm through the standard deviation of the performance and the maximum/minimum performance of the algorithm. An extreme version of k fold cross validation is when k is equal to the total number of training data instances, at this point it is known as leave one out cross fold validation. This provides a good measure of the performance on every instance of data however the computational costs of this method often make this infeasible for large datasets. The last common type of cross validation is Monte Carlo cross validation. This method involves randomly splitting the data into a test and training set and evaluating the performance, which is then repeated many times to obtain metrics about performance. The benefits of this method are that the split of the data is no longer linked to the number of iterations that are to be performed. The downsides are that there are chances that some data instances will never be used in the training or testing set and thus the performance might not be truly representative of the data.

3.9 Conclusions

Machine learning and data analytics are invaluable tools for condition monitoring and fault detection. Through a study of the literature it has been shown that various types of machine learning approaches and techniques have been applied to a wide variety of nuclear power station problems.

Machine learning can easily be split into three different categories based on learning. Reinforcement learning involves creating a learner than can interact with an aspect of its environment, with a target reward that aims to be optimised. Supervised learning involves learning a relationship between input data and associated labels. Unsupervised learning involves obtaining knowledge from a supply of data with no associated labels. Semi-supervised learning is a learning technique that is most commonly associated with supervised learning but borrows from the learning techniques of both supervised and unsupervised learning. Semi-supervised machine learning is still in its infancy stage with regards to nuclear condition monitoring.

However there have been some examples recently which have incorporated it for the purposes of fault detection and classification.

Through this study several techniques have been identified as being applicable to solving the problems associated with the UK's AGRs. The first problem being tackled is being able to investigate and model the relationship between FGLT through analysis and regression. Regression learning techniques have been shown to learn the relationship between inputs and outputs within machine learning.

Secondly detecting cracks in graphite bricks from FGLT data by comparing the performance of standard supervised and semi-supervised learning methods. Semi-supervised machine learning excels in situations where there are large amounts of unlabelled data but limited labelled data available. This situation occurs within graphite brick crack detection as there is limited ground truth data available but much more frequent monitoring data. Ground truth inspection data is costly to obtain as the reactor needs to offline, whereas the unlabelled data can be obtained while the reactor is a low power. Semi-supervised machine learning therefore is an appropriate approach to address the problem. These two problems are covered in much greater detail in the following chapters.

4 Estimating Fuel Channel Bore from Fuel Grab Load Trace Data

4.1 Introduction

There are two ways to obtain information about the condition of the fuel channels inside the graphite core. The first method is through inspection, which occurs every 18 months to 3 years during scheduled outages. This physical inspection is performed either by a Channel Bore Inspection Unit (CBIU) or a New In Core Inspection Equipment Mk 2 (NICIE 2) (Cole-Baker & Reed, 2005). The second method is through monitoring data obtained every 6-8 weeks during refuelling events, called the Fuel Grab Load Trace (FGLT) (West, et al., 2006). This data measures the mechanical load of the fuel stringer as it is being removed (discharge) and inserted (charge) into the core. FGLTs are required to be recorded and stored by the station for safety purposes, as they can be used to detect faults occurring during the refuelling process (West, et al., 2007). However, this data can also be exploited to provide health information as fuel channel dimensional change can be observed in the FGLT. Engineers examine the FGLT and look for evidence of anomalous behaviour, such as step changes in load between the brick interface peaks, the presence of additional unexpected peaks or changes in the magnitude of the interface peaks. Developing a model that robustly and accurately estimates fuel channel bore when provided with FGLT data would provide operators with additional supporting evidence concerning the health of the core.

This chapter will discuss a model that has been created which allows accurate estimations of the dimensions of the fuel channel to be obtained from FGLT data. One of the key benefits of this method is that FGLT data is already routinely recorded and stored. Therefore, no additional equipment or expertise is required to obtain this data. Additionally, the older recordings of FGLT data that were recorded on paper have been

digitised and are made available for analysis. The remainder of this chapter proceeds to explain the different components of FGLT data that produce the load value. The process of developing a model is then shown, including testing on a blind dataset. The dataset for testing was obtained from operational data from a nuclear power station in the UK.

4.2 Signal Decomposition

Within the signal processing domain there are numerous different techniques that are used to deconstruct signals. Some examples of this involve frequency domain analysis, wavelet transforms and singular spectrum analysis.

The most common approach to frequency domain analysis is performed with the Fourier transform. The Fourier transform decomposes a time signal into its constituent frequencies, allowing knowledge to be obtained about the frequencies that are present and filters can be designed that remove unwanted components. Benbouzid describes a classical approach of using a fast Fourier transform for detecting faults in induction machines in their review (Benbouzid, 2000). Nandi discusses the ability for frequency domain analysis of vibration signals for electrical motors in their review (Nandi, et al., 2005).

An issue that is encountered with frequency domain analysis is that aspects of the time domain are lost when only considering the frequency domain. This is addressed by performing time-frequency analysis instead, such as the short-time Fourier transform or wavelet transforms. Peng and Kunpeng have provided reviews of the applications of the wave transform for condition monitoring and fault diagnostics (Peng & Chu, 2004) and tool condition monitoring respectively (Kunpeng, et al., 2009). This method of signal deconstruction can be very useful for situations where certain signals change frequency over time.

Another example of signal decomposition is singular spectrum analysis (SSA), a method of time series analysis (Golyandina, et al., 2001). SSA splits a signal into a sum of its components which can be used for extracting trends in the data at different resolutions, smoothing, extraction of periodic components and extracting structure from data (Golyandina & Zhigljavsky, 2013). Kilundu has demonstrated the effectiveness of using SSA along with a bandpass filter to extract features that are used

to identify tool wear (Kilundu, et al., 2011). A method of fault diagnosis of roller element bearings using SSA has been implemented by Murugabatgan (Murugabatgan, et al., 2013).

The methods described so far could have success in isolating the different components but with some difficulties. The data that is recorded is independent of time and the sample rate that records the load value can vary from trace to trace. This would make it difficult to identify the different components based on their frequency.

The chosen method for the extraction of the desired signal is a knowledge-based approach as there is a good understanding of the underlying components within the FGLT data. Combining the knowledge of the fundamental design of AGR reactors, including the design of the fuel stringers, with the data has allowed for the extraction of key components. An understanding of the components of FGLTs is required to do this. Firstly, this is achieved by manually analysing the different contributions of load value, then when sufficiently understood, the process can be automated to allow for large scale signal deconstruction of FGLTs.

4.3 Four Components Analysis

It is understood that there are four main components that contribute to the load value of the FGLT (West, et al., 2014); this can be seen in Equation 4.

$$L(h) = K_L\Phi(h) + K_U\Phi(h+d) + F_G(h) + m \qquad \textbf{Equation 4}$$

Where L is the measured load at height h. K_L is a parameter which relates the friction to load caused by the lower stabilising brushes (LSB) due to the channel bore (Φ) at height h. K_U is a parameter which relates the friction caused by the upper stabilising brush (USB) to load due to the channel bore (Φ) at height h offset by d, the difference in height values between the LSB and USB. F_G is the summation of any aerodynamic effects to the load value at height h and m is the mass of the fuel stringer, also known as the deadweight, which is constant. When attempting to recreate an accurate bore measurement of the fuel channel it is important to understand the contributing components and to deconstruct the load into these four components. The ability to isolate the friction component from the LSB from the other components is the goal as

it contains a relationship of load to bore. The next sections will go into detail on the process of investigation for each of the four components and what their contributions are to the FGLT.

4.3.1 Deadweight Analysis

The deadweight of the fuel stringer is the largest contributing load component within the FGLT. It corresponds to the mechanical load that the weight of the fuel stringer exerts on the refuelling machine during the refuelling process. An assumption that has been made about the weight of the fuel stringer is that over the course of a single FGLT it will have a constant value. However, there is no assumption that the deadweight is the same for all traces, even paired charge and discharge traces may have differing deadweight values. Therefore, if the value of the fuel stringer can be calculated at any point of a single trace, it can be removed from that entire trace.

As mentioned in chapter 2, there is a region in the FGLT where both the LSB and the USB are in sections of the refuelling machine which is sufficiently wide that they are not in contact with the brushes. It is therefore assumed that within this region of the FGLT the recorded load value should be representative of only the deadweight and the aerodynamic effects of the carbon dioxide coolant gas. When considering exclusively the refuelling events that occur at offload inspections, the assumption is that there is only a minimal effect from any coolant gas. In this region an assumption is made that only the deadweight contribution in present.

The deadweight region of the FGLT occurs when the lower stabilising brush region of the fuel stringer is in the area between the top graphite brick and the guide tube. This region can be identified on the discharge trace as a constant load value area located between two large load peaks, after the LSB has been removed from the topmost brick layer. Meanwhile this same region can be identified in the charge trace as a flat plateau region which occurs before a large load value trough, both of which occur prior to the LSB entering the topmost brick layer of the core.

An example of the deadweight region of the discharge and charge trace can be seen in Figure 4.1 and Figure 4.2 respectively, with the deadweight region identified. The load values are shown in mV which can be directly related to the mass that is being exerted

on the load recording machine. The sample numbers indicate the frequency of the load recordings in this section with every sample being an individual recording.

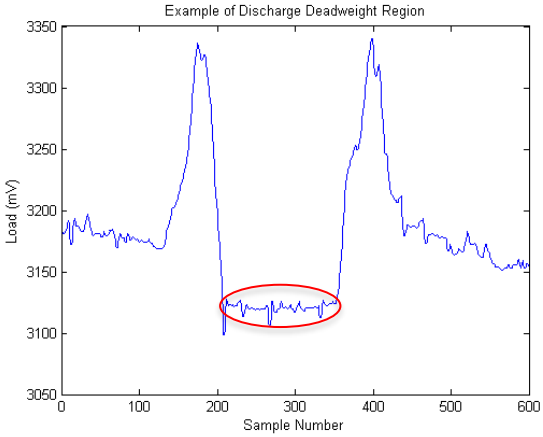


Figure 4.1 - Example of the discharge deadweight region identified on a FGLT.

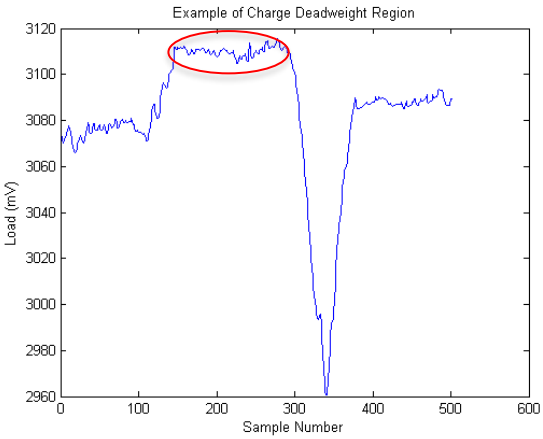


Figure 4.2 - Example of the charge deadweight region identified on a FGLT

These points can be identified by selecting the end point of the topmost brick layer in the trace and then extracting the region physically above it. In the case of the discharge trace this occurs after the last graphite brick and in the case of the charge brick it occurs prior to entering the topmost brick. With the sampling rate of the load measuring device and the speed of the fuel stringer, it requires approximately 600 samples to capture the entire double peak feature. By reducing the search region to a smaller size, it is possible to identify key points that can be explored. These have been indicated in Figure 4.3 for the discharge and Figure 4.4 charge traces. These points are characterised by turning points in the traces, but these are hard to identify due to the data being noisy. Therefore, before any changing gradient detection can be performed the data must be denoised by smoothing with a moving average/mean filter.

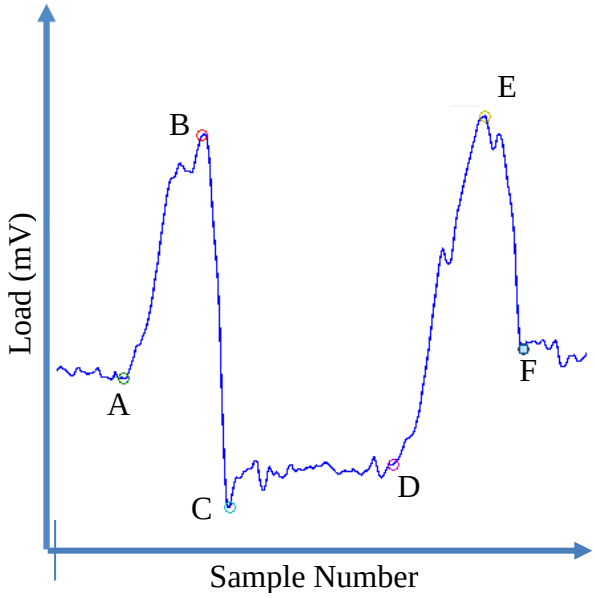


Figure 4.3 – Example of identifying the key points within the discharge deadweight region.

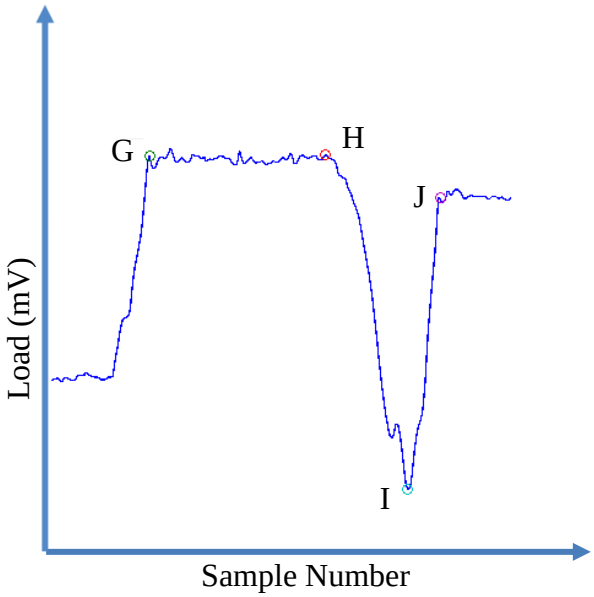


Figure 4.4 – Example of identifying the key points within the charge deadweight region.

In the case of discharge traces initially points B and E are identified by finding the largest points where there is a gradient change that are sufficiently separated from each other. By identifying the two peaks successfully this creates a window in which to search for the flat region of the trace. Again, gradient change point analysis is used to work from B and E to identify points C and D, which is the suspension region of the discharge trace.

In the case of the charge trace initially point I is identified by finding the lowest minimum value within the range. Points H and J can be identified by searching for the rapid changes in gradient either side of I. Point G can then be identified by searching backwards from point H for the data point after a prolonged positive gradient. The suspension region of the charge trace is found between points G and H.

Once the suspension regions have been identified they can have their mean values calculated. This was performed for all recorded traces for all the reactors during the years 2004-2015. The analysis was performed by identifying the double peak region in both the discharge and charge trace and extracting those regions. The results are

shown through Figure 4.5 to Figure 4.8 where the onload and offload data has been identified with different colours in the histogram. The purpose of doing this is to determine the variation between different fuel stringer's deadweights.

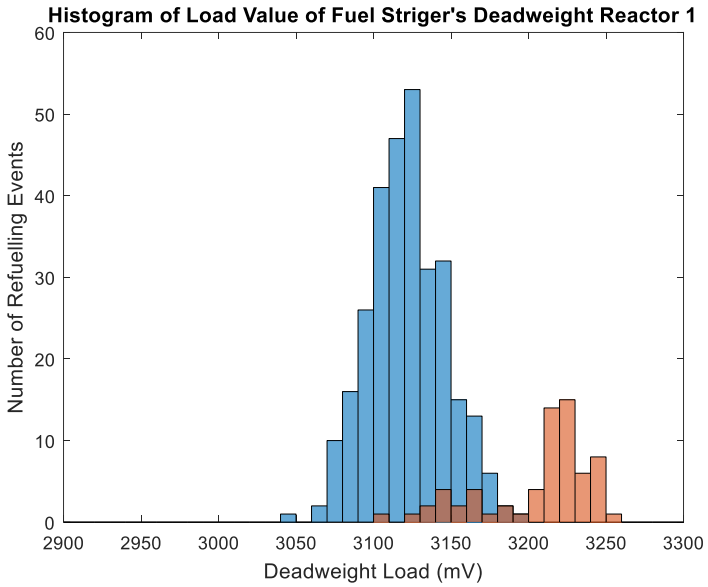


Figure 4.5 – Histogram showing the load value of the fuel stringer's deadweight for reactor 1.

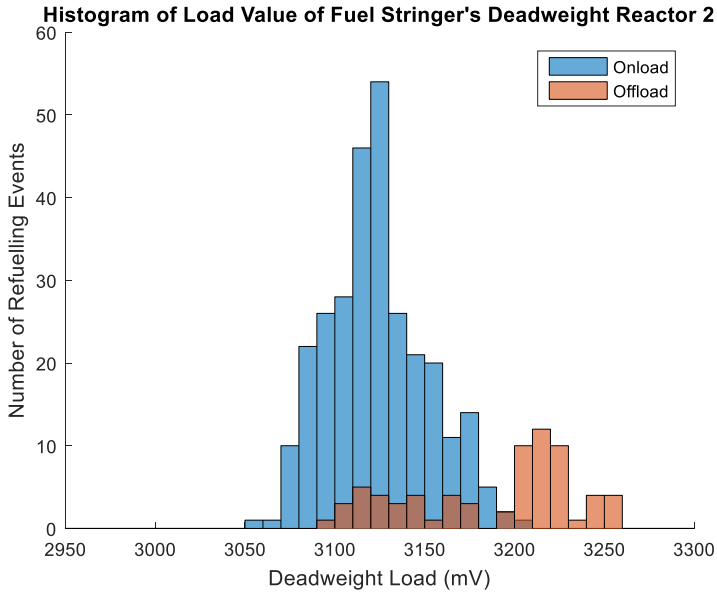


Figure 4.6 – Histogram showing the load value of the fuel stringer's deadweight for reactor 2.

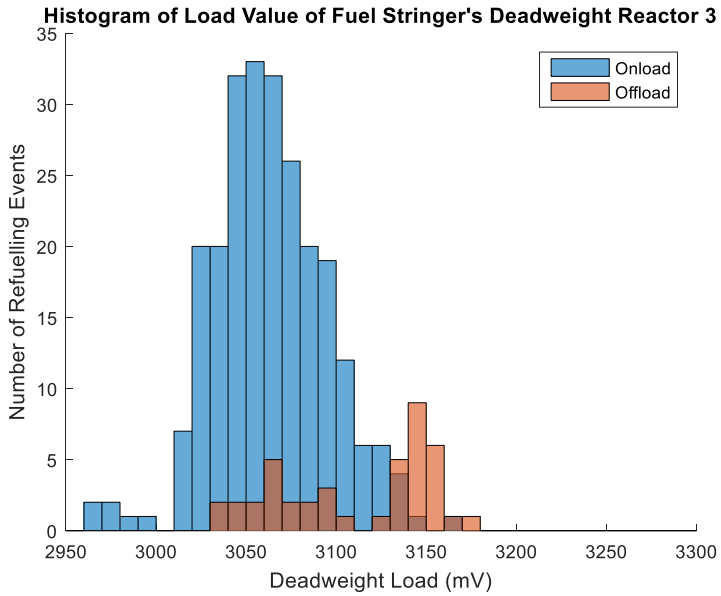


Figure 4.7 – Histogram showing the load value of the fuel stringer's deadweight for reactor 3.

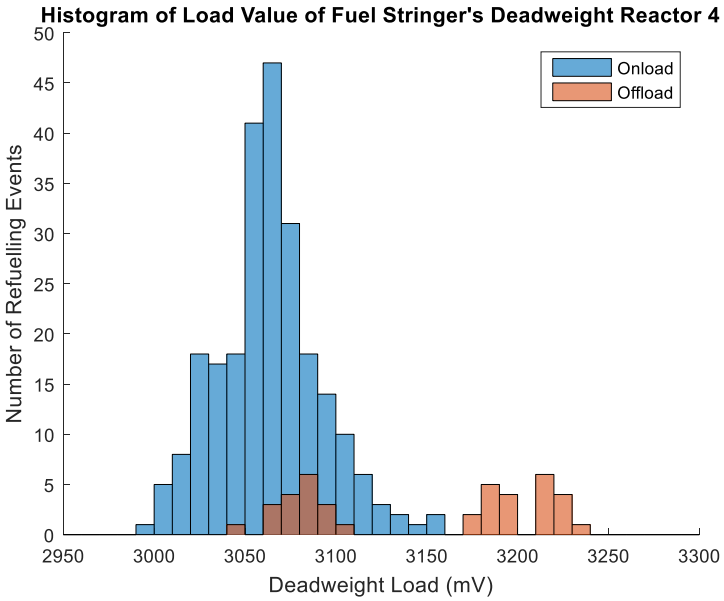


Figure 4.8 – Histogram showing the load value of the fuel stringer's deadweight for reactor 4.

It can be seen in Figure 4.5 through Figure 4.8 that there is a distinct distribution of higher load deadweight values. This confirms from a data analysis viewpoint that there

are differences in the load values between onload refuelling instances and offload refuelling instances. This was predicted as it is known that in this region of onload traces there should be contributions from aerodynamic gas effects. In the offload traces this should not be the case. However, in addition to the clearly distinct higher load value traces, there is a subset of the offload traces that have very similar values to the onload traces. This is attributed to the pressurised offload traces that also contain aerodynamic gas effects. Therefore, knowing this and the fact that the weight of the fuel stringer in terms of load can vary, it is important to be able to repeatedly identify the deadweight region, to understand its contributing value. Understanding the underlying aerodynamic effects better can provide an indication on the variation in the deadweight region.

4.3.2 Aerodynamic Analysis

The aerodynamic effects that are present in the FGLT are difficult to quantify since they are present and variable throughout the entire refuelling event. This is due to the unpredictable nature of the coolant gas flowing through the fuel assembly. At the beginning of the reactors lives the gas effects were modelled, however now that they are approaching the end of their initial design lives or have surpassed them the structure of the core is different. The original aerodynamic gas flow models for low power refuelling offload pressurised and offload depressurised refuelling can be seen in Figure 4.9.

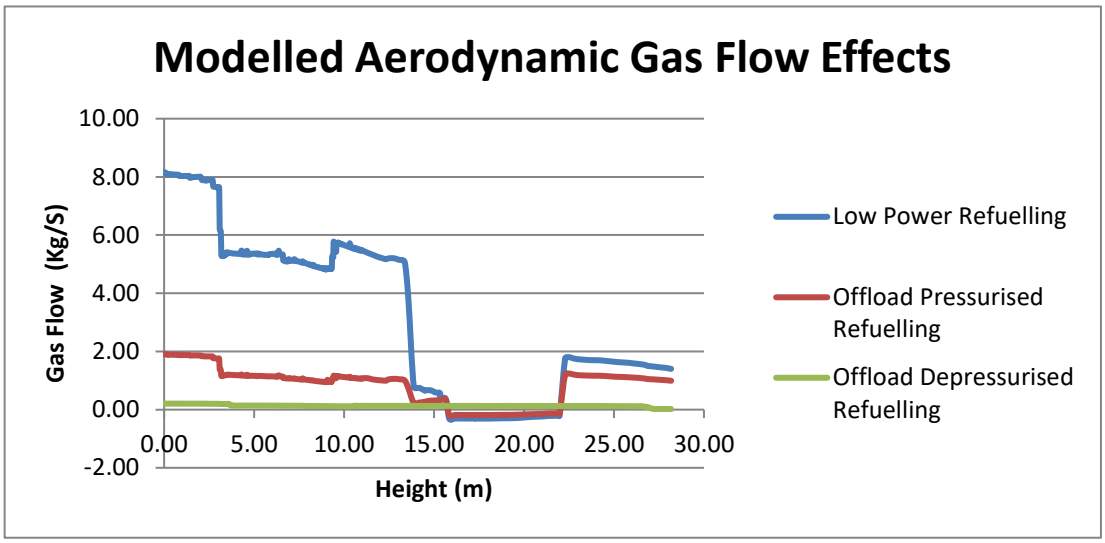


Figure 4.9 - Model of aerodynamic gas flow effects including low power refuelling, offload pressurised refuelling and offload depressurised refuelling.

The previous assumptions were, that during offline refuelling event there is only a minimal effect from any coolant gas and as such does not need to be accounted for. To analyse the contribution from aerodynamic effects both the discharge and charge traces are required.

$$L(h) = K_L\Phi(h) + K_U\Phi(h+d) - F_G(h) + m \qquad \textbf{Equation 17}$$

$$L(h) = -K_L\Phi(h) - K_U\Phi(h+d) - F_G(h) + m \qquad \textbf{Equation 18}$$

Based on the four-component model equation, shown in Equation 17 for discharge and Equation 18 for charge, by removing the deadweight from both traces and subtracting them from each other results in only the aerodynamic effects being present. An assumption that is made here is that the frictional response from the discharge and charge traces are equal but opposite contributions. Another assumption that is made is that the core conditions producing the aerodynamic effects are identical on both traces.

When applying this to some onload traces a selection of the results of the residual gas are shown in Figure 4.10 through Figure 4.12. This is done by observing the difference

between the charge and the discharge traces that have been produced by the same refuelling event. If the assumptions hold true, then the resulting trace should be twice the aero dynamic contribution. Halving this value should then provide the aerodynamic contributions over the course of a single FGLT. Although the shapes produced are very similar, they experience some variation from different channels.

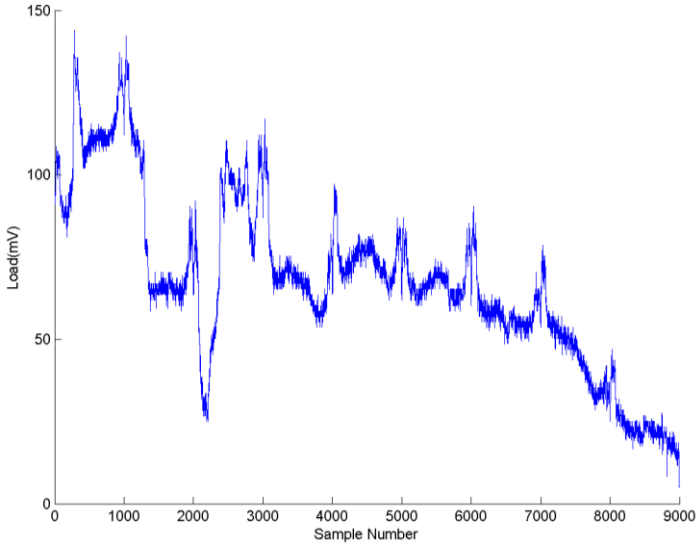


Figure 4.10 – An example of the aerodynamic contributions within the FGLT produce by taking the residual of the charge and discharge traces.

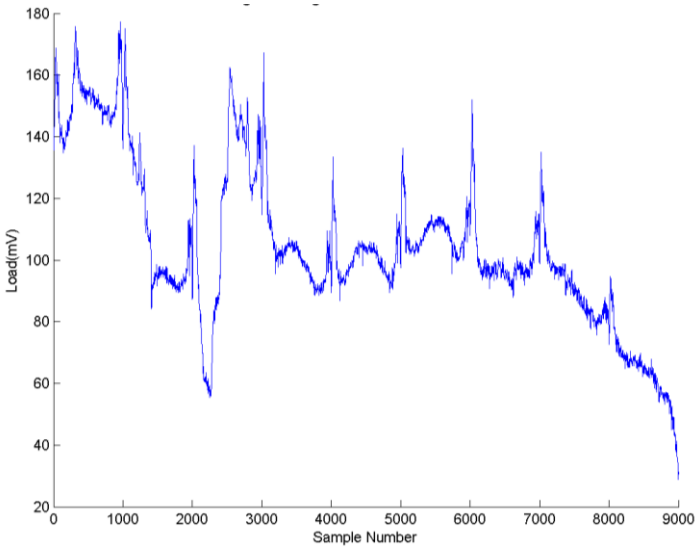


Figure 4.11 – An example of the aerodynamic contributions within the FGLT produce by taking the residual of the charge and discharge traces.

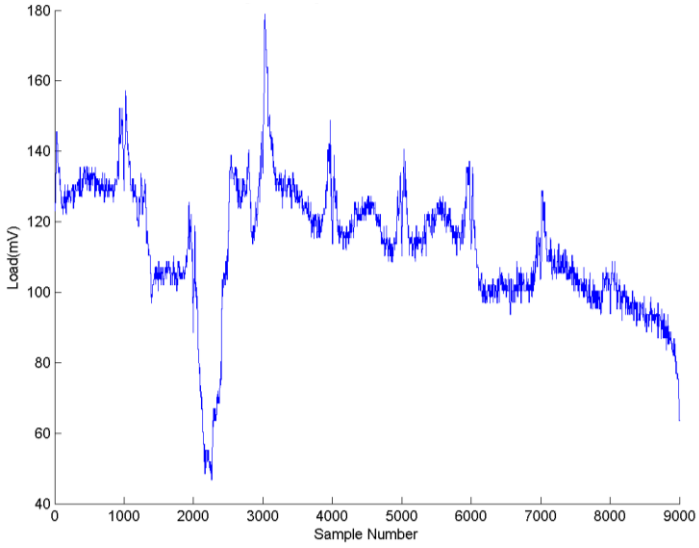


Figure 4.12 – An example of the aerodynamic contributions within the FGLT produce by taking the residual of the charge and discharge traces.

Another method used to analysis the aerodynamic force was to investigate the differences between onload and offload refuelling events on the same channel that

were performed within a short time period. It is assumed that if the time between the two FGLTs is small enough, then the bore measurements should be the same and that the deadweight of the fuel stringer should be the same. Therefore, the difference between the load values from these two traces should only be from differences in aerodynamic forces. From accessing the FGLT database, five of the channels were identified to have had both onload and offload refuelling events within two years. For these five channels the difference between onload and offload traces was calculated for both charge and discharge. This can be seen in Figure 4.13 for charge and Figure 4.14 for discharge where all five differences are on the same figure. The differences have been normalised around the double peak feature value where they were set to the same value for both onload and off load traces. It can be seen in both the charge and discharge traces that they are quite similar in value but even more similar in shape. There are two distinct regions which are two different contributions of aerodynamic forces. At the low sample numbers is the in core aerodynamic forces, there is then a transient effect and then the higher sample numbers are when the refuelling machine has become part of the core.

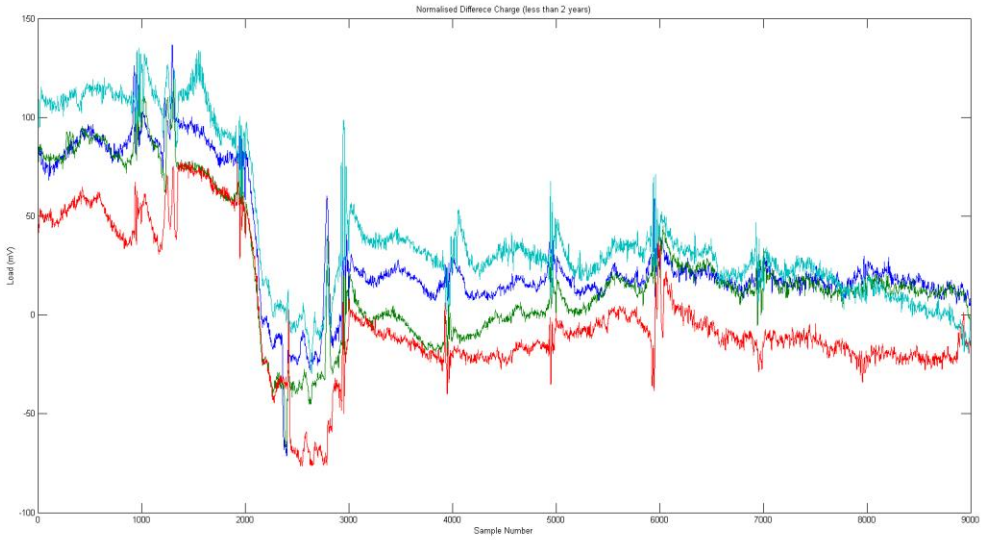


Figure 4.13 - The normalised aerodynamic contribution obtained from the difference between onload and offload charge FGLTs.

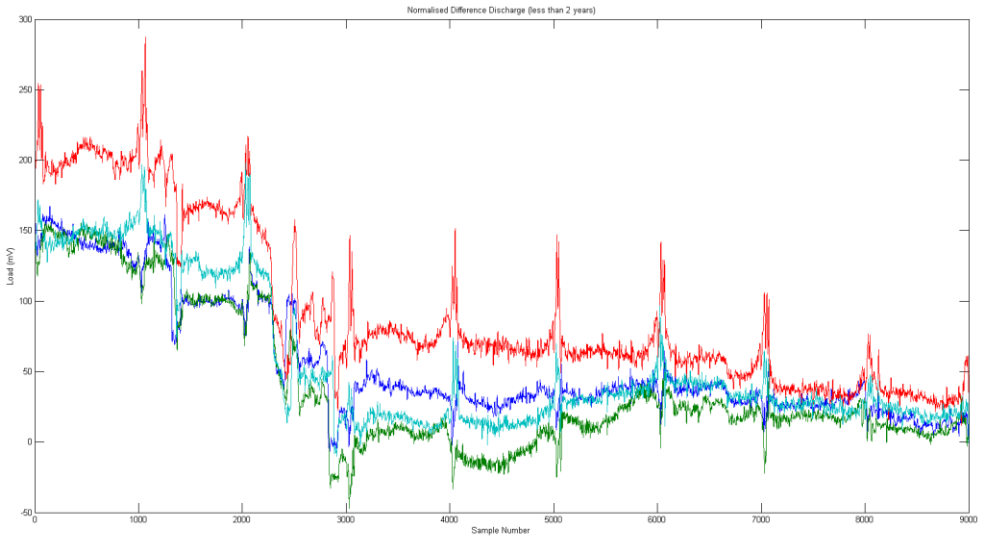


Figure 4.14 - The normalised aerodynamic contribution obtained from the difference between onload and offload discharge FGLTs.

These results are similar in shape to the gas flow profiles that were modelled over their core regions. However, the load contribution varies depending on the trace being investigated. This variation makes it very difficult to estimate what the aerodynamic contributions are for an individual trace.

4.3.3 Upper Stabilising Brush Contribution

The USB contribution is produced from the steel brushes that are located on the upper region of the fuel stringer. There are two regions of importance in the FGLT regarding the USB contribution. In the context of the discharge trace, the first region occurs when the fuel stringer is first lifted until the LSB enters the fourth brick layer. The second region occurs when the LSB is in the middle of brick layer 5 and corresponds with the USB passing the piston seal bore. These two regions have been identified on both a discharge and charge trace in Figure 4.15 and Figure 4.16 respectively.

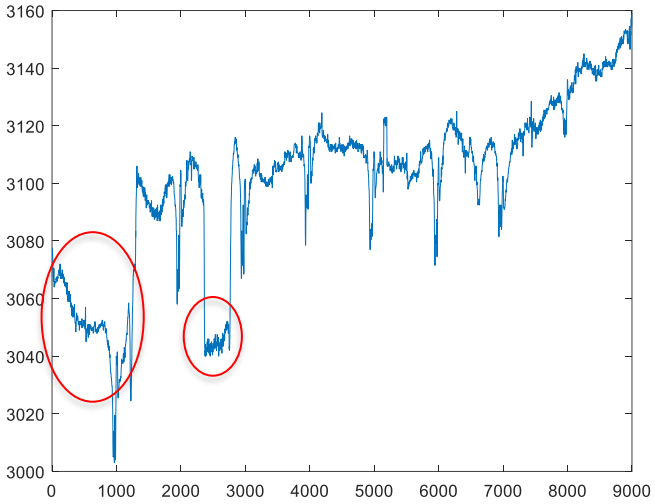


Figure 4.15 – An example of a charge trace with the upper stabilising brush regions identified.

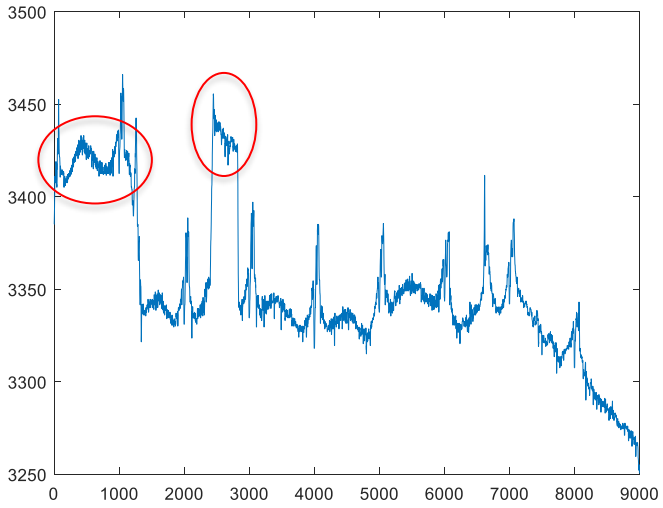


Figure 4.16 - An example of a discharge trace with the upper stabilising brush regions identified.

The main goal of the data exploration in this section was to understand the properties of the USB friction and to determine what its effects are on the load trace. The simplest

way of doing this was to investigate the regions where there is USB offset and determine the consistency between the other traces. This was first performed on layer 5 for both onload traces and offload traces. The layer 5 regions for both onload and offload traces are shown in Figure 4.17. In the offload trace the USB interaction can be seen clearly as an offset of C from the rest of the brick layers load values. This is not the case for onload, at this point there is a transient effect occurring where the aerodynamic forces change as essentially at this point the refuelling machine is becoming part of the core and matching the coolant gas flow. This can be seen in the onload trace where there is more of a stepwise increase in the load value in the equivalent region. The values of C, D and E were calculated for each of the onload traces and the C value was calculated for the offload traces. It was found that the values of E in the onload and C in the offload traces were consistently showing the same values. This can be seen in Figure 4.18, where the traces have been plot in order of newest traces to oldest with the red values being the offload C values and the blue values being the onload E values.

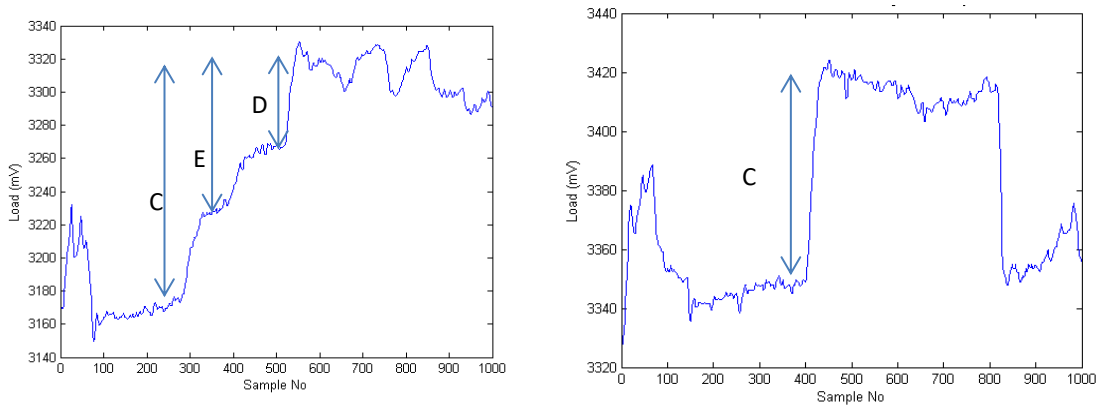


Figure 4.17 - Example of upper stabilising brush region in layer 5 for onload (left) and offload (right) discharge traces and respective features which were extracted.

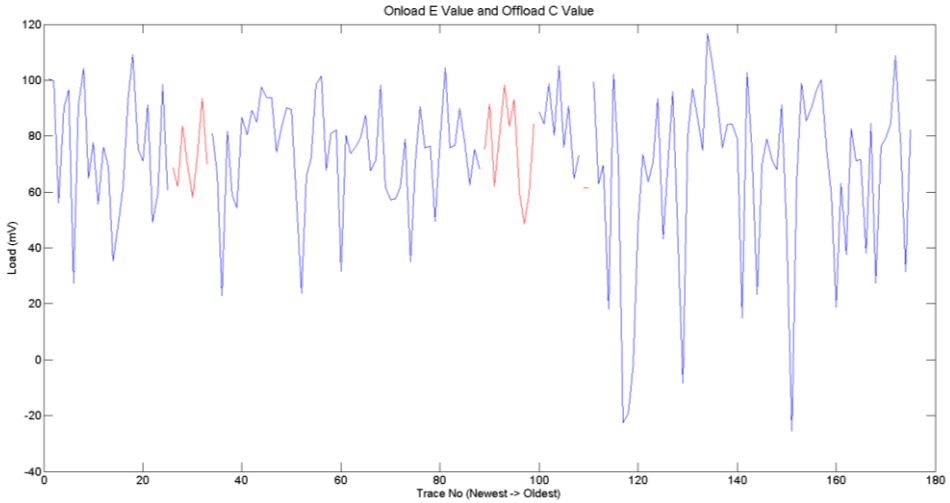


Figure 4.18 - Comparison of magnitude of layer five USB region for all traces where blue values are onload traces and red values are offload traces.

There are some missing values for the traces which are shown as gaps at some trace numbers. The method to obtain the values from the traces was not as accurate for some of the older traces and for the case of some onload traces an E value is not calculable as the two steps are combined and the onload D value is then detected instead.

This USB interaction produces a step change which is mainly located between the values of 60-100mV. As this step change is quite consistent it is possible to remove it from layer 5. The inaccuracies from determining the correct change in load can also be seen on the right side of Figure 4.18 where it the older traces from as early as 2002 are.

In comparison to the LSB, the USB has a smaller diameter and is at the upper end of the fuel stringer. When analysing the effects of the upper stabilising brush it is important to note that that the load value is referenced with respect to the location of the LSB. The USB is roughly 6 metres further up the channel; this means that any contribution from it will not have any relevance to the current brick. Another important consideration about the USB is the maximum size of bore that is possible to be detected is roughly 250mm as this is the diameter that the brush extends. This combined with the fact that only brick layers 3 through 11 are being modelled means that there are only a few regions that there are USB friction forces present. The first region it is

present in is from the start of the trace while the USB is in the guide tube, which has a diameter of 248.2 mm, until near the start of brick layer 4 where the guide tube widens. The second region is in brick layer 5 where the USB enters an area of restriction from the piston seal bore with diameter 247.65 mm. There is sometimes one other region with a contribution from the USB which can occur midway through layer 9. This is due to the USB "hunting" for the entrance to the stand pipe which has a diameter of 250.4mm. This last contribution from the USB is not always present as the brushes may not contact it at all.

The USB contributions have been segregated into the two individual contributions, the guide tube and piston seal bore interactions. Additionally, they have been separated based on the reactor the data has been obtained from. The results can be seen in Figure 4.19 to Figure 4.27. It is important to note that when generating these histograms some of the older traces which were digitised from paper copies have been too noisy to correctly identify and as such have been omitted.

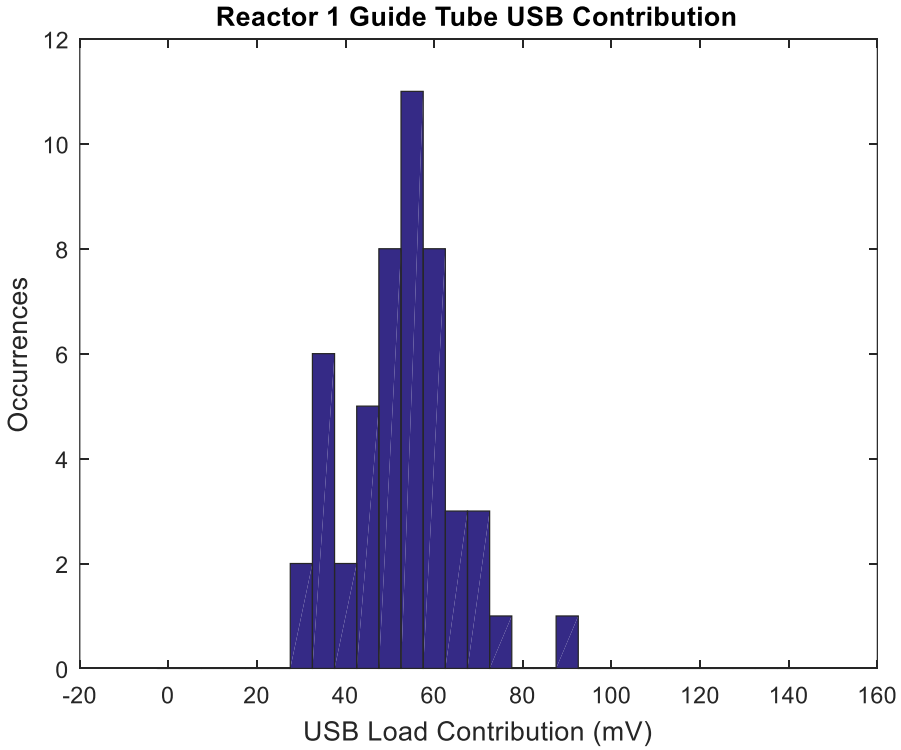


Figure 4.19 - Histogram of the guide tube USB contribution for reactor 1.

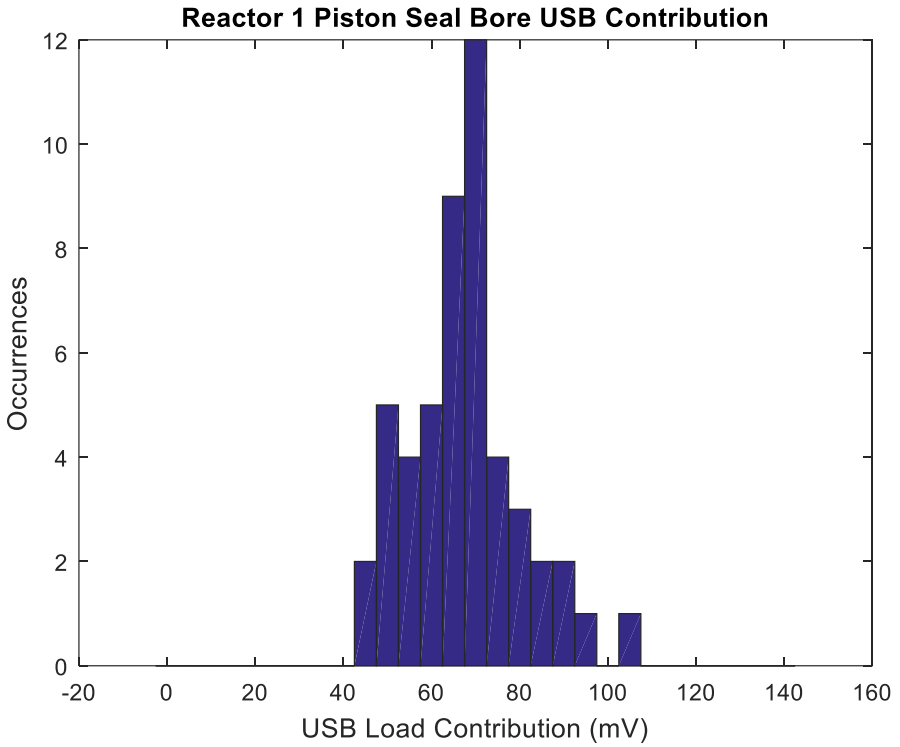


Figure 4.20 – Histogram of the piston seal bore USB contribution for reactor 1

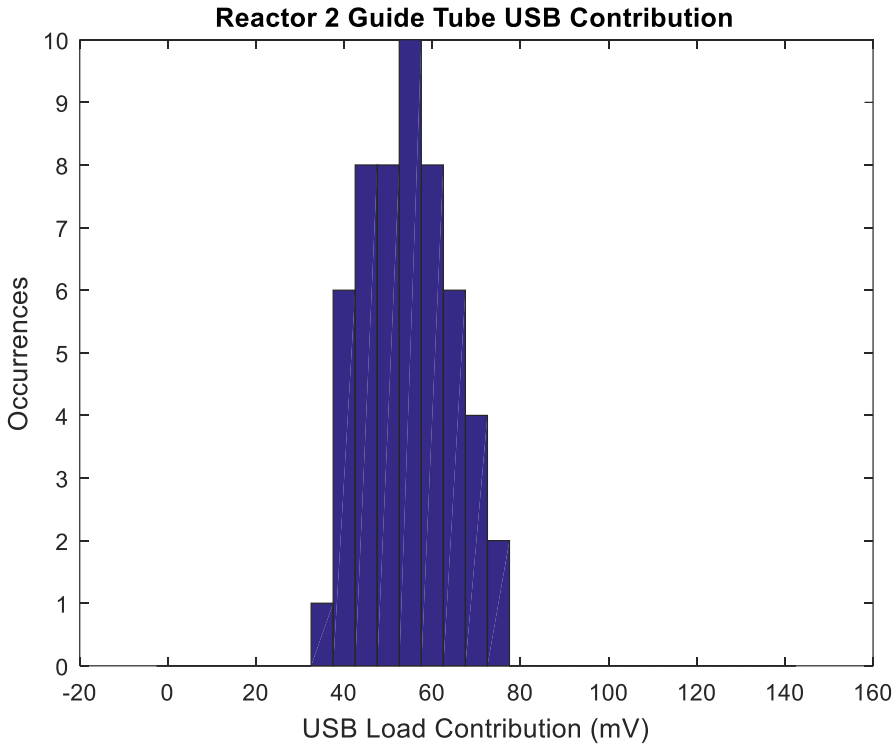


Figure 4.21– Histogram of the guide tube USB contribution for reactor 2.

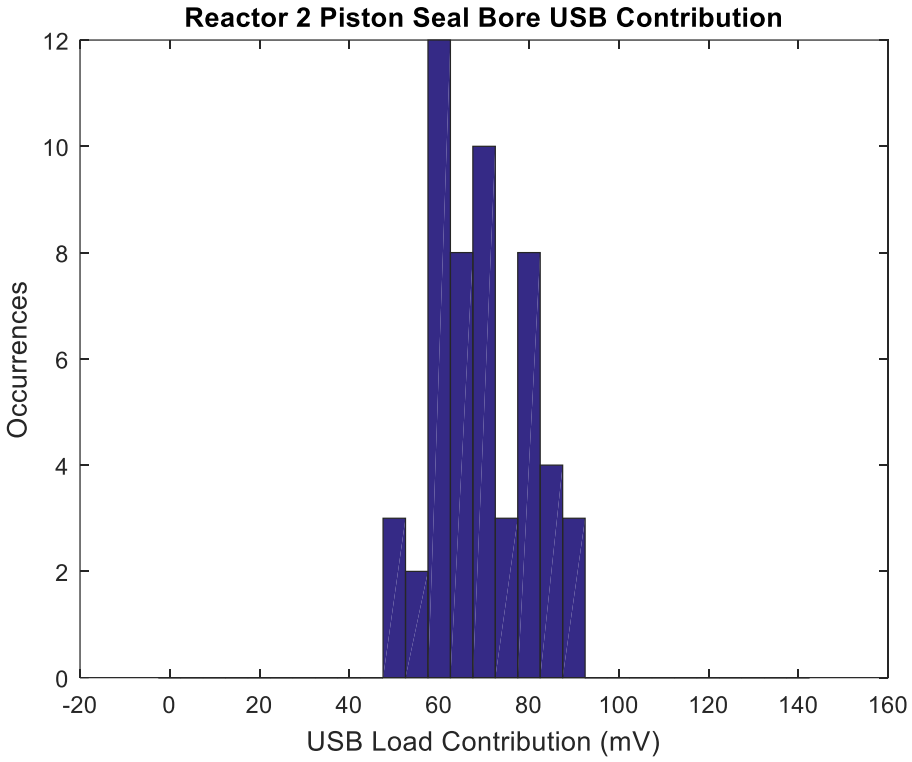


Figure 4.22– Histogram of the piston seal bore USB contribution for reactor 2.

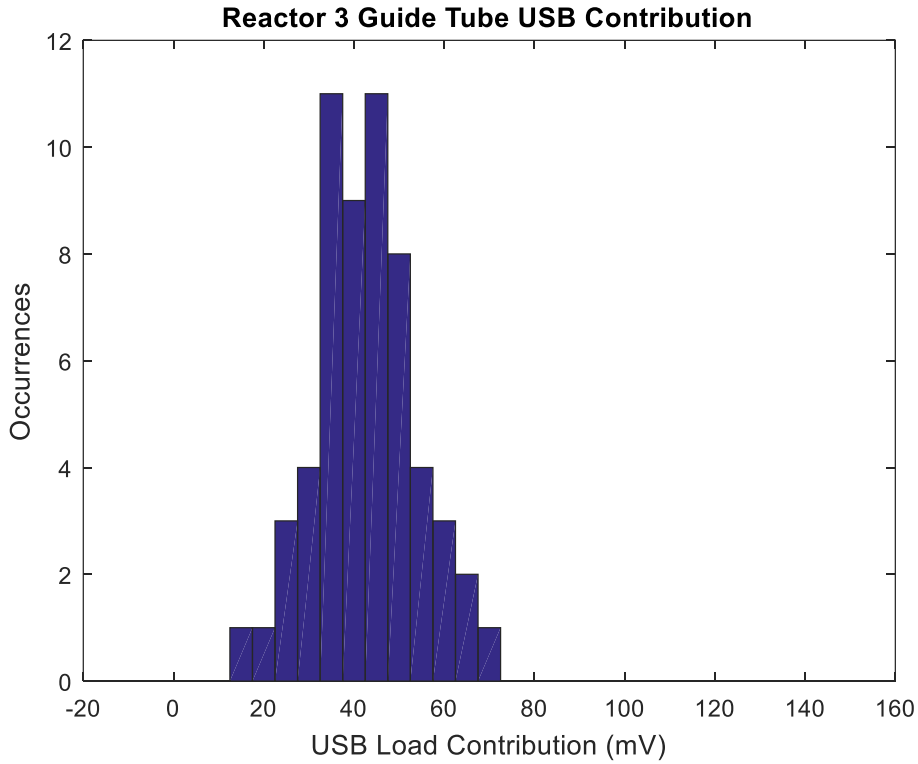


Figure 4.23– Histogram of the guide tube USB contribution for reactor 3.

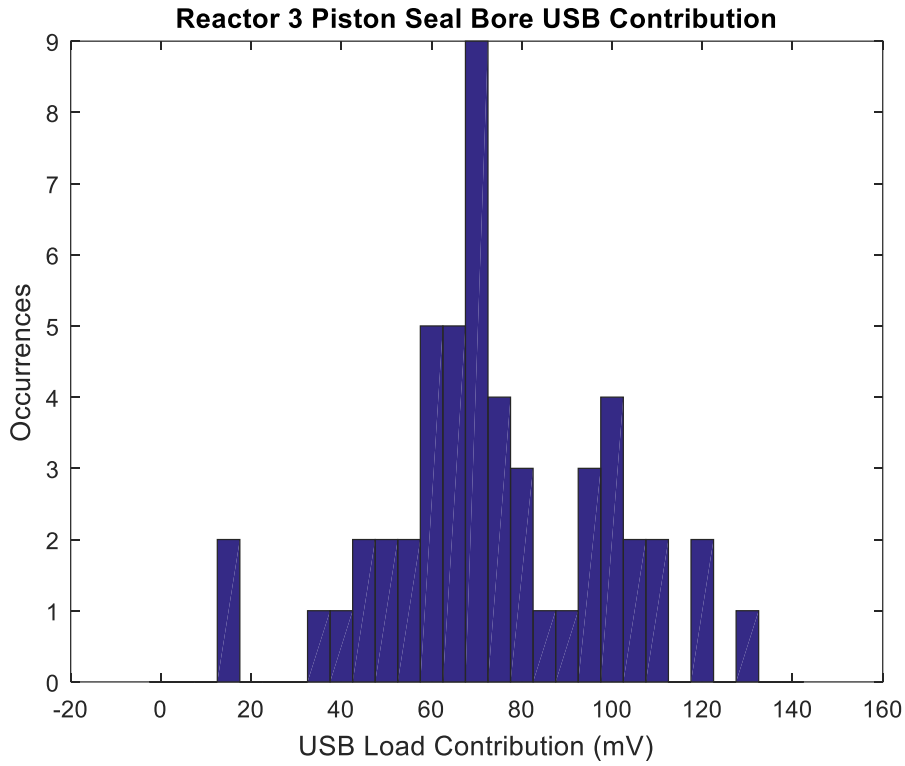


Figure 4.24 – Histogram of the piston seal bore USB contribution for reactor 3.

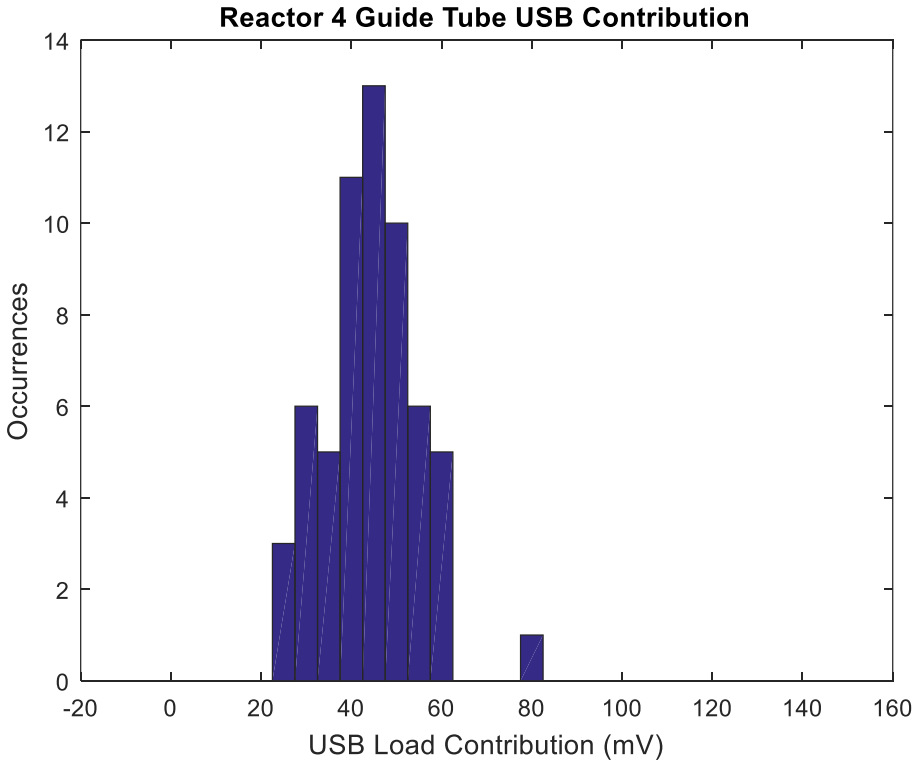


Figure 4.25– Histogram of the guide tube USB contribution for reactor 4.

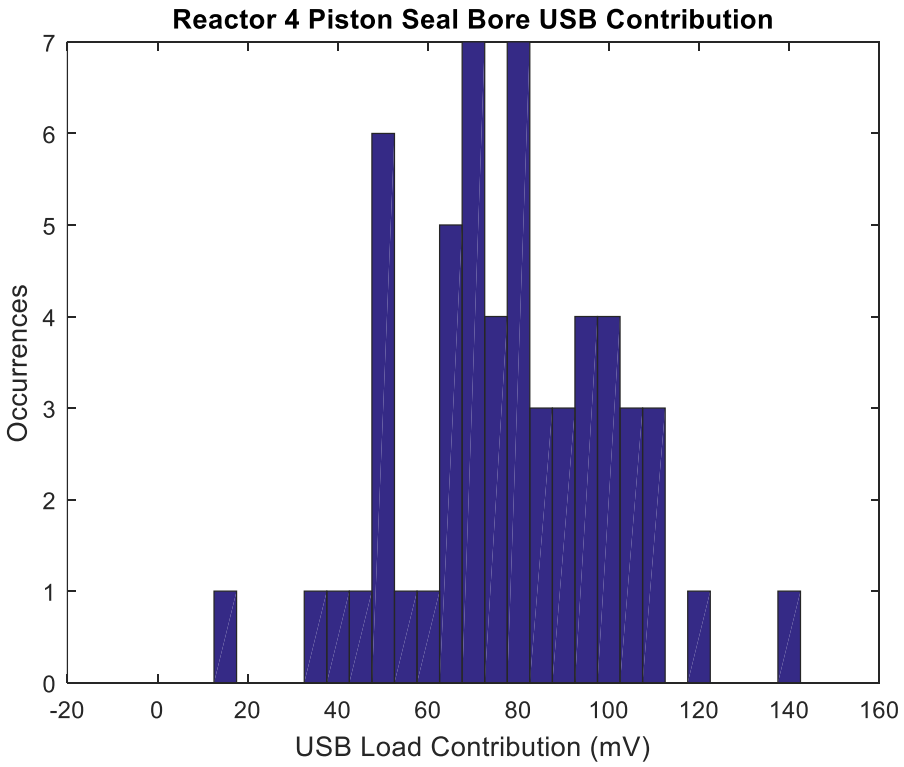


Figure 4.26– Histogram of the piston seal bore USB contribution for reactor 4.

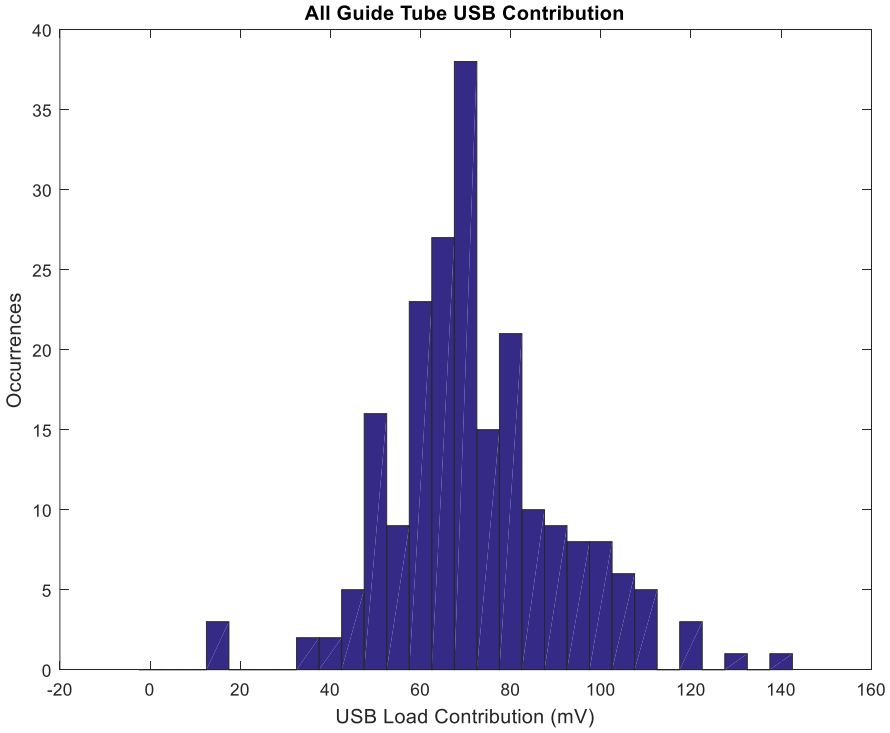
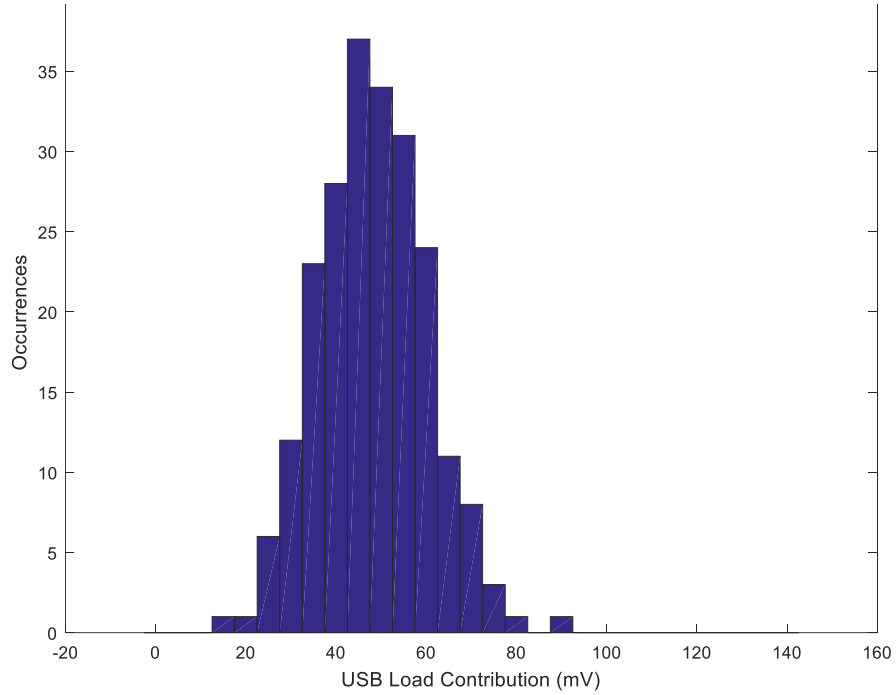


Figure 4.27– Histogram of the guide tube USB contribution for all reactors.

Figure 4.28 Histogram of the piston seal bore USB contribution for all reactors.



From these results, contributions from the USB are shown to have a range in values and appear to have a normal distribution associated with them. The variance from the USB contributions for both the guide tube (in brick layers 3 and 4) and for the piston seal bore (in brick layer 5), is too great to choose a single pre-set value to remove. Therefore, it is much more effective to detect the contributions on a trace by trace basis to remove them. This will result in a much more accurate value to remove from the trace in those regions.

It was expected that the offset values for the USB contributions would have consistent values over all the traces due to them having consistent dimensions within the reactor. It is unexpected that their dimensions would change over time as they should not be experiencing degradation as they are outside of the reactor and are made of steel. It is thought that the reason that there is some variance due to the different quality of the brushes that have resided within the core resulting in some changes in the friction they produce.

4.3.4 Lower Stabilising Brush Contribution

The lower stabilising brush contribution to the FGLT value is the most important contribution as it is directly related to the fuel channel bore diameter. The LSB consists

of three sets of brushes with stainless steel bristles and has a diameter of 266.2mm. This is larger than the initial diameter of the fuel channel, which at the beginning of its life was 263mm. This means that throughout the 11 brick layers of the core the LSB will be in constant contact with fuel channel walls. Therefore, if this component of the model can be isolated entirely it can be used to estimate the bore measurements of the channel. The contribution from the LSB is a frictional component and therefore it will always be impeding the motion of travel. In the discharge trace this manifests itself as an increase in load value from the deadweight value, with the load increasing even further when the channel bore measurement decreases. It is the opposite case for the charge trace (inserting the fuel stringer into the core), where the addition of friction results in a decrease in load from the deadweight value. When the channel bore diameter measurement decreases in this trace it will be seen in the lowering of the load value. Initially to visualise the data scatter graphs were used to show the relationship that exists between the bore and the load. The results from this visualisation can be seen in Figure 4.29 where pink is layer 11, dark blue is layer 10, cyan is layer 9, green is layer 8 yellow is layer 7 and red is layer 6. Layers 1-5 have been excluded from the as they all contain upper stabilising brush (USB) friction contributions in addition to LSB contributions. It is important to note that for these scatter graphs the load values for each layer of the trace have had their respective deadweight removed. This same

process was performed on the charge traces which can be seen in Figure 4.30 with the same respective colours for the different layers.

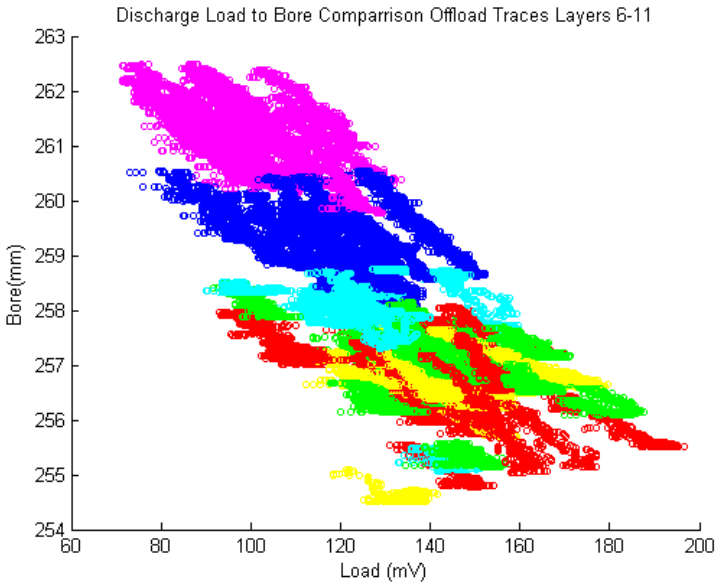


Figure 4.29- Discharge lower stabilising brush scatter analysis for layers 6-11 showing the relationship of load to bore.

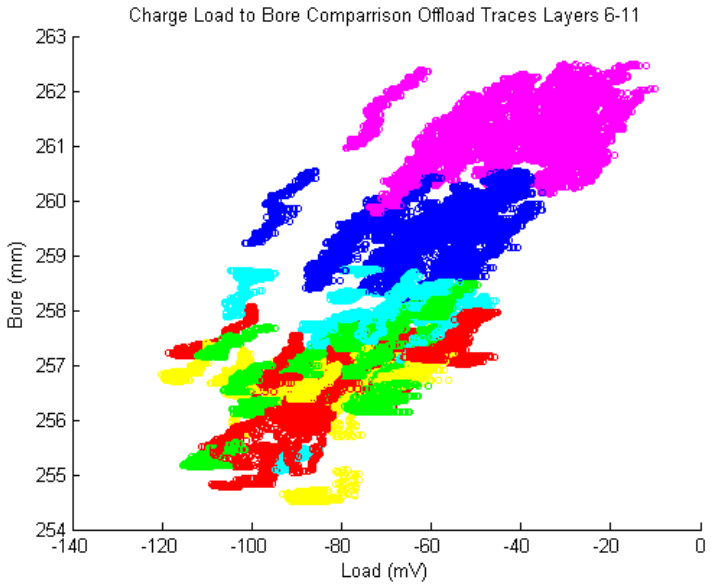


Figure 4.30 - Charge lower stabilising brush scatter analysis for layers 6-11 showing the relationship of load to bore.

From the results of the load to bore comparison there was a strong indication of a linear relationship over the data range provided and that there must be some underlying noise

or features explaining why there is some variation. One possible reason for the noise is that the lower stabilising brush is three sets of brushes. This will mean that load value responsible from friction will be more of an average over a larger area than the precise measurements from the inspection devices. Another reason could be that the fuel stringers are removed after different durations of time within the core and some fuel stringers are moved to different fuel channels routinely. Both reasons could account for different responses and noise present within the traces.

This assumption of linearity over the range of available bore measurements is different to the relationship which has been established through experimental work (Skelton, 2007) which has been obtained through performing refuelling on artificially machined graphite. The produced equation from the experimental work can be seen in Equation 19 where Fr is the frictional load value produced from the stabilising brushes in kg and ϕ is the bore diameter of the brick in mm. The region of bore measurements where this holds true is displayed in Figure 4.31.

$$Fr = -0.024986\phi^3 + 18.8256\phi^2 - 4729.3581\phi + 396226 \qquad \textbf{Equation 19}$$

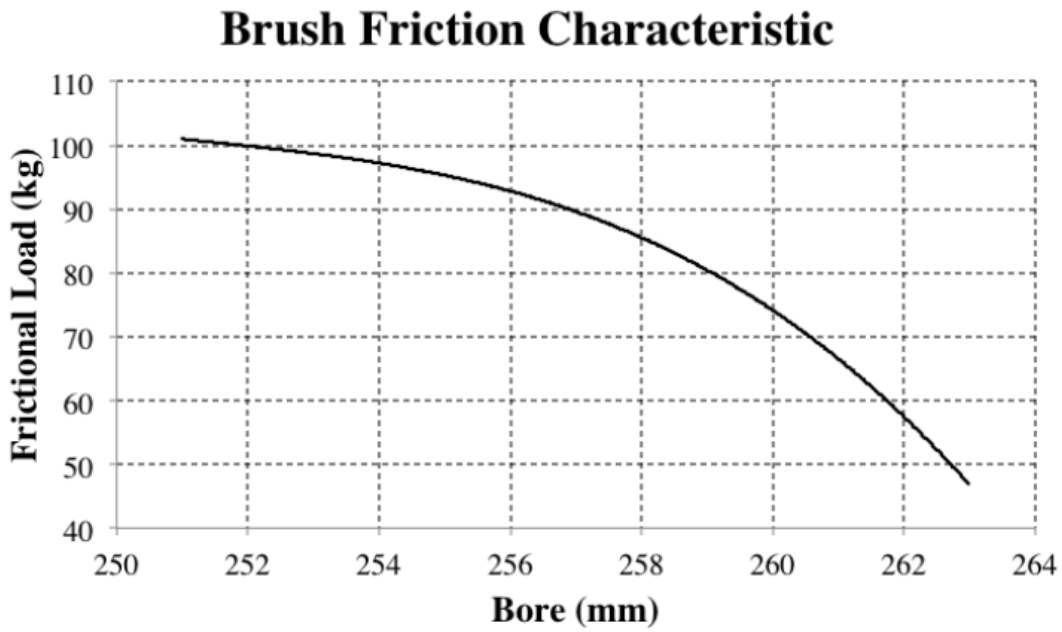


Figure 4.31 - Plot of variation of brush friction load against bore obtained from experimentation (West, et al., 2014).

Using the results from this, a model has been generated showing what effects changes to bore have on the frictional force found in the load trace (West, et al., 2014). The resulting model was a third-degree polynomial. It is important to note that the model had used results from as small bores as 251 mm, however from the real bore measurements there are only a few bricks less than 255 mm. The signals also contain a large amount of noise which could be attributed to the fact that there are multiple brushes contributing to the load and underlying core contributions which are not fully understood. Therefore, over this smaller range of bore a linear relationship may be observed as a starting point.

4.3.5 Four Component Summary

This subsection can be summarised as follows, the deadweight of the fuel stringer can be considered a constant load value throughout the core region of the trace. The aerodynamic effects of an onload trace can be described as two different load values with a transient region between them and for the offload trace the aerodynamic contributions are considered to be negligible. The USB contributes two regions to the trace that are experienced by two separate locations, the guide tube and the stand pipe. Lastly the LSB load contribution is directly related to the changes in friction caused by the narrowing and widening of the fuel channel. The simplified four contributions can be seen in Figure 4.32 for a discharge trace.

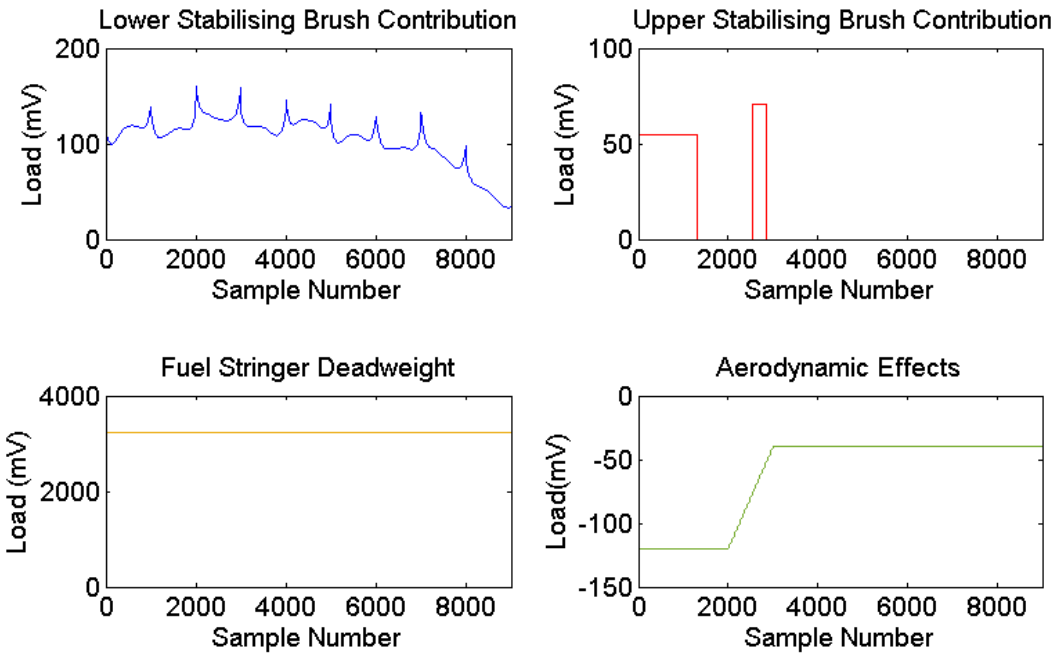


Figure 4.32 – Stylistic examples of the four contributions to the fuel grab load trace: the lower stabilising brush contribution, the upper stabilising brush contribution, the fuel stringer deadweight and the aerodynamic effects.

4.4 Lower Stabilising Brush Extraction

Using the knowledge gained from the analysis of the components it is possible to remove their contribution from the overall trace which results in only the friction contribution from the LSB. It is important to note that due to the quality of the traces and inspection measurements that the first two layers of the trace are unused and discarded.

The first step is to ascertain the extent of aerodynamic effects present within the trace. The process for uploading traces into the BETA database and in the future LoTAS (Load Trace Analysis Software) allows for metadata indicating the status of the reactor that the traces were obtained to be uploaded with the trace. Additionally, in the rare case where the metadata is missing the state of the reactor it is possible to categorise the traces as either onload or offload by using hierarchal clustering. The next stage is to calculate the deadweight contribution from the fuel assembly, as stated previously

the double peak feature can be identified and the load contribution can be extracted and removed from the entire trace.

At this point the contribution should only be from frictional forces from both sets of stabilising brushes. There are the two main USB friction contributions that are removed from the trace; these are the brick layer 5's piston seal bore and the guide tube of layer 3 and partly 4. The piston seal bore friction effect can be detected and suppressed by detecting the changes in load value in layer five at both sides. The beginning and end of the transition of contact of the USB is detected and used to calculate a value of load to be removed to suppress the USB. Then the transitional period when contact is made and when contact is lost is suppressed by connecting the start and end of the USB interaction to the rest of the trace. This same procedure is repeated to remove the guide tube interaction by detecting the step change at the lower end of brick layer 4. The only difference is that there is no beginning of this USB contribution and as such there can be greater inaccuracies in layer 3.

4.5 Load to Bore Model

With it being possible to extract the LSB frictional load contributions for the in-core region from brick layer 3 to 11, this can now be used to train a regression model. The purpose of the regression model is to understand the relationship between FGLT and bore. To do this training data is required that is representative to the operational data.

4.5.1 Training Data

The data used to train the regression model was obtained from inspections where both FGLTs and bore measurements are produced at the same refuelling event. At inspections, the fuel stringer needs to be fully removed to allow the inspection equipment to be inserted into the fuel channel. To safely remove the fuel stringer its mechanical load will be recorded for reactor trip protection. Therefore, every fuel channel inspection should be able to be paired up with its FGLT data. Over the four reactors that there are data for there is a relatively small amount of training data located in the database to be used to train and test the performance of the bore estimation algorithm compared to the amount of fuel channel within the reactors.

Due to the nature of the traces containing contributions from the USB the regions of the traces that can be used are limited. Therefore, brick layers that do not have any contributions from the USB are used to train the model. Additionally, an exception is also made for the topmost brick layer in the core as it is significantly shorter than the other brick layers and may not be consistent for training. For these reasons, only brick layers 6-10 are used for training purposes. In addition, the deadweight values are calculated for each trace and are removed from their respective bricks for training and testing. The last step of pre-processing of the data is to remove the brick interface points which contain the static friction components. Static friction occurs due to the narrowing of the fuel channel at the brick interface points, which requires a greater amount of force to move the stabilising brushes over them. Therefore, in these regions the frictional response will be different to the frictional forces present during the continuous motion of the fuel stringer, such as in the middle of graphite bricks. Due to the static friction present at brick interface points, the first 20% and the last 5% of the bricks are disregarded when analysing FGLT data for bore estimations. More data is removed at the beginning of the bricks as it takes an amount of distance before the frictional forces return to normal motion. This region has been defined for the discharge load traces and is different when considering the charge traces; this is due to the direction of travel of the fuel stringer.

4.6 Models for Mapping Load to Bore

There is some variance between the load values for the same bore value. It is believed that this is due to different conditions that are present in the core. The residual gas in the core and the degradation of the steel stabilising brushes are believed to be the primary causes of the variation.

Three models were trained using the ground truth data to allow for estimations of channel bore, these were a linear regression model, a quadratic regression model and a third order polynomial regression. The linear regression model was chosen as in the initial data exploration of the relationship of load to bore it was found that there was strong linear correlation between the two data sets in the regions without static friction.

The quadratic regression model was also chosen as it was believed that the friction relationship would exhibit a nonlinear relationship, particularly as the fuel channel becomes narrower. The third order polynomial was chosen as previous rig work that was undertaken by (Skelton, 2007) resulted in a cubic relationship.

4.7 Evaluation of Model

The evaluation of the model was undertaken by K-fold cross validation. K-fold cross validation is a method of cross validation whereby the data is split into K different groups. One group is used to test the performance of algorithms and the other groups are used as training data. This process is repeated for however many folds of data until all data has been tested. The results of the testing folds can then be used to obtain a generalised understanding of the performance of the models on future unseen data that has not been used to train the algorithms. The number of folds chosen for this application was 5-fold cross validation. Initially the different reactors are segregated in the training and testing process so that data is only trained and tested on the same reactor. The performances of the algorithms were evaluated on a per layer basis and this was done by calculating the mean absolute errors from the predictions and the bore data and the variance of estimation errors of the bore. These can be seen Table 4.1 for reactor 1 and on reactor 2. Additionally, the mean absolute error has been visually shown for reactor 1 in Figure 4.33 and for reactor 2 in Figure 4.34.

Table 4.1 – The mean absolute errors for each brick layer for reactor 1.

	Brick Layer 3	Brick Layer 4	Brick Layer 5	Brick Layer 6	Brick Layer 7	Brick Layer 8	Brick Layer 9	Brick Layer 10	Brick Layer 11
Linear Mean Absolute Error (mm)	1.1189	0.8922	0.8239	0.7563	0.7764	0.6962	0.7366	0.6328	1.0565
Linear Error Variance (mm)	0.6862	0.2882	0.2760	0.2607	0.2139	0.1805	0.2395	0.2155	0.4323
Quadratic Mean Absolute Error (mm)	0.7032	0.7075	0.6794	0.6210	0.6299	0.5541	0.6708	0.7127	1.0986
Quadratic Error Variance (mm)	0.1465	0.1964	0.1965	0.1863	0.1470	0.0960	0.2118	0.2652	0.4134
Cubic Mean Absolute Error (mm)	0.6972	0.6629	0.6494	0.5919	0.5868	0.5545	0.7211	0.7266	0.9203
Cubic Error Variance (mm)	0.2988	0.1961	0.1997	0.1838	0.1442	0.1120	0.2224	0.2564	0.4409

Table 4.2 – The mean absolute errors for each brick layer for reactor 2.

	Brick Layer 3	Brick Layer 4	Brick Layer 5	Brick Layer 6	Brick Layer 7	Brick Layer 8	Brick Layer 9	Brick Layer 10	Brick Layer 11
Linear Mean Absolute Error (mm)	1.1443	1.1806	1.0557	1.0463	0.9275	0.8479	0.8328	0.6962	1.3882
Linear Error Variance (mm)	0.7007	0.6576	0.6309	0.4558	0.4691	0.3960	0.4717	0.4081	0.5840
Quadratic Mean Absolute Error (mm)	0.7461	1.0270	0.8356	0.8762	0.7503	0.6254	0.7469	0.8297	1.3054
Quadratic Error Variance (mm)	0.3285	0.5774	0.5275	0.4464	0.3774	0.3091	0.4319	0.4557	0.7588
Cubic Mean Absolute Error (mm)	0.7712	1.0446	0.8417	0.8784	0.7595	0.6254	0.7893	0.8125	1.2134
Cubic Error Variance (mm)	0.3430	0.6033	0.5647	0.4724	0.3872	0.3263	0.4397	0.4762	0.7560

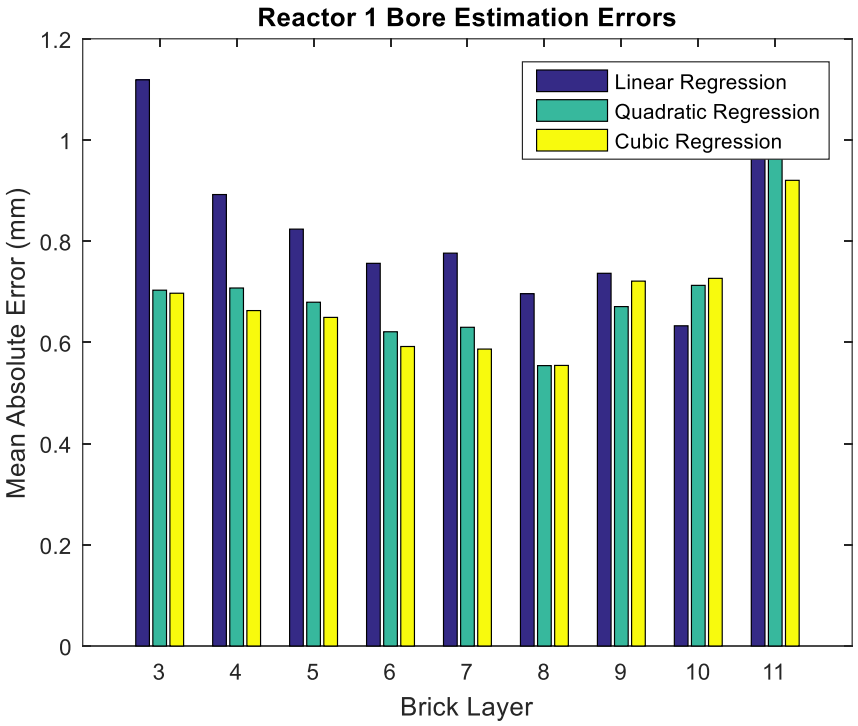


Figure 4.33– The mean absolute errors for each brick layer for reactor 1.

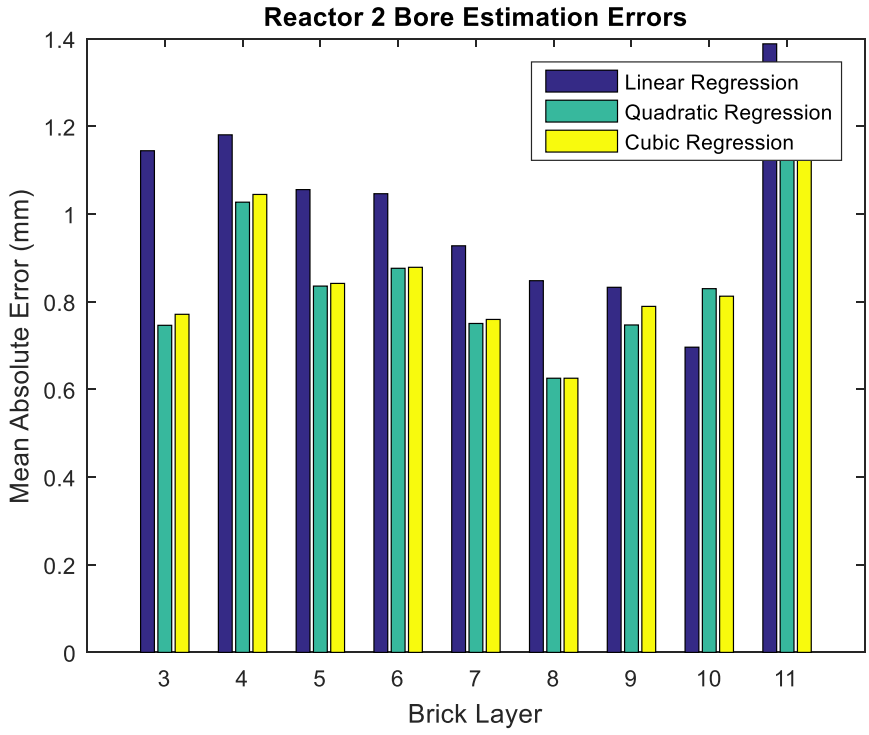


Figure 4.34 – The mean absolute errors for each brick layer for reactor 2.

It can be seen from the results of the K-fold cross validation testing that the performance of all the algorithms is better on reactor 2. It is believed that the reason for this is that typically in reactor 1 the bricks are similar to each other and there are less highly shrunken bricks. Additionally, for all models the errors on brick layer 11 are larger than most of the other brick layers, this is attributed to not using brick layer 11 as part of the training or testing data. It is worth noting however that these bricks are different in size and shape to the others and are subjected to the least amount of ageing mechanisms. Overall it can be seen that the quadratic and cubic regression models perform better than the linear model on both reactor 1 and reactor 2 however this is not the case for brick layers 10 and 11. Additionally the mean absolute error of the quadratic and cubic models has a mean absolute error of less than 1 mm for all brick layers outside of brick layer 11. This is seen as a success as this was more than the acceptable accuracy that graphite core engineers desired. There is very little difference between quadratic and cubic performance. The reason for this is due to the small region of bore that is available to be used for bore estimation training and testing resulting in the produced regression models being very similar over this region.

The recommendation going forward for graphite brick bore estimation is to be undertaken with a quadratic model as they exhibit a lower amount of variance. However, it should be noted that outside the current range of bore in the training and testing data that the performance cannot be correctly estimated and when new operational data from the extremes of bore diameter are obtained further training should be performed. This will allow the accuracy of the algorithm to be improved and maintained as the graphite core ages.

4.8 Conclusions

This chapter has shown that it is possible to estimate fuel channel bore from FGLT data. To do this the four different components of the FGLT were explored. The deadweight of the fuel stringer has been established in terms of mechanical load. This has been explored for both onload and offload refuelling events and it has been shown that there are differences in these values. From this work an algorithm was developed to identify the specific region on the trace automatically to extract the fuel stringer's

mass contribution and be able to remove it from the entire trace. The contributions of the USB have been able to be identified and analysed with respect to the load that they contribute to the overall trace. A method to automatically detect the change points that indicate the contribution has also been implemented for both the guide tube region and the piston seal bore regions of the trace. The aerodynamic effects have been explored and it has been found that it is a difficult task to be able to include this in a model due to their unpredictability. The lower stabilising brush has been shown to have linear properties which are why a linear regression model was originally proposed to model the relationship.

It has been shown by identifying important areas of interest in the FGLT, that it is possible to remove contributions to the load value, such as fuel stringer deadweight and USB friction, which allows the lower stabilising brush to be extracted. This might have been able to be achieved through signal decomposition techniques such as singular spectrum analysis but due to the available information about refuelling events the knowledge-based approach was chosen.

It was found that by comparing three different models of regression (linear, quadratic and cubic) that it is possible to accurately estimate fuel channel bore from FGLT data for 75% of the length of the brick. It was also found that out with one brick layer the mean average error for both quadratic and cubic models were within 1mm which is desirable for accurate predictions. The recommendation is to use either the quadratic or cubic model on new data but to retrain the model in the future with new training data that extends beyond the current limits of brick bore when it becomes available.

The bore estimation framework could have great value in the future of monitoring the bore changes in AGR graphite bricks in the UK. The implementation of this algorithm could allow for much more frequent data to be obtained. A process flow diagram of the framework can be seen in Figure 4.35. Initially, the raw FGLT data is obtained and both the deadweight and upper stabilising brush contributions are removed leaving only the lower stabilising brush contribution. The isolated trace is then subjected to a regression model which has been previously trained on the relationship between the lower stabilising brush component and bore inspection measurement, allowing accurate bore estimations to be produced.

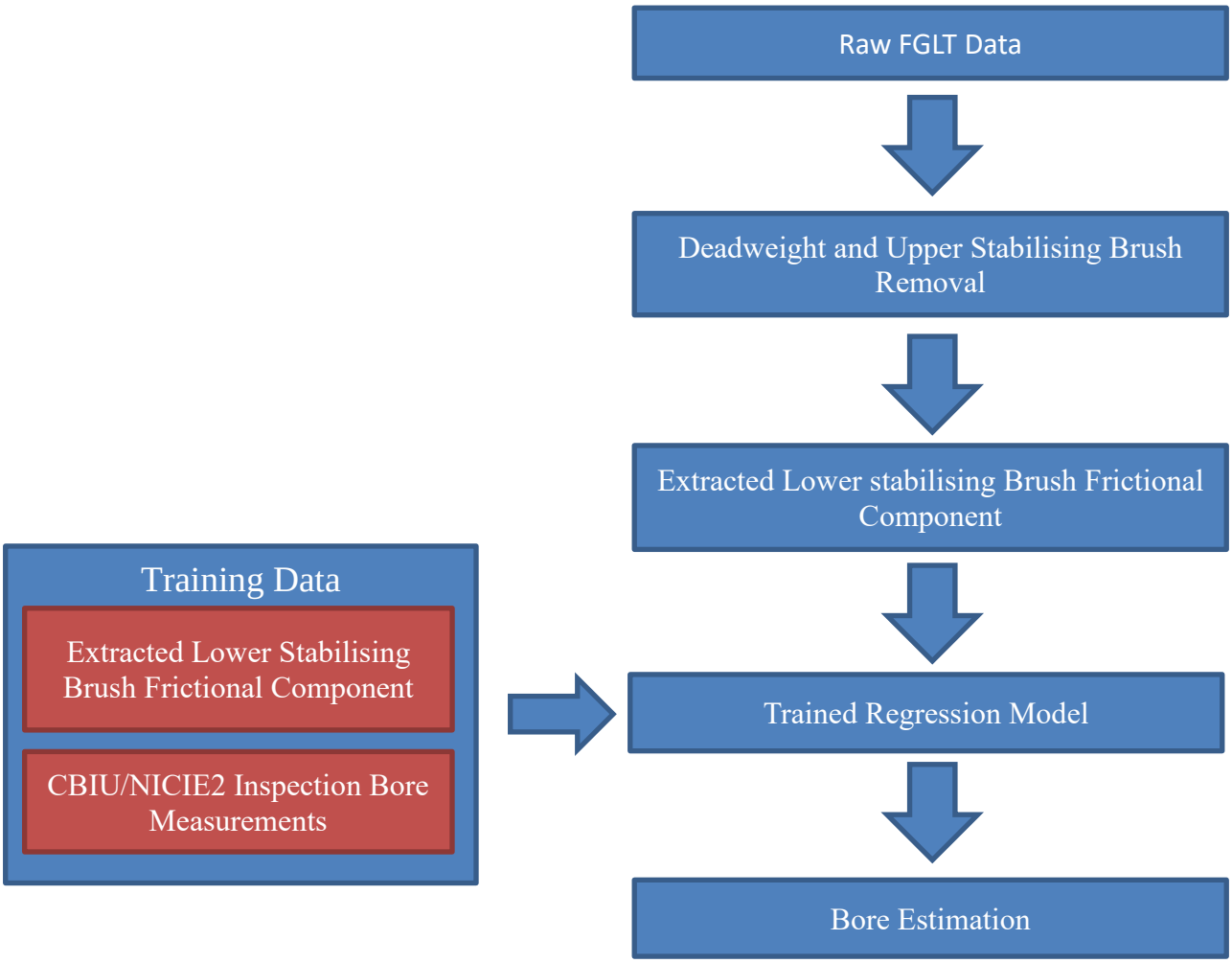


Figure 4.35 Bore estimation framework process flow diagram.

The performance of the regression model is good with many of the brick layers experiencing mean absolute errors of less than 1mm, which is desirable However, the bricks at the extremes range of the training data do not perform as well. This means that retraining should occur when more data is made available. As the reactors age, the graphite bricks will be more likely to be in the regions that are considered the extremes at the time of writing. Therefore, to maintain accuracy retraining should be incorporated when applicable.

5 Semi-Supervised Learning for Advanced Gas-Cooled Reactor Graphite Brick Crack Detection

5.1 Introduction

As stated previously graphite bricks change shape over the duration of their life within the core. These changes in the graphite brick structure result in the brick experiencing compressive and tensile stresses, which can result in cracks developing. It would be beneficial to detect these cracks from the FGLT monitoring data. A challenging aspect of developing a machine learning algorithm to detect cracks within graphite bricks is that there are an insufficient number of known cracks in the ground truth data to effectively train a classifier. This is due to only a small number of channels being inspected at statutory outages every 1-3 years. Whereas, monitoring occurs much more frequently, roughly every 6-8 weeks, and records data that has no ground truth information associated, known as unlabelled data. Under most supervised machine learning approaches unlabelled data is not used, but by utilising semi-supervised learning techniques this unlabelled data can be useful. Therefore, three different semi-supervised machine learning techniques have been applied to the graphite brick crack detection problem and their performance. The benefits of a semi-supervised approach are that a large amount of historical brick monitoring data, which can be used to supplement the smaller amount of labelled inspection data. This would circumvent the arduous, time-consuming task of the manual assessment of FGLT data, which requires specific domain expertise.

As with the previously discussed problem of monitoring the bore dimensions of the graphite bricks, monitoring data produced at refuelling events has been chosen to be used to detect the presence of cracks in the graphite bricks. This will allow more frequent information to be obtained that could aid in future inspections of the graphite cores as well as their operation. However, unlike the previous problem, which had 750 data samples per labelled brick to use to train each labelled brick in this instance is

only considered as 1 data sample. This is further justification to pursue semi-supervised machine learning approaches.

Within the nuclear condition monitoring, it is unlikely to have a large amount of available failure data for systems and components. It is also becoming increasingly common for large amounts of operational data to be stored and to remain unlabelled (Strachan, et al., 2007). This is due to the volume and velocity of data that is being stored as well as the cost of having limited specialised engineers that can interpret the data. Semi-supervised machine learning has been identified in other domains for improved performance by allowing both labelled and unlabelled data to be used as training, such as in vibration monitoring of compressors (Potočnik & Govekar, 2017), structural health information of bridges (Chen, et al., 2014) and applications for on line fault diagnostics (Yuan & Liu, 2013). Specifically, within the nuclear power condition monitoring domain there have been two main contributions of semi-supervised machine learning from Ma (Ma & Jiang, 2015) and Moshkbar-Bakhshayesh (Moshkbar-Bakhshayesh & Ghofrani, 2016).

5.2 Feature Extraction

Initially, the features that have been chosen as inputs to the classifier were the base calculations which are obtained from the M/W point features (West, et al., 2008). The M/W point features shown on stylised discharge brick trace which can be seen in Figure 5.1. They are named after the general shape that represents a single brick in the FGLT, where it can either be simplified to 5 points that look like the letters “M” and “W”. The points are calculated from different regions of the bricks with how their values are obtained shown in Table 5.1 where 0-100% indicates a relative location from the bottom of the brick to the top of the brick or left to right in the FGLT as indicated in Figure 5.1. These values were chosen to be able to obtain the key shape of the response of a single graphite brick in the trace. The base calculations then develop this further by performing operations on the bricks and their neighbours. These operations were originally chosen by graphite core engineers as a guideline to detect abnormal bricks and the calculations to obtain them can be seen in Table 5.2. In Table 5.2 the subscripts to the M/W points are indicating if they are obtained from the

neighbouring bricks specifically the upper or lower brick indicated by U and L respectively.

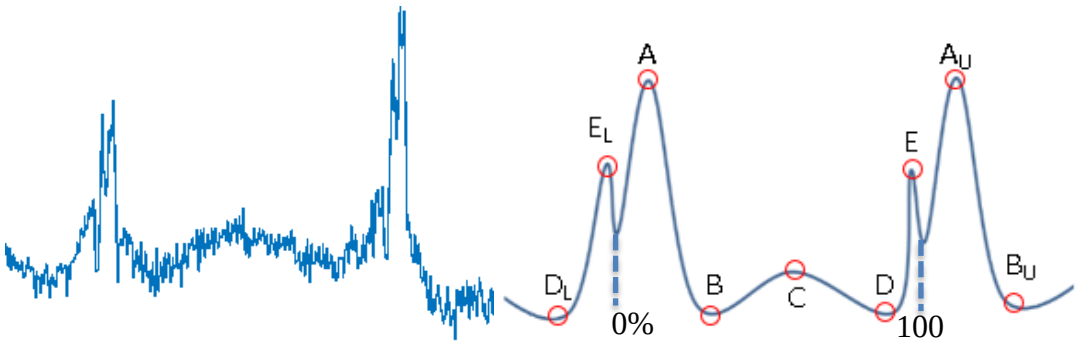


Figure 5.1 - Single brick trace including interface regions with accompanying illustration of the M/W point features on stylized representation of a single brick FGLT.

Table 5.1- Calculations to obtain W point features for a single brick from a discharge trace (West, et al., 2008)

A	The maximum load value between 0-10% of the brick
B	The minimum load value between 10-25% of the brick
C	The maximum load value between 25-75% of the brick
D	The minimum load value between 75-90% of the brick
E	The maximum load value between 90-100% of the brick

Table 5.2 - Basic calculations for feature selection obtained from W point data for discharge trace (West, et al., 2008)

BC1	Height of the upper brick interface peak for the previous brick layer (E_L-D_L)
BC2	Difference between the upper brick interface for the previous brick layer and the lower brick interface for the current layer (A-E_L)
BC3	Height of the lower brick interface peak for the current layer (A-B)
BC4	Difference in the peak base heights between the upper brick interface for the previous brick layer and the lower brick interface for the current brick layer (B-D_L)
BC5	Difference in the peak base heights between the upper brick interface for the current brick layer and the lower brick interface for the next brick layer (D-B_U)
BC6	Difference in the upper brick interface peak for the current brick layer (E-D)
BC7	Difference between the upper brick interface for the current brick layer and the lower brick interface for the next brick layer (E-A_U)
BC8	Height of the lower brick interface peak for the next brick later (A_U-B_U)
BC9	The difference between the mean load value of 45-55% and mean load value of 20-80% of the brick

In addition to these base calculations, addiontal graphite brick characterisation features have been chosen as additional inputs. These were chosen after discussions from graphite core engineers based on features they observe during manual analysis. These are the difference between the current bricks average load and the average load of the neighbouring bricks, both upper and lower. The last feature was the difference in load of the height of the highest peak in the middle 20-80% of the brick layer and the average load over that same region. These additional features can be seen in Table 5.3.

Table 5.3 Additional basic calculations to be included as features for crack detection algorithm.

BC10	The deference between the maximum load value of 20-80% of the brick and the mean load value of 20-80% of the brick
BC11	The difference between the mean load value of 20-80% of the brick and the mean load value of 20-80% neighbouring lower brick
BC12	The difference between the mean load value of 20-80% of the brick and the mean load value of 20-80% neighbouring upper brick

5.3 Comparison of Semi-Supervised Machine Learning Algorithms

5.3.1 Overview of Algorithms

The following section details the application of different semi-supervised machine learning algorithms. The semi-supervised learning algorithms that have been applied are Co-training, label propagation and transductive support vector machines.

1. Co-Training – A method which leverages two independent views and classifiers of the data to cooperatively learn about the unlabelled data from each other. This algorithm was originally developed by Blum and Mitchell (Blum & Mitchell, 1998)
2. Label Propagation – A learning graph-based method which allows labels to propagate through unlabelled data based on a similarity metric, typically Euclidian distance. This algorithm was originally developed by Zhu and Ghahramani (Zhu & Ghahramani, 2002).
3. Transductive Support Vector Machines (TSVM) or Semi-Supervised Support Vector Machines (S3VM) – A variation of support vector machines whereby, in addition to minimising the error due to the separating hyperplane, the margin between the two classes is maximised by incorporating low density regions in both the labelled and unlabelled data. This algorithm was originally theorised by Vapnik (Vapnik, 1998).

Co-training has been chosen as an algorithm to be investigated due to the nature of the available data. Co-training requires a minimum of two conditionally independent views of the same data to effectively operate (Krogel & Scheffer, 2004). Within the context of crack detection, the two views of the data are obtained from the discharge and charge traces of the same fuel channel. To the best of the author's knowledge, at the time of writing there has not been reported use of co-training within the nuclear condition monitoring domain. The requirement for the training data is two independent viewpoints of the problem which is not necessarily available in a lot of circumstances. However due to the unique situation of the FGLT data having a discharge and charge trace this makes co-training a valid option for graphite brick crack detection.

Co-training was originally developed and shown to have use in the classification of web pages by Blum in Mitchel (Blum & Mitchell, 1998). However, it has been used in other domains since. In bridge structural health monitoring co-training and tri-training are proposed as a structural health data analysis model (Yu, et al., 2012). A modification of the co-training algorithm, known as COPROG (co-training-based data-driven prognostic approach) was proposed by Hu (Hu, et al., 2011) (Hu, et al., 2012). The modification to the co-training algorithm is that instead of simply classifying a fault in the data it instead calculates a remaining useful life (RUL). This has been demonstrated in two applications in the literature, bearings and faulty fan. In these examples feedforward neural and radial basis networks were used as the base classifiers within the COPROG wrapper.

Label propagation is being investigated due to its demonstrated application on nuclear condition monitoring simulator data by Ma (Ma & Jiang, 2015). Ma has shown how a combinational algorithm of a combinational algorithm of spectral clustering and label propagation has managed to classify faults in a Canada deuterium uranium (CANDU) desktop simulator and a physical nuclear power plant simulator. This method was chosen as there was difficulty in obtaining a sufficient volume of training data of specific faults in nuclear power plant systems that would be able to train a supervised classifier. Ma combined operational unlabelled fault data with simulated fault data within the label propagation algorithm to allow the classification of faults. Outside of nuclear condition monitoring label propagation has been shown to have applications for structural bridge health monitoring (Chen, et al., 2014).

TSVM are being targeted as an algorithm to investigate due to its shown uses in transient classification in nuclear power station context by Moshkbar-Bakhshayesh (Moshkbar-Bakhshayesh & Ghofrani, 2016). This is achieved by using a previously created algorithm containing autoregressive integrated moving average (ARIMA) model which was combined with an error back propagation (EBP) algorithm (Moshkbar-Bakhshayesh & Ghofrani, 2014). The downside of this first step method is that it is unable to cluster the unknown and unlabelled transients and is rectified by the inclusion of the semi-supervised learning algorithm TSVM. This combination demonstrates the ability of TSVMs for classification purposes such as in graphite brick crack detection. Outside of nuclear condition monitoring, TSVMs have been shown to have applications for intelligent gear fault diagnosis by Shen (Shen, et al., 2012).

5.3.2 Supervised Competitors

This section of the chapter will describe the implementations of the supervised machine learning algorithms that are being compared to the semi-supervised techniques. The three supervised algorithms that are being compared to the semi-supervised learning algorithms are artificial neural networks (ANN), k-nearest neighbours and support vector machines.

5.3.2.1 Artificial Neural Network

Artificial Neural Networks are being chosen as a supervised machine learning to compare to Co-Training methods as they were the base classifier that was being trained in the process. The chosen ANN for supervised machine learning is the backpropagation multilayer perceptron.

One hidden layer was chosen to be used to detect cracks as the comparisons are meant to be to off-the-shelf applications of the supervised techniques and deep learning ANNs are not being considered.

The method of choosing the number of nodes in the hidden layer was by trial and error under the guidance of the fact that the number of nodes in should be between the size of the input and output layers. To aid in this process some general rule of thumb methods were used as a starting point to help identify the most consistent hidden layer size. These were defined by Heaton (Heaton, 2008) as follows:

- *The number of hidden neurons should be between the size of the input layer and the size of the output layer.*
- *The number of hidden neurons should be 2/3 the size of the input layer, plus the size of the output layer*
- *The number of hidden neurons should be less than twice the size of the input layer.*

Through this method the hidden layer size was found to have the most optimal performance with a size of 10.

5.3.2.2 Support Vector Machine

Support Vector Machines are compared to TSVM as they are the base classifiers that are adapted to be suitable for semi-supervised machine learning.
In addition to linear classification, nonlinear classification can be performed in much the same way by transforming the data into higher dimension space where it is linear separable.

$$K_{RBF}(x, x') = \exp\left(-\gamma \|x - x'\|^2\right) \qquad \textbf{Equation 20}$$

The radial basis function kernel was chosen to be used to allow the SVM to transform the data into a linearly separable space. The radial basis function equation can be seen Equation 20 where γ is a variable which gamma can be altered to define the influence of individual training examples with low values meaning they have further reaching influence and large values meaning they have much closer influence. Another parameter that is needed for SVMs is the C parameter that is associated with the cost of misclassification by the complexity of the decision surface. Large values of C indicate a complex surface with a greater number of samples as support vectors whereas small values create a smoother surface where less support vectors are used.
For this application a grid search was used to iteratively establish the best parameters between 10^{-3} and 10^{3}. The results of this process can be seen in Figure 5.2 which shows a heat map and gradient of the performance of the SVM on the cracked brick data set. From the result of this search it was found that the optimal value for the C parameter was 8 and for γ parameter the value was found to be 2.

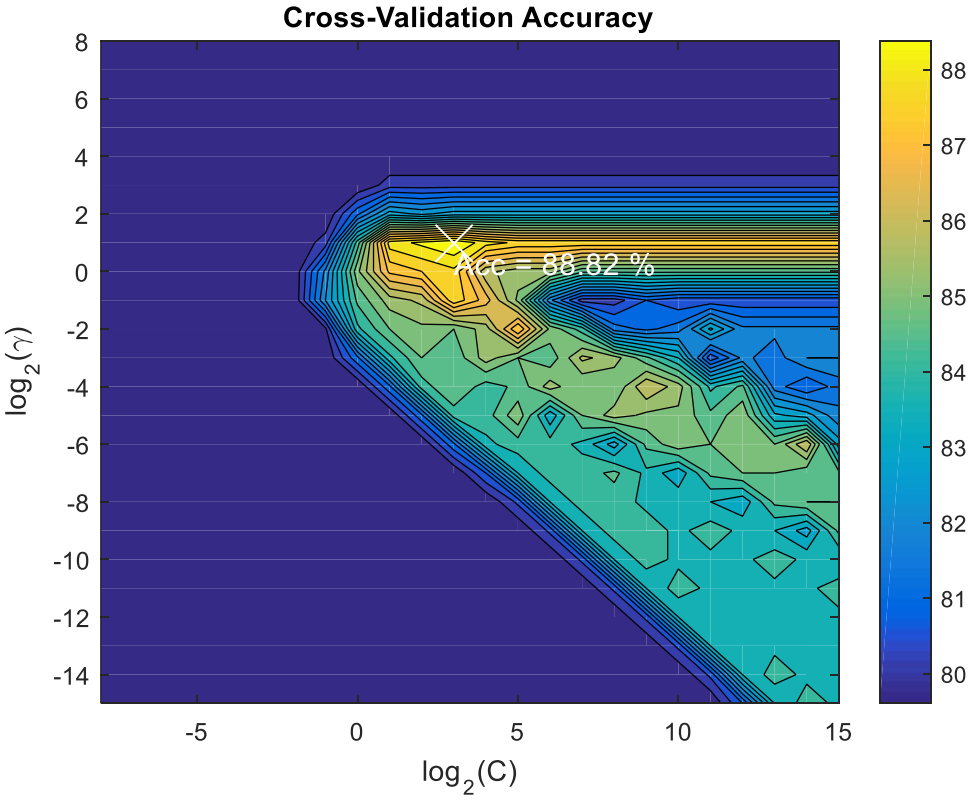


Figure 5.2 – A heat map of the search for C and γ to find the optimal values for the radial basis function.

5.3.2.3 K-Nearest Neighbours

K-Nearest Neighbours (kNN) was chosen as an appropriate algorithm to compare label propagation to. kNN and label propagation are similar as they are both graph based algorithms that provide labels to unlabelled data based on its similarity to other data. In k-nearest neighbours the unlabelled data is classified by a majority voting approach that is based on the other known labelled data. It does this by identifying the k nearest instances of data in the feature space. The class with the highest number of instances is then used to inform the classification of the data instance. This process can then be repeated for all the remaining data to allow for full classification of the unlabelled data. In the attempt for graphite brick crack classification the k value has been set to 3.

5.3.3 Training and Cross Validation

The data used for training the semi-supervised learning algorithms is the FGLT data. The labelled portion of the FGLT is obtained at offline inspections, where visual inspection equipment is used to assess the health of inspected fuel channels. Unlabelled instances of the FGLT data are obtained at low power refuelling events. Additional labelled data can be obtained by assigning labels to unlabelled data if they have labelled inspection data from the same channel. For example, if a crack in a brick is observed during an inspection, which will mean that all future instances of monitoring that brick can be labelled as being cracked.

It was possible to obtain 175 instances of labelled training data and an additional 524 instances of unlabelled training data from the bricks in the core from a single AGR reactor. Of the labelled training data 40 instances are of cracked bricks while the remaining 135 instances do not contain any cracks. Without time consuming analysis there is no way to establish the classification split of the unlabelled data. It is not possible to assume that the unlabelled data with have the same distribution as the labelled data as there could be a bias in the specific channels that are inspected. For example, the CHANSELA software tool (Watson, et al., 2011) supports the inspection regime by selecting channels which have the highest criteria to be inspected. This can lead to more inspections of the bricks which are already known to be cracked or regions of the core that are more prone to cracking.

Previous work (West, et al., 2008) has shown how, after discussion with domain experts, the general shape of a graphite brick can be extracted from a single FGLT by selecting five discrete points on the trace. Calculations are then performed on the value of the five points as well as the five characteristic points from both the bricks upper and lower neighbouring bricks on the same channel. By comparing the data relative to the neighbouring bricks this allows features to be extracted that are independent of offsets introduced by the trace or brick layer they are from.

5.3.4 Artificially Generated Bricks

Currently there is not a sufficient amount of cracked data to effectively test the algorithms due to a large class imbalance between the majority normal brick data and

the minority class crack data. Therefore, the decision was made to artificially generate new data exclusively for testing. To do this, two different types of cracks (bore initialled axial and circumferential) were chosen to be artificially generated for each of the layers being analysed.

Artificial circumferential cracks were created by extracting regions of cracks in the FGLT that are from known circumferential cracks and then overlaying them onto known healthy brick inspection data. Similarly, to this to generate the bore initiated axially cracked data the load value over a brick region is increased by 10mV. Through this process an additional eight cracks were introduced to be classified; one circumferential and one axial crack for each brick layer 7-10. Based on the theoretical understanding of how cracks affect FGLT data and visual inspection of the modified trace, these artificially generated brick would be identified as such by graphite core engineers. Figure 5.3 through Figure 5.10 shows the artificially generated crack FGLT data. The results of these artificially generated cracks will be discussed in the following section.

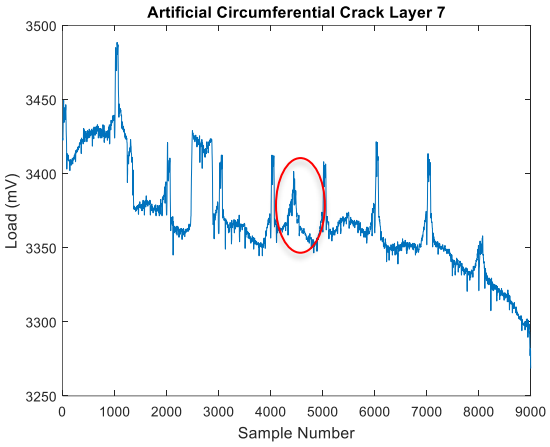


Figure 5.3 - Artificial circumferential crack on brick layer 7 of a discharge FGLT.

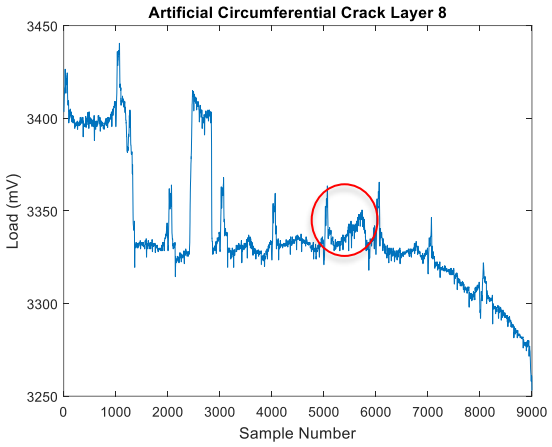


Figure 5.4 - Artificial circumferential crack on brick layer 8 of a discharge FGLT.

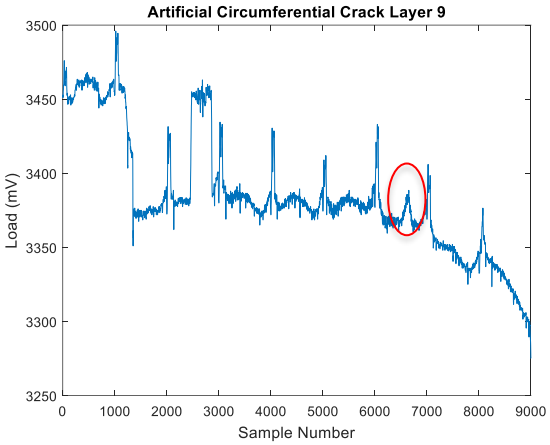


Figure 5.5 - Artificial circumferential crack on brick layer 10 of a discharge FGLT.

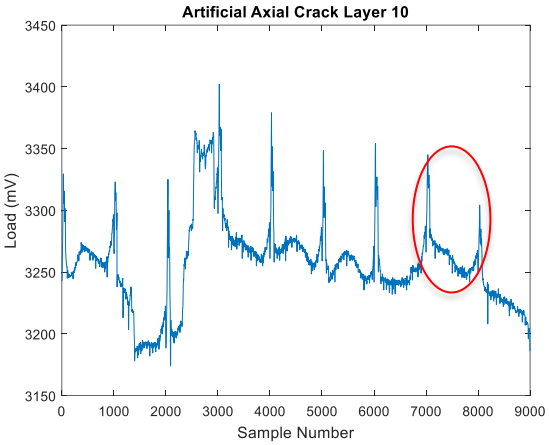


Figure 5.6 - Artificial circumferential crack on brick layer 9 of a discharge FGLT.

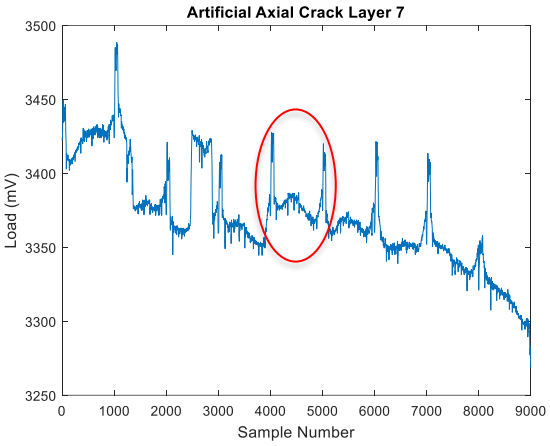


Figure 5.7 – Artificial axial crack on brick layer 7 of a discharge FGLT.

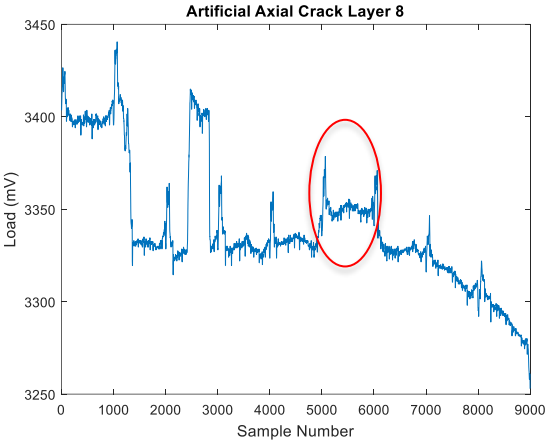


Figure 5.8 – Artificial axial crack on brick layer 8 of a discharge FGLT.

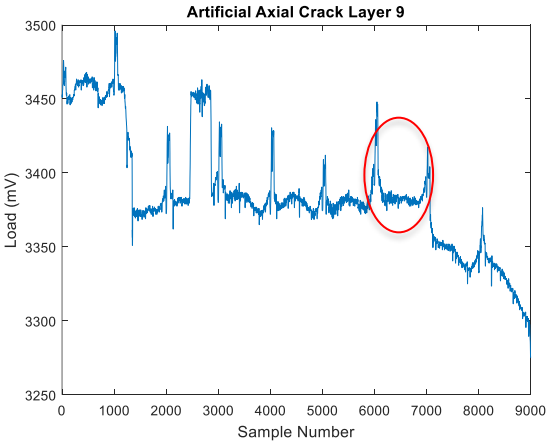


Figure 5.9 - Artificial axial crack on brick layer 9 of a discharge FGLT.

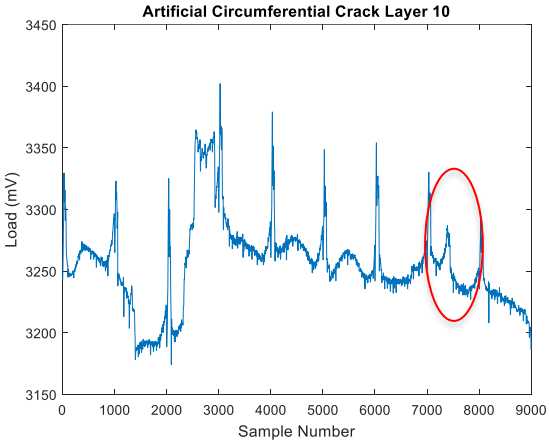


Figure 5.10 - Artificial axial crack on brick layer 10 of a discharge FGLT.

5.4 Results

To measure the performance of the three different semi-supervised learning algorithms (co-training, label propagation, TSVM) they were applied on the same AGR graphite brick monitoring data. Additionally, each of the semi-supervised learning algorithms has been compared to an equivalent supervised machine learning algorithm. Co-training is compared against its base classifier, an artificial neural network (ANN). Label Propagation is compared to the k- nearest neighbour algorithm. Lastly TSVM are compared to traditional supervised support vector machines (SVM). To evaluate the performance of the algorithms a Monte Carlo approach to cross validation is undertaken, whereby over 100 iterations the training set and test set are shuffled randomly using a 3:1 ratio. This allows metrics of performance to be calculated that are generalised to the entire available labelled test set. It is also worth considering the differences between the deterministic and stochastic natures of the algorithms. Therefore, a Monte Carlo approach has been taken compared to a k-fold cross validation approach. For example, label propagation will always perform the same if the training and test data remain the same however ANNs start with randomly initiated values so there will be variations in performance regardless of the training and test data being consistent.

The most basic metric of performance of the algorithms on the test data is the average performance on the test set; this can be seen in Equation 12. However, using accuracy alone is not a suitable measure of the performance of the algorithms due to there being a class imbalance problem. Therefore, it is necessary to consider both the recall, precision and resulting F1 score to determine the most suitable classifier. The recall value is important for graphite brick detection as it displays the mean percentage of crack bricks in the testing set that are correctly detected and can be calculated as shown in Equation 13. Recall provides more information than performance for detecting cracked graphite bricks due to the class imbalance. Normal graphite bricks that have not experienced any cracking outnumber those that have cracked. This means that basing analysis only on accuracy values result in bias towards correct classification on normal bricks. For this application however, it is important to understand the mean number of cracks that are detected by each algorithm. Additionally, another important metric of performance is the precision of the algorithms and can be calculated as shown in Equation

14. Precision, in the case of graphite bricks, is the mean percentage of the number of correctly detected cracked bricks out of all the cracked bricks detected. This allows for a confidence in classification to be established and provides an insight into the likelihood into whether detected cracked bricks are indeed cracked.

Lastly, the F1 score can be considered to bring both precision and recall together and can be calculated from Equation 15. F1 score can be more generally described as the F-score when precision and recall are not weighted similarly to each other, this can be seen in Equation 16 where β is the weighting factor (β values greater than 1 indicate greater priority on recall with β values less than 1 indicate greater priority on precision). When considering this specific application, recall should be prioritised over precision as it is safer to be conservative with regards to the safe operation the reactors. Given the option it would be better to know the maximum possible number of cracks in the reactor than the minimum number of cracked bricks.

Table 5.4 – Results of the three semi-supervised machine learning algorithms as well as the three supervised machine learning algorithms. Comparing their accuracy, recall, precision and F1 score.

Learning Algorithm	Accuracy Mean (%)	Accuracy Standard Deviation (%)	Recall Mean (%)	Recall Standard Deviation (%)	Precision Mean (%)	Precision Standard Deviation (%)	F1 Score (%)
Co-Training	81.67	4.17	53.52	16.15	56.61	12.89	53.28
ANN	75.91	4.77	73.51	9.88	45.35	9.09	55.34
Label Propagation	50.60	8.55	50.65	10.75	35.18	9.18	41.04
K-NN	84.14	3.44	30.66	9.64	79.26	12.25	43.51
TSVM	67.64	3.83	82.89	6.59	37.20	6.08	50.94
SVM	81.84	3.43	62.23	10.38	55.49	9.58	57.98

A possible reason that the precision of some of the semi-supervised learning algorithm falls while recall improves compared to their supervised counterparts is believed to be some overlap in the feature space. For example, some bricks which have experienced a large amount of shrinkage or alternatively have passed turnaround point and are now

growing could result in the FGLT data being perceived as a cracked brick. This indicates that either the input features selected are not sufficient at boundary regions to distinguish between the classes or the FGLT does not provide enough fidelity in its inferred measurements to allow the finer nuances to be learned. Another reason misclassification may have been introduced is that the input features used to represent the bricks are reliant on the neighbouring bricks. This could be a reason for the low value of precision in many of the algorithms cases as some uncracked bricks may be getting misclassified based on their neighbours.

The implemented algorithm with the highest accuracy was found to be the k-NN with a mean accuracy of 84.14% through a 100 iteration Monte Carlo cross validation.

Transductive Support Vector Machine does not have the highest accuracy, but it does have the highest average value of recall at 82.89%. However, the downside of this is that the precision is 37.20% which indicates that many bricks identified as cracked would be false positives.

K-NN has a very high accuracy but this is represented by a high precision of 79.26% but a lower recall of only 30.66%. This means that while the algorithm does not produce many false positive bricks it does misclassify a significant number of cracks.

Co-training has reasonable accuracy at a measure at 81.67% however this is largely due to the nature of assessing the performance of algorithms on an imbalanced data set. When considering the cracked bricks in the testing data the recall and precision of co-trained ANNs are 53.52% and 56.61% respectively.

The recall value of co-training is less than its supervised counterpart artificial neural networks with their recall values being 53.52% and 73.51% respectively. The reason for the poorer recall score of the co-training is due to the charge trace not being as good an indicator for cracked bricks as the discharge trace. Therefore, perhaps as opposed to using co-training a self-training method of machine learning should be considered. Alternatively, by changing the input features of the two base classifiers co-training may be possible. This could be achieved by using half the features with base classifier 1 and the other half with base classifier 2 however the downside of this would be that both classifiers would be trained on fewer features potentially losing valuable classification information.

Label propagation does not perform as well with detecting the presence of cracks as the other algorithms. It has an accuracy of 50.6% with a precision and recall values of 50.65% and 35.18% respectively. These low values would not be able to provide any confidence in the detection of cracked bricks and would identify less cracked bricks than simply guessing the classifications. The low values are most likely due to the data not being very easily separated within the feature space.

Recall in this situation has been identified as the most important metric for performance as the classification process is not meant to replace the human manual labelling process but help expedite it. High recall values will allow this as most of cracked bricks will have been identified which is the most important aspect to graphite core engineers that are currently analysing this data.

The recommendation for the most suitable algorithm for this problem is to use the semi-supervised TSVM algorithm primarily as it has the largest recall value. It could also be possible that the bricks that have been misclassified as cracked could be worth monitoring. Additionally, as the neighbouring bricks have been used to provide reference for some of the input features, this could be a reason why some bricks are getting misclassified. The bricks above and below them in the core could be cracked and as a result may influence the correct classification of the bricks being tested.

The results could see some improvements and optimisation around achieving a high level of precision while still maintaining an appropriate level of recall. These improvements and others will be discussed further in the following chapter as future work.

As mentioned previously in this chapter, artificially cracks were generated to provide some additional training data. Therefore, the eight cracks from Figure 5.3 through to Figure 5.10 underwent feature extraction and had each of the different crack detection algorithms applied to them. The results of the artificial cracks can be seen in Table 5.5 – The accuracy of the six machine learning algorithms when applied to the artificial cracked brick data. only showing accuracy as only cracked bricks were present in the test set.

Table 5.5 – The accuracy of the six machine learning algorithms when applied to the artificial cracked brick data.

Learning Algorithm	Accuracy (%)
Co-Training	24.5
ANN	70.25
Label Propagation	50
K-NN	25
TSVM	100
SVM	75

It can be seen from these results that they are similar to the results of the inspection test data when considering cracked bricks. This demonstrates the ability of the algorithms on prominent crack data and effectively represents the algorithms' recall performance. TSVM shows strong promise again on the artificial data where it correctly identified the cracks, something that none of the other algorithms managed.

5.5 Discussion

Co-training was chosen to be included in this comparison due to the availability of two independent views of the data in the form of the charge and discharge traces. If both individual algorithms were performed adequately on each trace, then the combination of both should have resulted in an overall better classifier. However, this was not the case as co-training performed worse than the performance of the base algorithm on the discharge trace. This is because it more difficult to identify cracked bricks on the charge traces compared to the discharge traces.

Co-training has large potential for use within the nuclear condition monitoring in situations where there are multiple sensors and unlabelled data. A benefit of this family of algorithms is that they are wrapper algorithms and as such can be implemented with a variety of base classifiers. This could allow for existing approaches to condition monitoring being improved without changing the current underlying way in which

their faults are detected. Additionally, in many circumstances due to the increase in the number of sensors monitoring assets which provide multiple viewpoint of the same equipment then co-training could see much greater use.

Label propagation has failed to provide adequate results for graphite brick crack detection. This is contrasted to the successful implementation of the algorithm on nuclear fault detection by Ma where it was utilised to classify faults with significantly fewer training data for correct classification (Ma & Jiang, 2015).

This could be for several reasons including the imbalanced data problem. In the case study provided the uncracked healthy bricks outnumbered the cracked bricks which can result in biases towards the more prominent class. In addition to the imbalanced data it is know that the classes overlap in the feature space which results in many instances of misclassifications. Optimisation of the hyperparameters within label propagation is a difficult task to perform an exhaustive search. A manual search was undertaken to establish the hyperparameters with the best outcome.

Additionally, Ma has combined the label propagation algorithm with a spectral clustering approach whereas in the cracked brick approach the data is only normalised before the algorithm is applied to the Euclidean distance of the data. The lack of a dimensionality reduction method in the crack detection may be responsible for the lower performance.

TSVMs have seen a strong performance in this application, similar to their strong performance as a nuclear transient identifier by Moshkbar (Moshkbar-Bakhshayesh & Ghofrani, 2016). In the context of this work TSVMs show an improvement over traditional SVMs in terms of recall accuracy. This may not be desirable in all situations as it can result in high amounts of false positives.

The biggest challenge with the data is the trade-off between precision and recall due to the overlap of classes in the feature space. In this case study regarding crack detection within graphite brick it is believed that having a high recall is important.

5.6 Conclusions

It has been shown that through extracting features from individual brick layers of the FGLT data can help in providing information of the presence of cracks within the

graphic bricks of AGRs. The points that have been extracted are chosen to characterise the M or W shape of the bricks represented in the FGLT in the discharge and charge traces respectively. It has been shown that these features can be identified through sets of rules which repeatedly achieve the correct identified points. When using this data to train a machine learning algorithm it has been identified that there is only a relatively small amount of available data which have labels of either cracked or non-cracked bricks, therefore semi-supervised machine learning has been used to make use of the large amount of available unlabelled data. This process is shown in a flow diagram in Figure 5.11 where the resultant output is the classification of the cracked status of the brick.

It has also been identified that there is a need to address the class imbalance problem by using the SMOTE algorithm to allow a balanced dataset. This algorithm increased the number of samples on the minority class to be equal to the majority class by synthetically generating new samples.

Six different algorithms, three semi-supervised and three supervised, were compared and it was found that the algorithm that produced the highest recall score was TSVM but did not have a high precision score. However due to recall having a larger impact than the precision of the algorithms it is the recommended algorithm to use for future cracked brick classifications. It has been identified that in these circumstances for high recall that semi-supervised machine learning can perform better than traditional supervised machine learning on this data set.

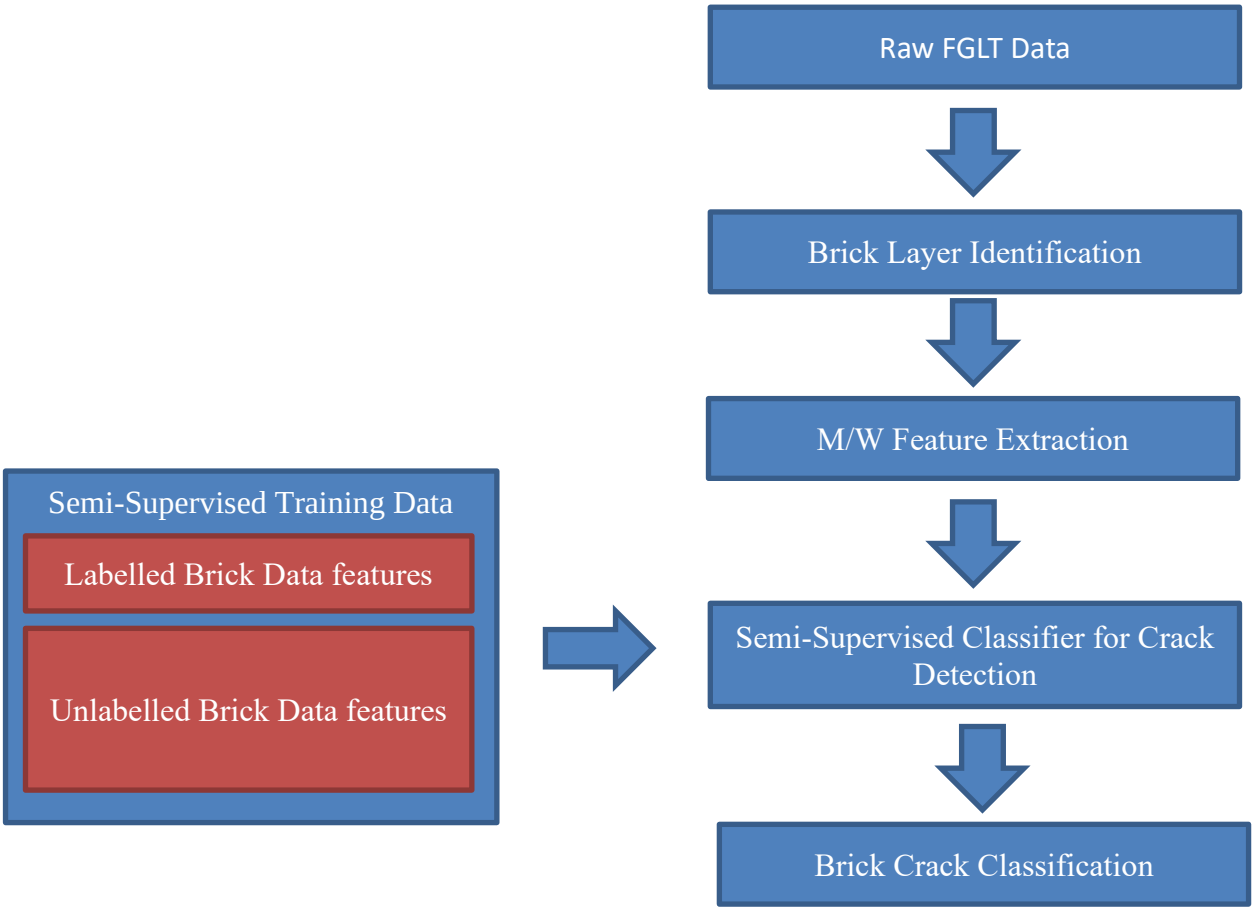


Figure 5.11 Semi-supervised cracked brick detection process flow diagram.

6 Conclusions and Future Work

6.1 Introduction

In the United Kingdom the Advanced Gas-cooled Reactors (AGR) are reaching the end of their initial design lives, by 2034 the current fleet of nuclear power stations will no longer be in operation. One of the key life-limiting components of an AGR is their graphite core. The graphite core cannot be repaired or replaced and experiences degradation from the harsh environment within the core which can affect its structure. Therefore, it is essential to perform inspections to understand the health of the cores to ensure that the AGRs are safe to operate. To support these activities additional monitoring of the cores can be undertaken to provide more frequent understanding of the state of the graphite within the core but with it comes the need to analyse the data. Two different approaches have been applied to understand the two different methods of ageing in the core. Both methods, which use the monitoring data known as the fuel grab load trace (FGLT), have been presented and evaluated. Both approaches use data analytics and machine learning to allow a more frequent representation of the health of the core than is currently available. The methods that have been presented are not looking to replace the manual analysis undertaken by graphite core engineers, instead they are looking to provide decision support, suggestions and reduction in the duration of manual analysis required. This work contributes to allowing the AGRs in the UK to continue operating to their current closure dates or potentially beyond by providing more frequent and meaningful information that can be used to ascertain core health.

6.2 Summary

In chapter 2 an overview of nuclear power generation was provided. This chapter focused on the design and the inspection/monitoring tools of the reactor core of AGRs. Furthermore, greater detail was given to the data that can be obtained during scheduled refuelling events of AGRs. This data is specifically known as the fuel grab load trace and its contributing physical component load values were discussed as well as highlighting them on an annotated copy of a full trace. An overview of work that has

been undertaken on condition monitoring and fault diagnosis of the UK's AGR fleet is also present within this chapter.

Chapter 3 detailed the approaches and methodologies of machine learning as well as the applications that they have been applied to within both the nuclear and wider condition monitoring domain. It focused on identifying what supervised, unsupervised and semi-supervised machine learning are and the situations where they are used. An overview of a several semi-supervised machine learning and more common machine learning algorithm was provided, identifying domains which they have seen use in.

Chapter 4 detailed one of the novel contributions presented in this work; the channel bore prediction model. The process by which the four components of the FGLT were analysed and generalised using 10 years of historical data was discussed. This allowed for the understanding of the contributions from the deadweight of the fuel stringer the aerodynamic effects of the coolant gas and the upper stabilising brush. The benefit of the extraction of the lower stabilising brush friction within the FGLT, which is directly related to the estimation of fuel channel bore is explained. The chapter then goes on to describe the machine learning training and testing of a regression model, which is used to estimate fuel channel bore from FGLT data. The regression model with the best performance was found to be the quadratic regression model which was able to accurately predict fuel channel bore dimensions.

Chapter 5 detailed the novel contributions that have been made to the evaluation of different semi-supervised machine learning algorithms for crack detection of graphite bricks in AGRs. It achieved this by comparing the performance of three semi-supervised machine learning techniques to their equivalent supervised implementations. This chapter included the process of identifying and extracting important features of the data, incorporating a method of class balancing and testing the methods in an accurate method. The three semi-supervised learning methods used for crack detection are the co-training algorithm, the label propagation algorithm and semi-supervised transductive support vector machines. Transductive support vector machines were identified as being the most suitable classifier to address the graphite brick crack detection problem where a high recall value is required.

The novel contributions detailed in this thesis are:

The development of an approach for extracting reactor health information from data not originally designed for condition monitoring purposes. This approach augments existing reactor inspection regimes by providing a supporting health indicator. This is demonstrated through an application of fuel channel bore estimation

- An in-depth analysis of the fuel grab load trace by using knowledge of the physical properties of the reactor core and a supply of historical data
- Development of a framework that can remove unwanted components present in the fuel grab load trace on a trace by trace basis
- Improving the understanding of the different contributions of the fuel grab load trace by using historical data to explore the relationship to channel bore measurements and allow it to be generalised
- Demonstration and evaluation of the performance of the algorithm on real nuclear power station fuel grab load trace data
- Recommendation of the application of the algorithm on operational data for monitoring and inspection regimes

An evaluation of semi-supervised learning approaches with respect to nuclear condition monitoring data, indicating improvements over traditional supervised machine learning classification techniques

- A critical review of three semi-supervised techniques exploring their suitability to nuclear condition monitoring data where obtaining labelled operational data of faults is a costly, time-consuming process
- The first demonstration of semi-supervised machine learning techniques being applied to the UK's fleet of nuclear power stations for condition monitoring purposes
- Recommendation of the semi-supervised machine learning algorithm for the application of crack detection and its implementation

6.3 Future Work

There are several opportunities for expanding and extending the work covered in this this thesis.

6.3.1 Bore Estimation from FGLT

As mentioned in chapter 4, the work that has been completed for the bore estimation from FGLT has only been trained and applied to only some of the reactors within the UK. Within the UK there are an additional 10 AGRs in operation where the bore estimation has not been applied to. Each set of stations is slightly different in design, such as having a slightly different number of bricks in the core, a different number of layers or different sizes of bricks. Therefore, due to the nature of how the bore estimation algorithm was implemented and trained it was not practical or able to apply its estimation to the other stations data. The next steps in the bore estimation algorithm would be to create a generalised algorithm that could be applied to the other stations. With further research into the designs of the other stations, it will be clear whether this could be as simple as changing where the USB interacts on the trace or whether a new regression model will have to be developed as the bricks and LSB exhibit a different frictional relationship.

Additional work that is still required on the bore estimation is the inclusion of a predictive gas model that can allow predictions when the reactor is refuelling at low power. The main difficulty associated with validating bore estimation for onload data is the lack of available ground truth data to validate. Currently there are only a very small amount of cases where there has been an inspection and monitoring event within a short space of time to allow this. Therefore, any estimation that is made using the current implementations of the models does not have confidence bounds on the performance of the algorithm with low power refuelling data. This problem could hopefully be rectified as more data is made available with a greater number of repeat refuelling events or more detailed information about the coolant gas floes at low power refuelling.

6.3.2 Semi-Supervised Crack Detection

It was found from the six machine learning methods tested that the best recall performance was from the TSVM semi-supervised implementation. However, one drawback of this method was that the precision of the implementation was low. This was not a problem due to the intended implementation of the algorithm, but future work should be to improve the precision and the overall accuracy of the implementation. One possible way that this improvement could be obtained is using ensemble techniques.

Ensemble learning is a method that combines the decisions of multiple learning algorithms to produce an output. The most common way of doing this is through a voting scheme, this will typically be a weighted or dynamically weighted voting scheme that allows the theoretical best output to be found. It has been found in the literature that under circumstances where errors of individual classifiers are uncorrelated then by using an ensemble instead of a single classifier the classification will be more accurate (Sewell, 2011).

The Department of Energy at the Polytechnic University of Milan have focussed on the applications of ensemble learning for identification with a focus on nuclear power station fault identification. Baraldi (Baraldi, et al., 2015) has investigated the use of data-driven techniques for signal reconstruction on data for fault detection through a comparison of three different methods. The three methods compared are Auto-Associative Kernel Regression (AAKR), Fuzzy Similarity (FS) and Elman Recurrent Neural Network (RNN). These methods were found to have inherent advantages and disadvantages for detecting faults in temperature data in a rotating machine. However, these disadvantages could be overcome by considering them as part of an ensemble. Ensembles of classifiers have been shown to be useful for classifications of faults in nuclear power stations previously. Razavi-Far (Razavi-Far, et al., 2012) has shown the use of incremental learning with an ensemble of classifiers for classification of faults in a feedwater system of a boiling water reactor allowing for new classes of faults to be added in to the ensemble as they are found during operation. Baraldi has shown how (Baraldi, et al., 2010) an ensemble of principal component analysis (PCA) models can be used to increase accuracy and robustness in signal reconstruction. The method

uses dynamic aggregation of the signals with the aggregation parameters being altered based on the data that is being reconstructed. It has shown to have been used in a pressurised water reactor (PWR) pressurizer to perform signal reconstruction on noisy of faulty sensors. Liu (Liu, et al., 2015) has shown how it is possible to use an ensemble of regression techniques to predict the remaining useful life of a reactor coolant pump from time series data. Many individual probabilistic support vector regression modules are trained and are weighted based on fuzzy similarity analysis between the training sets used for each module and the input data to generate the prediction.

Another improvement could be through the choice of the input features that are used to detect cracks. Currently the features are chosen based on the elicited features that human operators use. Future work could be undertaken to use a data driven approach to find suitable new features that may improve performance as there could be underlying relationships in the data that have yet to be discovered.

Currently the semi-supervised crack detection algorithm detects the presence of cracks within the graphite bricks from FGLT data. However, an improvement that should be pursued in future could be the identification of what type of crack is present. The best implementation should be to do a two-stage approach where only the bricks which have been identified as cracked should be labelled with a crack type. This could be beneficial to the operators of the UK's AGRs as not all cracks are of equal importance. For instance, keyway root cracks are going to be the most important cracks to monitor for the remainder of the AGRs' operational life.

6.3.3 Active Learning

Active or query-based learning can be defined as a machine learning methodology where the algorithms improve by asking questions of the data (Settles, 2012). One of the key features of active learning is that it is an interactive learning process between the algorithm and the labelling source, which is sometimes called an oracle. The labelling source could be a domain expert in the case of nuclear condition monitoring, who can manually label the data that is being queried.

There are three main methods of querying the oracle with samples to provide labels for them. These are query synthesis, stream-based selective sampling and pool-based sampling. Query synthesis involves the generation of theoretical samples which occur

inside the sample space. This method has the advantages of being able to generate a sample which can be very impactful to constructing the overall desired function. However, it may also produce meaningless data that is being queried such as situations that would not be able to occur or physically not make sense. Stream-based selective sampling involves the algorithm assessing every sample and deciding if that sample should be queried. The samples which are chosen to be queried are done so based on uncertainty criteria and that labelling that data will improve the overall algorithm. This situation may occur if the algorithm is having trouble assigning a class to the sample or conversely if it cannot be decided which class it belongs to out of multiple classes. Pool-based sampling on the other hand looks at a large sample of the labelled and unlabelled data and then decides based on the statistical properties of the distribution which samples will be most useful to query. The samples that would most likely be queried would be ones which are on or near the original boundaries decided by the initial labelled data. In a sense this method of active learning is very close to what semi-supervised learning is but with the ability to query the data.

A kernel principal component analysis (KPCA) active learning algorithm has been used in ship machinery condition monitoring to select optimal training samples (Li-Lin & Hai-chao, 2010). This was the solution to the problem which occurred where the training data was causing the algorithm to generalise too much to the results obtained from the test environment. Zhao has used a combination of semi-supervised learning and active learning to create a new algorithm that aims to optimise the samples that are chosen to be queried, the algorithm is called active Laplacian support vector machine (A-LapSVM) (Zhao, et al., 2011). This algorithm was shown to be an improvement over various other SVMs for the application of bearing diagnosis and gear fault detection when available labelled sample numbers are low. Zhang performs a comparison of two different types of support vector machine using active learning algorithms (Zhang, et al., 2006). These are active SVMs (ASVM) and probabilistic active SVMs (ProASVM). It is then shown that for rolling bearings faults then ProASVM performs better at diagnosing faults than ASVM.

Active learning could be applied in future work as a substitute to or in combination with the semi-supervised learning methods used in graphite brick crack detection. It

could be particularly useful in the co-training methods discussed in chapter 5, where the unlabelled data with the highest confidence is provided a label. Before that occurs, the designated data could be shown to an oracle (a human with specialist knowledge of the problem) to ensure that they are being given the correct classification. Alternatively, the choice of the data that is being queried could be changed in co-training to obtain better results. The queried data could instead be chosen as the data that the individual algorithms are the most unsure about, as classifying these instances will provide the greatest overall gain in information about the feature domain. A proposed method like this could be very useful in reducing the manual operators' time when analysing data. They would have to manually analyse a small section of the data as opposed to all the data.

6.3.4 Incremental Learning

Incremental machine learning is where an algorithm can learn new information from new instances of training data, without forgetting previously learned information and without requiring previously used training data (Polikar, 2000). This has the benefit of not requiring old historical training data to be merged with new training data when there is a need for the algorithm to be updated. This method of learning is mainly used when training data is obtained in a batch process. One particular benefit of incremental learning is that if the new training data has a new classification then it can be included and learned as part of the algorithm. This is particularly beneficial in nuclear power station fault detection and diagnosis due to the potential introduction of previously unseen faults as the power stations are operating long past their initial design lives.

Incremental learning has been shown to have use in condition monitoring for algorithms that need to be updated with new information over time. Vilakazi (Vilakazi & Tshilidzi, 2007) compares two incremental machine learning methods, Learn++ (Polikar, et al., 2001) and Fuzzy Adaptive Resonance Theory Mapping (FAM), on their performance of classifying high voltage bushing faults. It was found that FAM relied heavily on learning information in a correct order and implementing it with optimal parameters was challenging. Learn++ on the other hand, has issues of classifying new learned classes if there are only a small number of classifiers in the ensemble that have been trained to classify them. Until an updated version of the algorithm is used (which is discussed in the next section) the only solution is to match

the number of previously trained classifiers which are not trained on data containing that classification. This can lead to a highly computationally intensive algorithm when multiple new classes are required. Emmanouilidis (Emmanouilidis, et al., 2010) proposed a condition monitoring framework that would firstly use novelty detection that could identify conditions that had previously not been trained on. It would then indicate this as a potential new fault and could use that data as potential training data. Using incremental learning in cases like this show that diagnostic and prognostic approaches can be applied to problems that do not have an abundance of available data at the start of life but will progressively generate data to use. It would not be advised to use this method of novelty detection for training on safety critical systems, as it involves trying to detect faults appearing that have never been present before. Liu has developed an incremental learning algorithm that uses an Ellipsoid-ARTMAP classifier that is tested on a tool condition monitoring system that is used in milling (Liu, et al., 2015). The reason why incremental learning is applicable for monitoring the wear of machining tools is that there are such a large number of possible outcomes. It would not be practical to create training data for every type of wear. Therefore, the algorithm must adapt over time as wear occurs while retaining its original training information.

6.3.4.1 Learn++.NC Nuclear Condition Monitoring

Learn++.NC algorithm is an incremental learning ensemble classification algorithm which has been based on the works of the original Learn++ algorithm (Polikar, et al., 2001). The Learn++ algorithm is a supervised learning algorithm that allows for new classes of data to be added to the classifier without forgetting the original training of the classifiers. The main flaw with the Learn++ approach was that under many circumstances the number of new classifiers added would be required to be equal to the number of existing classifiers to be able to “outvote” them when a decision is being made. This is where the main difference comes in with the Learn++.NC algorithm (Muhlbaier, et al., 2009). The Learn++.NC algorithm introduces DW-CAV (dynamically weighted consult and vote) which allows a small number of classifiers to gain more weight in voting if they all share a confidence in a class that other classifiers have not been trained on.

The algorithm has two loops with one being nested inside the other. The outer loop determines the set of training data that is used (and what classes can be classified) and the inner loop is used to create a set number of classifiers on the current set of training data. Using weak classifiers, the ensemble is created. If all the classifiers are trained on all the data, it is just simply a weighted confidence calculation based on how accurate the results were on the validation training data. Once there are situations where new training data is being added that has unseen classes, DW-CAV calculates a value between zero and one for every class based on the number of classifiers that chose that class divided by the number of classifiers that could have chosen that class. The difference between this value and one is then used to multiply the weights of the classifiers that did not choose that class. Therefore, if the ratio is small it will not change the weights much but if the ratio is close to one, stating that all classes which can classify that class do, then the weights of the other classifiers will be reduced greatly.

The Learn++.NC algorithm was shown to have applications in nuclear power generation condition monitoring for classifying faults in a feedwater system of a boiling water reactor (Razavi-Far, et al., 2012). The premise for this is that when new characteristics of leaks appear, they can be trained and included in the ensemble for future classifications. The novel idea of this approach is that it also allows for detection of new classes based on the Normalized Weighted Average of the entire ensemble Outputs (NWAO). As the name of NWAO suggests each individual input data in the system is subjected to all the classifiers. Then the output of these classifiers is normalised allowing an average value to be produced. For each class a value of between zero and one will be produced. A predetermined high and low threshold will have been set based on how sufficiently different a new class is required to be. For a new class to be defined the highest NWAO class must be less than the high threshold while at least one more class has a value greater than the lower threshold. Once these criteria are met repeatedly, the algorithm can then use this new data as training data. This was shown to work on simulated data where the transients of leaks in the feedwater system were being classified. When data about a location of a leak that was not previously classified is applied to the system the NWAO criteria is matched. This provided continuous data about the new classification which then was used to update

the Learn++.NC framework allowing for future instances of that class to be correctly classified.

6.3.5 Future Framework

One of the main problems with the crack detection algorithm is that there is currently not enough available ground truth data for confident results. Therefore, an evaluation of incremental machine learning approaches with respect graphite crack detection in nuclear condition monitoring should be performed. This could be demonstrated to provide decision support for the continued operation of AGR nuclear power stations. Firstly, an investigation of the suitability of batch incremental learning compared to traditional retraining of all available historical data when new ground truth training data is available at statutory outages should be undertaken. Furthermore, the ability for incremental machine learning to learn new types of degradation or adapt to changing classifications of previously learned classes should be explored for its application on UK nuclear condition monitoring data. In this approach new types of AGR graphite bricks cracks can be learned while still retaining the knowledge of the historically known crack types. This would demonstrate that not all future faults need to be known at the inception of the algorithm. An example of this is that currently the historical supply of data contains a very limited supply of certain types of cracks, this is particularly apparent with keyway root cracks which until 2015 has seen very limited examples of in the data. This means that if there is a similar new classification found later in the reactors' life substantial retraining would not be required to account for it. The framework that would be involved in the incremental graphite brick detection can be seen in Figure 6.1. This framework would combine the current implementation approach of using semi-supervised machine learning and combine it with incremental learning.

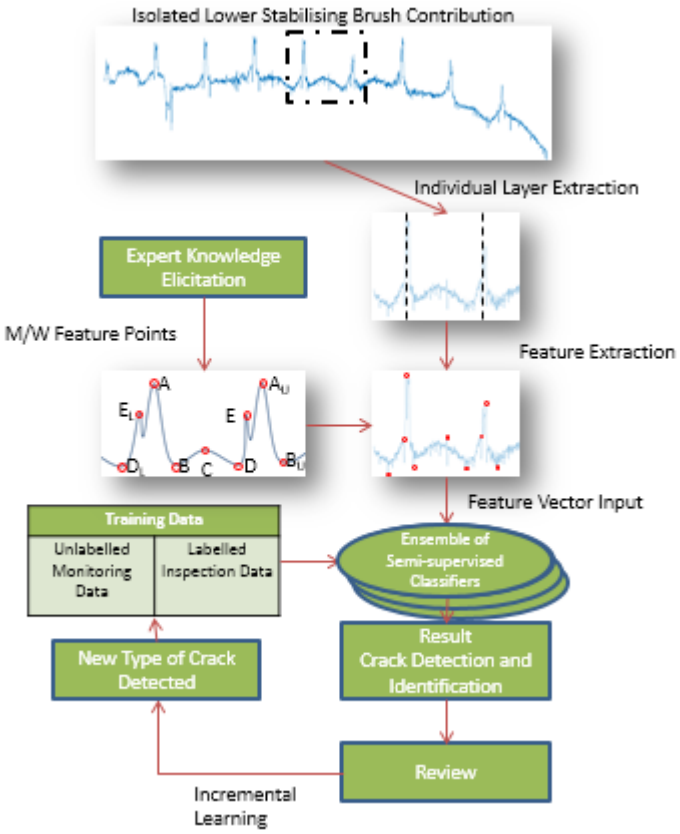


Figure 6.1 – Framework for future graphite brick crack identification which incorporated ensemble learning and incremental machine learning.

References

Abbasion, S., Rafsanjani, A., Farshidianfar, A. & Irani, N., 2007. Rolling Element Bearnings Multi-Fault Classification Based on the Wavelet Denoising and Support Vector Machine. *Mechanical Systems and Signal Processing,* Volume 21, pp. 2933-2945.

Agarwal, V. et al., 2013. *Online Monitoring of Plant Assets in the Nuclear Industry.* New Orleans, PHM Society.

Baraldi, P., Cammi, A., Mangili, F. & Zio, E. E., 2010. Local Fusion of an Ensemble of Models for the Reconstruction of Faulty Signals. *IEEE Transactions on Nuclear Science,* 57(2), pp. 793-806.

Baraldi, P., Di Maio, F., Genini, D. & Zio, E., 2015. Comparison of Data-Driven Reconstruction Methods for Fault Detection. *IEEE Transactions on Reliability,* 64(3), pp. 852-860.

Baraldi, P., Maio, F. D. M. & Zio, E., 2013. Unsupervised Clustering for Fault Diagnosis in Nuclear Power Plant Components. *International Journal of Computational Intelligence Systems,* 6(4), pp. 764-777.

Barbieri, F., Hines, J. W., Sharp, M. & Ventirini, M., 2015. Sensor-Based Degradation Prediction and Prognostics for Remaining Useful Life Estimation: Validation on Experimental Data of Electric Motors. *International Journal of Prognostics and Health Management,* 6(Special Issue Nuclear Energy).

Bartlett, E. B. & Uhrig, R. E., 1992. Nuclear Power Plant Status Diagnostics Using an Artificial Neural Network. *Nuclear Technology,* 97(3), pp. 272-281.

Benbouzid, M. E. H., 2000. A Review of Induction Motors Signature Analysis as a Medium for Faults Detection. *IEEE Transactions on Industrial Electronics,* 47(5), p. 984.

Blum, A. & Mitchell, T., 1998. *Combining Labeled and Unlabeled Data with Co-Training.* Madison, Wisconsin, s.n.

Bond, L. J., Ramuhalli, P. & Tawfik, M. S., 2011. *Prognostics and Life Beyond 60 Years for Nulear Power Plants.* Montreal, QC, s.n., pp. 1-7.

Booler, O. A., Slater, I. J., Baylis, S. & Steer, A. G., 2016. *Investigations Into Prompt Secondary Cracking in AGR Graphite Fuel Bricks Using the eXtended Finite Element Method.* Southampton, EMAS Publishing.

Bowler, N. & Liu, S., 2015. Aging Mechanisms and Monitoring of Cable Polymers. *International Journal of Prognostics and Health Management,* 6(Special Issue Nuclear Energy).

Bradford, M. B., 2015. *Reactor Ubspection and Monitoring and its Role in the Reactor Safety Case.* Nottingham, EMAS: Publishing.

Bradford, M. R., 2016. *Dealing with the "Known Unknowns": Improving Forecastingof Keyway Root Cracking.* Southampton, EMAS Publishing.

Brasier, S. & Rogers, S., 2011. *Solid-Body Simulation of the Dynamic Response of an Array of Graphite Bricks.* Birmingham, The Royal Society of Chemistry.

British Electricity International, 1992. *odern Power Station Practice. Volume J, Nuclear Power Generation : Incorporating Modern Power System Practice.* 3rd ed. Oxford: Pergamon.

British Energy, 2005. *10-Year Life Extension at Dungeness B Nuclear Power Station.* [Online]
Available at: https://web.archive.org/web/20060322013831/http:/www.british-energy.com/article.php?article=99
[Accessed 02 August 2018].

British Energy, 2007. *5 Year Life Extension at Hinkley Point B Nuclear Power STation.* [Online]
Available at: https://web.archive.org/web/20080804130343/http:/www.british-energy.com/article.php?article=219
[Accessed 2 August 2018].

British Energy, 2007. *Life Extension of Hunkley Point B and Hunterston B Power Stations.* [Online]
Available at: https://web.archive.org/web/20080804130421/http://www.british-energy.com/article.php?article=218
[Accessed 2 August 2018].

Brocklehurst, J. E. & Kelly, B. T., 1993. Analysis of the dimensional changes and structural changes in polycrystaline graphite under fast neutron irradiation. *Carbon,* 31(1), pp. 155-178.

Castro, E., 2005. *Investigation of Degradation of AGR Graphite Core Geometry using a Whole Core Scale Model.* Cardiff, The Royal Society of Chemistry.

Chapelle, O., Schölkopf, B. & Zien, A., 2006. *Semi-Supervised Learning.* Cambridge, Massachusetts: The MIT Press.

Chawla , N. V., Bowyer, K. W., Hall, L. O. & Kegelmeyer, W. P., 2002. SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research ,* Issue 16, pp. 321-357.

Chen, F.-C. & Jahanshahi, M. R., 2018. NB-CNN: Deep Learning-Based Crack Detection Using Convolutional Neural Network and Naïve Bayes Fusion. *IEEE Transactions on Industrial Electronics,* 65(5), pp. 4392-4400.

Chen, S., Cerda, F., Rizzo, P. & Bielak, J., 2014. Semi-Supervised Multiresolution Classification Using Adaptive Graph Filtering With Application to Indirect Bridge Structurak Health Monitoring. *IEEE Transactions on Signal Processing,* 62(11), pp. 2879-2893.

Cholette, M. E., Celen, M., Djurdjanovic, D. & Rasberry, J. D., 2013. Condition Monitoring and Operational Decision Making in Semiconductor Manufacturing. *IEEE Transactions on Semiconducto Manufacturing,* 26(4), pp. 454-464.

Coble, J. & Hines, J. W., 2011. Applying the General Path Model to Estimation of Remaining Useful Life. *International Journal of Prognostics and Health Management,* 2(1).

Coble, J. et al., 2015. A Review of Prognostics and Health Management Applications in Nuclear Power Plants. *International Journal of Prognostics and Health Management,* Volume 6, p. 22.

Cole-Baker, A. & Reed, J., 2005. Measurement of AGR Graphite Fuel Brick Shrinkage and Channel Distortion. In: G. B. Neighbour, ed. *Management of Aging Processes in Graphite Reactor Cores.* Cambridge: RSC Publishing, pp. 201-208.

Cowell, G. & Steer, A. G., 2016. *GCORE Seismic Analysis of AGR Cores: Development of Modelling and Analysis Methods for Later-In-Life Core Conditions.* Southampton, EMAS Publishing.

Crump, T. et al., 2016. *3D Dynamic Crack Propagation Within AGR Graphite Brick Slices.* Southampton, EMAS Publishing.

Dai, Y., Xue, Y. & Zhang, J. Z., 2015. Vibration-Based Milling Condition Monitoring in Robot-Assisted Spine surgery. *IEEE/ASME Transactions on Mechatronics,* 20(6), pp. 3028-3039.

Davies, B. C. et al., 2016. *Blackstone MTR Phase III: Characterisation of Graph Material Property Changes at Heysham 2 and Torness AGR Power Stations.* Southampton, EMAS Publishing.

Dean, J., Patterson, D. & Young, C., 2018. A new Golden Age in Computer Architecture: Empowering the Machine-Learning Revolution. *IEEE Micro,* 38(2), pp. 21-29.

Dekdouk, B., Chapman, R., Brown, M. & Peyton, A. J., 2012. Evaluating the conductivity distribution in isotropic polycrystalline graphite using spectroscopic eddy current technique for monitoring weight loss in advanced gas cooled reactors. *NDT&E International,* Volume 51, pp. 150-159.

Dihoru, L., Crewe, A., Taylor, C. & Horgan, T., 2011. *Shaking Table Experimental Programme.* Birmingham, The Royal Society of Chemistry.

Dihoru, L. et al., 2016. *Shaking Table Testing of an Advanced Gas Cooled Reactor Core Model with Degraged Components.* Southampton, EMAS Publishing.

Dihoru, L. et al., 2016. *Shaking Table Testing of an Advanced Gas-Cooled Reactor Core Model with Degraded Components.* Southampton, EMAS Publishing.

EDF Energy, 2012. *EDF Energy Announces Seven Year Life Extension to Hinkley Point B and Hunterston B Nuclear Power Stations.* [Online]
Available at: https://www.edfenergy.com/media-centre/news-releases/edf-energy-announces-seven-year-life-extension-hinkley-point-b-and
[Accessed 2 August 2018].

EDF Energy, 2012. *EDF Group - 2011 Full-Year Financial Results.* [Online]
Available at: https://www.edfenergy.com/media-centre/news-releases/edf-group-2011-full-year-financial-results
[Accessed 2 August 2018].

EDF Energy, 2015. *EDF ENergy Announces 10 More Years for Dungeness B.* [Online]

Available at: https://www.edfenergy.com/media-centre/news-releases/edf-energy-announces-10-more-years-dungeness-b
[Accessed 2 August 2018].
EDF Energy, 2016. *EDF Group Results 2015 - Highlights for EDF Energy in UK.* [Online]
Available at: https://www.edfenergy.com/media-centre/news-releases/edf-group-results-2015-%E2%80%93-highlights-edf-energy-uk
[Accessed 2 August 2018].
EDF Energy, 2018. *Monitoring the Core.* [Online]
Available at: https://www.edfenergy.com/energy/monitoring
[Accessed 2 August 2018].
Emmanouilidis, C. et al., 2010. *Condition Monitoring Based on Incremental Learning and Domain Ontology for Condition-Based Maintenance.* Cernobbio, s.n.
Girard, S. et al., 2013. Sensitivity Analysis and Dimension Reduction of a Steam Generator Model for Clogging Diagnosis. *Reliability Engineering & System Safety,* Volume 113, pp. 143-153.
Golyandina, N., Nekrutkin, V. & Zhigljavsky, A. A., 2001. *Analysis of Time Series Structure SSA and Related Techniques.* London: CHAPMAN & HALL/CRC.
Golyandina, N. & Zhigljavsky, A., 2013. *Singular Spectrum Analysis for Time Series.* London: Springer.
Gouriveau, R., Ramasso, E. & Zerhouni, N., 2013. Strategies to Face Imbalanced and Unlabelled Data in PHM Applications. *Chemical Engineering Transactions,* Volume 33, pp. 115-120.
Hacker, P. J., Neighbour, G. B. & McEnaney, B., 2000. The Coefficient of Thermal Expansion of Nuclear Graphite with Increasing Thermal Oxidation. *Journal of Physics D: Applied Physics,* Volume 33, pp. 991-998.
Hadad, K., Mortazavi, M., Mastali, M. & Safavi, A. A., 2008. Enhanced Neural Network Based Fault Detection of a VVER Nuclear Power Plant With the Aid of Principal Component Analysis. *IEEE Transactions on Nuclear Science,* 55(6), pp. 3611-3619.
Heaton, J., 2008. *Introduction to Neural Networks with Java.* 2nd ed. St. Louis: Heaton Research, Inc.

He, H. & Garxia, E. A., 2009. Learning from Imbalanced Data. *IEEE Transactions on Knowledge and Data Engineering,* 21(9), pp. 1263-1284.

Hodge, V. J., O'Keefe, S., Weeks, M. & Moulds, A., 2015. Wireless Sensor Networks for Condition Monitoring in the Railway Industry: A Survey. *IEEE Transactions on Intelligent Transportation Systems,* 16(3), pp. 1088-1106.

Hodgkims, A. et al., 2008. *A Consideration of the Effect of Radiolytic Oxidation on Fracture Behaviour of Reactor Core Graphites.* Nottinghma, The Royal Society of Chemistry.

Hodgkins, A. D. et al., 2011. *Quasi-Brittle Fracture Concepts to Improve Structural Integrity Assessments of High Oxidation Weight Loss Graphite Components.* Birmingham, The Royal Society of Chemistry.

Hodgkins, A. et al., 2005. *Crack Propagation Resistance and Damage Mechanisms in Nuclear Graphite.* Cardiff, The Royal Society of Chemistry.

Horgan, T. & Webster, E., 2014. *Shaking Table Experiment for Graphite Core Plant Lifetime Extension.* Nottingham, EMAS Publishing.

Hu, C., Youn, B. D. & Kim, T., 2011. *Semi-Supervised Learning with Co-Training for Data-Driven Prognostics.* Washington, s.n.

Hu, C., Youn, B. D. & Kim, T., 2012. *Semi-Supervised Learning with Co-Training for Data-Driven Prognostics.* Denver, s.n.

Hu, C., Youn, B. D., Kim, T. & Wang, P., 2015. A Co-Training-Based Approach for Prediction of Remaining Useful Life Utilizing Failure and Suspension Data. *Mechanical Systems and Signal Processing,* Volume 62-63, pp. 75-90.

International Atomic Energy Agency, 2008. *On-Line Monitoring for Improving Performance of Nuclear Power Plants Part 2: Process and Component Condition Monitoring and Diagnostics.* Vienna: International Atomic Energy Agency.

Jin, X. et al., 2011. Anomaly Detection in Nuclear Power Plants via Symbolic Dynamic Filtering. *IEEE Transactions on Nuclear Science,* 58(1), pp. 277-288.

Kilundu, B., Dehombreux, P. & Chiementin, X., 2011. Tool Wear Monitoring by Machine Learning Techniques and Sigular Spectrum Analysis. *Mechanical Systems and Signal Processing,* 25(1), pp. 400-415.

Knowles, M., Baglee, D. & Wermter, S., 2011. *Reinforcement Learning for Scheduling of Maintenance.* London, s.n.

Kralj, B., Humphreys, S. J. & Duncan, B. G. J., 2005. *Seismic Modelling of an AGR Nuclear Reactor Core.* Cardiff, The Royal Society of Chemistry.

Krogel, M. A. & Scheffer, T., 2004. Multi-Relational Learning, Text Mining, and Semi-Supervised Learning for Functional Genomics. *Machine Learning,* 57(1-2), pp. 61-81.

Kunpeng, Z., San, W. Y. & Soon, G. H., 2009. Wavelet Analysis of Sensor Signals for Tool Condition Monitoring: A Review and Some New Results. *International Journal of Machine Tools & Manufacture,* 49(7-8), pp. 537-553.

Kyaw, S. T. et al., 2014. Modelling Crack Growth within Graphite Bricks due to Irradiation and Radiolytic Oxidation. *Procedia Materials Science,* Volume 3, pp. 39-44.

Law, K., West, G., Murray, P. & Lynch, C., 2017. *Towards Extracting 3-D Structural Representations of AGR Core Fuel Channels from 2-D In-Core Inspection Videos.* Gyeongju, s.n.

Leguillon, D. et al., 2016. *Modelling Crack Initiation on AGR Graphite Bricks.* Southampton, EMAS Publishing.

Li-Lin, C. & Hai-chao, Z., 2010. *The Study of KPCA Active Learning Method based on the ROC Curve.* Shanghai , s.n.

Liu, C., Wang, G. & Li, Z., 2015. Incremental Learning for Online Tool Condition Monitoring using Ellipsoid ARTMAP Network Model. *Applied Doft Computing,* Volume 35, pp. 186-198.

Liu, D., Mingard, K., Lord, O. & Flewitt, P. E. J., 2016. *Overview of the Micromechanical Testing of Graphites.* Southampton, EMAS Publishing.

Liu, J., Seraoui, S., Vitelli, V. & Zio, E., 2013. Nuclear Power Plant Components Condition Monitoring by Probabilistic Support Vector Machine. *Annals of Nuclear Energy,* Volume 56, pp. 23-33.

Liu, J., Vitelli, V., Enrico, Z. & Saraoui, R., 2015. A Novel Dynamic-Weighted Probabilistic Support Vector Regression-Based Ensemble for Prognostics of Time Series Data. *IEEE Transactions on Reliability,* 64(4), pp. 1203-1213.

Livshitz, A., Chudnovsky, B. H. & Bukengolts, B., 2004. *On-Line Condition Monitoring and Diagnostics od Power Distribution Equipment.* s.l., s.n.

Loutas, T. H., Roulias, D. & Georgoulas, G., 2013. Remaining Useful Life Estimation in Rolling Bearings Utilizing Data-Driven Probabilistic E-Support Vectors Regression. *IEEE Transactions on Reliability,* 62(1), pp. 821-832.

Loyd's Register Foundation, 2014. *Forseight Review of Big Data,* London: Loyd's Register Foundation.

Ma, J. & Jiang, J., 2011. Applications of Fault Detection and Diagnosis Methods in Nuclear Power Plants: A Review. *Progress in Nuclear Energy,* 53(3), pp. 255-266.

Ma, J. & Jiang, J., 2015. Semisupervised Classification for Fault Diagnosis in Nuclear Power Plants. *Nuclear Engineering and Technology,* 47(2), pp. 176-186.

Maul, P. R., Robinson, P. R. & Steer, A. G., 2008. *Understanding AGR Graphite Brick Cracking Using Physical Understanding and Statistical Modelling.* Nottingham, The Royal Society of Chemistry.

Mazzoleni, M., Previdi, F., Scandella, M. & pispola, G., 2019. Experimental Development of a Health Monitoring Method for Electro-Mechanical Actuators of Flight Control Primary Surfacesin More Electric Aircrafts. *IEEE Access,* Volume 7, pp. 153618-153634.

McCarter, D., Shumaker, B., McConkey, B. & Hashemian, H., 2015. Nuclear Power Plant Instrumentation and Control Cable Prognostics Using Indenter Modulus Measurements. *International Journal of Prognostics and Health Management,* 6(Special Issure Nuclear Energy).

McLachlan, N., Reed, J. & Metcalfe, M. P., 1995. *AGR Core Safety Assessment Methodologies,* Berkeley: Nuclear Electric plc..

Meyer, R. M., Bond, L. J. & Ramuhalli, P., 2012. Online Condition Monitoring to Enable Extended Operation of Nuclear Power Plants. *Nuclear Safety and Simulation,* 3(1), pp. 31-50.

Moshkbar-Bakhshayesh, K. & Ghofrani, M. B., 2014. Development of a Robust Identifier for NPPs Transients Combining ARIMA Model and EBP Algorithm. *IEEE Transactions on Nuclear Science,* 61(4), pp. 2383-2391.

Moshkbar-Bakhshayesh, K. & Ghofrani, M. B., 2016. Combining Supervised and Semi-Supervised Learning in the Design of a New Identifier for NPPs Transients. *IEEE Transactions on Nuclear Science,* 63(3), pp. 1882-1888.

Muhlbaier, M. D., Topalis, A. & Polikar, R., 2009. Learn++.NC: Combining Ensemble of Classifiers With Dynamically Weighted Consult-and-Vote for Efficient Incremental Leaning of New Classes. *IEEE Transactions on Neural Networks*, 20(1), pp. 152-168.

Murphy, K. P., 2012. *Machine Learning A Probabilistic Perspective.* Cambridge, Massachusetts: The MIT Press.

Murugabatgan, B., Sanjith, M. A., Krishnakumar, B. & Murty, S. A. V. S., 2013. Roller Element Bearing Fault Diagnosis Using Singular Spectrum Analysis. *Mechanical Systems and Signal Processing*, 35(1-2), pp. 150-166.

Mu, Y. et al., 2014. Study on Fault Diagnosis Method for Nuclear Power Plant Based on Fuzzy Rough Sets and Decision Tree. *Advanced Materials Research,* Volume 981, pp. 701-705.

Nandi, S., Toliyat, H. A. & Li, X., 2005. Condition Monitoring and Fault Diagnosis of Electrical Motors - A Review. *IEEE Transactions on Energy Conversion*, 20(4), pp. 719-729.

Neighbour, G. B., 2000. Modelling of dimensional changes in irradiated nuclear graphites. *Journal of Physics D:Applied Physics,* 33(22), pp. 2966-2972.

Office for Nuclear Regulation, 2017. *Graphite Core Ageing.* [Online] Available at: http://www.onr.org.uk/civil-nuclear-reactors/graphite-core-ageing.htm [Accessed 26 3 2018].

O'Leary, D. E., 2013. Artificial Intelligence and Big Data. *IEEE Intelligent Systems*, 28(2), pp. 96-99.

Pearce, C. J. & Kaczmarczyk, Ł., 2016. *Computatuinal Modelling of Continuous Crack Propagationin Nuclear Graphite.* Southampton, EMAS Publishing.

Peng, Z. K. & Chu, F. L., 2004. Application of the Wavelet Transform in Machine Condition Monitoring and Fault Diagnostics: A Review with Bibliography. *Mechanical Systems and Signal Processing,* 18(2), pp. 199-221.

Petheric, J., 2016. *Updating the FEAT-DIFFUSE6 Brick Boundary Conditions for Graphite Weight Loss Predictions of Dungeness B.* Southampton, EMAS Publishing.

Pimentel, M. A., Clifton, D. A., Clifton, L. & Tarassenko, L., 2015. A Review of Novelty Detection. *Signal Processing,* Volume 99, pp. 215-249.

Polikar, R., 2000. Incremental Learning. In: *Algorithms for Enhancing Pattern Separability, Optimum Feature Selection and Incremental Learning with Applications to Gas Sensing Electronic Nose Systems.* s.l.:Iowa State University, pp. 144-216.

Polikar, R., Udpa, L., Udpa, S. S. & Honavar, V., 2001. Learn++: An Incremental Learning Algorithm for Supervised Neural Networks. *IEEE Transactions on Systems, Man, and Cybernetics- Part C: Applications and Reviews,* 31(4), pp. 497-508.

Potočnik, P. & Govekar, E., 2017. Semi-Supervised Vibration-Based Classification and Condition Monitoring of Compressors. *Mechanical Systems and Signal Processing,* Volume 93, pp. 51-65.

Qiao, W. & Lu, D. L., 2015. A Survey on Wind Turbine Condition Monitoring and Fault Diagnosis - Part1: Components and Subsystems. *IEEE Transactions on Industrial Electronics,* 62(10), pp. 6536-6545.

Razavi-Far, R., Baraldi, P. & Zio, E., 2012. Dynamic Weighting Ensembles for Incremental Learning and Diagnosing New Concept Class Faults in Nuclear Power Systems. *IEEE Transactions on Nuclear Science,* 59(5), pp. 2520-2530.

Razavi-Far, R., Zio, E. & Baraldi, P., 2012. Dynamic Weighting Ensembles for Incremental Learning and Diagnosing New Concept Class Faults in Nuclear Power Systems. *Nuclear Science, IEEE Transactions on (Volume:59 , Issue: 5),* pp. 2520-2530.

Reddy, M. J. B. et al., 2017. An on-line geographical information system-based condition monitoring system for 11-kv distribution line insulator. *IEEE Electrical Insulation Magazine,* 33(3), pp. 26-32.

Riley, H. et al., 2016. *GCORE Validation for Large Arrays of Radially Keyed Bricks: Cimparison of the Experimental Measurements and Computational Predictions.* Southampton, EMAS Publishing.

Robinson, P. C., Hill, B. J., Burrow, J. F. & Maul, P. R., 2016. *Statistical Modelling of Keyway Root Cracking in AGR Graphite Bricks.* Southampton, EMAS Publishing.

Robinson, P. C. & Maul, P. R., 2011. *Statistical Modelling of Graphite Properties.* Birmingham, The Royal Society of Chemistry.

Roscow, A. R., Skelton, J. & McLachlan, N., 2008. *Continuing Development of the Quarter Scale Core Model to Support the Safe Performance of AGR Cores.* Nottingham, The Royal Society of Chemistry.

Salih, H. & McLachlan, N., 2016. *SALCOR Simulation of Local AGR Core Region with Multiple Singly and Doubly Cracked Fuel Bricks and Potential Key Disengagement.* Southampton, EMAS Publishing.

Salih, H. & McLachlan, N., 2016. *SALCOR Solid Modelling and Analysis Methodology for an Ageing AGR Core: Progress and Benefits.* Southampton, EMAS Publishing.

Şeker, S., Ayaz, E. & Türkcan, E., 2003. Elman's Recurrent Neural Network Applications to Condition Monitoring in Nuclear Power Plant and Rotating Machinery. *Engineering Applications of Artificial Intelligence,* Volume 16, pp. 647-656.

Seo, I. et al., 2011. Detection of Sensor Degradation Using K-means Clustering and Support Vector Regression in Nuclear Power Plant. *Nuclear Science and Technology,* Volume 1, pp. 312-315.

Settles, B., 2012. *Active Learning.* s.l.:Morgan & Claypool Publishers.

Sewell, M., 2011. *Ensemble Learning,* London: Depertment of Conputer Science, UniversityCollege London.

Shaw, D. J., Salih, H. & McLachlan, N., 2005. *Development and Validation of Whole Core Modelling Methology for AGR Graphite Cores.* Cardiff, The Royal Society of Chemistry.

Shen, Z., Chen, X., Zhang, X. & He, Z., 2012. A Novel Intelligent Gear Fault Diagnosis Model Based on EMD and Multi-Class TSVM. *Measurement,* 45(1), pp. 30-40.

Shi, C. et al., 2018. Using Multiple-Feature-Spaces-Based Deep Learning for Tool Condition Monitoring in Ultraprecision Manufacturing. *IEEE Transactions on Industrial Electronics,* 66(5), pp. 3794-3803.

Shumaker, B. et al., 2012. *Cable Condition Monitoring for Nuclear Power Plants.* Gatlinburg, TN, s.n.

Skelton, J., 2007. *Station Core Monitoring - Experimental Work to Relate HPB/HNB Frab Trip Load to Bore Shrinkage,* s.l.: Amec.

Steer, A. G., 2005. AGR Core Design, Operation and Safety Functions. In: G. B. Neighbour, ed. *Management of Ageing Processes in Graphite Reactor Cores.* Cambridge: RSC Publishing, pp. 11-18.

Stephen, B. et al., 2009. The Use of Hidden Markov Models for Anomaly Detection in Nuclear Core COndition Monitoring. *IEEE Transactions on Nuclear Science,* 56(2), pp. 453-461.

Strachan, S. M., Stephen, B. & McArthur, S. D., 2007. *Practical Applications of Data Mining in Plant Monitoring and Diagnostics.* Tampa, IEEE Power Engineering Society General Meeting.

Tan, E., Fahad, M., Hall, G. N. & Warren, N., 2016. *Prediction of Keywar Root Cracking in the Graphite Core of Advanced Gas-cooled Nuclear Reactors.* Southampton, EMAS Publishing.

Tan, E., McNally, K. & Warren, N., 2011. *Statistical Analysis of Bore Cracking in AGRs.* Birmingham, The Royal Society of Chemistry.

Tanner, D. W. J., Becker, A. A., Sun, W. & Hyde, T. H., 2011. *A Review of Current Finite Element Models for Irradiation Creep and Failure of Nuclear Graphite.* Birmingham, The Royal Society of Chemistry.

Tavner, P., 2008. Review of Condition Monitoring of Rotating Electrical Machines. *IET Electrical Power Applications,* 2(4), pp. 215-247.

Taylor, J. E. L., Davies, M. A. & Jones, A. N., 2016. *X-Ray Tomography Analysis of the Microstructure and Porosity Distributions of Irradiated AGR Graphite.* Southampton, EMAS Publishing.

Teng, H. et al., 2016. *Cracked Graphite Brick Interaction Analysis: An Application to the AGR Graphite Core.* Southampton, EMAS Publishing.

Tzelepi, A. et al., 2016. *Micromechanistic Model of Graphite Irradiation Behaviour.* Southampton, EMAS Publishing.

Tzelepi, N. et al., 2005. *The Development of Measurement Techniques for Mechanical Properties Applicable to Small Reactor Graphite Samples.* Cardiff, RSCPublishing.

Vapnik, V. N., 1998. *Statistical Learning Theory.* New York: Wiley.

Vilakazi, C. B. & Tshilidzi, M., 2007. *Online Incremental Learning for High Voltage Bushing Condition Monitoring.* Orlando, s.n.

Vishwakarma, M., Purohit, R., Harshlata, V. & Rajput, P., 2017. Vibration Analysis & Condition Monitoring for Rotating Machines: A Review. *Materials Today: Proceedings*, 4(2), pp. 2659-2664.

Wallace, C. J., West, G. M., J, J. G. & McArthur, S. D. J., 2010. *Control Rod Monitoring of Advanced Gas-Cooled Reactors.* Las Vegas, s.n.

Watson, C. E., Robinson, P. C. & Maul, P. R., 2011. *The Role of Data Visualisation in COre Inspection Decision Making.* Birmingham, s.n.

Watson, C. E., Robinson, P. C. & Maul, P. R., 2011. *The Role of Data Visualisationin Core Inspection Descision Making.* Birmingham, The Royal Society of Chemistry.

Wechsler, A. et al., 2012. Condition Monitoring of DC-Link Capacitors in Aerospace Drives. *IEEE Transactions of Industry Applications*, 48(6), pp. 1866-1874.

Welz, Z. et al., 2015. Improved Heat Exhanger Lifecycle Prognostic Methods for Enhanced Light Water Reactor Sustainability. *International Journal of Prognostics and Health Management*, 6(Special Issure Nuclear Energy).

West, G. M. et al., 2006. Data Mining Reactor Fuel Grab Load Trace Data to Support Nuclear Core Condition Monitoring. *IEEE Transactions on Nuclear Science*, 53(3), pp. 1494-1503.

West, G. M., Jahn, G. J., McArthur, S. D. J. & Reed, J., 2007. Graphite Core Condition Monitoring through Intelligent Analysis of Fuel Grab Load Trace Data. In: G. B. Neighbour, ed. *Management of Ageing Processes in Graphite Reactor Cores.* Cambrisge: RSC Publishing, pp. 232-239.

West, G. M., McArthur, S. D. & Towle, D., 2008. *BETA: A System for Automated Intelligent Analysis of Fuel Grab Load Trace Data for Graphite Core Condition Monitoring.* Nottingham, The Royal Society of Chemistry.

West, G. M., McArthur, S. D. & Towle, D., 2008. *BETA: A System for Automated Intelligent Analysis of Fuel Grab Load Trace for Graphite Core Condition Monitoring.* Nottingham, s.n.

West, G., Murray, P., Marshall, S. & McArthur, S., 2015. Improved Visual Inspection of Advanced Gas-Cooled Reactor Fuel Channels. *International Journal of Prognostics and Health Management*, 6(Special Issue Nuclear Energy PHM).

West, G. M., Wallace, C. J. & McArthur, S. D., 2014. Combining Models of Behavior with Operational Data to Provide Enhanced Condition Monitoring of AGR Cores. *Nuclear Engineering and Design,* Volume 272, pp. 11-18.

West, G. M., Wallace, C. & McArthur, S. D., 2014. Combining models of behavior with operational data to provide enhanced condition monitoring of AGR cores. *Nuclear Engineering and Design,* Volume 272, pp. 11-18.

Wickham, A. J., 2016. *Revisiting AGR Graphite Weight Loss Predictions: A First Principles Review of the Basis and Application of the FEAT-DIFFUSE Code.* Southampton, EMAS Publishing.

Wilkinson, S., Tzelepi, A., Davies, B. C. & Steer, A. G., 2016. *Innovate UK Project: Influence of Creep and Geometry on Strength ofIrradiated Graphite Components.* Southampton, EMAS Publishing.

World Nuclear Assocation, 2015. *Nuclear Power in the United Kingdom.* [Online] Available at: http://www.world-nuclear.org/info/Country-Profiles/Countries-T-Z/United-Kingdom/

[Accessed 10 January 2016].

Wu, X., Zhu, X., Wu, G.-Q. & Ding, W., 2014. Data Mining with Big Data. *IEEE Transactions on Knowledge and Data Engineering,* 26(1), pp. 97-107.

Yang, D., Usynin, A. & Hines, J. W., 2006. *Anomaly-Based Intrusion Detection for SCADA Systems.* Albuquerque, American Nuclear Society.

Yuan, J. & Liu, X., 2013. Semi-Supervised Learning and Condition Fusion for Fault Diagnosis. *Mechanical Systems and Signal Processing,* 38(2), pp. 615-627.

Yu, C. et al., 2012. *Classification With Cooperative Semi-Supervised Learning Using Bridge Structural Health Data.* Yantai, s.n.

Zhang, Z., Lv, W. & Shen, M., 2006. Active Learninf of Support Vector Machine for. *Advances in Neural Netowrks,* Volume 3973, pp. 390-395.

Zhao, X., Li, M., Xu, J. & Gangbing, S., 2011. An Effective Procedure Exploiting Unlabeled Data to Build a Monitoring System. *Exper Systems with Applications,* Volume 38, pp. 10199-10204.

Zhou, Z.-H. & Li, M., 2005. Tri-Training: Exploiting Unlabelled Data Using Three Classifiers. *IEEE Transactions on Knowledge and Data Engineering,* 17(11), p. 1529.

Zhu, X. & Ghahramani, Z., 2002. *Learning from Labeled and Unlabelled Data with Label Propagation.* s.l.:s.n.

Zhu, X. & Goldberg, A. B., 2009. *Introduction to Semi-Supervised Learning.* San Rafael, CA: Morgan & Claypool Publishers.

Zio, E., Baraldi, P. & Pedroni, N., 2006. Selecting Features for Nuclear Trnsients Classification by Means of Genetic Algorithms. *IEEE Transactions on Nuclear Science,* 53(3), pp. 1479-1493.

Zio, E., Baraldi, P. & Popescu, I. C., 2009. A Fuzzy Decision Tree Method for Fault Classification in the Steam Generator of a Pressurized Water Reactor. *Annals of Nuclear Energy,* Volume 36, pp. 1159-1169.