Comparison of different classification algorithms for fault detection and fault isolation in complex systems

Jung, Marcel; Niculita, Octavian; Skaf, Zakwan

Published in:
Procedia Manufacturing

DOI:
10.1016/j.promfg.2018.01.016

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
Jung, M, Niculita, O & Skaf, Z 2018, Comparison of different classification algorithms for fault detection and fault isolation in complex systems. in *Procedia Manufacturing.* vol. 19, Elsevier B.V., pp. 111-118. https://doi.org/10.1016/j.promfg.2018.01.016

Available online at www.sciencedirect.com

ScienceDirect

Procedia Manufacturing 19 (2018) 111–118

Procedia MANUFACTURING

www.elsevier.com/locate/procedia

6th International Conference on Through-life Engineering Services, TESConf 2017, 7-8 November 2017, Bremen, Germany

Comparison of Different Classification Algorithms for Fault Detection and Fault Isolation in Complex Systems

Marcel Jung [a]*, Octavian Niculita [a], Zakwan Skaf [b]*

[a]*Glasgow Caledonian University, Glasgow, G4 0BA, UK*
[b]*Cranfield University, Cranfield, MK43 0AL, UK*

Abstract

Due to the lack of sufficient results seen in literature, feature extraction and classification methods of hydraulic systems appears to be somewhat challenging. This paper compares the performance of three classifiers (namely linear support vector machine (SVM), distance-weighted k-nearest neighbor (WKNN), and decision tree (DT) using data from optimized and non-optimized sensor set solutions. The algorithms are trained with known data and then tested with unknown data for different scenarios characterizing faults with different degrees of severity. This investigation is based solely on a data-driven approach and relies on data sets that are taken from experiments on the fuel system. The system that is used throughout this study is a typical fuel delivery system consisting of standard components such as a filter, pump, valve, nozzle, pipes, and two tanks. Running representative tests on a fuel system are problematic because of the time, cost, and reproduction constraints involved in capturing any significant degradation. Simulating significant degradation requires running over a considerable period; this cannot be reproduced quickly and is costly.

Peer-review under responsibility of the scientific committee of the 6th International Conference on Through-life Engineering Services.

Keywords: Maintenance, Monitoring, Machine Learning

* Corresponding author. Tel.: +441412731816.
E-mail address: mjung200@caledonian.ac.uk

2351-9789
Peer-review under responsibility of the scientific committee of the 6th International Conference on Through-life Engineering Services.
10.1016/j.promfg.2018.01.016

1. Introduction

In the past, industries have always focused on quality, efficiency, and cost effectiveness. With technological advancements over the past decade, industries have added one very important aspect to their top priorities: the ability to detect and isolate faults from the onset of evolving. These advancements have made it possible to intelligently measure conditions of equipment, as well as the transport of mediums, in a cost-effective way. Data from sensors placed at specific locations can reveal instrumental information about the health of the system. This information can be of high value to an operator if processed accurately and can lead to actions that ensure processes run smoothly, reduce the risks of unexpected downtime, and deliver expected throughput, maximizing the overall availability of an asset/system/process or a plant. Obtaining real datasets to be used for development and testing of fault detection and fault isolation algorithms is always challenging. Running representative tests on a fuel system are even more problematic because of the time, cost, and reproduction constraints involved in capturing any significant degradation.

Three different approaches to a solution can be taken. Firstly, accelerated testing, which can be achieved by increasing the duty of the components or by manufacturing them using less durable materials. Secondly, by knowing the degradation modes to be investigated, the components can be machined to represent the degraded mode – often referred to as 'seeded fault' testing. The second solution represents only one snapshot in the wear process but can be repeated gradually to increase the effect. Thirdly, by emulating some degradation modes, e.g. a filter replaced by a valve, so that a clogged filter failure mode can be emulated by gradually closing the valve. The latter of these three approaches is adopted in this paper.

Sensor set optimization in the context of fault diagnosis accuracy is a major challenge for specific industry sectors like aerospace as every single additional sensor has an impact on the overall asset availability. A widely-accepted process to identify the minimum number of sensors capable of meeting the fault detection and isolation (FDI) requirements is not available and very often, OEMs ended up developing bespoke processes and tools suitable to specific applications/systems. It was demonstrated, using the same fuel system testbed, how such a process supporting instrumentation optimization can be implemented by using a quantitative model-based approach [1]. This paper continues such work and evaluates the impact on FDI accuracy when using data from optimized vs. non-optimized sensor set solutions.

In general, fault diagnostic approaches in the literature can be categorized into model-based approaches and data-driven approaches, based on the process knowledge that is required a priori [2]. The hierarchy of fault diagnosis approaches is shown in Fig. 1. The model-based approaches depend on a fundamental understanding of the physics of the process. Among thesis model-based approaches are parameter estimation methods [3], parity relation methods [4] and fault tree methods [5].

In data-driven methods, it is given to have a significant amount of available historical process data. In these methods, features can be extracted to present the historical process data as a priori knowledge to a diagnostic system. The feature extraction methods can be divided into qualitative methods, such as expert systems [6] and qualitative trend analysis [7], or quantitative methods. Also, the quantitative feature extraction methods can be categorized into statistical and non-statistical methods. Among the quantitative feature extraction methods are principal component analysis/partial least squares [8], neural network [2], and statistical classifier methods [9].

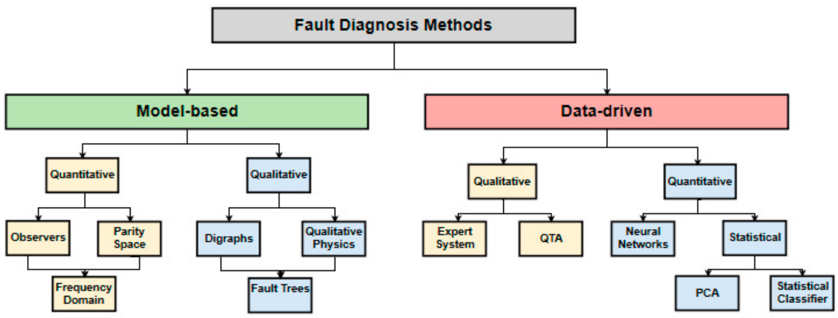

Fig. 1. Classification of fault diagnosis methods [2].

Besides the methods mentioned, there also exist many hybrid methods in the realm of fault diagnosis. [10] used a neural network method and fuzzy logic method to diagnose faults in dynamic systems. [11] proposed a fuzzy/Bayesian formulation for detecting faults in dynamical systems. [12] used the genetic algorithm and SVM for power transformer fault diagnosis. In this work, the SVM, WKNN and DT approaches are used to distinguish between healthy and faulty data of a fuel system test rig.

2. Brief Description of Classification Methods

2.1. Support Vector Machine (SVM)

The Support Vector Machine algorithm (SVM) [13] is one of the common powerful learning methods for statistical pattern recognition, with an effective implementation in a wide range of industrial applications [14]. The basic principal is the separation of two groups of data through hyperplanes/boundaries with the maximum margin of separation. The optimal separating hyperplane is defined as the one that maximizes the distance between the hyperplane and the nearest points of both groups of data.

Fig. 2 demonstrates the fundamental idea behind the SVM binary classification method in the simplest case i.e. linear separation where g(x)= 0 represents the hyperplane (solid line), g(x)= +1 belongs to one group of data (crosses), g(x)= −1 to another group of data (circles), and r stands for the distances between the hyperplane and the support vectors (blue solid line).

.

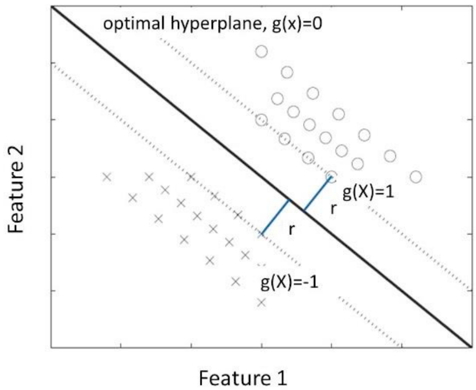

Fig. 2. SVM in the linear case.

2.2. Distance-Weighted K-Nearest-Neighbor (WKNN)

K-Nearest-Neighbor (KNN) is a non-parametric classification algorithm that is first proposed by Fix and Hodges in an unpublished paper, written in 1951, that is reviewed 28 years later by Silverman and Jones [15]. The KNN method is commonly used for statistical pattern recognition but has been implemented in a wide range of industrial applications because of the simplicity and effectiveness [16]. The KNN method is based on finding k-nearest samples in some reference set by taking a majority vote among the classes of these k samples.

A modification of KNN has been presented to improve the misclassification rate to that of the standard KNN approach. The distance-weighted version of the KNN algorithm is one of these modifications that implement the weighting of distances rather than the standard voting which is based on majority. Lei in [17] shows that the WKNN method yields to more effectiveness compared to the standard KNN, particularly when dealing with not well-separated data.

2.3. Decision Tree (DT)

The Decision Tree (DT) method is a classifier that represents its classification knowledge in a tree structure as a recursive partition of the data space. Decision trees operate on a divide and conquer approach. This means that a tree can be formed from a dataset and then be broken down into smaller subsets until no further splitting can be implemented (unless a limit for splitting is set). Ultimately, the goal is to achieve a tree structure that is based on decision nodes that break further down into branches and leaf nodes that are the decisions, i.e. classifications. The node that is furthest on top is called the “root” and is the best predictor of a given dataset. All subsets, i.e. predictors that are less important for classification, are branching off from this node and build the decision tree.

For example, the structure of a simple decision tree is shown in Fig. 3. Pressure was determined as the root. Normal, FM1 and FM2 express the leaf nodes. Clearly, classification rules can then be defined as follows:

- IF 1.5 bar < Pressure <2.0 bar Then Normal
- IF 1.5 bar > Pressure Then FM2
- IF 2.0 bar < Pressure Then FM1

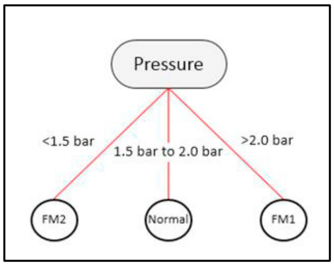

Fig.3. Structure of a simple DT.

3. Experiment and Results

3.1. Fuel System Test Rig

This section describes the fuel system tested briefly with different failure modes [1, 18]. The fuel system test rig is a laboratory testbed representing an Unmanned Aerial Vehicle (UAV) fuel system. The test rig is specifically designed to replicate some component degradation faults with high accuracy and repeatability so that it can produce benchmark datasets to evaluate and assess the developed algorithms. The rig consists of the following representative components: Main and sump tank, external gear pump, filter, polyurethane tubing, direct proportional valves (DPVs), and instrumentation. The modified schematic that is based on five valves can be seen in Fig. 4a. A photograph of the test system (valves circled in red) is displayed in Fig. 4b.

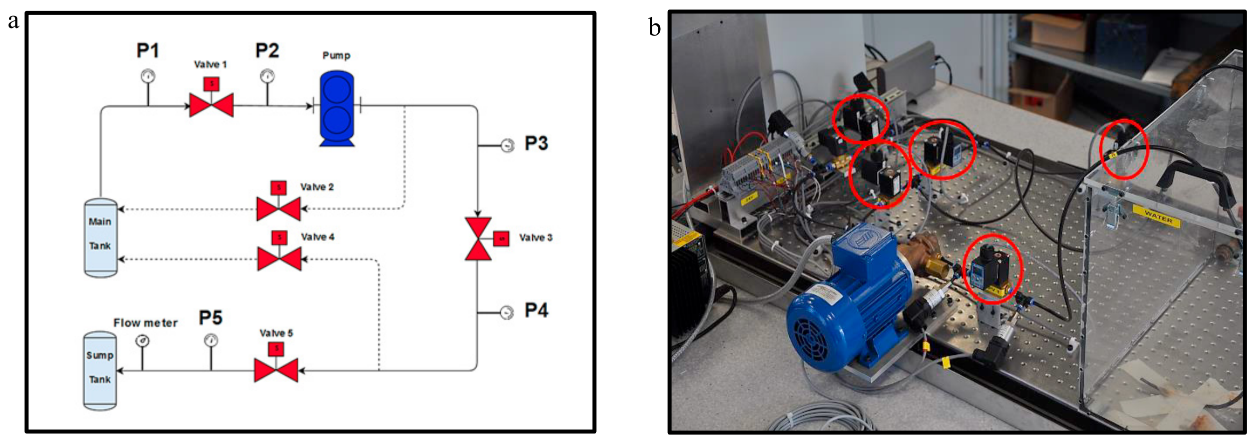

Fig. 4. (a) Schematic of modified fuel system; (b) Fuel system test rig.

National Instruments (NI) LabVIEW virtual instrumentation has been utilised to control the fuel system test rig. The main GUI for controlling the fuel rig has three layers, as shown in Fig. 5a. The user has control over valve position and pump speed in the first top layer. The second layer enables the user to monitor the fault injection mechanisms at the component level. The monitoring is done via knobs that are setting the position of the five DPVs. The third layer allows the user to control the fault injection mechanism at the sensor level.

The degradation of five different components is emulated by closing or opening one of the five DPVs. From the hardware and software point of view, the control system accommodates five failure modes in a plug and play manner: a clogged filter [F1], a degrading pump [F2], a stuck valve [F3], leaking pipe [F4], and a clogged nozzle [F5]. These five faults are emulated by using valve 1 to 5, respectively (Table 1). For example, the filter was replaced by DPV1 (valve 1) and set to be initially fully open (this setup captures the healthy scenario as the pressure drop across the valve was identical to the pressure drop across the filter). By gradually closing this valve, the system replicates a behavior of a clogged filter as shown in Fig. 5b.

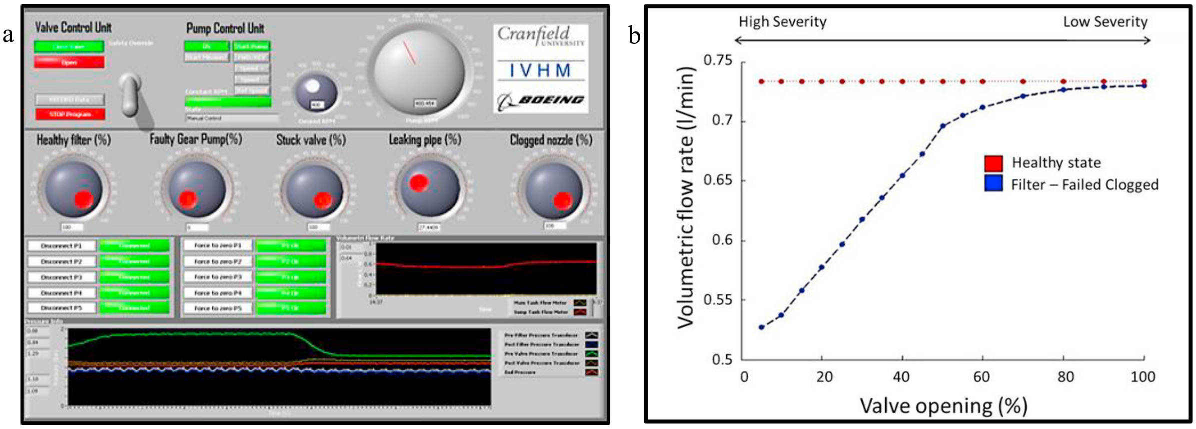

Fig. 5. (a) Fuel system – GUI for controls; (b) Output of simulated failure mode for a clogged filter vs. the healthy state.

Table 1. Fault emulation via DPVs.

Faults	Valve	Initial Position
Clogged filter [F1]	DPV1	Open
Degrading pump [F2]	DPV2	Closed
Stuck valve [F3]	DPV3	Open
Leaking pipe [F4]	DPV4	Closed
Clogged nozzle [F5]	DPV5	Open

3.2. Test and Results

In this work, all faults have been tested under three different degrees of severities. For example, a 30% severity in leakage is equivalent to an opening of valve 2/valve 4 by 30% (70% closed). While, a 30% severity in clogged filter, means that valve 1 is closed by 30% (70% open). Similarly, the stuck valve and clogged nozzle, where a 30% fault severity represents a 30% closed valve. The experiment is therefore constraint to three fault severities. Experimental measurements that lie in between these values are classified to the nearest class of data and depend on the decision of the classification algorithm. The algorithms were trained by the healthy condition and the three fault severities i.e. low severity (LS), medium severity (MS) and high severity (HS), see Fig. 6.

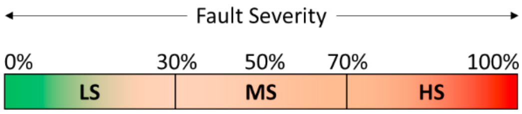

Fig. 6. Fault severity notation used during the experiment.

Fig. 7a and 7b show the selected features i.e. system pressure against the flow rate of the system for two different sensor sets. The first sensor set has five pressure sensors (P1-P5). While the second sensor set has three pressure sensors (P2-P4). All solutions, obtained during the classification performance comparison, will consider these two sensors sets. Both figures represent five mean values of five seconds of data (sample rate 1kHz).

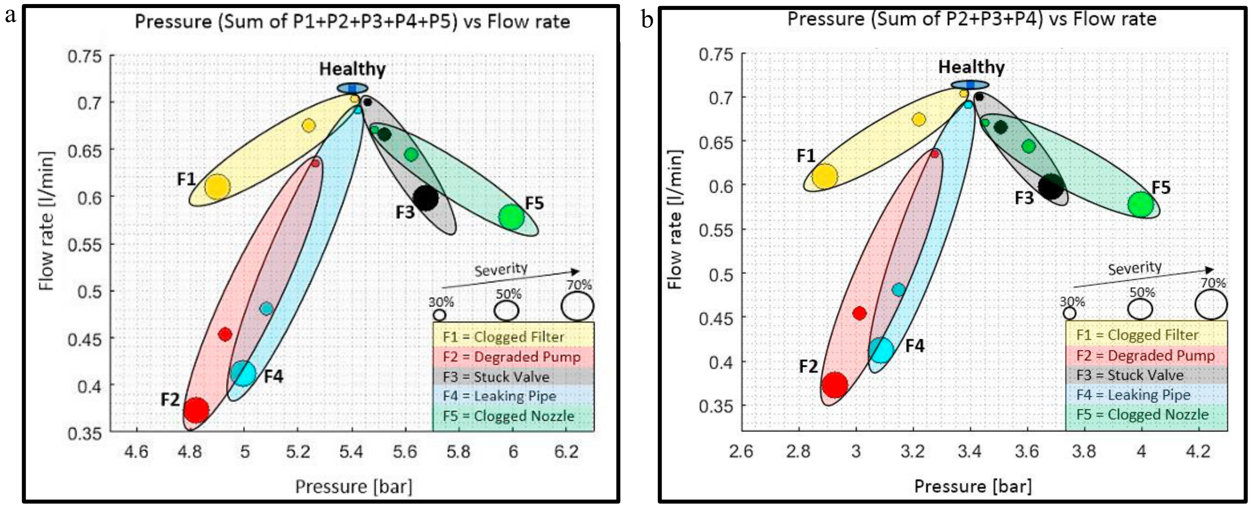

Fig. 7. (a) System Pressure (Sum of P1+P2+P3+P4+P5) vs Flow rate; (b) System Pressure (Sum of P2+P3+P4) vs Flow rate.

3.3. Classification Performance Evaluation

All classifiers used in this study are trained on k-fold-cross validation. Each classifier was trained on 10-fold-cross validation. A well common method for comparing the performance of different classifier techniques is the confusion matrix. The confusion matrix assesses the performance of a classifier by comparing the actual with the predicted classification.

Fig. 8a shows the confusion matrix for DT with five mean values (sample rate 1 kHz), representing five seconds' worth of data per condition (F1[LS], F1[MS] etc.). There are ten misclassifications during the training. These misclassifications exist between the healthy state and the low severe leaking pipe, as well as between a severe stuck valve and clogged nozzle.

Fig. 8b shows that SVM and WKNN achieve ideal validation and testing accuracies for both sensor sets. DT acquires lower accuracies during the training with fewer samples, compared to SVM and WKNN, for both sensor sets. The DT seems to be less robust as it uses only one variable (pressure) as the determining factor. The way the DT branches off differs with sample sizes. The lower sample size of 80 samples could not reach a validation accuracy above 87.5%, leaving the DT as the weakest classification algorithm in this experiment.

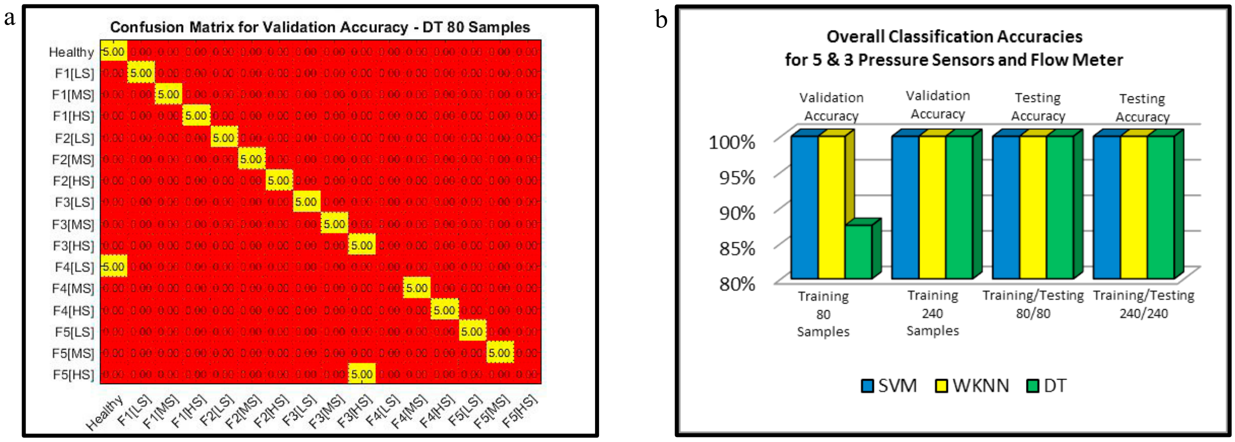

Fig. 8. (a) Confusion matrix for DT with 80 samples [19]; (b) Overall classification accuracies for 5 & 3 pressure sensors and flow meter.

The classification algorithms are trained and tested in a two-step process. Firstly, an Excel file is created, that captures five/fifteen seconds’ worth of data from the measurement data sets (gathering the pressure readings at five/three different locations at the same time with the volumetric flow rate). Additionally, filtering is applied to minimise the false alarm rate. The size of 5000/15000 samples (sample rate 1 kHz) per condition and sensor is reduced by taking the mean values of each second of data. This forms five/fifteen values worth five/fifteen seconds per condition and sensor. Table 2 reflects the time taken to create these files. The files are then used as input for the classification process. The differences between the sample sizes and classification algorithms can be seen in Table 3.

Table 2. Creation of processed pressure and flow rate files.

	Sensors: Sum of (P1, P2, P3, P4, P5) vs flow meter		**Sensors: Sum of (P2, P3, P4) vs flow meter**	
Time (sec)	**80 samples**	**240 samples**	**80 samples**	**240 samples**
	39.511	39.837	37.303	38.101

Table 3. The impact of different sampling sizes on training and testing times.

	Sensors: Sum of (P1, P2, P3, P4, P5) vs flow meter				**Sensors: Sum of (P2, P3, P4) vs flow meter**			
	80 samples		**240 samples**		**80 samples**		**240 samples**	
Classifier	Time of training (sec)	Time of testing (sec)	Time of training (sec)	Time of testing (sec)	Time of training (sec)	Time of testing (sec)	Time of training (sec)	Time of testing (sec)
SVM	20.903	0.117	20.989	0.151	20.909	0.108	20.918	0.141
WKNN	2.498	0.042	2.537	0.048	2.591	0.046	2.617	0.053
DT	2.535	0.053	2.563	0.061	2.592	0.055	2.723	0.058

Table 3 shows that the training times for both WKNN and DT are very similar; while the training times for the SVM classifier need about eight times longer.

All comparisons in this work are dependent on the hardware specifications of the computer that was used (Intel® Core™ i5-3317U CPU @ 1.7GHz and 4GB RAM).

4. Conclusion and Future work

This paper compared the performance of three classifiers through an experiment of a fuel system test rig under different failure modes. In terms of performance, both the SVM and the WKNN classifier perform similarly. Comparatively, the performance of the DT classifier is reasonably worse in dealing with lower sample sizes used for the training phase, but is still acceptable. The SVM is slower compared to WKNN and DT in case of training and testing times. In terms of the effects of the size of sensor sets, both the first sensor set (which contains pressure sensors P1 to P5 and one flow meter) and the second sensor set (which contains pressure sensors P2 to P4 and one flow meter) perform similarly in terms of classification accuracy and time taken for training/testing.

Future work will focus on extracting features that can accurately reflect the health status of the system under various faults with different degrees of severity, especially in the case of faults with very low degrees of severity. In addition, the experiments conducted in this paper are under steady state conditions. Future work can also include investigations on FDI accuracy with the system running under transient state conditions.

References

[1] Niculita O., Irving P., & Jennions, I. K. (2012). Use of COTS Functional Analysis Software as an IVHM Design Tool for Detection and Isolation of UAV Fuel System Faults, In Proceedings of the Prognostic and Health Management Society Conference.

[2] Venkatasubramanian,V., Rengaswamy, R., Yin, K., and Kavuri., S. N., A review of process fault detection and diagnosis part i: Quantitative modelbased methods. Computers and Chem. Eng. 27, 293—311 (2003).

[3] Rajaraman, S., Hahn, J., & Mannan, M. S. (2004). A methodology for fault detection, isolation, and identification for nonlinear processes with parametric uncertainties. Industrial & engineering chemistry research, 43(21), 6774-6786.

[4] Chen, J., & Patton, R. J. (2012). Robust model-based fault diagnosis for dynamic systems (Vol. 3). Springer Science & Business Media.

[5] Aslansefat, K., Latif-Shabgahi, G., & Kamarlouei, M. (2014). A strategy for reliability evaluation and fault diagnosis of Autonomous Underwater Gliding Robot based on its Fault Tree. International Journal of Advances in Science Engineering and Technology, 2(4), 83-89.

[6] Wu, K., An, S., Ma, G., & Tao, Y. (2012, March). Research and application of fuzzy expert system on transformer fault diagnosis. In Computer Science and Electronics Engineering (ICCSEE), 2012 International Conference on (Vol. 1, pp. 378-382). IEEE.

[7] Villez K., and Rengaswamy R., "A generative approach to qualitative trend analysis for batch process fault diagnosis," 2013 European Control Conference (ECC), Zurich, 2013, pp. 1958-1963.

[8] Chen, A., Zhou, H., An, Y., & Sun, W. (2016). PCA and PLS monitoring approaches for fault detection of wastewater treatment process. In Industrial Electronics (ISIE), 2016 IEEE 25th International Symposium on (pp. 1022-1027). IEEE.

[9] Ocak H., “Fault Detection, Diagnosis and Prognosis of Rolling Element Bearings: Frequency Domain Methods and Hidden Markov Modeling”. PhD thesis, Case Western Reserve University (2003).

[10] Lee, K. W., & Lan, H. N. (2014). Fault Detection and Identification using Neural Network and Fuzzy Logic. In IFAC Workshop (pp. 185-190).

[11] D’Angelo, M. F., Palhares, R. M., Takahashi, R. H., & Loschi, R. H. (2011). Fuzzy/Bayesian change point detection approach to incipient fault detection. IET control theory & applications, 5(4), 539-551.

[12] Li, J., Zhang, Q., Wang, K., Wang, J., Zhou, T., & Zhang, Y. (2016). Optimal dissolved gas ratios selected by genetic algorithm for power transformer fault diagnosis based on support vector machine. IEEE Transactions on Dielectrics and Electrical Insulation, 23(2), 1198-1206.

[13] Salcedo-Sanz, S.; Rojo-Álvarez, J.L.; Martínez-Ramón, M.; Camps-Valls, G. Support vector machines in engineering: An overview. Data Mining and Knowledge Discovery; JohnWiley & Sons: Hoboken, NJ, USA, 2014; Volume 4, pp. 234–267.

[14] Salcedo-Sanz, S.; Rojo-Álvarez, J.L.; Martínez-Ramón, M.; Camps-Valls, G. Support vector machines in engineering: An overview. InWiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery; JohnWiley & Sons: Hoboken, NJ, USA, 2014; Volume 4, pp. 234–267.

[15] Silverman, B.W. and Jones, M.C. (1989), ‘An Important Contribution to Nonparametric Discriminant Analysis and Density Estimation: Commentary on Fix and Hodges (1951)’, International Statistical Review / Revue Internationale De Statistique. 57(3), pp. 233-238.

[16] Gou, J., Xiong, T. and Kuang, Y. (2011), ‘A Novel Weighted Voting for K-Nearest Neighbor Rule’, Journal of Computers, 6(5).

[17] Lei, Y. (2017), ‘4 - Clustering algorithm–based fault diagnosis’, Intelligent Fault Diagnosis and Remaining Useful Life Prediction of Rotating Machinery’, Butterworth-Heinemann, pp. 175-229.

[18] Niculita O., Skaf Z., & Jennions, I. K. (2014). The Application of Bayesian Change Point Detection in UAV Fuel Systems. 3rd International Conference on Through-life Engineering Service. 2014; Volume 22, Pages 115-121

[19] Deoras, A. (2016). Customizable Heat Maps - File Exchange - MATLAB Central. [online] Uk.mathworks.com. Available at: https://uk.mathworks.com/matlabcentral/fileexchange/24253 [Accessed 6 Aug. 2017]