

# Development and Realisation of Changepoint Analysis for the Detection of Emerging Faults on Industrial Systems

Sepehr Maleki, Chris Bingham, Yu Zhang  
School of Engineering, University of Lincoln, UK  
Email: {smaleki, cbingham, yzhang} @lincoln.ac.uk

**Abstract**—An online 2-D changepoint detection algorithm for sensor-based fault detection, is proposed. The methodology consists of a *differential detector* which looks for characteristics across datasets at a particular instant, and a *standard detector* which when combined can identify anomalies and meaningful changepoints while maintaining low rates of false-alarm generation. A key aspect of changepoint detection methodologies is the setting of relevant thresholds which are typically based on empirical trial and error. Here, a statistical methodology is adopted which provides the engineer with a trade-off between correct detection and false-alarm rates, thereby informing decision making at the design stage. The efficacy of the techniques is demonstrated through application to two industry case studies of fault detection on Industrial Gas Turbines, and are shown to readily provide an early warning indicator of impending failures.

**Keywords**—Fault detection, Changepoint detection, FDI.

## NOMENCLATURE

Notation	Definition
$N$	number of data streams
$\mathbf{x}$	new measurement sequence of dimension $N$
$\mathbf{S}$	vector of $N$ variables used for calculation of incremental variance
win length	fixed length of the window
win <sub><math>i</math></sub>	$i$ -th window of length win length, $i = 1, 2, \dots, N$
$L$	number of datapoints inside win <sub><math>i</math></sub>
mean <sub>win</sub>	vector of $N$ mean values for datapoints within their respective windows
mean <sub>global</sub>	vector of $N$ mean values for the whole data-set so far
mean <sub>old</sub>	vector of $N$ mean values for the whole data-set at the previous timestamp
mean <sub>new</sub>	auxiliary vector for computation of mean <sub>global</sub>
std <sub>global</sub>	vector of $N$ standard deviation values for the whole data-set at current time
std <sub>score</sub>	vector of $N$ values indicating number of standard deviations which an observation is above or below the corresponding mean <sub>global</sub>
$SE$	standard error of the data set at current time
signal <sub>size</sub>	size of the data-set at current time
$\oslash$	element-wise division
$\sqrt{\phantom{x}}$	element-wise square root
$\circ$	Schur product operator

TABLE I. NOTATIONS

Copyright (c) 2009 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

## I. INTRODUCTION

**D**EMAND for reducing down-time and facilitating flexible maintenance scheduling on industrial machinery has driven increased attention towards the development and use of Fault Detection and Isolation (FDI) schemes. Although many techniques have been reported, they can be broadly categorised into model-, knowledge- or signal-based solutions (see e.g. [1], [2], [3]). When dealing with complex systems, however, it is often the case that only a limited understanding of the underlying relationship between many physical measurements exist to classify normal operation from that of operation subject to an emerging fault condition. Typically therefore, the temporal correlation of consecutive physical measurements provide an opportunity for machine learning and data mining techniques to be employed for FDI [4].

Novelty detection is one of the most commonly used techniques [5], [6] whereby a description (or fingerprint) of normal operation is learned/identified and subsequent operating data is then compared with the fingerprint to assess the degree of correlation; a significant disparity providing an indication of abnormal operation.

The accommodation of high rates of data capture and total data volume generated by complex sensor networks that typically monitor industrial systems pose one of the main challenges for online novelty detection. A desirable feature when designing diagnostic and prognostic algorithms is therefore to ensure they can be executed in appropriate timeframes and important historic information/features are not discarded.

Changepoint detection [7] is a well-established class of novelty detection schemes where the objective is to detect whether the general distribution of a sequence of measurements have remained steady (in some statistical sense) or have exhibited a degree of change (in whole or in part) that may be considered as abnormal. Given a data sequence that includes abrupt changes, a desirable changepoint detection algorithm must therefore be able to distinguish between “important” and “unimportant” changes in sets of measurements in order that false alarms can be minimised (the interpretation of “important” in this sense is of course application dependent).

## II. CHANGEPPOINT DETECTION

From a statistical perspective, abrupt variations that change the probability distribution of a stochastic process or time series are referred to as changepoints. Often, such variations are

important in indicating an event of interest e.g. the emergence of a fault in this case, or perhaps unimportant when indicating an expected variation due to load or set-point changes. Here changepoint detection is specifically concerned with identifying when a statistical variation between measurements is considered important in the above sense.

Typically, a statistical-based formulation of changepoint detection considers probability distributions of measurements from the past and present, with the aim of identifying if the two distributions differ by some a-priori defined degree e.g., using CUSUM (cumulative sum) [7] and GLR (generalized likelihood ratio) [8]. Using both on-line, realtime, and off-line (batch) variants, changepoint detection has previously been reported and demonstrated to be effective in application fields such as process control [9], EEG analysis [10], [11], [12], DNA sequencing [13], finance [14] and healthcare monitoring [15]. In what follows, an effective and reliable method for online changepoint detection is proposed with a focus on fault detection in industrial systems, although it is expected that the underpinning methodology is much more widely applicable to other systems.

Due to the requirement for rapid calculation, in what follows Welfords method [16] is used to calculate the variance of all on-line, real-time measurement sets i.e. using the iterative formulation:

$$M_{1j} = M_{1j-1} + \frac{1}{j}(x_j - M_{1j-1})$$

$$S_{1j} = S_{1j-1} + (j-1) \left( x_j - M_{1j-1} \right) \left( \frac{x_j}{M_{1j-1}} \right), \quad (1)$$

where

$$T_{ij} := \sum_{k=1}^j x_k, \quad M_{ij} := \frac{1}{j-i+1} T_{ij}, \quad S_{ij} = \sum_{k=1}^j (x_k - M_{ij})^2,$$

with  $M_{11} = x_1$  and  $S_{11} = 0$ .  $T_{ij}$  and  $M_{ij}$  are used to denote the sum and the mean of the data points from  $x_i$  to  $x_j$  respectively. Each iteration consists of updating  $M_{ij}$  by the addition of a single data point. The desired value of  $S$  is ultimately obtained as  $S_{1n}$  and therefore, the sample variance is calculated as  $S/(n-1)$ .

#### A. Online 2-D Changepoint Detection

Consider  $N \geq 2$  streams of measurements that are assumed to be correlated to some degree. A 2-D changepoint detection algorithm is proposed which identifies changepoints in real-time whilst maintaining lower rates of false-alarm than would normally be expected using traditional techniques. The algorithm is considered 2-D as it combines characteristics of both classical statistical time-based features, with those taken from the statistical ensemble of the collective sensor measurement dataset. Specifically, this is accomplished through the integrated use of two detection schemes, viz. a differential detector which considers changes across the different data

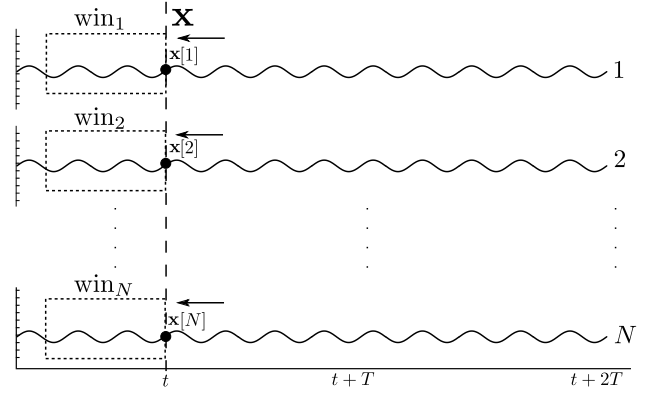


Fig. 1. Window of  $N$  measurement sequences

streams at a particular instant, and a standard detector (from the term standard score used to determine the difference between current and previous measurement sets) which aims to identify possible changepoints within individual streams. Put simply, if either detector independently flags an anomaly, then it is considered as an early warning of a possible emerging fault. However, if the integrated use of both detection schemes indicate an anomaly, it is concluded that a changepoint has definitely occurred and a fault may be emerging. A history of identified change points is then used to further reduce false-alarm rates.

Consider the sequence of measurements  $\mathbf{x} = x[1], x[2], \dots, x[N]$  at time  $t$ , (see Figure 1). The differential detector determines the standard deviation of the  $i$ -th signal from the mean of  $\mathbf{x} - \{x[i]\}$ ,  $i \in \{1, 2, \dots, N\}$  and compares it with an a-priori determined threshold (see Algorithm 1). If  $x[i]$  exceeds the threshold ( $\text{Threshold}_1$ ), an intermediate decision flag is generated.

**while** Receiving data **do**

**mean**  $\text{vs} = \frac{1}{(N-1)} \times (\text{Sum}(\mathbf{x}) \cdot \text{ones}(N) - \mathbf{x})$

**dev**  $\text{score} = \frac{1}{2} |\mathbf{x} - \text{mean vs}|$

**end**

**Algorithm 1:** Differential detector algorithm

For the development of the standard detector, a sequence of sliding windows is used which stores the  $L$  most recent measurements, considered here as the *sample*. Lengths of the windows are fixed, and are updated in a last-in/last-out pipeline manner. Measurements taken up-to the current time step are collectively referred to as the *population* (see Figure 1).

In this case, for simplicity, the standard detector is constructed by calculating the distance from the sample mean ( $M$ ) to the population mean ( $\mu$ ) in units of standard error:

$$\text{standard score} = \frac{M - \mu}{SE}.$$

**Initialise:** signal size = 0

**for**  $i = 1$  **to**  $N$  **do**

    | win <sub>$i$</sub>  = window of fixed length win length

**end**

$S = \mathbf{zeros}(N)$

**std** score =  $\mathbf{zeros}(N)$

**mean** win =  $\mathbf{zeros}(N)$

**mean** global =  $\mathbf{zeros}(N)$

**while** Receiving data **do**

**Let:**  $x$  = new data

**increment:** signal size

$L = \min(\text{win length}, \text{signal size})$

**Let:**  $\text{mean}_{\text{old}} = \text{mean}_{\text{global}}$

$\text{mean}_{\text{new}} = \text{mean}_{\text{old}} + \frac{(x - \text{mean}_{\text{old}})}{\text{signal size}}$

$S = S + (x - \text{mean}_{\text{new}}) \circ (x - \text{mean}_{\text{old}})$

$\text{std}_{\text{global}} = \sqrt[3]{\frac{S}{\text{signal size}}}$

$\text{SE} = \frac{\text{std}_{\text{global}}}{\sqrt{L}}$

**Let:**  $\text{mean}_{\text{global}} = \text{mean}_{\text{new}}$

**for**  $i = 1$  **to**  $N$  **do**

        | win <sub>$i$</sub> .append( $x[i]$ )

**end**

$\text{mean}_{\text{win}} = \text{Mean}(\text{win}_i)$

$\text{std score} = |\text{mean}_{\text{win}} - \text{mean}_{\text{global}}| \oslash \text{SE}$

**end**

**Algorithm 2:** Standard detector algorithm

This is commonly also referred to as the *standard score* and leads to the given name of the associated detector detailed here.

Welford's update method (1) is used to rapidly calculate the standard score in this case. If the standard score exceeds an a-priori fixed threshold (Threshold<sub>2</sub>), an intermediate decision flag is generated. A description of the method is given by Algorithm 2.

The overall procedure for the proposed changepoint detection

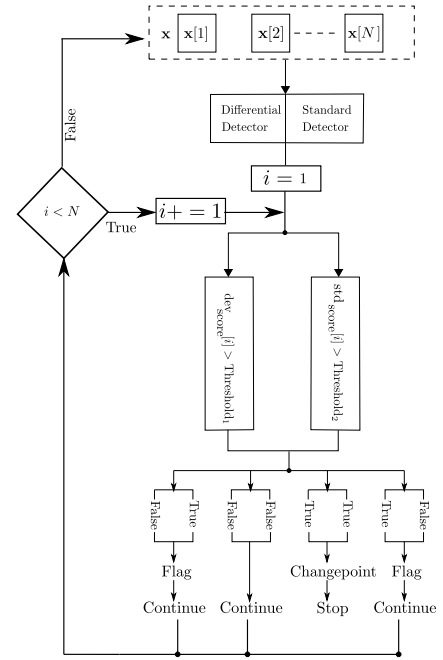


Fig. 2. Changepoint detection flowchart

methodology is summarised in the flowchart of Figure 2.

### III. EXPERIMENTAL CASE STUDIES

To provide an industrial focus and demonstrate the value of the study, the problem of fault detection on the burners of industrial gas turbines is considered. Failure to detect emerging faults early can lead to structural damage to the burner cans (cracking and deformation), Figure 3, and subsequent unscheduled shutdown of the unit. The specific engines considered here are sub-15MW Siemens IGTs with a formation of 6 circumferential equidistant burners, Figure 4. Measurements used in the study are taken from operation of the IGTs in the field. Considering the relative proximity of the burners, it is expected that the designated sensors will possess a reasonable degree of correlation during normal operation. However, importantly, burner temperature measurements can contain abrupt changes that are not a characteristic of an impending failure, but are a result of noisy measurements and changes of load/set-point and shutdowns, which are considered “normal” behaviour. The example application therefore provides a realistic industrial platform for showing how the proposed methodology discriminates between normal and abnormal changepoints without triggering the false-alarms which would be expected using alternative techniques.

#### A. Threshold Selection

**Definition 1:** [18] Consider a random variable  $X$  with the distribution function  $\mathcal{F}$  and let  $0 < p < 1$ . A value  $x_p$  is called a *quantile of order  $p$*  if

$$\mathbb{P}\{X < x_p\} \leq p \leq \mathbb{P}\{X \leq x_p\},$$

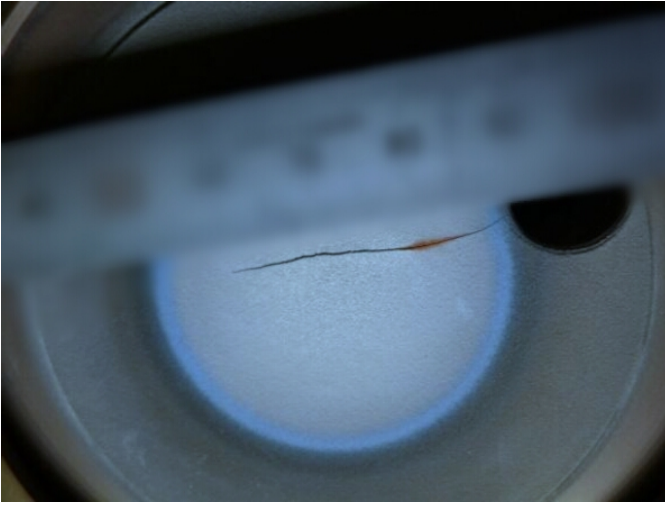


Fig. 3. The result of an undetected failure showing a crack in one of the burners

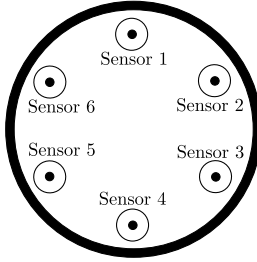


Fig. 4. Annular array of 6 burners in the combustion chamber

or equivalently

$$\mathcal{F}(x_p - 0) \leq p \leq \mathcal{F}(x_p) .$$

Specifically, a quantile of order  $p$  is the value where CDF equals the probability  $p$ . Percentiles are a relative measure of quantiles, normalised to 100. To select appropriate thresholds the algorithm is applied to what is considered as “normal” operational data for both detectors, and standard and deviation scores are consequently calculated. To determine the high quantiles of the observations, the quantile function (inverse of cumulative distribution function) is used and the score corresponding to the 99th percentile is selected as the threshold (see Figures 5 and 6). In the case where a lower figure, e.g., 95%, is chosen, then it would be statistically expected to raise the false alarm detection rate. Alternatively, if a higher confidence level is chosen then an emerging failure may take longer to detect. The process of determining the threshold is outlined in Algorithm 3.

### B. Case 1: Emerging Burner Failure

The first scenario considers an operating turbine with measurements being collected over 31 days, Figure 7. Applying

```

Initialise: std stack = [ ]
              dev stack = [ ]

while Reading data do
    read(new data line)
    Run: Algorithm 1
    get(std score )
    Run: Algorithm 2
    get(dev score )

    for  $i = 1$  to  $N$  do
        std stack.append (std score[ $i$ ])
        dev stack.append (dev score[ $i$ ])
    end
end

```

*# quantile( $q$ ) computes the  $q$ -th quantile.*

**std** threshold = **std** stack.**quantile**(.99)

**dev** threshold = **dev** stack.**quantile**(.99)

**Algorithm 3:** Algorithm for determining the thresholds

Standard Score Threshold	9.12
Deviation Score Threshold	37.23

TABLE II. EXPERIMENT PARAMETERS

the measurement data to the proposed changepoint detection algorithm results in a rolling sequence of deviations and standard score, as given respectively, in Figures 8 and 9. From Figure 8, it can be readily seen that early measurements from sensors 4 and 6 indicate different trends from the other burner sensor measurements, and from Figure 10 it can be seen that the differential detector repeatedly flags a possible emerging anomaly; although this is not apparent from the standard detector output until day 9 when it then flags a possible emerging fault from sensors 3 and 4. Having flagged possible anomalies, closer inspection of the measurements show that the temperature of burner 6 is marginally higher and burner 4 marginally lower than the rest during early periods of measurement capture. Both detectors then provide fault confirmation between days 14 and 15. Notably, the unit remained in operation until day 31 when a forced shutdown commenced. Although not directly obvious from early manual consultation of the measurements, it is clear in this instance that “early warnings” can be identified in the data some 4 weeks prior to final forced shutdown, with a definite fault condition being identified at least 2 weeks before ultimate shutdown.

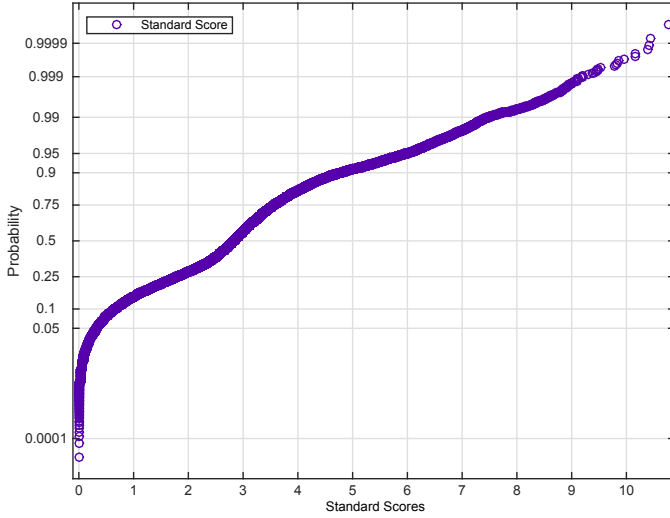


Fig. 5. Cumulative Distribution Function plot for the standard detector threshold selection

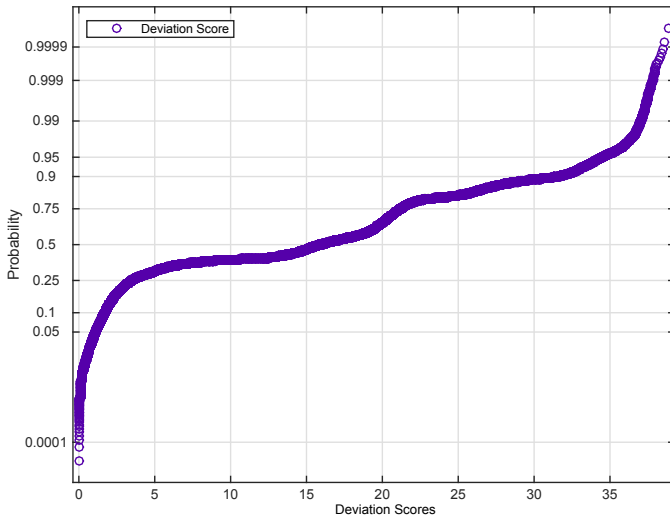


Fig. 6. Cumulative Distribution Function plot for the differential detector threshold selection

### C. Case 2: Identifying and Rejecting False Alarms

An important feature of novelty detection schemes is the ability to filter out false-alarms. Here, an example is given using measurements of burner tip temperatures over a period of 1 week (see Figure 11). It is known that no failure occurred during this period although the data contains abnormalities that would typically be perceived/detected as characteristic of a failure or emerging fault. Specifically, it can be seen that on day 2, all temperatures abruptly fall, potentially indicating a fuel system fault. The sudden change is captured by the

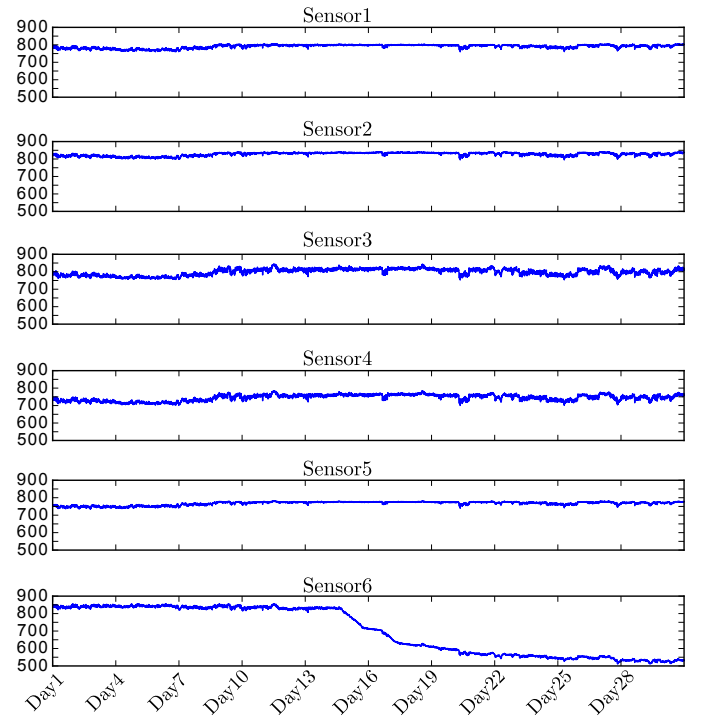


Fig. 7. Burner temperatures (degrees °C) over 31 days. Sensor 6 indicates a malfunction on day 15

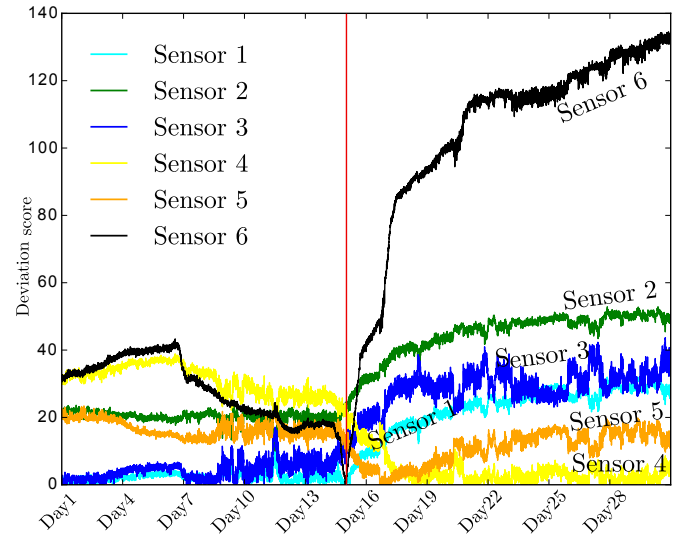


Fig. 8. Directional graph of the rolling deviation score calculated for burner sensor measurements

standard detector although it is effectively ignored by the differential detector since the characteristic is common to 5 sensors –thus, whilst an early warning is indicated, an alarm is not raised and the engine continues to operate. Further, day 4 shows a fall in temperature on most of the burners while sensor 6 reads a higher temperature for around 2 days prior to returning to normal levels by day 6. By consulting the rolling

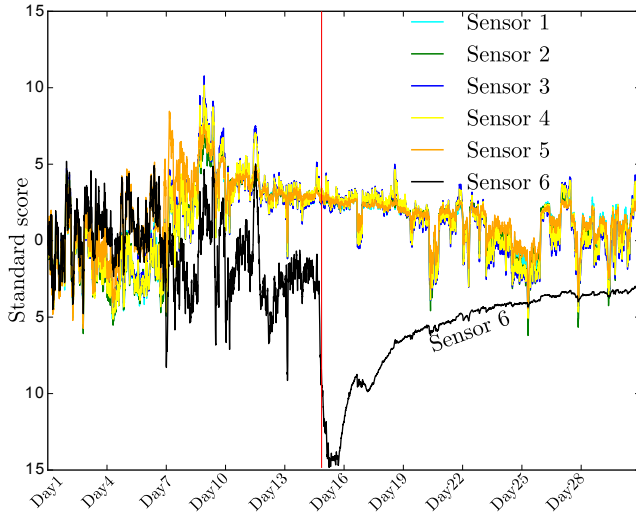


Fig. 9. Graph of the rolling standard score calculated for burner sensor measurements

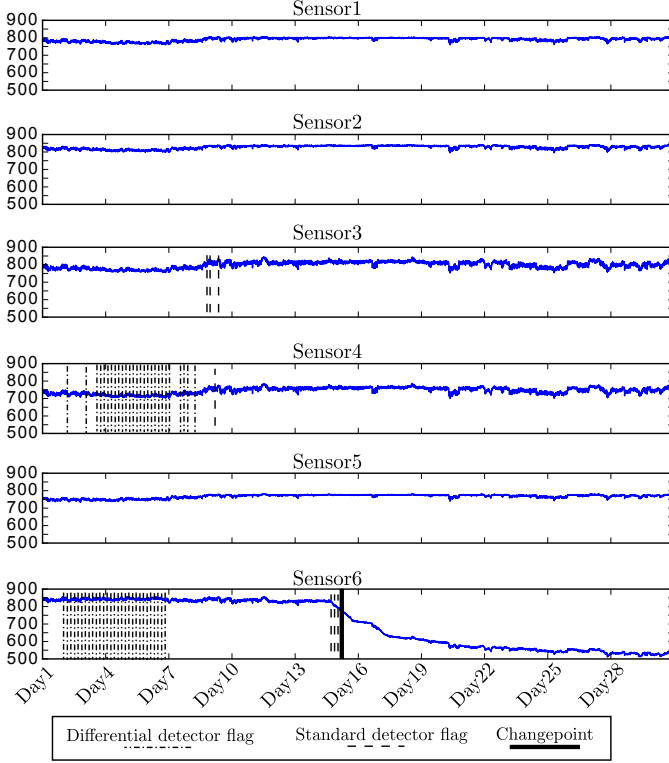


Fig. 10. Changepoints identified by the detectors

deviations and standard score value, respectively, Figures 13 and 14, it can be seen that sensor 6 does indeed show an uncharacteristic trend and that the abrupt change in day 2 is captured. In both cases, therefore, significant behaviour characteristics are identified as being present but the integrated changepoint detection methodology has correctly rejected them as evidence of an impending failure.

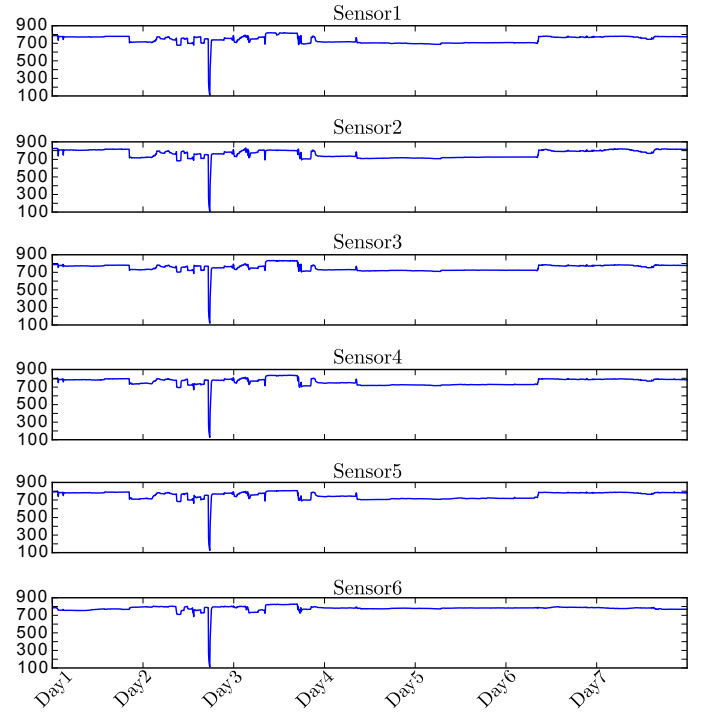


Fig. 11. Burner temperatures (degrees °C) over a period of 7 days

Figure 12 shows the times when the abnormalities in the data are flagged by either of the detectors.

#### D. Benefits afforded by use of the integrated detector

Results from each of the individual detectors can be compared with those from integrated detection scheme to show the benefits thus accrued.

1) *Differential Detector*: Consider the scenario depicted in Section III-B. However, in this instance assume that a failure occurs on day 15 with the result that the measurement characteristics of all sensors show a similar trend to that of sensor 6. Even though a failure has evidently occurred the measurements remain highly correlated and therefore the differential detector used alone would not readily detect the failure.

2) *Standard Detector*: Consider the scenario described in Section III-C. As can be seen from Figure 15, a changepoint is detected on day 2. However, this abrupt change is a result of what is considered as normal operation in this instance and now a precursor to an impending failure. Used independently therefore, the Standard detector would initiate a false alarm.

## IV. CONCLUSIONS

An integrated online 2-D changepoint detection algorithm for fault detection and identification is proposed for use with multiple correlated sensor datasets. Through use of experimental trials on industrial gas turbine units it is shown that

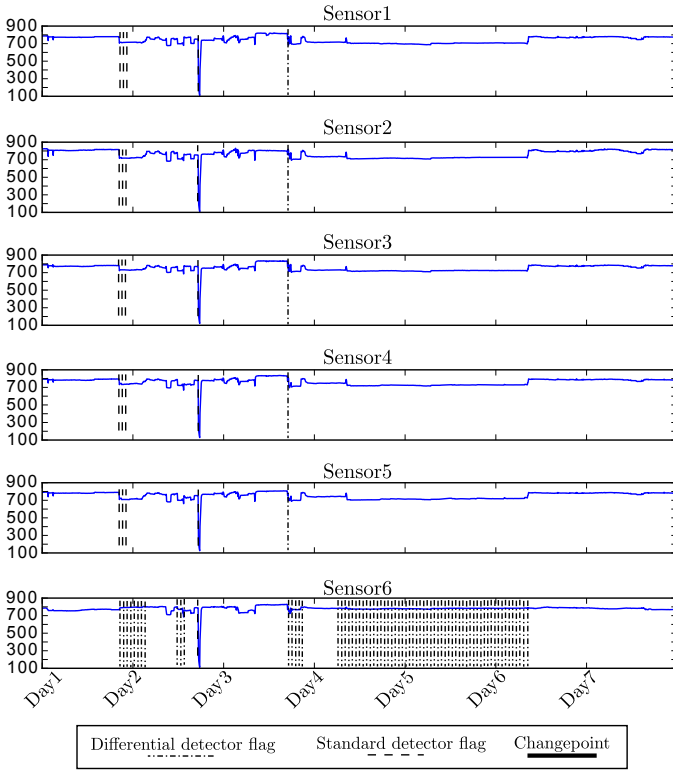


Fig. 12. Changepoints identified by the detectors.

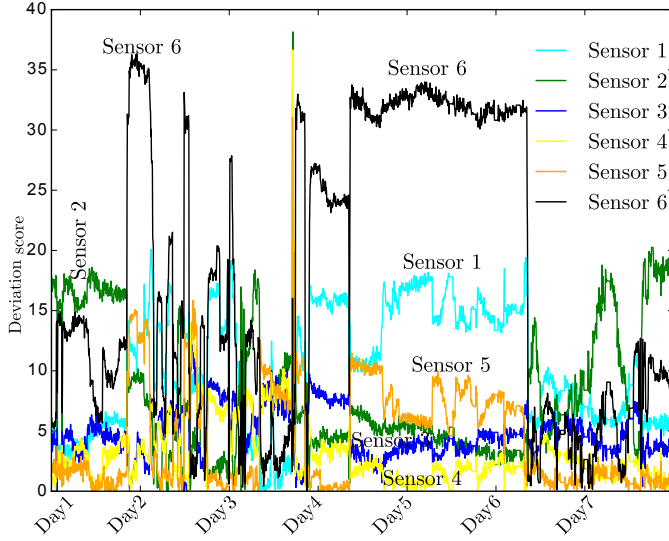


Fig. 13. Directional graph of the deviation score computed for each sensor.

the integrated detector both discriminates between changes in behaviour during normal operation and those due to the possible emergence of fault conditions –in the case given, an indication of emerging failure was apparent around 4 weeks in advance of ultimate unit shutdown. Moreover, the detector is also shown to reject behavioural characteristics that would typically generate false alarms, and thereby prevent unnecessary

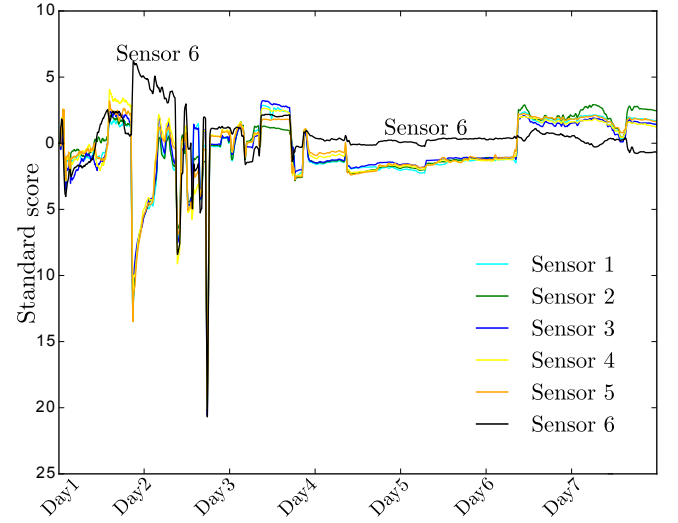


Fig. 14. Graph of the standard score computed for each sensor

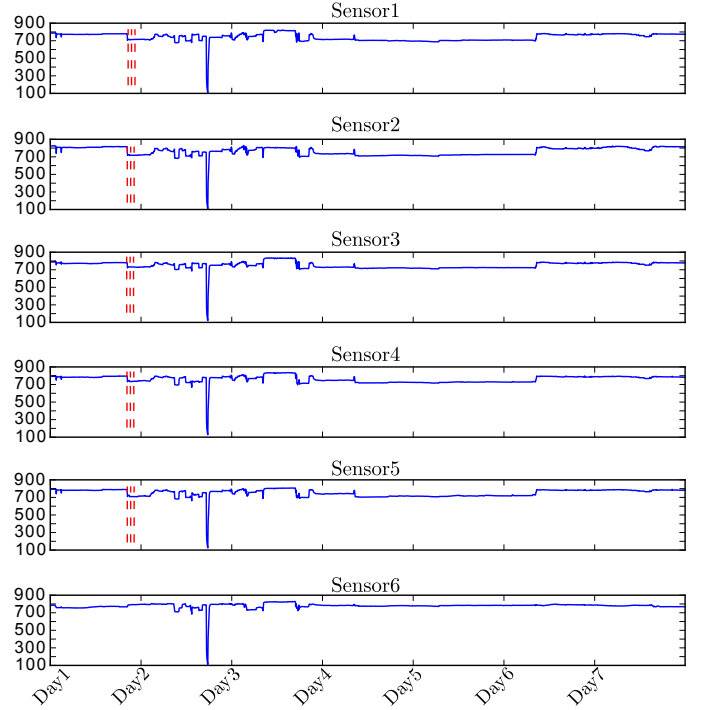


Fig. 15. Standard detector individual performance

unit stoppages and associated downtime. The demonstrated efficacy of the methodology now means that it is currently being commissioned as an “early warning” tool to monitor a global fleet of Siemens gas turbines.

#### ACKNOWLEDGEMENT

The authors would like to thank Siemens Industrial Turbomachinery, Lincoln, U.K., for providing access to real-time data to support the research outcomes.



## REFERENCES

- [1] S. Maleki, P. Rapisarda, and E. Rogers, *Failure identification for linear repetitive processes*, Mult. Syst. & Sig. Proc., Vol. 26, 2015, pp. 1037-1059.
- [2] M.G. Singh, *Fault Detection & Reliability: Knowledge Based & Other Approaches*, International Series on Systems and Control, Elsevier Science, 2013.
- [3] C.d. Costa, M. Kashiwagi, and M.H. Mathias, *Rotor failure detection of induction motors by wavelet transform and Fourier transform in non-stationary condition*, Case Studies in Mechanical Systems and Signal Processing, vol. 1, 2015, pp. 15-26.
- [4] Z. Zhang, and B.D. Rao, *Sparse Signal Recovery With Temporally Correlated Source Vectors Using Sparse Bayesian Learning*, IEEE J. of Select. Topics in Sig. Proc., vol 5, 2011, pp. 912-926.
- [5] M. A. F. Pimentel, D. A. Clifton, L. Clifton, and L. Tarassenko, *A review of novelty detection*, Signal Processing, vol. 99, 2014, pp. 215-249.
- [6] B. Sofman, B. Neuman, A. Stentz, and J. Bagnell, *Anytime online novelty and change detection for mobile robots*, J. Field Robot. vol. 28, 2011, pp. 589-618.
- [7] A. Tartakovsky, M. Basseville, and I.V. Nikiforov, *Sequential Analysis: Hypothesis Testing and Changepoint Detection*, Taylor & Francis, 2014.
- [8] F. Kiasi, J. Prakash, S. Patwardhan, and S. L. Shah, *A unified framework for fault detection and isolation of sensor and actuator biases in linear time invariant systems using marginalized likelihood ratio test with uniform priors*, Journal of Process Control, vol 23, no. 9, 2013, pp. 1350-1361.
- [9] L. Jian, F. Tsung, and C. Zou, *Directional changepoint detection for process control with multivariate categorical data*, Naval Research Logistics (NRL), vol 60, no. 2, 2013, pp.160-173.
- [10] M. Rakesh, G. P. Kalamangalam, and B. Aazhang, *Online Bayesian change point detection algorithms for segmentation of epileptic activity*, IEEE Conference on Signals, Systems and Computers, 2013, pp.1833-1837.
- [11] L. Vernon, S. Kerick, and K. A. Robbins *Detecting alpha spindle events in EEG time series using adaptive autoregressive models*, BMC neuroscience 14, no. 1, 2013.
- [12] L. Zhichao, X. Li, J. Xing, J. Lv, X. Jiang and D. Zhu, and S. Zhang et al., *Exploring functional brain dynamics via a bayesian connectivity change point model*, IEEE 11th International Symposium on Biomedical Imaging (ISBI), 2014, pp. 600-603.
- [13] S. Jeremy and N.R. Zhang, *Change-point model on nonhomogeneous Poisson processes with application in copy number profiling by next-generation DNA sequencing*, The Annals of Applied Statistics, vol. 6, no. 2, 2012, pp. 476-496.
- [14] H. Abeer, W. Ning, and A.K. Gupta, *An Information-Based Approach to the Change-Point Problem of the Noncentral Skew t Distribution with Applications to Stock Market Data*, Sequential Analysis, vol. 33, no. 4, 2014, pp. 458-474.
- [15] T. A. Kass-Hout, Z. Xu, P.I. McMurray, S. Park, D.L. Buckeridge, J. S. Brownstein, L. Finelli, and S.L. Groseclose, *Application of change point analysis to daily influenza-like illness emergency department visits*, Journal of the American Medical Informatics Association, vol. 19, no. 6, 2012, pp. 1075-1081.
- [16] B. P. Welford, *Note on a method for calculating corrected sums of squares and products*, Technometrics, vol. 4, 1962, pp. 419-420.
- [17] D. E. Knuth, *Art of Computer Programming*, vol. 2, 3rd ed., Addison-Wesley, 1997, pp. 232.
- [18] M. Ahsanullah, V.B. Nevzorov, and M. Shakil, *An Introduction to Order Statistics*, Atlantis Studies in Probability and Statistics, Atlantis Press, 2013, pp. 26.



**Sepehr Maleki** obtained the M.Sc. degree in Wireless Communications in 2011 and a Ph.D. in Electronics and Electrical Engineering in 2015 from the University of Southampton, UK. He is currently a research fellow at the Intelligent Systems research group of the Engineering School of the University of Lincoln. His research interests are in the areas of machine learning, statistical learning, anomaly detection, fault detection and isolation, and multi-dimensional systems.



**Chris Bingham** received the B.Eng. degree in electronic systems and control engineering, from Sheffield City Polytechnic, Sheffield, U.K., in 1989, the M.Sc. (Eng.) degree in control systems engineering from the University of Sheffield, Sheffield, U.K., in 1990, and the Ph.D. from Cranfield University, Swindon, U.K., in 1994, where his research focused on control systems to accommodate nonlinear dynamic effects in aerospace flight-surface actuators. From 1994 to 2010, he held academic positions at the University of Sheffield, UK. From 2010 he has held the position of Professor of Energy Conversion at the University of Lincoln, Lincoln, UK, where he is also the College of Science Director of Research.



**Yu Zhang** received Ph.D. degree from Department of Civil Engineering, University of Nottingham, Nottingham, U.K., in 2011. She is currently a Senior Lecturer in the School of Engineering, University of Lincoln, Lincoln, U.K. Her research interests include intelligent systems and machine fault diagnosis.