McGhee, Mark J. and Catterson, Victoria M. and Brown, Blair (2017) Prognostic modelling utilizing a high fidelity pressurized water reactor simulator. IEEE Transactions on Systems Man and Cybernetics: Systems. ISSN 2168-2216 , http://dx.doi.org/10.1109/TSMC.2017.2662478

Prognostic Modelling Utilizing a High Fidelity Pressurized Water Reactor Simulator

Mark J. McGhee, Victoria M. Catterson, *Member, IEEE*

and Blair Brown

***Abstract*— Within power generation, ageing assets and an emphasis on more efficient operation of power systems and improved maintenance decision methods has led to a growing focus on asset prognostics. The main challenge facing the implementation of successful asset prognostics in power generation is the lack of available run-to-failure data. This paper proposes to overcome this issue by use of full scope high fidelity simulators to generate the run-to-failure data required. From this simulated failure data a similarity based prognostic approach is developed for estimating the Remaining Useful Life of a valve asset. Case study data is generated by initializing prebuilt industrial failure models within a 970MW Pressurized Water Reactor simulation. Such full scope high fidelity simulators are mainly operated for training purposes, allowing personnel to gain experience of standard operation as well as failures within a safe, simulated operating environment. This research repurposes such a high fidelity simulator to generate the type of data and affects that would be produced in the event of a fault. The fault scenario is then run multiple times to generate a library of failure events. This library of events was then split into training and test batches for building the prognostic model. Results are presented and conclusions drawn about the success of the technique and the use of high fidelity simulators in this manner.**

***Index Terms*— Model-based prognostics, remaining useful life, high fidelity simulation, power generation,**

Acronyms and Abbreviations

PWR	Pressurized Water Reactor
RUL	Remaining Useful Life
MSS	Main Steam System
RMS	Root Mean Square
PR	Predicted RUL
TR	True RUL
CSV	Comma Separated Values

This work was supported by GSE Systems, which provided the high fidelity simulator and knowledge of the simulation environment.

M. J. McGhee, is with the University of Strathclyde, Glasgow, Scotland, G1 1XW. (e-mail: mark.j.mcghee@strath.ac.uk).

V. M. Catterson, is with the University of Strathclyde, Glasgow, Scotland, G1 1XW. (e-mail: v.m.catterson@strath.ac.uk).

B. Brown, is a simulation engineer with GSE Systems, Glasgow, Scotland, G1 1XW. (e-mail: blair.brown@gses.com).

I. Introduction

Within electrical power utilities there is a growing demand for condition monitoring methods capable of accurate prediction of mission critical systems. This increasing demand is driven by the need to improve maintenance decisions, decrease maintenance costs, and improve the safety conditions and operation of in service plant. The goal of prognostic systems is to perform accurate prognosis of an asset's future condition by predicting an estimation of remaining useful life (RUL) [1]. The field of prognostics has made great advances in certain sectors with high requirements on safety and dependability, such as the aerospace industry. However, within the power generation field, the implementation of prognostic systems has not been achieved to the same degree [2]. The main challenge is the gathering of sufficient data to enable robust testing and validation. This data issue arises as power generation systems are rarely allowed to run-to-failure and release of such data is commercially sensitive [3].

Implementation of prognostic systems within the power generation environment will have large impacts on ensuring successful economic operation and improving maintenance decisions [4]. When failures occur within power plants there results a loss of operation, unplanned downtime, higher maintenance costs, and a lack of supply to the electrical grid [5]. Therefore it is essential to implement predictive maintenance to reduce unplanned downtime utilizing prognostic maintenance policies in place of time based approaches [6] [7]. However, building physical test systems from which to generate and gather run-to-failure data within power generation is prohibitively expensive. Additionally, gathering, understanding, and transforming data provided by on-site industrial facilities into reliable models is both a costly and difficult, human resource intensive process [8].

One approach to overcoming the challenges presented by data gathering is to utilize a simulation approach to gathering the required run-to-failure data. This paper proposes the use of high fidelity full scope industrial simulators to generate degradation data within the plant environment to build, test, and validate a data driven prognostic model. A high fidelity full scope industrial pressurized water reactor (PWR) simulator is used to generate the degradation data and used to build a similarity based prognostic model. The full methodology is displayed in Fig 1.

The remainder of the paper is organized as follows. Section II discusses the use of high fidelity industrial simulators, the method of extracting data and the generation of failure data. Section III describes the creation of a similarity based prognostic model. In Section IV, the results are analyzed and the technique verified. In Section V the results are discussed. In the final section conclusions are presented.

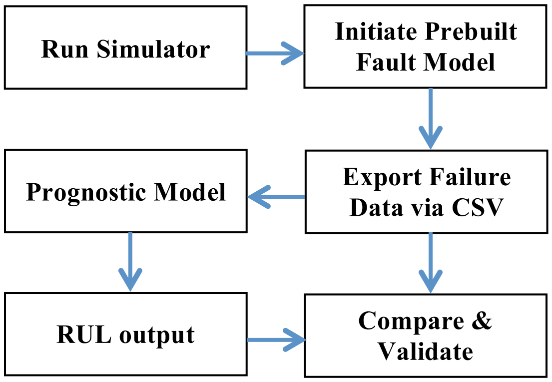

Fig. 1. Overview of the prognostic process utilizing a high fidelity simulator

II. High Fidelity Industrial Simulators

A. High Fidelity Full Scope Simulators

Within power generation, high fidelity full scope industrial simulators have been extensively deployed, particularly within the nuclear sector, for training purposes with an emphasis on improving operation and operational safety [9]. Nuclear regulations require training to take place on a high fidelity simulator which corresponds to the plant where the operator will be stationed [10][11] therefore essentially requiring all nuclear power stations to maintain such a simulator. As the benefits of the training received in this way are being observed, such high fidelity simulators are being widely deployed across all generation types and are also expanding into sectors such as refineries and chemical processing plants. However, when not in use for training, the simulator will lie idle. Utilizing the downtime of the simulator for maintenance and prognostic purposes offers the opportunity for further exploitation of the simulator asset.

A full scope simulator models all the functionality of the plant being simulated; encompassing all systems and components of the real plant they are modeled upon. Such simulators are bespoke systems built from on-site plant data and schematics to represent the whole plant operation. The simulated system mirrors the condition, components, and control system present on the actual plant. This allows operators to be trained to recognize and handle a number of scenarios, from normal operation to various rare events.

High fidelity industrial simulators are required to operate with high precision performance. The minimum requirement for precision is for critical parameters to be within 1% of real plant data. This high precision is driven by the need for operators trained on such simulators to transfer the plant behavior and knowledge gained from the simulated environment directly into the real plant environment.

These industrial simulators are certified as high fidelity tools and thereby the model and sensor data are within industrially accepted tolerances of real plant values. This research can take advantage of such high fidelity full scope simulators to remove the need for the creation of real physical test systems, whilst retaining an industrial acceptance and robustness to any simulated data generated [12].

While the simulation tools model all aspects of the plant (e.g. electrical systems, rotating plant), a critical part of both nuclear and conventional generation is the operation of the steam system. The simulation tools for modelling flow and pressures are capable of:

- Modelling two-phase behavior (both liquid and gas) and non-thermal equilibrium behavior
- Accounting for the mass, momentum, and energy balances for each phase
- Comprehensive heat transfer correlations for all heat transfer regimes (natural and forced convections, nucleate boiling, condensation)
- Encompassing laminar and turbulent flow regimes
- Separate temperature modelling for solids, including metal/concrete walls, heat exchanger tubes, turbine metals, etc.

While it is possible to place virtual measurement points at any location in the plant simulation, this research focuses only on the sensors and data currently available to plant operators. This additional constraint works as an advantage for ease of transfer to practical use, as it means the prognostic technique developed here needs no new sensors to be added to the plant. Currently sensor maintenance comprises a substantial amount of the maintenance requirements, so additional sensors would limit the benefit of the prognostic approach.

The high fidelity simulator used in the research is a full scope 970 MW PWR. Additionally the fault modelling and failures used in the research are prebuilt models present within the simulator systems. Within any given industrial simulator there can be hundreds of such failures on components and systems throughout the entire simulated plant. These prebuilt failures are driven by industry (i.e. a commercial customer has requested their inclusion), are typical of events seen within plant operation, and are of high importance, either for safety or operation of the running of the plant.

B. Degradation Model Implementation

A common source of faults within the power generation environment arises from valve failures. Valves are critical components within a power generation system [13] and are essential assets for safety and operational control. Valve failures are responsible for at least 7% of mechanical failures [14] [15], and within a typical generation environment there can be many hundreds of valves present [16]. As such valves are an asset which can be seen to benefit from the application of a prognostic method for improving both maintenance decisions and operational safety.

The degradation model used in this paper represents a valve steam leak in one of the steam generator paths in the main steam system (MSS) of the PWR. The chosen fault model is a prebuilt model within the simulator environment. The leak is initiated with the PWR running at base load. What can be observed once the failure has been initiated is the PWR control system responding to the failure. The control system attempts to compensate for the malfunction and maintain steady output, producing a complex response [17]. If no operator action is taken, the effect of the initiated leak is a unit trip, where the plant stops supplying power to the grid. The point at which the unit trip occurs is classed as the time of failure.

It should be noted that the plant can still operate with some degradation of a valve, and the fault takes hours to develop to the point of failure. Therefore, having an accurate prediction of time to failure would allow scheduling of the repair or replacement to fit in with other events. It would be possible to continue running the plant with some confidence that the valve will remain operational until the next outage, or, if the failure is expected to occur earlier, the downtime can at least be scheduled instead of the plant being forced offline.

For this research, the pressure and flow parameters within the MSS were analyzed and relevant parameters tracked in virtual time. For each run of the failure model, this parameter data was recorded as a time series in a comma separated values (CSV) file. This data was later used to generate a prognostic model.

The parameters chosen for analysis are focused mainly on flows and pressures on and around the degrading component. This allows an understanding of the effect the degradation is having within a broader context of the MSS system. The main parameter chosen for prognostic purposes is the pressure across the main pathway on which the leak is initiated, as this is both the most relevant and would be the easiest for the operator to observe to locate the fault.

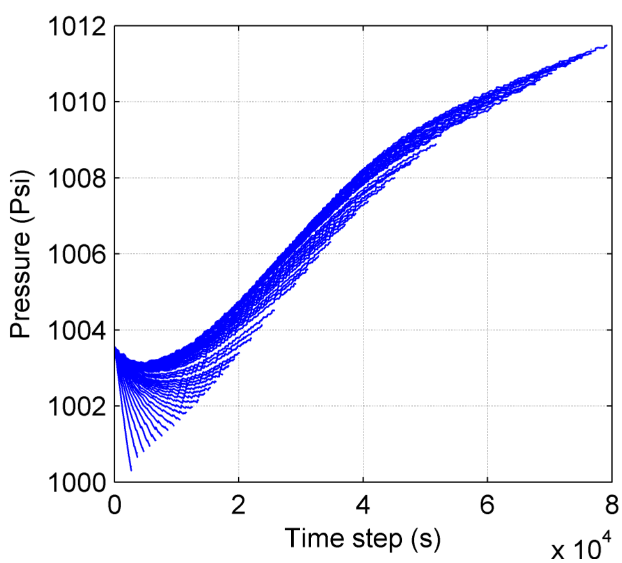

Fig. 2. 53 run to failure training events

The fault model was initiated multiple times to build up a library of run-to-failure events. The library covers degradation times from 1 up to 24 hours. For this study, the training data comprised 53 sets of time series pressure values, corresponding to the effect of the steam loss on the steam path, displayed in Fig. 2. The simulator runs at a rate of 50Hz with the data recorded at a rate of once per second. Additionally, instrument noise is enabled on all parameters to represent realistic data as would be seen on a real PWR.

III. Similarity Based Prognostics

The prognostic model used in this research is a similarity based approach derived from [18]. Similarity based prognostics has three main requirements for successful implementation as proposed by [19]:

- Multiple recordings of run-to-failure data are available,
- The recorded data ends when the point of failure is reached,
- The data covers a representative set of components.

This similarity based prognostic approach has particular resonance with this application – it requires relatively large numbers of run-to-failure events, which are provided here by use of high fidelity simulation. With use of the high fidelity simulators, the run-to-failure data required to implement a similarity based approach can be generated with ease and low cost. Additionally, as well as being able to generate data for a single component, the use of full scope simulators can enable the use of this technique across many types of failure as well as many different types of asset. Thus a similarity based prognostic method was chosen as a promising companion to the use of full scope high fidelity simulators.

A. Curve Fitting

Having defined the point of failure as the occurrence of a unit trip, the run-to-failure events, from fault inception to failure, can be seen to exhibit nonlinear behavior. The data is not-monotonic and does not have a single pressure threshold of failure. The data is arranged such that time zero is the failure threshold (Fig 3).

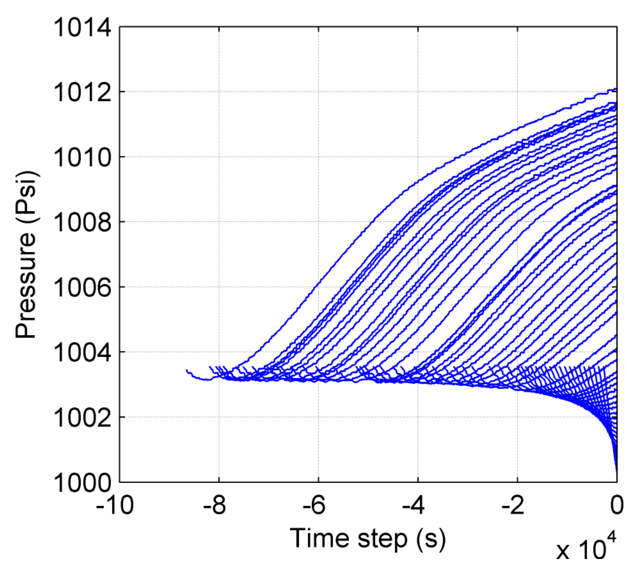

Fig. 3. 53 run to failure events rearranged with time zero being point of failure

A polynomial fitting function is then applied to each training case. Many orders of polynomial fitting were tested, as well as a Gaussian fit, to find the lowest complexity of fitted function whilst maintaining a reasonable fit. In the case study presented the fitting function is a 5^{th} order polynomial fit with mean squared error of 0.0253, compared with 0.0351 for the next best fit.

$$f(x) = a_1x^5 + a_2x^4 + a_3x^3 + a_4x^2 + a_5x + a_6 \qquad (1)$$

where a_{1-6} are the model parameters. This polynomial curve is fitted to each training case for every run-to-failure event with the least squares fitting approach.

B. RUL Calculation

For testing the performance of the model, a batch of five run-to-failure events was generated. These test events were randomly generated across the 24 hour time range from a uniform distribution and were initiated following the same process as was used for generating the training data.

To determine the RUL of the test runs, a sample of data from each test section was chosen to represent known information about the system's condition. In the following case studies each of the five test runs are analyzed five times under different selected known information conditions. The five test sections representing the known condition of the asset are varying length sections, all starting at fault inception, and ending at one of 1 hour, 2 hours, 3 hours, 4 hours, and 5 hours from the fault inception, t = 0 in Fig. 2. An example of test sections of 1 hour, 3 hours, and 5 hours for one fault are shown in Fig 4.

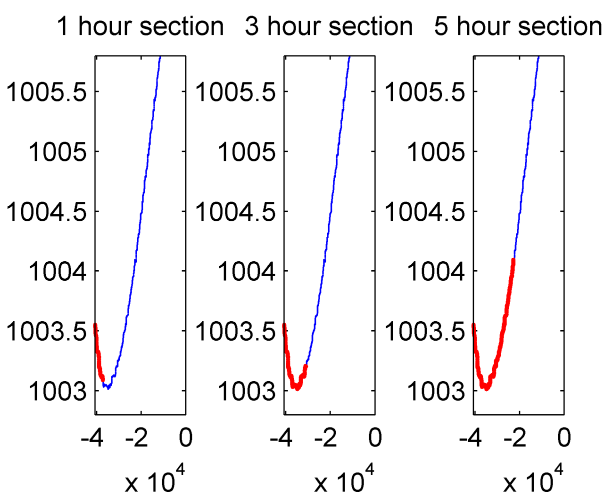

Fig. 4. Test sections lasting 1 hour, 3 hours and 5 hours. The red sections represent the section of data used for testing. The blue line represents the whole run to failure event

These test sections are then compared against every similar sized segment of each training data polynomial fit. The fit with the closest match to the test segment is then selected. The distance estimation is determined by:

$$d(\tau, Y, i) = \sum_{j=1}^{r} \frac{(y_j - f_i(-\tau - r + j))^2}{\sigma_i^2} \qquad (2)$$

where d is the distance of the test data from the training data sample, Y is the position of the test data (time step number), f_i is the polynomial curve fitted to the ith training data sample, r is the length of the test segment, τ is the number of time steps Y is shifted from 0 and σ is the root mean square (RMS) error from the polynomial fit.

Once the distance between the test run and all possible segments of all training events is calculated, the training cases are then ranked in order of the distance estimation. The relevant training case is chosen by selecting the training data with the smallest distance (i.e. the most similar run-to-failure event). The RUL for the test run is then output by evaluating the length between the matched section of the training data and the end of the training event.

The error between the predicted RUL and the true RUL of the test cases is displayed as a percentage error. The percentage error calculation is given by:

$$Error(\%) = \frac{EstimatedRUL - TrueRUL}{TrueRUL} \times 100\% \qquad (3)$$

A positive percentage represents a prediction beyond the true RUL and a negative percentage describes an early prediction of failure.

IV. Results

The results of the five test cases using the five known condition segments are displayed in Table 1. In Table 1, PR and TR represent the predicted RUL and true RUL respectively and all measurements are in minutes. Fig 5 displays the error from Table 1.

TABLE I

Case Study Results for 5 test runs using 5 different test sections

Test Case		*1 hour*	*2 hours*	*3 hours*	*4 hours*	*5 hours*
1	PR	*246*	*186*	*142*	*66*	*n/a*
	TR	*241*	*181*	*138*	*61*	*n/a*
2	PR	*487*	*426*	*415*	*205*	*245*
	TR	*505*	*445*	*402*	*325*	*265*
3	PR	*590*	*632*	*629*	*622*	*515*
	TR	*776*	*716*	*672*	*596*	*536*
4	PR	*667*	*754*	*709*	*711*	*651*
	TR	*943*	*883*	*839*	*763*	*703*
5	PR	*669*	*1224*	*1180*	*1103*	*1040*
	TR	*1229*	*1169*	*1126*	*1049*	*989*

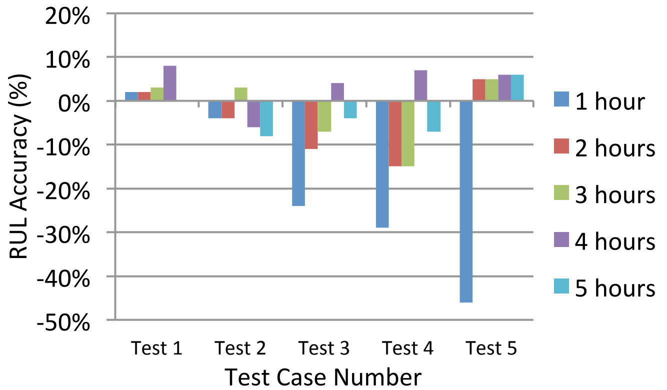

Fig. 5. Accuracy of the similarity based approach for each test case across the five testing times

The test cases are arranged by length of RUL with the shortest RUL first. In Test Case 1, at 5 hours, the valve has arrived at the failure threshold, so there is no RUL to predict. It can be seen that the shorter cases (case 1 and 2) have high accuracy, within 5%, after the first hour segments and the longer lasting cases (cases 3, 4 and 5) have poorer performance at the 1 hour segment. This difference is due to the difference in behavior of the failure. The shorter cases have less complex behavior whilst the longer cases have higher order complex behavior which drives the requirement for the 5th order polynomial fitting described. However as more data is included in the test segments the accuracy of the predicted RUL improves, after the 3 hour test segment all predicted RULs are within 10% of the true RUL.

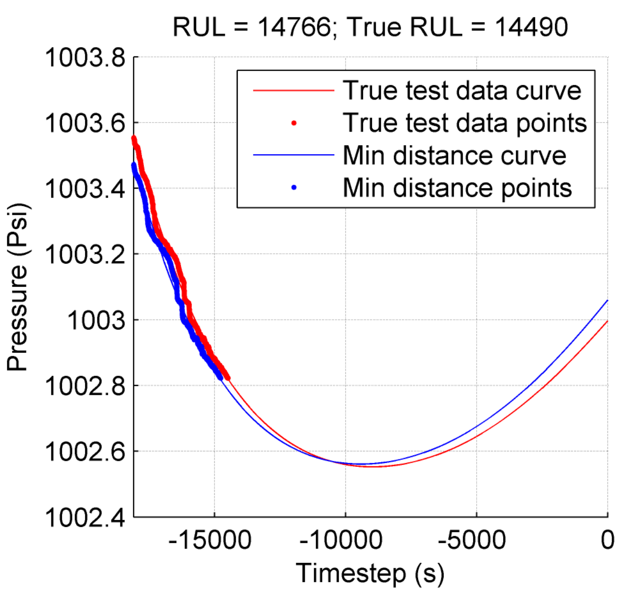

Fig. 6. Test Case 1, 1 hour, estimated RUL = 14766s, true RUL = 14490s

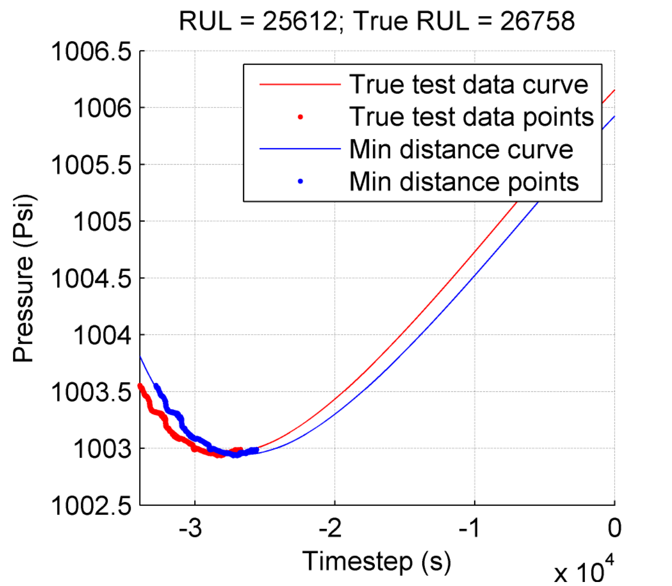

Fig. 7. Test Case 2, 2 hours, estimated RUL = 25621s, true RUL = 26758s

Fig 6-10 give a method of visualizing how the similarity based prognostic model has arrived at its estimated RUL values. Each graph shows the model representing the closest matched fit to the test segment, represented in blue. The test segment and true RUL curve of the test cases are displayed in red. The RULs are displayed in seconds, as this is the time base of the recorded data, and converted into minutes for Table 1 for ease of understanding. Fig 6-10 provide examples of each test case with different lengths of test segment, but all cases could be visualized in the same manner.

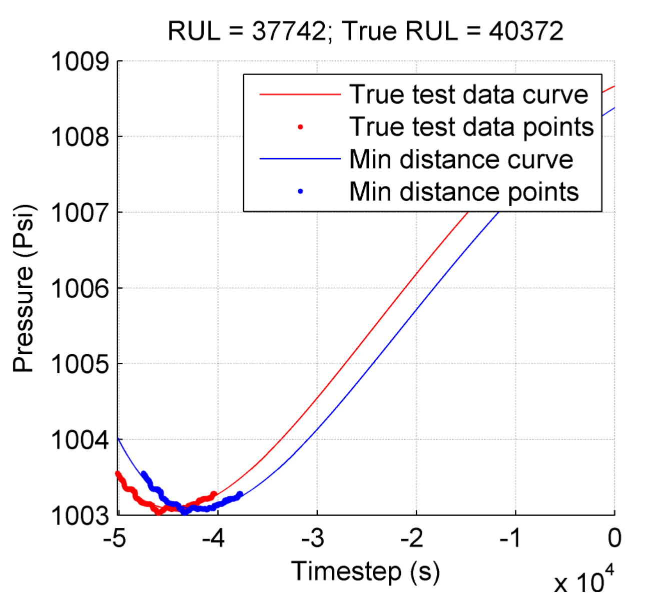

Fig. 8. Test Case 3, 3 hours, estimated RUL = 37742s, true RUL = 40372s

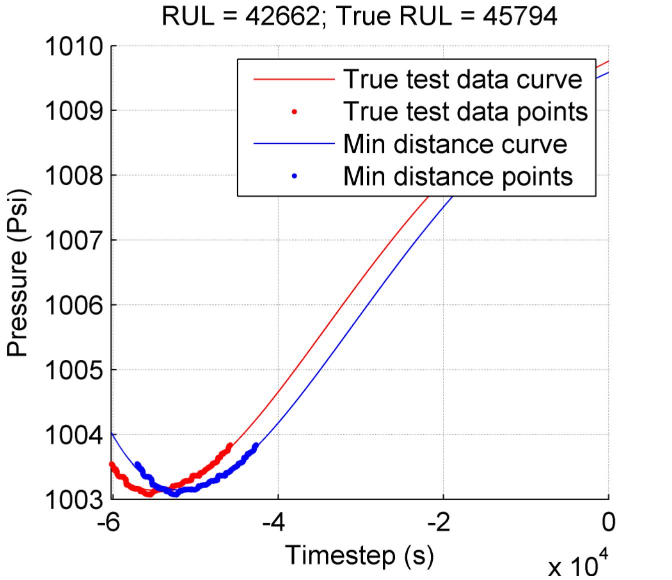

Fig. 9. Test Case 4, 4 hours, estimated RUL = 42662s, true RUL = 45794s

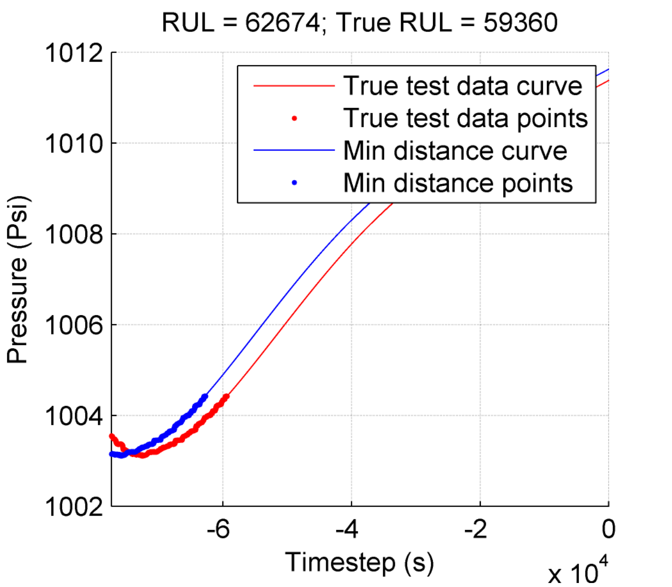

Fig. 10. Test Case 5, 5 hours, estimated RUL = 62674s, true RUL = 59360s

V. Discussion

For a prognostic approach to be acceptable the primary issues for successful implementation of the approach are a consistent accuracy as well as providing enough time for corrective or maintenance decisions to be enacted. As can be seen by the results displayed in Table 1, the overall accuracy of the similarity method is acceptable by these standards as 19 out of 25 RUL predictions lay within 10% of the true RUL, as well as all RUL predictions after the first three hour sections being within 10% of the true RUL. The high accuracy of the RUL predictions of the test cases is likely to be a result of the large number of training events for building the method. This is however the main benefit of utilizing a simulation approach for failure data generation. Although the RUL is calculated as the time until the occurrence of a unit trip, operator actions could be carried out in advance of this failure threshold. As mentioned previously, the plant can continue operation with some degradation of the valve, and this fault condition can continue for multiple hours before failure. Appropriate maintenance actions can be taken within the time between fault inception and failure, informed by the RUL output. In particular the accuracy of the longer failure events such as Test Cases 4 and 5 would provide operators with a reliable indication of the time window in which to perform corrective action or a safe, scheduled shut down of the plant.

The presence of both early and late predictions, shown in Fig 5, are a result of the similarity based prognostic approach trying to find the closest fit with the test case lying approximately halfway between two training cases. This could be resolved by weighting the RUL predictions more strongly in favour of earlier predictions when finding the closest match.

An additional benefit of generating data within the simulated plant environment is the ability to observe the plant control responses and behaviour. As can be seen in the case study this can result in complex behaviour which would also be observed by the plant operators as the failures evolve. This allows prognostic systems to be built incorporating the data available to the operators.

With the full scope high fidelity simulator, plant conditions can be varied and reset for multiple fault runs. Utilizing the high fidelity simulators in this manner will allow relevant conditions and failures to be modelled for a wide range of assets and failures. Since various failure models are built for operator training scenarios, this full range of faults can be used to generate datasets. This method of utilizing full scope high fidelity simulators can therefore be considered a promising method for generating a large range of failure data, overcoming one of the main obstacles for implementing successful prognostic applications within the power generation domain.

A potential disadvantage of the technique is that it is restricted to apply only to observed failure events, for which there is data. This is alleviated by the use of the failure models present in the full scope simulator, as these failures are driven by industrial concerns and simulation of the outcomes is certified as robust and accurate across the plant systems.

Future work will look at effective sizes of training data libraries whilst retaining accurate predicted RUL. This may require more robust and extensive distance evaluation metrics.

VI. Conclusion

The similarity based prognostic approach described in this paper provided acceptable results based on the system requirements when estimating the RUL of a steam leak within a PWR. The research utilized a high fidelity full scope PWR simulator to generate the failure data by enabling the creation of a large library of data from failure events. The prebuilt failure model simulates a steam leak in the MSS of the PWR, which results in a unit trip if no action is taken.

The research provides evidence of the success of utilizing preexisting full scope industrial simulators for generation of the relatively rare failure case data. As hundreds of full scope high fidelity simulators have already been deployed, this paper validates one solution of overcoming the main challenge to the creation of prognostic systems – a lack of failure data – by utilizing such full scope high fidelity simulators for prognostic purposes.

ACKNOWLEDGMENT

The authors thank Dr Graeme West for discussions held in the preparation of this paper.

REFERENCES AND FOOTNOTES

[1] S. Sankaraman, M. J. Daigle, K. Goebel, "Uncertainty Quantification in Remaining Useful Life Prediction Using First-Order Reliability Methods" *IEEE Trans. Reliability,* vol. 63, no. 2, pp 603 – 619, June 2014

[2] V. M. Catterson, J. J. A. Costello, G. M. West, S. D. J. McArthur and C. J. Wallace, "Increasing the Adoption of Prognostic Systems for Health Management in the Power Industry", *Chemical Engineering Transactions* vol. 33, 2013

[3] A. Heng, A. C. C. Tan, J. Mathew, N. Montgomery, D. Banjevic, and A. K. S Jardine, "Intelligent Condition-Based Prediction of Machinery Reliability," *Mechanical Systems and Signal Processing*, vol. 23, no 5, pp 1600–1614, July 2009

[4] P. Johnson "Building Asset Monitoring and Prognostics Systems using Cost Effective Technologies for Power Generation Applications", in *IEEE PHM,* Gaithersburg, MD, June 2013

[5] C. Shuping, W. Ruitao, L. Yukun and G. Ming "The Application of Failure Prognostic System in State Monitoring of Power Plant Generation Equipments" in WAC, Puerto Vallarta, Mexico, June 2012

[6] G. Vachtsevanos, F. L. Lewis, M. Roemer, A. Hess, and B. Wu, "Prologue", in *Intelligent fault diagnosis and prognosis for engineering system*. Hoboken, NJ, John Wiley & Sons, 2006, pp xv-xxii

[7] B. Sun, S. Zeng, R. Kang and M.G. Pecht, "Benefits and Challenges of System Prognostics", *IEEE Trans. Reliability*, vol. 61, no. 2, pp 323 – 335, June 2012

[8] W. Wang and M. Carr, "A Stochastic Filtering Based Data Driven Approach for Residual Life prediction and Condition Based Maintenance Decision Making Support", *IEEE PHM*, Macao, January 2010

[9] S. Harrison, "The Case for Simulation and Visualisation Based Training", *MECSS,* Amsterdam, Netherlands, October 2013

[10] *Regulatory Guide 1.149 - Nuclear Power Plant Simulation Facilities for Use in Operator Training and License Examinations,* USNRC Power Reactors (Division 1): Regulatory Guides 1.141 - 1.160, 2011

[11] The United Kingdom's National Report on Compliance With The Obligations of the International Convention on Nuclear Safety, ONR, Revision 2, 2001

[12] M. J. McGhee, V. M. Catterson, S. D. J. McArthur and E. Harrison, "Using a High Fidelity CCGT Simulator for building Prognostic Systems", *EuroTechCon,* Glasgow, UK , November 2013

[13] J. B. Coble, P. Ramuhalli, L. J. Bond, J. W. Hines and B. R. Upadhyaya, "Prognostics and Health Management in Nuclear Power Plants: A review of Technologies and Applications", Pacific National Northwest Lab., Richmond, WA, Rep. PNNL – 21515, July 2012

[14] J. Latcovich, T. Åstrom, P. Frankhuizen, S. Fukushima, H. Hamberg and S. Keller, "Maintenance and Overhaul of Steam Turbines", *International Association of Engineering Insurers 38th Annual Conference*, Moscow, Russia, Rep. IMIA – WGP 42, September 2005

[15] G. Radu, D. Mladin and I. Prisecaru., "Analysis of potential common cause failure events for Romania-TRIGA 14 MW reactor", *Nuclear Engineering and Design*, vol. 265, pp. 164-173, December 2013

[16] Westinghouse Nuclear, "AP1000 PWR Nuclear Reactor Brochure", 2013

[17] B. R. Upadhyaya, K. Zhao and B. Lu, "Fault Monitoring of Nuclear Power Plant Sensors and Field Devices" *Progress in Nuclear Energy* vol. 43, pp. 337-342, 2003

[18] M. J. McGhee, G. Galloway, V. M. Catterson, B. Brown and E. Harrison, "Prognostic Modelling of Valve Degradation within Power Stations", *Annual Conf of the PHM society*, Fort Worth, TX, September 2014

[19] T. Wang, J. Yu, D. Siegel and J. Lee, "A Similarity Based Prognostics Approach for Remaining Useful Life Estimation of Engineered Systems", *Annual Conf of the PHM society*, Denver, CO, October 2008

Mark J. McGhee is a PhD student within the Institute for Energy and Environment at the University of Strathclyde, Scotland, UK. He received his MSci in Applied Physics from the University of Strathclyde in 2012.

His PhD focuses on condition monitoring and prognostics for power plant systems, in collaboration with GSE Systems, a provider of high fidelity industrial simulation technology and training solutions.

Victoria M. Catterson is a Lecturer within the Institute for Energy and Environment at the University of Strathclyde, Scotland, UK. She received her B.Eng. (Hons) and Ph.D. degrees from the University of Strathclyde in 2003 and 2007 respectively.

Her research interests include condition monitoring, diagnostics, and prognostics for power engineering applications.

Blair Brown is a Simulation Engineer with GSE Systems, Glasgow, UK.