Machine learning model for event-based prognostics in gas circulator condition monitoring

Jason J. A. Costello, Graeme M. West, *Member, IEEE*, Stephen D. J. McArthur, *Fellow, IEEE*

***Abstract*—Gas circulator (GC) units are an important rotating asset used in the Advanced Gas-cooled Reactor (AGR) design, facilitating the flow of CO_2 gas through the reactor core. The ongoing maintenance and examination of these machines is important for operators in order to maintain safe and economic generation. GCs experience a dynamic duty cycle with periods of non-steady state behavior at regular refuelling intervals, posing a unique analysis problem for reliability engineers.**

In line with the increased data volumes and sophistication of available technologies, the investigation of predictive and prognostic measurements has become a central interest in rotating asset condition monitoring. However, many of the state-of-the-art approaches finding success deal with the extrapolation of stationary time series feeds, with little to no consideration of more-complex but expected events in the data.

In this paper we demonstrate a novel modelling approach for examining refuelling behaviors in GCs, with a focus on estimating their health state from vibration data. A machine learning model was constructed using the operational history of a unit experiencing an eventual inspection-based failure. This new approach to examining GC condition is shown to correspond well with explicit remaining useful life (RUL) measurements of the case study, improving on the existing rudimentary extrapolation methods often employed in rotating machinery health monitoring.

***Index Terms*—Condition monitoring, prognostics, machine learning, support vector machines, logistic regression.**

ACRONYMS AND ABBREVIATIONS

AGR	Advanced gas-cooled reactor
DE	Drive end
GC	Gas circulator
GMM	Gaussian mixture model
LPR	Low power refuelling
ML	Machine learning
NDE	Non-drive end
PHM	Prognostics and health monitoring
PWR	Pressuirzed water reactor
RCP	Reactor coolant pump
SVM	Support vector machine
RUL	Remaining useful life

J. J. A. Costello, G.M. West and S.D.J McArthur are with the Institute for Energy and Environment, Department of Electronic & Electrical Engineering, University Of Strathclyde, 204 George Street, Glasgow, G1 1XW, UK e-mail: jjacostello@gmail.com

Manuscript received January 1, 2016

NOMENCLATURE

$\mathbf{x}$	Training input (from ML model)
$\mathbf{x\prime}$	Test input (from ML model)
y	Output (from ML model)
θ_i	i^{th} parameter
$\underline{\theta}$	Parameter set
$h_\theta(\mathbf{x})$	Hypothesis of model (given test $\mathbf{x\prime}$)
t	Time (in asset life-cycle)
$RUL(t)$	Remaining useful life at t

I. INTRODUCTION

Decision support and monitoring systems have seen wide application in engineering condition monitoring areas [1] [2], with automated diagnostics being used to better inform the processes of maintenance professionals. Increased data availability coupled with the rising performance expectations in disciplines such as aerospace, defence and energy has paved the way for the development and deployment of data-driven intelligent system techniques: algorithms which utilize cutting-edge computational approaches to provide useful information and insights regarding data features and behaviors from the growing volume of available operational data. The ongoing maintenance of rotating turbomachinery is no different, with a variety of automated systems being applied to the condition monitoring of such assets as steam turbine generators [3], gas turbines [4] and aero engines [5].

The nuclear industry employs rotating machinery in a variety of different scenarios: from primary cycle assets which propagate fluid in radioactive conditions to supporting auxiliary motors. An important example is the gas circulator (GC) asset class: an induction motor-based gas propagation rotating machine which maintains CO_2 flow through the reactor core of the UK designed Advanced Gas-cooled Reactor (AGR). GCs are subject to extensive data interrogation, archiving and analysis procedures in order to ensure their continued operation. With numerous GCs per reactor and multiple reactors under the auspices of the nuclear operator, GC health monitoring represents a large data-driven maintenance requirement of utmost importance to safe generation.

Recent developments in the field of condition monitoring have seen a marked rise in interest and investment in the creation of predictive, or *prognostic*, reliability metrics [6], [7]. The ability to assign a probabilistic view of potential future states for an asset is a powerful target for the health monitoring industry, allowing for future failures to be mitigated or avoided entirely. Many of these techniques extrapolate steady-state behaviors into the future in order to ascertain the likely time

t when a failure criteria is met. This often cannot be applied to rotating machinery systems in power systems due to their typically non-steady-state duty cycles: often exhibiting regular changes in conditions due to the everyday requirements of operating and maintaining generation.

Less investigation in prognostics has been made into systems which have duty cycles characterized by regular dynamic events. The AGR GC is one such asset which experiences periods of non-stationery operation: the reactor design allows for online refuelling periods where intermittent states of high and low generation correspond with fuel channel replenishment. Circulators experience a wide dynamic range during these periods, with a distinctive vibration response to the changing conditions. These periods of substantial change have the potential to yield previously unexamined data-based signals regarding the vibration response, and implicitly the health, of the GC units.

This paper investigates the associated GC vibration response exhibited during refuelling events, with a focus on extracting useful health and prognostic measures from this data. A machine learning approach is presented, which first classifies the states making up a typical refuelling campaign and then estimates the temporal position with respect to potential end-of-life for each of these identified states. After demonstrating the effectiveness of this approach to GC prognostics on a real machine example, the paper then discusses the next steps anticipated with the development of the technique towards a full predictive system for use in operational reliability scenarios.

II. Machine learning in reliability engineering

Machine learning (ML) refers to a broad family of algorithmic techniques which take advantage of historical data to learn behaviours, patterns and functions to provide useful inference in a variety of scenarios: including decision-making. The defining feature of a problem space apt for the application of ML approaches is *data availability*: the increased existence of historical reliability data regarding rotating plant, engines, and other key engineering assets has prompted interest in such techniques amongst the more general class of data-driven [8], [9] approaches to reliability engineering.

For example, a variety of ML-driven methods have seen success in as diverse engineering analysis disciplines as wind generation [10], [11], systems monitoring [12], transportation [13], and electrical machine [14] health surveillance. Many condition monitoring methods undertaken manually or through rudimentary data analyses are now seeking improvement from data-driven solutions [15]. Along with reliability engineering as a whole, interest in the nuclear [16] domain has grown in recent years. Less research exists in prognostics spefically for the AGR GC asset, which forms part of the motivation behind this study.

Historically, physics-based models [6], [17] have been dominant for important assets where failure data is less likely to exist. However, the depth of expertise required for such approaches, along with the continually increasing availability of monitoring data, has meant empirical techniques continue to see investigation in [16], [18] in the diagnostics and prognostics for high reliability, critical systems like the AGR GC.

Much of the existing successes of ML in reliability have been in application to steady-state assets: held at relatively constant duty cycles without notable change. Kan *et al.* [19] correctly identify that non-stationary properties characterise the operating conditions of many rotating machines, including those found in the generation industry. Reasoning about future health in the context of changes of state presents a complex challenge to the prognostics and health monitoring discipline, and this is one of the major problems approached in this paper.

Kernel methods such as the support vector machine (SVM) are often applied [20], [21], [22] to reliability problems, either as the primary method or as part of a combination of techniques. Specifically in nuclear energy, [23] found success in utilising kernel method-based techniques for prognostics applied to reactor coolant pumps (RCPs) for the pressurized water reactor (PWR) at the component level. While the RCP operates in differing conditions from AGR GC units and the case study differs in nature (the system in [23] specifically concerns leakage), their function and importance as critical primary cycle coolant machines is similar.

Logistic regression (LR) has been used to estimate the likelihood of data belonging to the later stages of a failure progression [24] and the use of this measure as an implicit view of long-term machine health. A combination of kernel methods and logistic regression was also explored by the authors in [25] as a combined prognostic system development on machinery bearing monitoring.

III. Gas circulators

A. Overview

In the UK, the major design of civil nuclear reactor is the Advanced Gas-cooled Reactor (AGR). Deployed on seven sites with fourteen units, the design marks continued efforts of the UK nuclear industry to utilize the properties of graphite in core structure and moderation [26]. Propagated by eight gas circulator (GC) units, the coolant flowing in the AGR is CO_2 gas pressurized to 40 bar. These induction-based motors maintain safe operating temperatures throughout the reactor core and transfer the heat energy from the nuclear fuel assemblies through to the boiler units. GCs therefore represent a key rotating plant item within the overall AGR system, responsible for both safe and effective operation. A schematic of the gas circulator function in relation to the full core is provided in Fig. 1.

GCs are dynamic and operate at a variety of modes; experiencing a wide range of operational conditions corresponding to reactor events and maintenance. For example, the rate of CO_2 can be tuned by the inlet guide vane angle parameter, allowing for circulator output to be tailored to a particular target power output from the reactor.

Accordingly, GCs are carefully monitored, being subject to extensive health analysis throughout their operational lifetime in order to avoid unplanned outages. The modernization of the condition monitoring discipline has seen a rise in the storage,

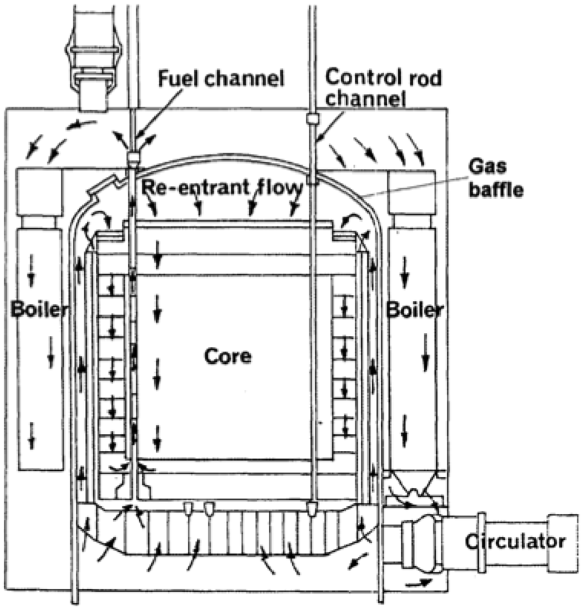

Fig. 1. Illustration of AGR core with single gas circulator unit. Diagram reproduced from [26]

archiving and analysis of data incident from these machines, with large volumes of historical vibration data surrounding the GCs now being available for study. Historically, the health monitoring of these units has shared a large portion in common with general rotating machinery monitoring - based upon the low-level, rudimentary analysis of vibration signals. Accordingly, alarm-driven strategies are used to identify changes in asset behavior and states, with verification of prompted notifications providing the majority of required analysis for the condition monitoring engineering staff on a day-to-day basis. Similar issues have been encountered in steam turbine generator units, with automated data-driven decision support systems being developed to address the problem [27].

B. Low power refuelling (LPR)

GC units are also operational during regular low power refuelling (LPR) events, at which point they experience a dynamic load duty cycle corresponding to periods of fuel replenishment in reactor channels. These refuelling events are characterized by 'castling': intermittent periods of low and moderate generator load to allow online refuelling of individual channels. This is useful from an operational perspective as it maintains partial generation during periods normally associated with outages in other reactor designs.

An example load regime and corresponding vibration response is provided by the time series' in Fig. 2, illustrating the three distinct levels of load operation associated with refuelling. For nomenclature purposes in this paper, each of these LPR states are referred to as `Online` (full load), `Upper` (approx. 70% load) and `Lower` (approx. 30% load). These proportions are a feature of the LPR itself: with `Lower` corresponding to the periods of fuel replenishment, `Upper` corresponding to the intermittent periods of raised generation between replenishment periods and `Online` corresponding to the normal steady-state GC operation state which bookends each LPR. This three behaviour segmentation of the operational time series is domain-specific: defined by the operational and reliability engineering staff managing generation and overseeing the health monitoring of the circulator units.

Typically, the horizontal- and vertical-axis vibration is monitored at both the drive end (DE) and non-drive end (NDE) of circulators to identify the machine's response to the changing conditions. The relationship between operational changes like those seen in the varying load profile and the resulting response captured by vibration transducers forms the basis for much of the health monitoring of the GC units.

Previous investigations [28] into the refuelling behavior of circulators suggests that the LPR is a rich data source for state estimation metrics and key indicators of long-term asset health. Building a representative model of the LPR is therefore presented to be an important area for understanding GC condition.

IV. Refuelling model

A. Motivation and strategy

The LPR is considered to be a useful data view for GC health monitoring for two key reasons:

- LPR events drive the circulators over a large dynamic and transient range, comparable to the run-up and run-down conditions experienced by rotating assets in general which are already widely examined [29], [8] by reliability experts.
- LPR events occur on a regular basis through the lifetime of a GC, providing a repeatable stressor event to the

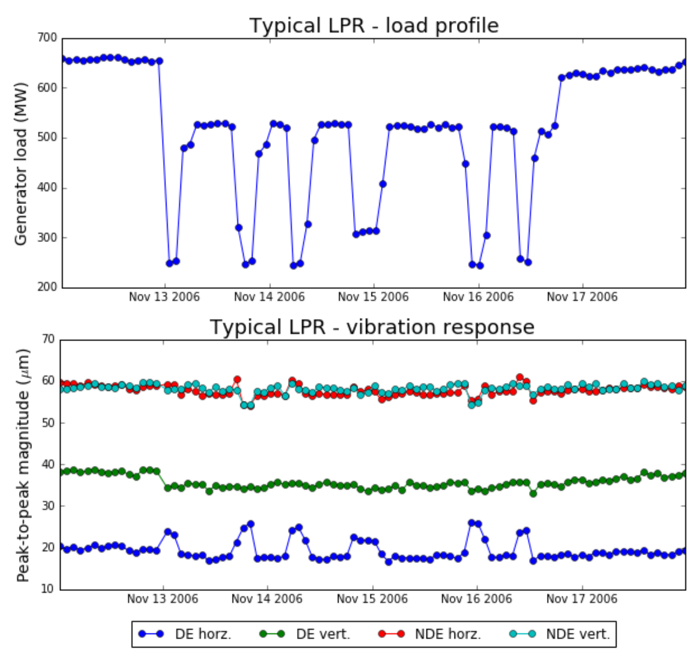

Fig. 2. Typical load and vibration response patterns exhibited during refuelling

machine with a corresponding vibration response rich in potential data-based features.

Taking advantage of these features represents an advance from the typical steady-state analysis associated with many existing diagnostic and prognostic systems in vibration monitoring contexts. A data-driven model mapping the latent relationship between the driving load behavior and the resultant vibration response over subsequent refuelling instances was built to explore this.

The ML model has two major aims:

- Construct a data-driven view of the individual LPR events which accurately classifies GC state from the vibration data,
- Identify if there is a meta-evolution of sequential LPR event models that provides a quantitative feature mapping to the RUL of the circulator.

The techniques are intended to be used for the investigation of predictive metrics and RUL-correlative features in batch analysis, not for online or real-time monitoring purposes.

B. Modelling overview

The approach outlined in this paper broadly tackles the health estimation problem with the following steps:

1) Twenty-one labelled LPR events are selected from a four year operational period,
2) A classifier is built to identify the LPR state {`Online`, `Upper`, `Lower`} from the vibration response,
3) The classifier is then applied to the entire operational period to determine LPR-type vibration state for all historical data,
4) The data is segmented into each of the {`Online`, `Upper`, `Lower`} classes, ordered by timestamp and segmented into four temporal slices: {`early`, `mid1`, `mid2`, `late`},
5) A classifier is built to identify the likelihood of a particular LPR state data being in the `late` temporal slice. The probabilistic output from this acts as the implicit RUL measure.

A schematic overview of the entire approach is provided in Fig. 3, showing the flow of the datasets and dependencies of input throughout each of the stages. Note that the greyed out interactions on the flowchart represent the ML input/output functionality of the trained model. The full approach forward-chains the labelled data into the temporal model for prognostics, but the LPR model can be used to classify the LPR state of any given vibration profile input independent of long-term health considerations.

The following sections discuss the process of selecting and evaluating the ML approaches for each of these stages to provide the most accurate prognostic model with the GC case study.

C. Dataset

A single circulator unit which experienced an eventual inspection-based failure (i.e. the decision was made post-inspection to replace the GC in question) was used, with time series data taken from various periods during 2006 - 2010. This time window includes steady-state online behavior, numerous LPRs, outages and ad hoc operational condition adjustmests. A total of twenty-one labelled LPRs were identified in this dataset and form the basis of the training data for constructing the model. The atomic format of the data is (`timestamp`, `load`, `DE horizontal`, `DE vertical`, `NDE horizontal`, `NDE vertical`). Fig. 4 provides the full load-based behavior of the circulator with the LPRs used for training.

The corresponding vibration response to these periods of variable load form the basis of the data-driven approach to modelling the event. As discussed in Section III-B, there are three distinct generator load levels which correspond to the elements of a refuelling campaign. Defining the mode value empirically for each of these is achieved by using a $k = 3$ Gaussian mixture model (GMM) to cluster the distribution of the mean LPR load profile, the output of which is provided in Fig. 5. Note that this value of k is selected as the three behaviours map directly to domain knowledge in the GC monitoring discipline. Other k values would not correspond to the standard behavioural groups understood by relaibiltiy experts in the field. Determining these values, which are expressed in (1), allows for a labelled training dataset dependent on the load values to be defined at the local minima of the GMM density (values of 386MW and 603MW respectively). This enables the resultant vibration response to be classified into each of the states using a supervised machine learning approach.

$$f_{State}(x) = \begin{cases} Online, & : x > 603 \\ Upper, & : 386 \leq x \leq 603 \\ Lower & : x < 386 \end{cases} \tag{1}$$

D. Technique selection

Four supervised learning techniques were evaluated in building the data-driven model:

- Perceptron-based linear model,
- Logistic regression,
- Linear SVM ($L1$-regularized),
- Linear SVM ($L2$-regularized).

Linear models were selected after examining the dataset and the typical training times associated with more complex, higher-order classifiers: it was decided that a primary study building discriminant functions with no non-linearity would provide useful results without the need for excessive computation. The ML modelling techniques themselves were selected due to a combination of their historical application to reliability engineering problems, and the functionality in multi-class classification scenarios. Other candidate algorithms include the relevance vector machine (RVM) [30], nearest neighbour models and extreme learning machines [31]: comparison of these approaches would be a worthwhile future development.

As a general overview for linear classification problems, consider a linearly separable binary classification problem, defined by a training set $\{(\mathbf{x}_i, y_i)\}_{i=1,\dots,m}$ where m is the number of training tuples and $y \in \{-1, 1\}$. y_i is the label for the i-th multidimensional input pattern $\mathbf{x}_i$.

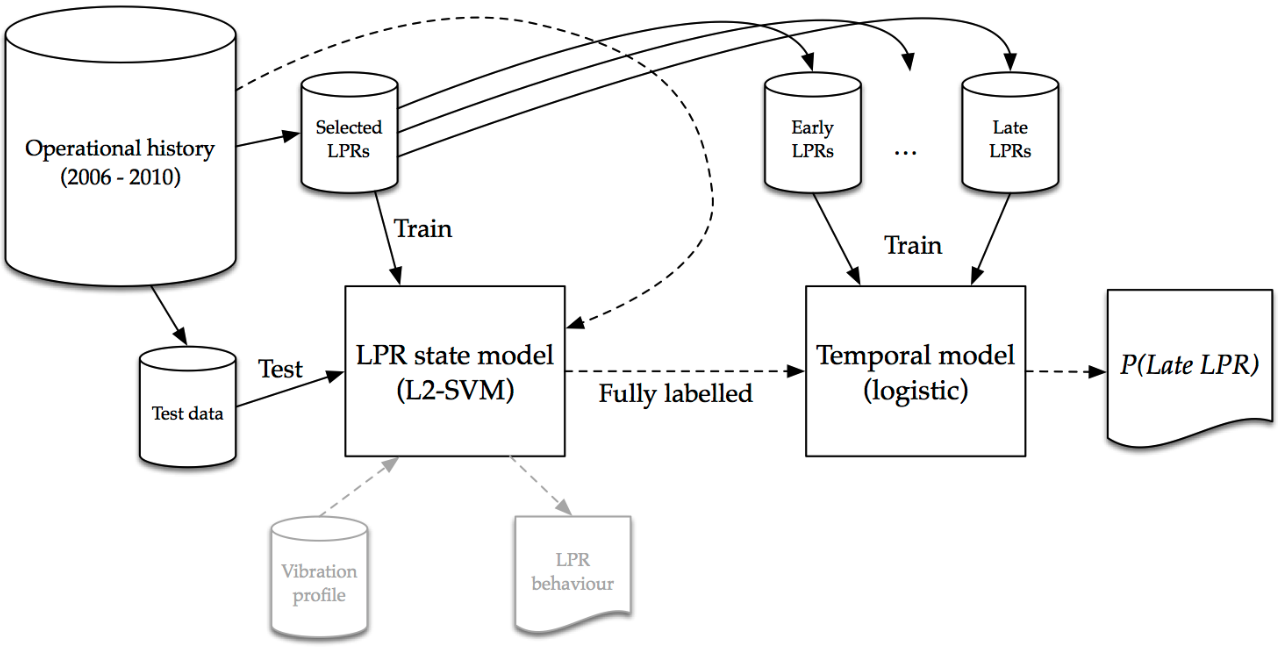

Fig. 3. Overview of the system, with solid lines illustrating data selection and dashed lines showing inputs and output of particular models

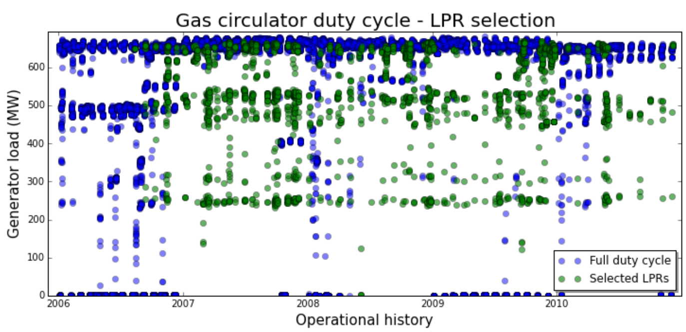

Fig. 4. Load duty cycle of GC used in the training of the circulator model, with labelled LPRs highlighted

The parameterization of a particular model can be denoted $\underline{\theta}$, which represents a finite array of parameters or weights. Linear decision boundaries in two dimensional examples (such as the LPR phase space) take the generic parametric form $\theta_1 \mathbf{x} + \theta_0$. Classification from a successfully learned separating hyperplane on a test input pattern $\mathbf{x}\prime$ is achieved by examining the decision function:

$$y(\mathbf{x}\prime|\underline{\theta}) = sgn(\underline{\theta}^T \mathbf{x}\prime + \theta_0) \tag{2}$$

where $sgn(.)$ denotes a *sign* or *threshold function*, defined as:

$$sgn(x) = \begin{cases} 1, & x \geq 0 \\ -1, & x < 0 \end{cases} \tag{3}$$

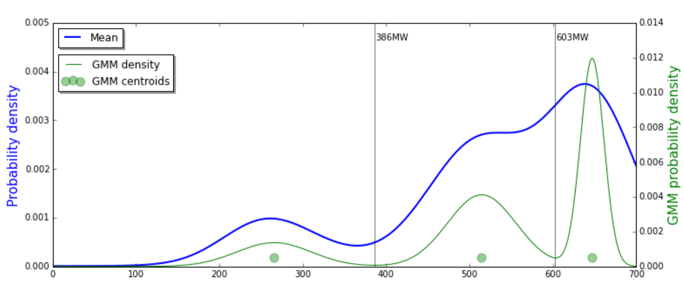

Fig. 5. Kernel density estimation of the mixture of load behaviors exhibited by the circulator over the full dataset, with $k = 3$ Gaussian mixtures annotated

The LPR classification problem is multi-class (with three state labels {`Online`, `Upper`, `Lower`}). Each of the trained models herein follow the *one-vs-all* strategy for binary classification: with each class being attributed a defining hyperplane which denotes membership or non-membership.

1) Perceptron-based linear: Also referred to as a single-layer perceptron, linear modelling using the perceptron learning rule is a binary classification technique used to create a discriminant function between two classes of behavior. The approach can also be extended to multi-class scenarios: which is relevant in the LPR state classification problem, where there are three classes of interest.

The perceptron learning approach can be outlined as follows:

- A primary state is initialised (typically $\underline{\theta}, \theta_0 = 0$),
- With values for $\underline{\theta}, \theta_0$, the function is examined for each input example - comparing the hypothesised output with each target label. In the instance $\underline{\theta}, \theta_0 = 0$, every example is labelled $y = +1$ due to (3),
- When a disparity between an output and target label exists (when $y = +1$ is hypothesis against an example labelled

$y = -1$, for example), the perceptron amends the values of $\underline{\theta}, \theta_0$ to better reflect the training examples.

The values of $\underline{\theta}$ are updated at each iteration using the rule: for the i^{th} parameter:

$$\theta_i \leftarrow \theta_i + \alpha(y - h_\theta(\mathbf{x})) \times x_i \tag{4}$$

where α is a selected parameter known as the *learning rate*; the selection of which impacts the magnitude with which updates are made to the values of $\underline{\theta}$. Perceptron-based linear classifiers can been proven to converge [32] when presented with a linearly separable data domain.

2) Logistic regression: Where the perceptron linear model described previously had a discrete decision function characterized by a hard threshold (training examples close to the decision boundary are treated 'as incorrect' as those far away), logistic regression models make the threshold continuous by smoothing over the boundary with a logistic function. This is particularly useful for noisy datasets, where an absolute linear separation is not possible.

The decision boundary can be recast from the linear model example as:

$$y(\mathbf{x}\prime|\underline{\theta}) = \frac{1}{1 + e^{\underline{\theta}^T \mathbf{x}\prime}} \tag{5}$$

with a corresponding learning update rule:

$$\theta_i \leftarrow \theta_i + \alpha(y - h_\theta(\mathbf{x})) \times h_\theta(\mathbf{x})(1 - h_\theta(\mathbf{x})) \times x_i \tag{6}$$

3) Support Vector Machine (SVM): Support vector machines (SVMs) seek to create a maximum margin between data classes by searching for the most optimal separating hyperplane. In contrast to iterating through a family of parametrized discriminant functions through means of a θ_i update rule (as with the perceptron-based linear and logistic regression examples), SVMs are an example of kernel methods: where the training domain is recast into a new feature space by applying a kernel function to each data point. A generic kernel mapping function can be defined as:

A generic mapping form can be expressed as:

$$k(\mathbf{x}_1, \mathbf{x}_2) = \phi(\mathbf{x}_1).\phi(\mathbf{x}_2) \tag{7}$$

where $\phi(.)$ is a defined kernel mapping function. Kernel space views of training data can be made utilizing a variety of kernel mapping selections, with the most rudimentary kernel function being the linear kernel $\phi(\mathbf{x}) = \mathbf{x}$. With this alteration to the feature space, the decision function for the linear classification problem can be now defined as:

$$h(\mathbf{x}\prime|y) = sgn(\sum_i \theta_i y_i(\mathbf{x}\prime.\mathbf{x}_i) - \theta_0) \tag{8}$$

SVMs seek to maximize the margin around the selected hyperplane, which is demonstrated to be the minimization of a term containing $||\underline{\theta}||$ in some form [33]. For $L1$-SVMs, the optimization goal is:

TABLE I
CROSS-VALIDATION AND TEST SCORES FOR EACH CLASSIFIER

	CV results			*Test acc.*
Classifier	*Hyperparameter*	*Value*	*CV score*	*Score*
Linear perceptron	α	0.1233	88.5%	76.05%
Linear logistic	α	8E^{-5}	87.8%	71.55%
Linear $L1$-SVM	C	10.975	87.8%	75.69%
Linear $L2$-SVM	C	58.57	87.9%	88.23%

$$\texttt{minimize} \quad \frac{1}{2}||\underline{\theta}||^2 + C\sum_{i=1}^{m} \xi_i \tag{9}$$

while for $L2$-SVMs, this is altered to:

$$\texttt{minimize} \quad \frac{1}{2}||\underline{\theta}||^2 + \frac{C}{2}\sum_{i=1}^{m} \xi_i^2 \tag{10}$$

where C is the margin hyperparameter and ξ is the slack variable. Training of both $L1$- and $L2$-SVM approaches in this study focus on the C margin setting (with the slack variable remaining constant), as this denotes the balance between margin maximization and classification error total. Finding a suitable C often determines the efficacy of an SVM to generalize across a domain of application.

4) Training methodology: In order to build a robust set of candidate models for each of the algorithms, k-fold cross validation [34] was used throughout the training and testing process, with $k = 21$ in line with the number of LPR instances in the dataset. A grid search over a range of hyperparameters for each of the selected model types identified the best estimator for each: the results of which are illustrated in the diagrams presented in Fig 6 and summarized in Table I. Hyperparameters are algorthmic variables set as initial conditions which dictate the learning behavior for machine learning models in the training phase: in this example, the perceptron and logistic models depend on learning rate α and the SVMs [35] depend on C, which manages the balance between classifier accuracy and misclassification errors in the training dataset.

From the results of the grid search, the trained $L2$-SVM classifier was selected with its stand-out accuracy of 88.23% on the test data. From a generic time series classification perspective, this can be considered a good rate of accuracy: slightly over one-in-ten points will be misclassified, with the temporal nature of LPR behaviors further mitigating this. (i.e. questionable results could be examined in the context of nearby results in time to verify which state the given point most likely belongs to).

E. Identifying LPR state

The predicted classes for the full dataset for both the DE and NDE orientations of the circulator are shown in phase space in Fig. 7, showing clear clustering in DE space of the vibration response values. This is useful as it generalizes the learned behaviors in each of the labelled LPR instances to the full operational dataset and helps confirm the supposition that the vibration responses for each of the LPR states can be segmented and examined independently.

The NDE phase space view example shows less distinct class separation, suggesting that the DE features are the most indicative when describing the vibration response at each of the LPR states. This difference in response is likely due to the physical orientation of the GC: with the NDE end separated from the driving mechanisms housed at the DE. Presenting GC vibration data in the context of the learned typical behavior from a sufficiently trained model allows for the comparative analysis of individual or group refuelling sessions with new data points and the learned hyperplanes of the model.

To summarize, the trained model has the ability to define LPR state from a vibration-only input, with no operational values required. This is useful as it allows for the engineer to determine periods of behavior which correspond to historical normality and identify those which do not, entirely from the response data defined by the GC unit. Taking this entirely data-driven view of the circulator allows for a deeper investigation of long-term condition metrics: as described in the next section.

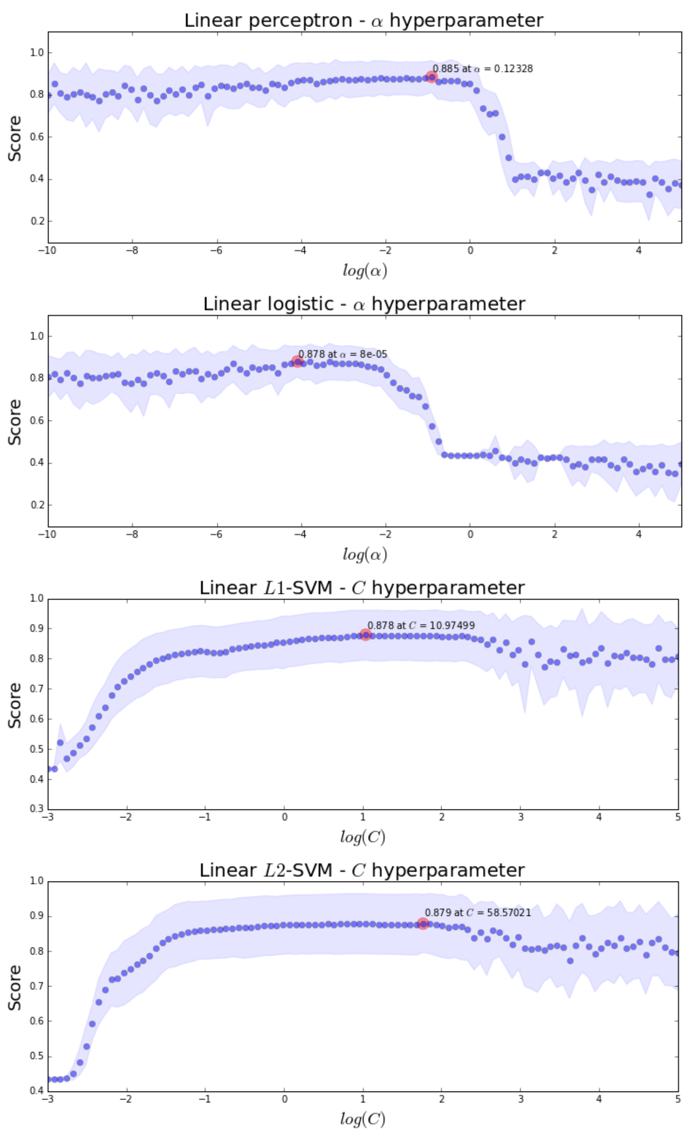

Fig. 6. Grid search results for each of the trained machine learning techniques, with the most accurate hyperparameter setting for each highlighted

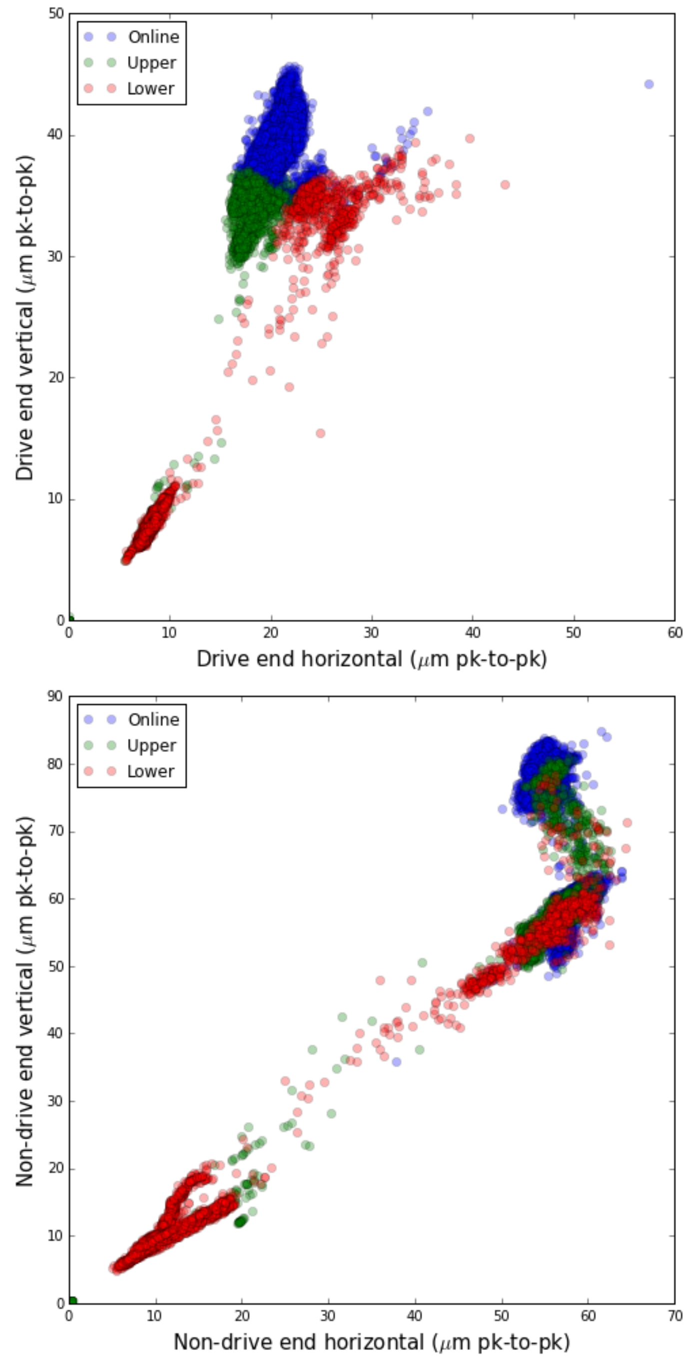

Fig. 7. State prediction from the $L2$-SVM of the full dataset for both the DE and NDE time series

V. Health Monitoring

Providing a quantitative measure of normality for a given point on the lifecycle for a GC, the model output labels can be used to both segment large volumes of historical circulator vibration response into the LPR state labels and examine any temporal component to changes in these groupings. This presents a number of opportunities for advanced condition monitoring techniques when considering the evolution of the learned model across different periods in the circulator history. This section highlights the potential to calculate the remaining useful life of a GC using the approach.

A. State estimation

To evaluate membership to states indicative of machine health in the history of the GC, each of the LPR classes of `Online`, `Upper` and `Lower` were examined independently (defined via the output of the successful $L2$-SVM model). These were subsequently divided into four consecutive temporal states: {`early`, `mid1`, `mid2`, `late`} after ordering the data points by timestamp. Note that the assumption is that as the GC in the example progresses through its lifetime that some latent damage variable is trending towards a failure criteria condition, which triggered the eventual de-comissioning of the asset.

Each of these LPR state-to-temporal state dataset pairs were used as input to a further logistic regression classification model (selected due to its continuous value class boundary, allowing for probabilistic membership values). Given an input of vibration response data from `Online`, `Upper` or `Lower`, this classifier provides a predictive output to membership of each of the temporal states. This allows for a prediction to be made about the likely period of the GC life-cycle the data comes from.

The three plots in Fig. 8 show the estimated class membership of the ordered data for each of the LPR behaviors; demonstrating that the `Online` and `Upper` labelled data have the strongest 'ordering' of the temporal class i.e. {`early`, `mid1`, `mid2`, `late`} are largely in order, when compared with the less ordered `Lower` results. This suggests that these two classes from the LPR modelling approach have a potential implicit mapping to the long-term health of the GC in this example.

B. Remaining useful life (RUL) estimation

For each point in the operational history of the GC time series, the RUL can be defined as:

$$RUL(t) = t_{Failure} - t \tag{11}$$

where $t_{Failure}$ is the end of the time series corresponding to the point of failure and t is any timestamp with vibration data. Explicit RUL measurements are unavailable during operation, so the goal of a prognostic system is to find some implicit metric of the explicit RUL which best approximates it.

Since the `late` temporal state corresponds closest to the exhibited vibration behavior at the end of GC life, strong evidence for membership to this class can be interpreted as an indicator that the selected data is nearing the failure criteria. Fig. 9 compares explicit RUL with the membership likelihood to the `late` state with each of the LPR model behaviors. Similar to the strong temporal segmentation shown by the states, the `Online` and `Upper` data subsets correlate well with the explicit RUL. This is quantified by the Pearson's correlation values provided in II, which highlight that probability of membership to the `late` temporal class for vibration data classified as part of the `Upper` LPR behaviour is strongly *inversely* correlated with the true RUL data. (For comparison, the correlation values for the `mid2` classes are also provided without their associated diagrams). The results suggest that GC vibration response data corresponding to operating conditions seen at both full and 70% generation have indicative properties relating to machine degradation with continued operation of the circulator. In particular, membership to the `late` temporal state of the vibration data labelled as part of the `Upper` LPR behaviors appears to follow the explicit RUL best among the three options: from both a visual and quantitatively correlative perspective.

TABLE II
CORRELATION BETWEEN LATE TEMPORAL CLASS & TRUE RUL

Temporal class	*LPR behaviour*	*Pearson's corr.*
`late`	Online	-0.838
`late`	**Upper**	**-0.938**
`late`	Lower	-0.472
`mid2`	Online	-0.041
`mid2`	Upper	0.542
`mid2`	Lower	-0.146

In terms of in-scenario use, these results point to vibration data labelled as belonging to the `Upper` LPR state as the most indicative of long-term damage trends. An engineering system

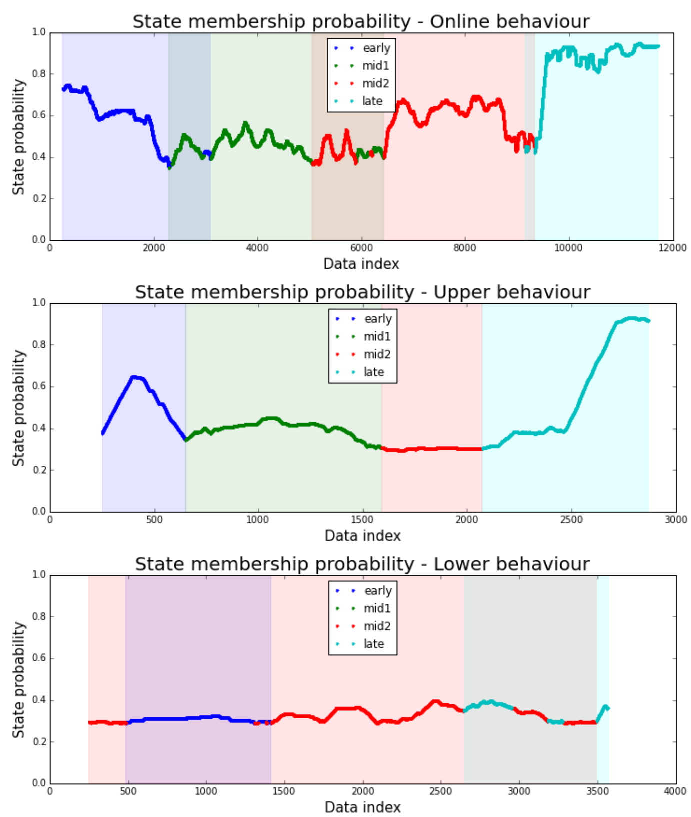

Fig. 8. Most probable temporal state for each of `Online`, `Upper` and `Lower` behaviors

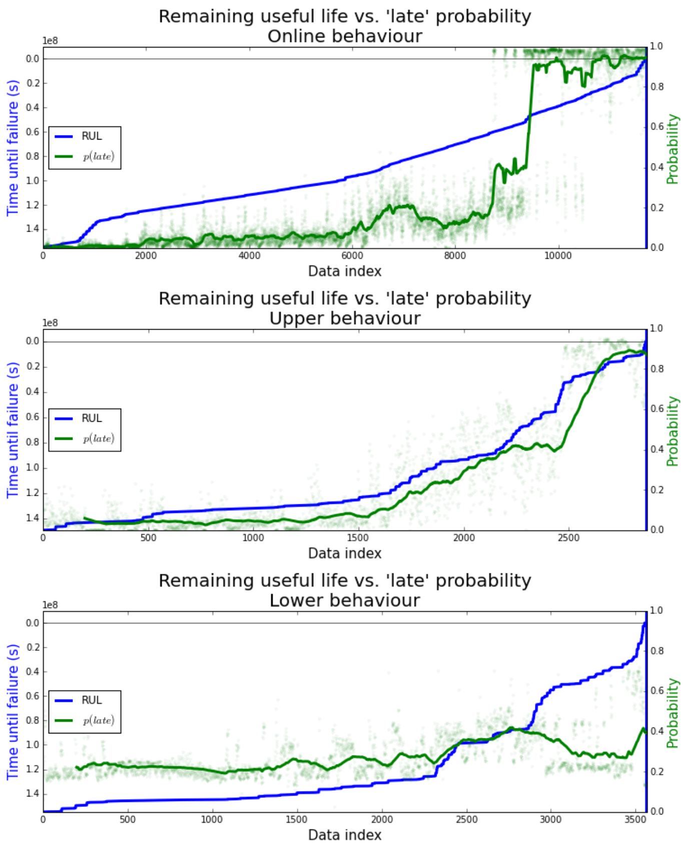

Fig. 9. Explicit RUL compared with probabilistic state membership for `late`

to query the most probable temporal class of incoming data from this LPR state could be used to survey LPR health on a day-to-day basis, with further investigation of the machine required when the probabilistic output of the `late` class membership begins to rise.

VI. Discussion & next steps

As demonstrated, the modelling approach identified that data belonging to the `Upper` LPR state class provides a monotonically increasing probabilistic output when examining `late` state class membership from the temporal model. This strongly suggests that the vibration response exhibited by GCs during these operational periods contains important information regarding the health of the unit: a segmentation of the data that has been previously disregarded in favor of generic measures. The evidence presented in Fig. 9 in particular shows a direct mapping between explicit RUL and the derived temporal measures from this data. From an engineering perspective, there are two options in extracting value from this discovery: utilize the modelling approach as is in making RUL predictions, or focus more investigative effort on building metrics from data during `Upper` LPR conditions.

It should be acknowledged that the ML techniques investigated herein are fairly mature and the combination of kernel methods with logistic regression in particular is not novel [25] in application to relibaility monitoring. However, the combination of event state identification (classification of the LPR states using the vibration data) on non-stationary periods of operation and subsequent health state estimation utilising the probablistic membership of late lifetime from specific event behaviours is a new approach in AGR GC monitoring. The nature of the LPR event itself is a defining element of the prognostic problem, and serves as the platform for enhanced predictive capabilities.

The procedure outlined was built on the operational history of a single circulator undergoing degradation with continued use as a means of exploring new long-term health monitoring metrics in the vibration signals. While robust testing and cross-validation was done using as much of the data as possible, the classification models for both the LPR state and temporal state only apply to this machine example. A wider domain of circulator data should be investigated in order to construct a general model for GC refuelling vibration response, taking into consideration multiple examples of both normal and degradation-type behavior. This would allow for previously unexamined (from the perspective of this health monitoring approach) circulators to be benchmarked against an aggregate degradation pattern which covers an aggregate of numerous circulators.

Another potential improvement could be in empirically investigating the optimal number of temporal states employed by the logistic regression classifier, which was heuristically set to 5 consecutive time slices in the degradation history. This could be further bolstered by the exploration of the underlying ML algorithms themselves: with potential improvements likely to come from alternative kernel methods [30], or alternatively the ensembling of each of the investigated algorithms.

VII. Conclusion

This paper has presented a new data-driven modelling approach for the health monitoring of AGR GC units, using a combination of $L2$-SVM and logistic regression machine learning techniques. The method moves away from steady-state monitoring and selects key data points from a particular period of GC operation: the semi-regular LPR event. Evolution of this data segmentation has been shown to follow the progression of GC operational history and the explicit RUL towards de-commissioning for an example unit which went through an eventual inspection-based failure.

VIII. Acknowledgements

This work was funded by EDF Energy and EPSRC. The views presented by the authors do not necessarily represent the views of EDF Energy. The authors would like to express thanks to EDF Energy for the use of their data for this study.

References

[1] R. C. M. Yam, P. W. Tse, L. Li, and P. Tu, "Intelligent Predictive Decision Support System for Condition-Based Maintenance," *The International Journal of Advanced Manufacturing Technology*, vol. 17, no. 5, pp. 383–391, feb 2001.

[2] G. M. West, S. D. J. McArthur, and D. Towle, "Industrial implementation of intelligent system techniques for nuclear power plant condition monitoring," *Expert Systems with Applications*, vol. 39, no. 8, pp. 7432–7440, jun 2012. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S095741741200125X

[3] P. J. Tavner, "Review of condition monitoring of rotating electrical machines," vol. 2, no. 4, pp. 215–247, 2008.

[4] A. D. Kenyon, V. M. Catterson, S. D. J. McArthur, and J. Twiddle, "An Agent-Based Implementation of Hidden Markov Models for Gas Turbine Condition Monitoring," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 44, no. 2, pp. 186–195, 2014.

[5] D. A. Clifton, P. R. Bannister, and L. Tarassenko, "A framework for novelty detection in jet engine vibration data," *Key Engineering Materials*, vol. 347, pp. 305–310, 2007. [Online]. Available: http://www.scientific.net/KEM.347.305

[6] A. Heng, S. Zhang, A. C. C. Tan, and J. Mathew, "Rotating machinery prognostics: State of the art, challenges and opportunities," *Mechanical Systems and Signal Processing*, vol. 23, no. 3, pp. 724–739, apr 2009. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0888327008001489

[7] R. Kothamasu, S. H. Huang, and W. H. VerDuin, "System health monitoring and prognostics a review of current paradigms and practices," *The International Journal of Advanced Manufacturing Technology*, vol. 28, no. 9-10, pp. 1012–1024, jul 2006. [Online]. Available: http://www.springerlink.com/index/10.1007/s00170-004-2131-6

[8] J. Lee, F. Wu, W. Zhao, M. Ghaffari, L. Liao, and D. Siegel, "Prognostics and health management design for rotary machinery systems - Reviews, methodology and applications," *Mechanical Systems and Signal Processing*, vol. 42, no. 1-2, pp. 314–334, 2014.

[9] K. L. Tsui, N. Chen, Q. Zhou, Y. Hai, and W. Wang, "Prognostics and Health Management : A Review on Data Driven Approaches," *Mathematical Problems in Engineering., Hindawi Publishing Corporation*, vol. 2015, 2014.

[10] H. Jafarnejadsani, J. Pieper, and J. Ehlers, "Adaptive control of a variable-speed variable-pitch wind turbine using RBF neural network," *2012 IEEE Electrical Power and Energy Conference*, vol. 21, no. 6, pp. 216–222, 2012. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6474954

[11] G. Santamaría-bon, A. Reyes-ballesteros, and C. Gershenson, "Wind speed forecasting for wind farms : A method based on support vector regression," vol. 85, pp. 790–792, 2016.

[12] M. N. Nguyen, C. Bao, K. L. Tew, and S. D. Teddy, "Ensemble Based Real-Time Adaptive Classification System for Intelligent Sensing Machine Diagnostics," *IEEE Transactions on Reliability*, vol. 61, no. 2, pp. 303–313, 2012.

[13] J. Li, L. Yan-Fu, and E. Zio, "A SVM framework for fault detection of the breaking system in a high speed train," *Mechanical Systems and Signal Processing*, vol. 87, pp. 401–409, 2017.

[14] F. Jia, Y. Lei, J. Lin, X. Zhou, and N. Lu, "Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data," *Mechanical Systems and Signal Processing*, vol. 72-73, no. November, pp. 303–315, 2016.

[15] E. Zio, "Some Challenges and Opportunities in Reliability Engineering," *IEEE Transactions on Reliability*, vol. PP, no. 99, pp. 1769–1782, 2016.

[16] J. B. Coble, P. Ramuhalli, L. J. Bond, W. Hines, and B. Upadhyaya, "A review of prognostics and health management applications in nuclear power plants," *International Journal of Prognostics and Health Management*, vol. 6, no. SP3, pp. 1–22, 2015.

[17] D. An, N. H. Kim, and J. H. Choi, "Practical options for selecting data-driven or physics-based prognostics algorithms with reviews," *Reliability Engineering and System Safety*, vol. 133, pp. 223–236, 2015. [Online]. Available: http://dx.doi.org/10.1016/j.ress.2014.09.014

[18] X. S. Si, W. Wang, C. H. Hu, and D. H. Zhou, "Remaining useful life estimation - A review on the statistical data driven approaches," *European Journal of Operational Research*, vol. 213, no. 1, pp. 1–14, 2011. [Online]. Available: http://dx.doi.org/10.1016/j.ejor.2010.11.018

[19] M. S. Kan, A. C. C. Tan, and J. Mathew, "A review on prognostic techniques for non-stationary and non-linear rotating systems," *Mechanical Systems and Signal Processing*, vol. 62, no. March 2015, pp. 1–20, 2015.

[20] H.-Z. Huang, H.-K. Wang, Y.-F. Li, L. Zhang, and Z. Liu, "Support vector machine based estimation of remaining useful life: current research status and future trends," *Journal of Mechanical Science and Technology*, vol. 29, no. 1, pp. 151–163, 2015. [Online]. Available: http://link.springer.com/10.1007/s12206-014-1222-z

[21] A. Soualhi, K. Medjaher, and N. Zerhouni, "Bearing Health Monitoring Based on Hilbert Huang Transform , Support Vector Machine , and Regression," vol. 64, no. 1, pp. 52–62, 2015.

[22] X. Chen, Z. Shen, Z. He, C. Sun, and Z. Liu, "Remaining life prognostics of rolling bearing based on relative features and multivariable support vector machine," *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 227, no. 12, pp. 2849–2860, 2013. [Online]. Available: http://pic.sagepub.com/lookup/doi/10.1177/0954406212474395

[23] J. Liu, V. Vitelli, E. Zio, and R. Seraoui, "A Novel Dynamic-Weighted Probabilistic Support Vector Regression-Based Ensemble for Prognostics of Time Series Data," *IEEE Transactions on Reliability*, vol. 64, no. 4, pp. 1203–1213, 2015.

[24] J. Yan and J. Lee, "Degradation Assessment and Fault Modes Classification Using Logistic Regression," *Journal of Manufacturing Science and Engineering*, vol. 127, no. November 2005, pp. 912–914, 2005.

[25] W. Caesarendra, A. Widodo, and B. S. Yang, "Application of relevance vector machine and logistic regression for machine degradation assessment," *Mechanical Systems and Signal Processing*, vol. 24, no. 4, pp. 1161–1171, 2010.

[26] E. Nonbol, "Description of the Advanced Gas Cooled Type of Reactor (AGR)," *Riso National Laboratory Report*, no. November 1996, 1996.

[27] J. J. A. Costello, G. M. West, S. D. J. McArthur, and G. Campbell, "Self-tuning Routine Alarm Analysis of Vibration Signals in Steam Turbine Generators," *IEEE Transactions on Reliability*, vol. 61, no. 3, pp. 731–740, 2012.

[28] ——, "Investigation of gas circulator response to load transients in nuclear power plant operation," in *Eighth American Nuclear Society International Topical Meeting on Nuclear Plant Instrumentation, Control and Human-Machine Interface Technologies*, 2012.

[29] C. Zuluaga-Giraldo, D. Mba, and M. Smart, "Acoustic emission during run-up and run-down of a power generation turbine," *Tribology International*, vol. 37, no. 5, pp. 415–422, 2004.

[30] E. Zio and F. Di Maio, "Fatigue crack growth estimation by relevance vector machine," *Expert Systems with Applications*, vol. 39, no. 12, pp. 10 681–10 692, 2012. [Online]. Available: http://dx.doi.org/10.1016/j.eswa.2012.02.199

[31] K. Javed, R. Gouriveau, and N. Zerhouni, "A new multivariate approach for prognostics based on extreme learning machine and fuzzy clustering," *IEEE Transactions on Cybernetics*, vol. 45, no. 12, pp. 2626–2639, 2015.

[32] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*, 2012.

[33] K. Yoshiaki and S. Abe, "Comparison of L1 and L2 Support Vector Machines," in *Proceedings of the International Join Conference on Neural Networks*, 2003, pp. 2054–2059.

[34] R. Kohavi, "A study of cross-validation and bootstrap from accuracy estimation and model selection," in *International Joint Conference on Artificial Intelligence (IJCAI)*, 1995.

[35] Chih-Wei Hsu, Chih-Chung Chang and C.-J. Lin, "A Practical Guide to Support Vector Classification," 2003. [Online]. Available: http://www.csie.ntu.edu.tw/%7B %7Dcjlin/papers/guide/guide.pdf

Jason J. A. Costello received the BSc. (Hons) in Applied Physics in 2009 from the University of Strathclyde. He is a data scientist at snap40, and was an Eng.D candidate at the Institute for Energy and Environment at the University of Strathclyde. His research interests include machine learning and data science.

Stephen D. J. McArthur (M93-SM07-F15) received the B.Eng. (Hons.) and Ph.D. degrees from the University of Strathclyde, Glasgow, U.K., in 1992 and 1996, respectively. He is co-Director of the Institute for Energy and Environment at the University of Strathclyde and Director of the EPSRC Centre for Doctoral Training in Future Power Networks and Smart Grids. His research interests include distributed intelligence for power system applications, intelligent systems and data analytics.

Graeme West (M10) received the B.Eng. (Hons.) and Ph.D. degrees from the University of Strathclyde, Glasgow, U.K., in 1998 and 2002, respectively. He is a Lecturer in the department of Electronic and Electrical Engineering at the University of Strathclyde. His research interests include intelligent system applications in condition monitoring, diagnostics and prognostics, particularly applied to the nuclear power generation industry.