# Strathprints Institutional Repository

West, G.M. and Jahn, G.J. and McArthur, S.D.J. and McDonald, J.R. and Reed, J. (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. ISSN 0018-9499

http://strathprints.strath.ac.uk/

# Data Mining Reactor Fuel Grab Load Trace Data to Support Nuclear Core Condition Monitoring

Graeme. M. West, Gordon. J. Jahn, S.D.J. McArthur *Member, IEEE*, James. R. McDonald *Member, IEEE* and Jim Reed

*Abstract*— **A critical component of an advanced-gas cooled reactor (AGR) station is the graphite core. As a station ages, the graphite bricks that comprise the core can distort and may eventually crack. As the core cannot be replaced the core integrity ultimately determines the station life. Monitoring these distortions is usually restricted to the routine outages, which occur every few years, as this is the only time that the reactor core can be accessed by external sensing equipment. However, during weekly refueling activities measurements are taken from the core for protection and control purposes. It is shown in this paper that these measurements may be interpreted for condition monitoring purposes, thus potentially providing information relating to core condition on a more frequent basis. This paper describes the data-mining approach adopted to analyze this data and also describes a software system designed and implemented to support this process. The use of this software to develop a model of expected behavior based on historical data, which may highlight events containing unusual features possibly indicative of brick cracking, is also described. Finally, the implementation of this newly acquired understanding in an automated analysis system is described.**

*Index Terms*—**Fission Reactor Control and Monitoring, Data Processing, Power Generation, Nuclear**

## I. INTRODUCTION

WITHIN the UK, the Advanced Gas-cooled Reactor (AGR) nuclear power stations are approaching the end of their predicted operational lives. The major factor that dictates the life of a station is the condition of the graphite reactor cores, which distorts over time with the prolonged exposure to heat and radiation. Currently, it is proposed that the operational lifetime of the plants could be extended if the distortions of the reactor cores are not as severe as initially predicted, and that it can be demonstrated that the reactors are still safe to operate. The reactor core is composed of a large number of interlocked hollow graphite bricks that form channels into which the uranium fuel and the control rods are inserted. Over time, heat and radiation cause stresses to build up within these bricks causing them to shrink and distort, and they may eventually crack. To date, these distortions have been within the expected and predicted limits. The continued operation of the plants have been supported through rigorous monitoring and inspection processes, including core channel diameter measurements, ovality measurements and chemical analysis of trepanned core samples, which take place during planned outages approximately every three years. However, as the core becomes older, and the distortions become more severe, the need for increased information relating to core condition becomes greater.

One proposed source of information is the fuel grab load trace (FGLT) data, which is routinely gathered during reactor refueling. Although not originally intended for core condition monitoring purposes, the FGLT data contains a contribution from a frictional interface between the fuel channel wall and the fuel assembly. Changes in the fuel channel shape are reflected in the FGLT data and if this relationship can be understood then it may be possible to derive information relating to the condition of the core from this data. Also, refueling is undertaken on a weekly basis and the FGLT data is kept as part of station records, a requirement of the nuclear generating license, providing a more frequent, all be it less comprehensive, source of information relating to the core condition than that obtained during outages. In addition, the records stored at station would allow historic analysis of this data to be performed.

The key challenges of this work include developing an understanding of data that was not originally intended for monitoring the condition of reactor core bricks. This understanding has then been used to develop a method for assessing new refueling data for the presence of severe distortions, and ultimately, cracks.

Graeme West is with the Advanced Electrical Systems Research Group, Institute of Energy and Environment, University of Strathclyde, Glasgow, G1 1XW (phone: +44 (0)141 548 5857, fax: +44 (0)141 548 4872, e-mail: graeme.west@strath.ac.uk)

Gordon Jahn is with the Advanced Electrical Systems Research Group, Institute of Energy and Environment, University of Strathclyde, Glasgow, G1 1XW (e-mail: gordon.jahn@eee.strath.ac.uk)

Stephen McArthur is with the Advanced Electrical Systems Research Group, Institute of Energy and Environment, University of Strathclyde, Glasgow, G1 1XW (e-mail: s.mcarthur@eee.strath.ac.uk)

James McDonald is with the Advanced Electrical Systems Research Group, Institute of Energy and Environment, University of Strathclyde, Glasgow, G1 1XW (e-mail: j.mcdonald@eee.strath.ac.uk)

Jim Reed is with British Energy PLC, Barnett Way, Barnwood, Gloucester, GL4 3RS (e.mail: jim.reed@british-energy.com)

This paper describes a data mining approach adopted to meet this need of understanding the new domain of FGLT data analysis for core condition monitoring. Data mining includes the process of analyzing and visualizing data in order to discover previously unknown, or confirm previously suspected, patterns and knowledge. A data mining approach for examining the data, based on [1] was adopted. This allowed the raw data to be explored for patterns, without necessarily having the understanding of how the load data directly related to the core condition. Subsequent discussions with plant experts allowed a deeper understanding of the nature of certain channel features and distortions, and how they relate to the fuel grab load data, to be developed.

The detail of how the resulting analysis system was developed and delivered to the desktop of the core monitoring engineers is also described. The structure of the paper is as follows: firstly, an overview of AGR core design is provided, then the approach adopted to understand fuel grab load trace data for condition monitoring purposes is discussed. An overview of the resultant system is then provided followed by a discussion on how this system is used to develop a greater understanding of core condition. Finally, the encapsulation of this knowledge within an automated reactor core brick anomaly detection system is proposed.

## II. BACKGROUND

A brief overview of the construction of the core, refueling process and expected modes of brick cracking is provided.

### A. AGR Refueling

The core of an AGR is constructed from thousands of interlocking cylindrical graphite bricks. These bricks are arranged into layers of about 300 fuel moderator bricks, which form the fuel channels into which the uranium fuel is inserted, combined with interstitial bricks that provide channels into which boron control rods can be inserted to control the rate of fission or shut down the reactor. Typically the core comprises of 11 or 12 layers, resulting in fuel channels of around 10m deep. On top of the core are guide tubes and standpipes for each channel resulting in a channel of approximately 30m through which the fuel assembly must travel during refueling.

Nuclear fission is used to generate heat in order to produce steam to generate electricity from a turbine. In order to sustain a constant power output, the uranium dioxide fuel needs to be periodically replaced. This process, termed reactor refueling, is carried out in batches of around 10 fuel assemblies on a monthly basis. Exact times between refueling operations and numbers of channels refueled in a batch are station dependent. A fuel assembly consists of two main elements: the fuel stringer which houses the uranium dioxide fuel and the fuel plug unit which provides shielding from heat and radiation and is also used to lock the fuel assembly in the channel during operation.

### B. Fuel grab load measurements

During this refueling process measurements of the load, and sometimes height, of the fuel assembly being extracted or inserted into the core are taken for control and protection purposes.

Load cells on the refueling machine directly measure the apparent load of the fuel assembly, as it is being lowered into, and raised out, of the reactor core. A number of factors contribute to the value of the net load. The significant ones are:

- The weight of the fuel assembly
- Frictional forces caused by the interaction between stabilizing brushes on the fuel assembly and the channel wall
- Up thrust effects of the gas circulating through the core supporting some of the weight of the fuel assembly

The output has historically been recorded on a flatbed plotter and kept as part of the station records. In addition, some of the stations record this signal digitally along with a height measurement. There is a project currently ongoing to capture and store this data digitally at all stations. This research deals with both the digital and digitized data sources.

With the knowledge that over a period of 3-4 years every channel in the reactor would be refueled at least once it was hypothesized that the load trace data could give an indication of the condition of the core. Characterizing this relationship would allow analysis of the load trace data to provide information relating to the condition of the core on a more frequent basis.

### C. Brick Cracking

There are two main modes of brick cracking: circumferential and axial (primary) cracking. These represent the extremes of graphite brick distortions and although the nature and causes of the cracking are beyond the scope of this paper, a brief overview of what they are and how they may appear on a load trace is given. Fig. 1 illustrates both circumferential cracking, on the left, and primary cracking, on the right.


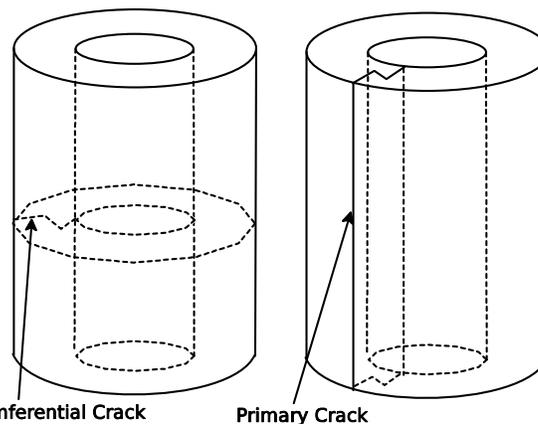
Circumferential Crack          Primary Crack
Fig. 1.      Types of crack

Circumferential cracking occurs around the diameter of the brick. Deformations of graphite associated with the crack result in a reduction in diameter at the crack location. This reduced diameter will result in more friction between the fuel assembly and the channel, which will appear as a peak on a

discharge load trace, and as a trough on a charge trace. The frictional component depends upon the direction of travel of the fuel assembly. During discharge (removal of the old fuel assembly from the core), friction opposes the movement of the assembly and therefore a narrowing of the channel will result in an increase in the apparent load of the fuel assembly. During charge (new fuel insertion) the frictional force supports the weight of the fuel assembly, resulting in a decrease in the apparent load from a narrowing of the fuel channel.

Primary brick cracking is where a brick cracks along its entire length, which may cause the brick to open up, resulting in a larger than expected internal diameter. It is also possible that the bricks may doubly crack and shear, causing a reduction in the effective internal diameter across the brick. Either of these effects would be likely to result in a step-change in the friction and hence load across a brick layer. The magnitude and direction of this step change would depend upon the severity of the dimension change and the direction of travel of the fuel assembly.

An aim of this research is to prove the hypothesis that the analysis of fuel grab load trace data can provide useful and meaningful results in terms of identifying and classifying cracks within AGRs. It is currently the belief of the experts that when deformations occur, they are more likely to occur on an individual brick basis, rather than affecting the whole channel. In particular, circumferential cracks occur within a single brick and affect the load trace within that brick layer without affecting those around it.

In order to detect primary brick cracking, the data needs to be split from full trace data into layer data so their relative loads can be tested. It is also thought that any variation due to effects other than change in the channel diameter, such as gas up thrust, will be reduced when considering a smaller, single brick layer data trace as opposed to the full trace. In addition to being able to detect these two main expected types of anomaly there is a requirement to be able to identify any trace that does not conform to normal or expected behavior.

## III. RELATED WORK

To date, no research has been undertaken in the field of using fuel grab load trace measurements to determine AGR core condition. An earlier project in a related area, resulted in the development of the ALTA system [2], which examines the final part of a load trace in an attempt to automatically determine whether fuel has been set-down correctly in the reactor using a combination of k-means clustering, Kohonen network and rule induction techniques to assess elements of the load trace. These are combined with a rule-based system to provide an assessment of fuel set-down.

In the application of intelligent system techniques to nuclear power plants, there are many reported examples in the areas of plant operation, emergency response, maintenance and some in design activities [3]. One application that deals with examining a time-series set of data for condition monitoring of motor operated valves in a nuclear power plant is reported in

[4]. This system deals with extracting known features from a set of response data from motorized valves. However, in this application, various faulty conditions are simulated to obtain expected responses.

## IV. DATA MINING AND ANALYSIS

The initial stage of developing a system that will automatically assess new FGLT data is to develop an understanding of the domain and how it relates to the core condition.

### A. Available data and knowledge

As stated earlier, it has only recently been considered that FGLT data could provide information relating to the core condition. There is detailed knowledge of the physical core construction from both design and construction of the stations as well as predicted models of behavior of the fuel assembly approaching the touchdown phase of refueling. There is, however, little documented knowledge of the meaning of the FGLT data as it traverses through the core region. What does exist is the FGLT data from every single refueling event from every refueling operation carried out at each AGR station within the UK, as this data is required to be logged as part of station records. Unfortunately the majority of this data exists as paper traces, though a considerable number of man-hours have been spent converting a representative proportion of this data into a suitable electronic format. As mentioned earlier, there is also a project in place to capture this data, along with height measurements, electronically. This will provide more accurate data, though the digitized data will still provide a useful historical perspective and will provide a benchmark against which the newer, digital data can be compared.

In addition to the analysis of the raw FGLT data, there are three areas in which work has been carried out in order to develop understanding of brick cracking. Firstly, there is a large body of work involved in investigating how graphite ages under exposure to heat and radiation such as the process of radiolytic oxidation [5][6]. This attempts to explain the underlying process through theoretical examination at the molecular level and provides a theoretical basis for graphite brick deformation. Secondly, during outages detailed measurements and photographs of the core are taken for a limited number of channels. This sample set of channels provides the station with an appreciation of the current state of the core by extrapolating the detailed analysis of individual channels to the reactor core as a whole. This provides information relating to the actual condition of the core. The measurements taken during outages have been directly compared to the FGLT data gathered from the same channel and particular instances of anomalous behavior in the detailed measurements have been corroborated with the FGLT data to demonstrate specific instances of anomalies. This has been undertaken on a very limited number of channels but supports the proposal for use of FGLT data for detecting abnormalities in the core. Thirdly, rig work has been carried out where a physical model of a section of a channel was constructed using

spare bricks. A series of experiments were conducted where a test brick had a crack machined into it and the crack was forced apart over a range of distances. This work has been carried out by an external partner and reported upon directly to the station operator. These experiments are designed based on input from the above two areas to generate likely cracking scenarios, but also have limitations in terms of emulating the real situations (i.e. the graphite is not irradiated, not at high temperature and does not have gas flows through it).

*B. Approach*

The approach adopted to the analysis of this data follows the general KDD process proposed by [1] and is shown in Fig. 2.
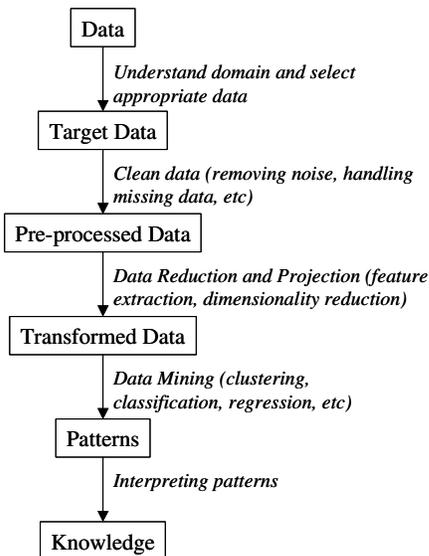


Fig. 2. Stages of the knowledge and data discovery process

The following section describes the selection and cleaning of the data in order to have a database of data that can be readily compared, regardless of whether it originated from paper or electronic sources. This database can then be explored for patterns and a model of expected behavior can be derived.

*C. Data pre-processing*

There were a number of key tasks involved in the preparation of the FGLT data before it could be analyzed. Due to the volume of data involved, these needed to be automated where practically possible. These included:

- Cleaning data of digitized data: the digitization process is not perfect and can introduce spurious data points. The paper traces are scanned into an electronic format, which is then turned into a set of data points using a proprietary software package. Specks of dirt and/or ink can sometimes be incorrectly interpreted as data points by the conversion process. These spurious points are identified by a large step change in load over one or two data points and returning to the previous level. It is important that as many of these artificially introduced data points as possible are identified and removed.

- Selection of relevant data: the FGLT covers data from the entire refueling process whereas the area of interest in this research is restricted to the core region. Identification and isolation of this data region is important

- Removal of stoppages: during refueling sometimes the fuel assembly is halted so that seals can be inspected before re-insertion into another channel. This causes a discontinuity in the FGLT data that can be removed from the digital data through examination of the associated height data. Unfortunately, there is currently no method for removing stoppages from the digitized data.

- Rescaling of data: in order to compare FGLT data streams from both paper and digital sources, a common sampling rate was required. That is to say that the $i^{th}$ data point in a digital data set and the $i^{th}$ data point in a digitized data set both referred to the same, known physical location in the core.

Throughout this process it is important to record any changes made to the data, such as removal of points, to ensure that the appropriate Quality Assurance requirements can be maintained and that the original data can be returned to if desired. This is particularly important in this domain where, in future, the results of a comparison with an expected envelope of behavior may be used to guide the selection of which channels to inspect during an outage.

*1) Cleaning data*

The process of digitizing the paper traces is not perfect. During the process, specks of dirt or ink can be interpreted as data points. A filter was developed to automatically screen out single data points that differed significantly in value from adjacent points in the data stream. Every point that was removed in this way was logged to ensure that the QA grade of the data was maintained.

With the digital traces, where height data was available, stoppages were removed by automatically scanning the height data for regions of constant height using a simple algorithm, which fits a straight line to a window of data. When the gradient of this fitted straight line tends to zero, this indicates a stoppage region. The window was chosen so that only stoppages over five seconds were detected and removed. In practice, these stoppages last several minutes.

*2) Data selection*

A set of FGLT data contains data from the fuel assembly's entire journey starting from inside the fuelling machine, down through the guide tubes, through the reactor core to touchdown at the bottom of the core. A sample plot of FGLT data is shown in Fig. 3. Only a portion of this data is required, the region relating to the journey through the core, so means of identifying and extracting this data is required. Through knowledge of the physical construction of the core, it is known that there is a restriction caused by a piston seal bore, which is used to lock the fuel assembly in place. From the load trace, this can be seen as a very distinctive double peak feature, highlighted as A in Fig. 3. This feature occurs at the same physical height in the core and was selected to be the common datum between all the data traces. This feature could be

identified from the associated height measurements from the digital data, whereas with the digitized data, this feature needed to be automatically detected within the load signal. An algorithm was developed to automatically identify this feature within a stream of digitized data, and the data relating to the core region and below could be identified.
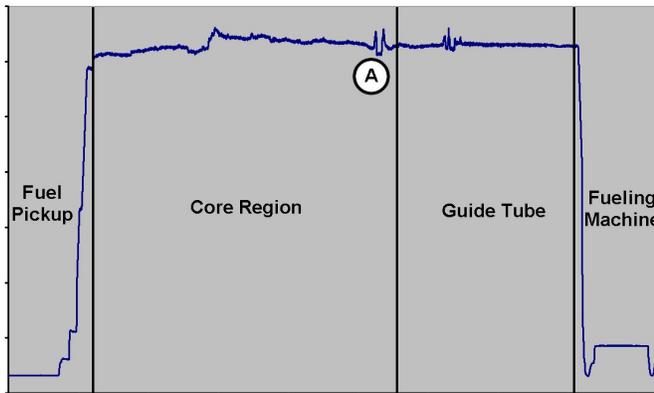


Fig. 3.     Regions of an event on a FGLT data plot. The common datum feature is highlighted as A on the graph.

### 3) Data Re-scaling

Graphs of the raw data aligned to this datum, A, were plotted. Examination of the digital traces showed that all the traces were aligned. However, though some of the digitized traces aligned with each other, there were a number of clusters of aligned traces. In addition, these did not align with the digital data. This could have been for a number of reasons, for example a change in data plotter at the station or a change in the digitization process. In order to perform direct point-by-point comparisons between the digitized data and the digital data, the digitized data would need to be re-scaled to be compatible with the digital data. In order to undertake this re-scaling a second datum needed to be identified within every trace and the data converted to the same number of data points as the digital trace over the same core region.

The data could have been left in its raw format and the rescaling undertaken at runtime only when analysis on that data set was required. However, each time an analysis was undertaken, this re-scaling would be required and it was therefore decided to be more efficient in the long term to do all the pre-processing up front.

### 4) Segmentation and storage of data

Once all the data was in a common format, the data was segmented into the individual brick layers and stored in a relational database. The segmentation into individual brick layers was possible as it was known that the fuel assembly moves at a constant speed through the core region, and that the bricks have a known height.

It was important to store the event information along with the raw data in the database to allow searching and selection of meaningful subsets of the data. The event information that was stored included the station, reactor, channel identifier, date of event, the original data source (paper or electronic) and the electrical load condition during refueling (whether it was refueled while the reactor was on- or off-line). It was thought that some of these might affect the behavior, and therefore

comparisons between data sets from different reactors, different time periods and different channel locations, such as those on the periphery of the core, could be undertaken.

Fig. 4 summarizes the pre-processing steps undertaken to turn FGLT data from both paper and digital sources into a common useable data format to allow the data to be compared using bulk processing techniques. It should be noted that this database grows with time, and as new refueling events are undertaken, the data can be entered into the database, provided the same pre-processing steps are taken.
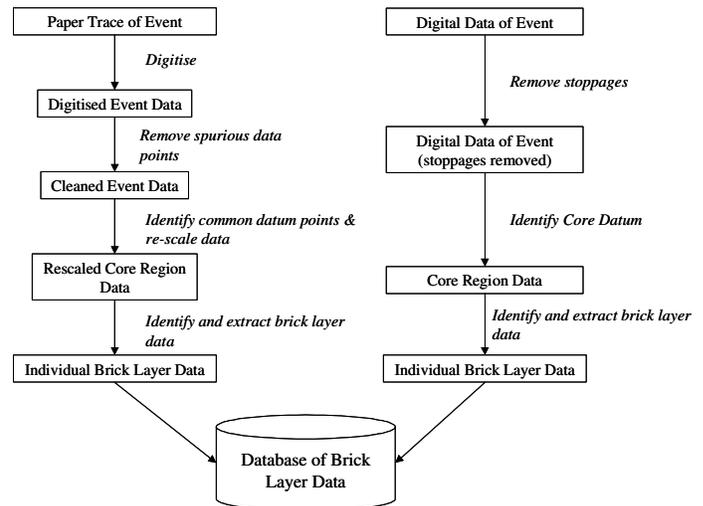


Fig. 4.     Summary diagram showing both paper and electronic route to common data set

The database is initially populated with over 7,000 individual brick layer data sets taken from two different stations. The ratio of digitized to digital data is roughly 8 to 1.

In addition, a new set of data was derived for each event, comprising the average brick layer values for each brick in the event. This allows a normal set of relative brick layer average values to be derived, from which step changes in load across a brick layer, indicative of primary brick cracking, can be detected.

## V.  Data Exploration and visualisation

It is known that the majority of the data available represents normal behavior, although a small percentage of the data may represent anomalous behavior. It is not known, however, if normal behavior is restricted to a single shape per brick layer, or if it is possible for a brick layer to have more than one acceptable normal profile. A powerful method of exploring this concept is visualization of the raw data. This will provide the user with an appreciation of the situation and allow outlying events to be identified, should there only be a single mode of normal behavior. This visualization is a subjective process, so in addition a repeatable process was required for analyzing and assessing the data and therefore led to a method of assessing new refueling events for anomalies.

The method will be described then its applicability to a number of data sets to demonstrate its usefulness will then be

described.

*A.   Approach*

This approach is applicable to any set of refueling data though the focus here is on individual brick layer data. In the context of this work, a trace is defined as a single stream of FGLT data. The purpose of this approach is to provide a measure of how similar each trace within a set of traces is to the rest of the traces. Firstly, a trace which is representative of all the traces in the data set is established, then each individual trace is assessed against this representative trace. This can be used to determine whether the majority of the traces follow a single profile, to identify traces that differ significantly from the majority of the others (if they do all follow a single profile) and to define an envelope of expected behavior from all the available traces. Due to the volume of data involved, it was important that this was encapsulated in software that would allow the rapid and repeated application of the method to explore different data sources. This software implementation is described in detail in a later section, with the analysis approach described here. Fig. 5 shows the how the software is used to analyze a collection of load traces. The steps contained within the "Profiling Software" section represent the automated functions performed by the system, and each of these steps is explained in detail next. The output of using the software in this manner is to establish, through a similarity metric, a set of suitable traces to be used to define acceptable behavior. This could then be implemented within an automated analysis system.
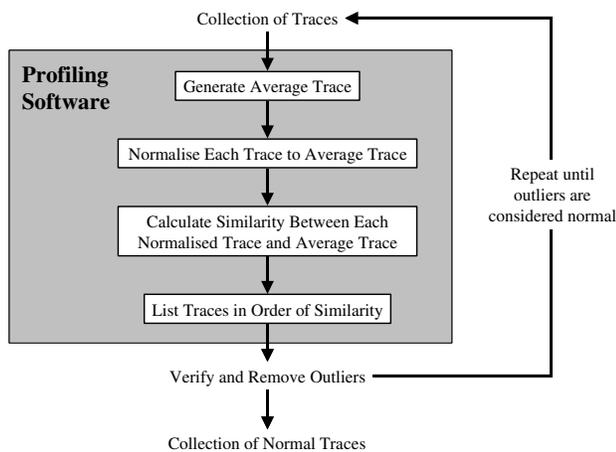


Fig. 5.    Analysis of a set of traces to determine a collection of normal traces to be used in automated anomaly detection.

*1)   Generation of an Average Trace Shape*

The first step in this process is to define the average shape for all the traces in the current selection. The underlying assumption at this stage is that the majority of the data has a common shape. Through the pre-processing stage, it has already been established that each trace in the data set can be examined on a point-by-point basis.

The aim of this stage is develop a reference trace against which all the other traces could be compared. It was decided that the most representative trace would be a point-by-point

average of all the individual traces. If $x_{i,j}$ is taken as the $j^{th}$ point of the $i^{th}$ trace and there are M traces all of length N, then the $j^{th}$ point of the average trace, t is given as:

$$t_j = \frac{1}{M} \sum_{i=1}^{M} x_{i,j} \quad for\, j = 1,2,...,N$$

This new trace *t* is the considered to be the average trace for that dataset and is considered to be a representative trace of all the available data.

An alternative that was considered was selecting an exemplar trace to be used as the reference. This required some prior knowledge of the data to make the selection of the exemplar, which would also mean the user was required to make a choice, rather than leaving it as an automated process applicable to any data set.

*2)   Normalizing each data set*

For the analysis of brick layer data it was also considered desirable to remove any offset from each individual trace. The comparisons undertaken were on the shape of the brick layer FGLT data rather than their absolute values. This offset is caused by a combination of factors including the varying $CO_2$ flow rate through the core and differences in absolute weights of the different assemblies. This removal of offset, or normalization, is achieved by ensuring that all the traces have the same mean value. The chosen value is the mean of the average trace defined above and each trace is normalized to this value by adding a constant value to each point in the trace. The constant calculated is unique to each trace and is the difference between the mean of the trace and the mean of the average trace.

An alternative would be to normalize all the traces at a single point. However, the selection of this point would require prior knowledge of a suitable point and how the traces might vary. For example, the traces could have the offset removed based on the first point of the brick layer. However, this is known to be part of the brick layer interface and can vary from trace to trace. By selecting the average of the data set the effect of variation of one point or section of the data is minimized.

*3)   Similarity to Average Trace*

Each normalized trace now needs to be compared to the average trace. This allows the original set of traces to be ordered from those that are most similar to those that are least similar to the average.

There are a number of statistical measures that can be used to provide a goodness of fit between a set of observables (in this case a brick layer FGLT) and a predicted model (in this case, the average FGLT). Some example measures include the Root Mean Squared Error (RMSE) and the Mean Absolute Percentage Error (MAPE).

As the notion of similarity is very subjective beyond two data sets being the same or not, no definitive measure could ever be defined. However, statistical measures, such as those mentioned can provide a means of providing a measure that is both repeatable and provides a useable measure of similarity for the purposes of assessing brick layer shapes. In the

software implementation, the RMSE was selected as a suitable measure of similarity and is calculated for every trace in the data set and used to order the traces from least to most similar. Users are directed to those traces with particularly high RMSE values, as these tend to indicate abnormal traces, as will be shown later.

*4) Selection of datasets*

The approach presented in this section is applicable to a range of datasets. There may indeed be a number of different reference traces depending upon the application. For example, a reference trace for channels towards the centre of the core could be developed and a separate reference trace for those on the periphery of the core using datasets taken only from the appropriate regions. Likewise, data from different dates could be used to define different reference traces for those time periods. For initial analysis, a single reference trace for each brick layer using data from the whole reactor, regardless of channel location, was developed. This allows a baseline against which subsequent refueling event data can be compared to be defined. In future, should it be discovered that the channels distort differently depending upon their location within the reactor core, more specific envelopes could be developed.

*B. Software implementation of data profiling system*

As a result of the data-mining process approach identified through the research, and described in Section V, Part A, a software toolset was designed, developed and implemented to support British Energy engineers. The overall system is summarized in Fig. 6 and comprises a relational database management system (RDBMS) containing pointers to the individual data files along with the trace metadata. The web based front end, written in PHP and hosted by the Apache HTTPD web server, allows registered users to quickly search for required data, group data into datasets and submit those datasets for analysis.

An analysis engine application monitors the database for new analyses entered and schedules them for running. This analysis engine is executed on high-powered, centralized machine mitigating the need to distribute what can range from several to hundreds of megabytes across a network. Internally, the analysis engine performs the "Generate Average Trace", "Normalize Each Trace to Average Trace", "Calculate Similarity Between Each Normalized Trace and Average Trace" and "List Traces in Order of Similarity" stages shown in Fig. 5.

When the analysis engine completes its work, the resulting charts and tables are written back to the RDBMS and the job is marked as complete. At this time, the web-based front end is able to display these results on an analysis results page.

The process of "Identifying the Least Similar Traces (Outliers)" is still carried out manually, with user input, but there are currently plans to include this within the website and allow traces marked as "abnormal" by one user to be removed from all other analyses automatically. By tracking the dependencies, all analyses that are tainted by these changes

can then be re-run automatically and the user notified automatically. This resulting system, populated with an initial seed of over 7000 data files, is currently undergoing evaluation at British Energy's engineering offices.
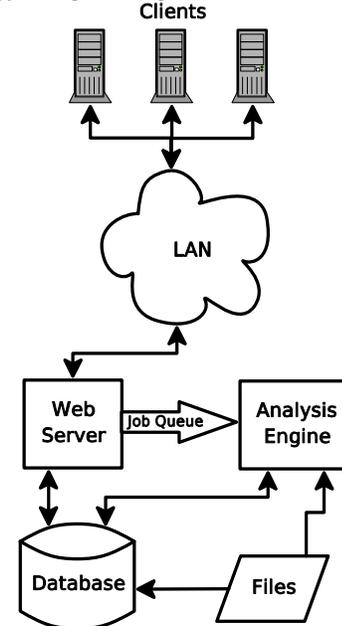


Fig. 6.    Overview of the system architecture

*C. Case Studies*

The development of the software has allowed the data to be explored in a number of ways. Through the parameter-search front end, a dataset containing any of the stored events can be analyzed. Two examples of use of the software are now presented.

In the first example, a set of traces can be visually examined to see if they contain one or more modes of normal operation. Fig. 7 shows FGLT data for a single brick layer taken from many different channels. In Fig. 7 it can be seen that there are two separate shapes, or modes of operation. The majority of traces follow a shape similar to the thick black line of the average trace as expected. However, there is a smaller subset of traces that are present, which contain a shorter initial slope. This can be more clearly seen when the two data sets are plotted separately as in Fig. 8 and Fig. 9. Compare length of slopes for both the first rapid and third major of loads in both Fig. 8 and Fig. 9. The resultant calculations of RMS error for the combined data set show an equivalent step change in error when the spread of the RMS error was examined. This is as expected as there are much more traces in one dataset, which influences the shape of the average trace to a much greater extent. As a result, the traces in the second data set are dissimilar to this average and as a result have a much larger RMS error. Correlation with the time these events occurred and subsequent discussions with station engineers revealed that the change in response was caused by the electrical loading condition of the reactor. The design of the AGR stations permits refueling to be undertaken while the station is still generating, termed on-load refueling, as well as if it is off-load during an outage. An underlying physical reason, likely

to be the flow of gas within the reactor core, has a direct impact on the FGLT for this brick layer. Therefore, two reference envelopes are required for this layer and additional consideration as to the operational status of the station is needed, in order that the correct envelope is selected for comparison to a new fuel trace.
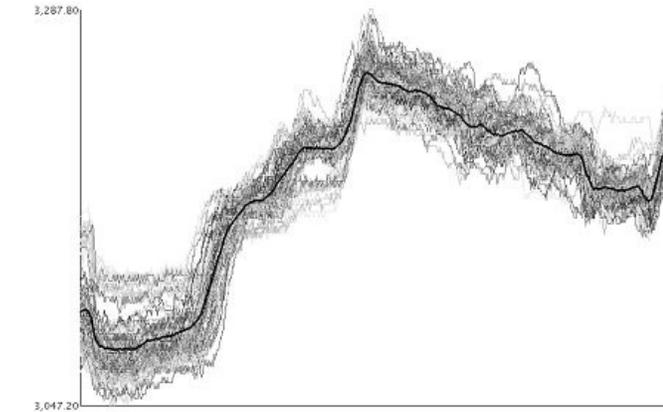


Fig. 7.     Data for a single brick layer taken from a number of channels. Two distinct modes of operation can be seen.
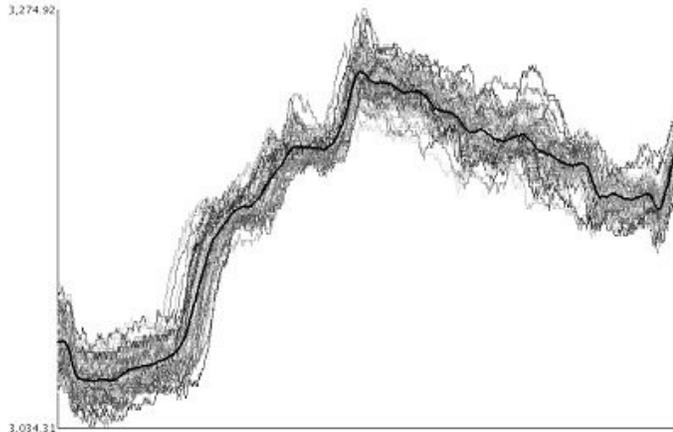


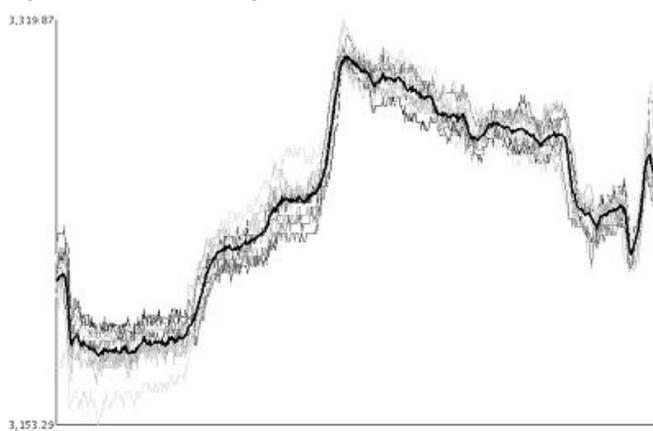Fig. 8.     Data from the larger subset of data.



Fig. 9.     Plot containing the smaller subset of data

The second example demonstrates the use of the profiling system to identify an outlying trace from a given set of data for a single brick layer, then subsequently define an envelope of expected behavior from the remaining traces and follows the process outlined in Fig. 5. This is a simplified example where

data from ten channels were analyzed. Nine of the traces contain raw data taken directly from the load measurement device. The tenth trace has a peak simulating a circumferential crack inserted into it. The raw data can be seen in Fig. 10 with the brick layer containing the simulated crack shown as a distinct peak in the data. It should be noted that the offset caused by the different absolute weight of fuel assemblies has been removed from all the traces. In addition, Fig. 10 also contains the average trace constructed from the 10 individual traces, and is represented by the bold black line.
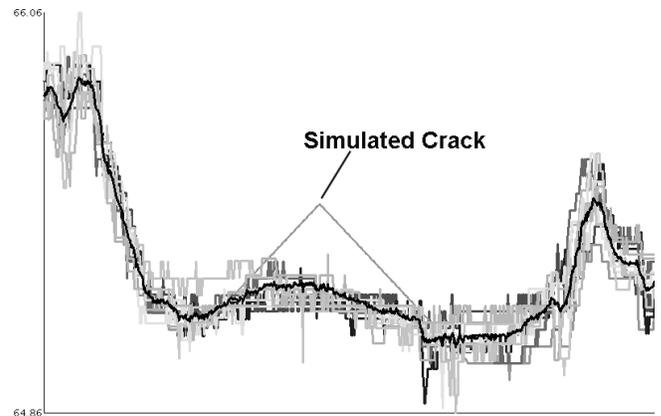


Fig. 10.     Raw data, including average, for brick layer with offset removed

Each trace is compared to the average trace and a measure of similarity, in this case the RMS error from the average trace, is calculated. Viewing the traces in order of RMS error, it can quickly be seen that the trace containing the crack has a significantly larger RMS error than the others.

TABLE 1
RMS ERROR VALUES

| Trace ID | RMS error |
|---|---|
| Cracked Trace | 0.1 |
| Normal 9 | 0.06 |
| Normal 8 | 0.06 |
| Normal 3 | 0.06 |
| Normal 4 | 0.06 |
| Normal 1 | 0.05 |
| Normal 5 | 0.05 |
| Normal 7 | 0.05 |
| Normal 6 | 0.05 |
| Normal 2 | 0.05 |

The spread of RMS error values depends upon the amount of variation in the traces that comprise the set of traces that are analyzed. Provided that the majority of the traces follow a similar pattern and that the abnormal traces are significantly different, then they should be detectable by this method. The actual values of RMS are dependant upon the dataset and so one cannot assign a limit such that RMS values above a cut-off value are abnormal, while all those below are normal. Instead the spread of RMS errors must be examined. Examination of a plot of the squared errors, shown in Fig. 11, supports the claim that the trace containing the crack is significantly different than

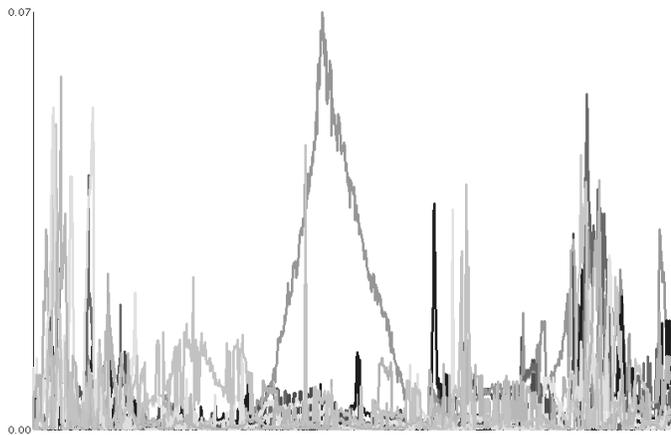those containing normal behavior.



Fig. 11.    Plot of squared errors from the average trace

Having identified this outlying trace, the user makes a note of the channel it refers to in order to perform more detailed examination of it at a later date, and removes the trace from the current set. The remaining traces are then re-analyzed to determine the next outlying trace. The remaining nine traces, with the average trace, are shown in Fig. 12.
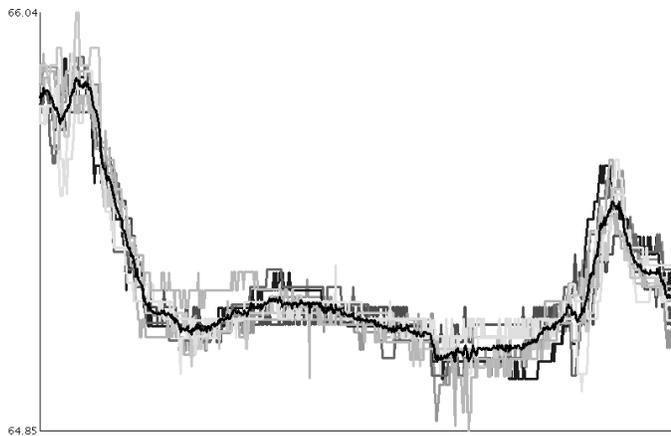


Fig. 12.    Raw data, including average, for brick layer with offset removed and outlier removed

Examining the RMS errors, the highest is 0.06 and the lowest is 0.04. At this stage the user confirms that all the data in this set refers to normal operation and can be used to determine an expected envelope. The resulting envelope is shown in Fig. 13. This envelope could then be implemented within an intelligent system in order to establish whether a new load trace conforms to the expected behavior for that particular brick layer and this process is repeated for all the brick layers within the trace. It should be noted that this example contains far fewer traces than is used in practice, which would utilize all the available data, approximately 300 traces, one from each channel, as the initial data set.
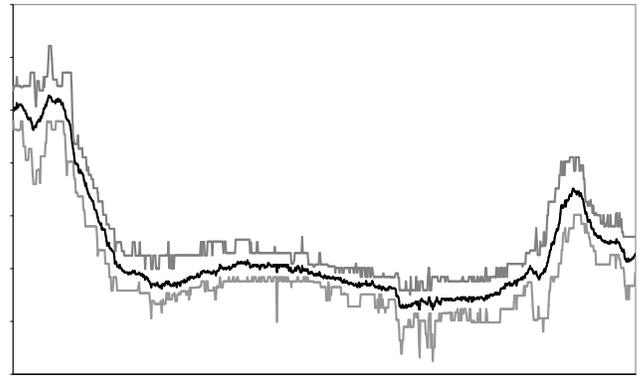


Fig. 13.    Expected envelope: Maximum, minimum and mean values

An important aspect of this research is the transparency and repeatability of any analyses undertaken. A manual, subjective assessment of the data is not desirable in this domain and so a defined set of assessment criteria, which can be repeated and will provide the same outcome offers an acceptable alternative. These analyses could be used to guide the selection of which channels to examine in detail during an outage, therefore it is important to have a consistent and common means of assessing the FGLT data.

## VI. FUTURE WORK

As mentioned previously, this work focuses on deriving understanding of how FGLT data relates to graphite core condition. The next stage of this work will see this process automated, so that the defined envelopes of acceptable behavior can be used to automatically assess any new refueling event data.

There are a number of additional areas where further research is planned. One area is the application of the methodology to other sets of data related to the condition monitoring of reactor cores including reactor control rod operational response data and reactor core temperature data. There is scope for the historical analysis of tie-bar load trace data with the ultimate goal of deriving a model of how graphite core bricks distort over time. In addition to this quantitative model, a physical model of the refueling process could be incorporated into a model-based reasoning approach in order to provide fully automated analysis of FGLT data.

## VII. CONCLUSIONS

This paper has discussed an approach for analyzing fuel grab load trace data for the purposes of leveraging core condition information from an existing core measurement system not originally designed to supply condition monitoring information. In order to allow the interpretation of new traces, a base of knowledge relating the existing core condition to the response shown in the FGLT data was delivered through a software system. This system contains pre-processed data originating from both electronic and paper capture devices and permits the user to identify anomalous behavior such as brick cracking, within various FGLT data sets. This software is currently deployed within the engineering offices at British

Energy and is undergoing evaluation, though already the use of the software has been valuable in defining expected envelopes of operation in a field where no explicit knowledge existed. In addition it has allowed the interpretation of a large volume of data, which would have been far to great to undertake manually. It is anticipated that this system will be used in conjunction with existing, but less frequent, core inspection and monitoring regimes to provide the station operators with a greater amount of useful information relating to the reactor core, thus providing additional confidence to support continued station operations.

It is hoped that this will provide the necessary knowledge to allow a system to automatically detect anomalous behavior in new refueling events, and indeed will form the basis for a fault classification system if and when more anomalous traces become available.

### REFERENCES

[1] Usama Fayyad, Gregory Piatetsky-Shapiro, and Padhraic Smyth. "From Data Mining to Knowledge Discovery in Databases" AI Magazine 17(3): Fall 1996, 37-54

[2] J.A. Steele, L.A. Martin, S.D.J. McArthur, A. Moyes, J.R. McDonald, D. Howie, R. Elrick, I.Y. Yule "An Intelligent System for Interpreting the Nuclear Refuelling Process within an AGR". Proceedings of the Institution of Mechanical Engineers Part A, Journal of Power and Energy, Vol. 217, 2003

[3] Beck, C.E. and Behera, A.K., "A bibliography of technical papers relating expert systems to nuclear power plants" IEEE Transactions on Energy Conversion, Vol 8, No. 1, pp130-137, 1993

[4] Mukhopadhyay, S., Chandhuri, S. "A feature-based approach to monitor motor operated values in nuclear power plants", IEEE Transactions on Nuclear Science, Volume 42, No. 6, pp2209-2220

[5] P. Leary, C.P. Ewels, M.I. Heggie, R. Jones, P.R. Briddon, "Modelling Carbon for Industry: Radiolytic Oxidation", Physica Status Solidi (b), Wiley Interscience, Volume 217, Issue 1, pp 429-447

[6] Marsden B.J. with Zou Z., Fok S.L, and Oyadiji S.O, "Failure predictions for nuclear graphite using a continuum damage mechanics model", Journal of Nuclear Materials, Volume 324, pp 116-124, 2003.

**Dr Graeme West** is a research fellow at the Institute for Energy and Environment. He received his BEng (Hons) and PhD degrees from the University of Strathclyde in 1998 and 2002 respectively. His research interests include intelligent system applications and data mining within power engineering.

**Gordon Jahn** received the M.Eng degree in 2001 from the University of Strathclyde, Glasgow, UK. He is a Research Assistant in the Institute for Energy and Environment, University of Strathclyde. His research interests include agents, intelligent systems, alarm processing, Internet/Intranet applications, knowledge engineering and AI-based incident identification techniques.

**Dr Stephen McArthur** is a Senior Lecturer within the Institute for Energy and Environment. He received his B.Eng. (Hons) and PhD degrees from the University of Strathclyde in 1992 and 1996 respectively. His research interests include Intelligent System Applications in Power Engineering, Condition Monitoring and Multi-agent Systems.

**Professor James R. McDonald** holds the Rolls-Royce Chair in Power Engineering at the University of Strathclyde. Professor James R. McDonald received his BSc, MSc and PhD degrees from the University of Strathclyde. He is Head of the Electronic and Electrical Engineering Department, Chairman of the Institute for Energy and Environment. He has published over 300 technical papers and two books.

**Dr Jim Reed** is the Head of the Graphite Core Team with British Energy. He received his BSc (Hons) and PhD degrees from the University of Kent in 1978 and 1981 respectively. He is responsible for the safety assessments of the graphite cores and is therefore interested in a wide spectrum of graphite research which ranges from atomic behaviour of irradiated graphite through to the dynamic response of large structures.