Energy Efficiency

Performance Degradation Identification for Compressed Air Systems Using a Hybrid Machine Learning and Knowledge Based Fault Detection and Diagnostic System
--Manuscript Draft--

Manuscript Number:	
Full Title:	Performance Degradation Identification for Compressed Air Systems Using a Hybrid Machine Learning and Knowledge Based Fault Detection and Diagnostic System
Article Type:	Original Research
Keywords:	Fault detection; Clustering; Compressed Air; Expert Systems; Mode Identification; Energy Efficiency
Corresponding Author:	Sean Martin Hayes University College Cork Cork, IRELAND
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	University College Cork
Corresponding Author's Secondary Institution:	
First Author:	Sean Martin Hayes
First Author Secondary Information:	
Order of Authors:	Sean Martin Hayes
	Ken Bruton, PhD
	D. T. J. O'Sullivan, PhD
Order of Authors Secondary Information:	
Funding Information:	MaREI	Mr Sean Martin Hayes

Abstract:	A rule based expert system for compressed air system fault detection and operational performance management is presented. The expert rule set takes a qualitative model based approach to fault detection, relying on fundamental engineering principles to detect when the compressed air system is not performing as expected. K-means clustering on the compressor power consumption is employed for mode identification and intelligent efficiency monitoring. The system is formally coded in MATLAB with a GUI for visualisation. The tool has been trialled on a compressed air system which is representative of installations in global industry. Successful determination of the system's operating mode was achieved. The rule set of 15 rules indicated that 9 faults were recurrent in the test case data set. Each of the symptoms which resulted in the rules highlighting a fault had two potential causes or diagnoses. Manual verification of which fault is actually present is currently required, with future refinement of the system planned allowing definitive diagnoses of faults detected.

Suggested Reviewers:	DIONYSIOS P XENOS Imperial College London d.xenos@imperial.ac.uk
	Lijun Wang North Carolina Agricultural and Technical State University lwang@ncat.edu
	Giorgos Kopanos Cranfield University g.kopanos@cranfield.ac.uk

Performance Degradation Identification for Compressed Air Systems Using a Hybrid Machine Learning and Knowledge Based Fault Detection and Diagnostic System

Seán Martin Hayes · **Ken Bruton** · **D.T.J. O'Sullivan**

Seán Martin Hayes · Ken Bruton · D.T.J. O'Sullivan
School of Engineering
University College Cork
Ireland
Tel.: +353-21-4902913
Fax: +353-21-4276648
E-mail: sean.m.hayes@umail.ucc.ie

Performance Degradation Identification for Compressed Air Systems Using a Hybrid Machine Learning and Knowledge Based Fault Detection and Diagnostic System

Blinded Manuscript

Abstract A rule based expert system for compressed air system fault detection and operational performance management is presented. The expert rule set takes a qualitative model based approach to fault detection, relying on fundamental engineering principles to detect when the compressed air system is not performing as expected. K-means clustering on the compressor power consumption is employed for mode identification and intelligent efficiency monitoring. The system is formally coded in MATLAB with a GUI for visualisation. The tool has been trialled on a compressed air system which is representative of installations in global industry. Successful determination of the system's operating mode was achieved. The rule set of 15 rules indicated that 9 faults were recurrent in the test case data set. Each of the symptoms which resulted in the rules highlighting a fault had two potential causes or diagnoses. Manual verification of which fault is actually present is currently required, with future refinement of the system planned allowing definitive diagnoses of faults detected.

1 Introduction

In 2012 industry consumed 2,542 Mtoe of energy globally, which represented over 28% of the 8,980 Mtoe of global final energy consumption (IEA (2012)). In an Irish context, industry consumed 2.26 Mtoe of energy in 2012, representing almost 22% of Irelands 10.3 Mtoe of final energy consumption. The business case for energy efficiency improvements in industry is clear, particularly in the high-tech sector

Blinded Manuscript

(Mills et al (2007)). An economic efficiency model for energy using equipment has previously been proposed which quantifies the cost-benefit of improved efficiency (Blum (2014)). Within the category of industrial energy, compressed air is recognised as an energy intensive utility (Wang and Brown (2013)), accounting for 10% of the 86.6 Mtoe of industrial electricity in the European Union (Saidur et al (2010), Eurostat (2014)). Energy costs typically account for 78% of the total life cycle cost of a compressed air system (Radgen (2006)). Compressed air is known colloquially in industry as the fourth utility (Fairest (2014)), due to the high electrical cost associated with generation. Compressed air systems are typically running at 19% overall system efficiency (Saidur et al (2010)), due to energy losses largely due to lost heat of generation and leakages. Oversized systems are also prevalent in industry, leading to energy wastage (Laurijssen et al (2013)). It is therefore desirable to manage the performance of air compressors in order to minimise their associated energy consumption, which is the goal of this paper.

Compressed air has been recognised as having significant energy saving potential in industry (Wang (2014)), not least through measures such as retrofitting of variable speed drives, inlet air temperature reduction, waste heat recovery, and pressure and leakage reduction (Wang (2008)). Savings in the region of 20-50% can often be expected by the adoption of an integrated management policy for compressed air SEAI (2007a). If the lower end of this saving opportunity scale were to be applied to the European Union figures above, it would represent 1.7 Mtoe of savings potential by effectively managing compressed air operations.

This paper is structured as follows. Section 2 defines the current landscape in terms of compressed air installations in industry and how their performance may be managed. Section 3 outlines the method used for operational management used in this work, namely the deployment of a knowledge

based fault detection and diagnostic (FDD) system. Section 4 explains the means by which the operational mode of the compressed air system may be determined using K-means clustering. Section 5 gives the means by which the hybrid system operated on an operational compressed air system at University College Cork. Finally Section 6 gives the conclusions as to the effectiveness of the methodology deployed, and outlines future work which is planned.

2 Determining the operational performance of an air compressor

A wide range of configurations and types of compressed air systems are installed in industry. In many cases there exist systems which are running sub-optimally, either due to unsuitability for the task at hand or running in a faulty condition SEAI (2007a). Given that compressed air represents such a dense form of energy transport, with over 90% of the energy used in generation being converted to heat (SEAI (2016)), it is beneficial in terms of long and short term overall energy efficiency goals to manage the performance of air compressors. Performance management is typically achieved through means such as those in Table 1. The effectiveness of measures such as energy audits has been documented (Fleiter et al (2012)), as well as the varying effectiveness of different maintenance management programs (Duffuaa and Raouf (2015)). The key disadvantages of existing methods are either that they are manual and periodic in nature, or that they require the intervention of a human expert in compressed air systems to be effective. In the case of maintenance contracts and periodic audits, there is also the potential for unnecessary work to be carried out, as both these measures are typically carried out on a timed basis. The intervention of a human expert can also lend itself to an inefficient method of performance measurement, as an expert may be particularly well versed with one type of system, but not another. The disparate range of compressed air systems can lead to an expert restricting themselves to one type of system, preventing possible lessons learned to be applied in other suitable cases. If a big data approach is correctly taken lessons learned in one facility may be effectively applied in others (ODonovan et al (2015b)).

In order to test the method developed during this research, a real world compressed air system was utilised, comprising two rotary tooth air compressors with a heated desiccant dryer, with the layout given in Figure 1. These machines are rotary tooth type machines, which are widely deployed across industry for applications with medium pressure and capacity requirements, as shown in Figure 2 (SEAI (2007b)). The various types of compressors typically used in industry are shown in Figure 3. The three most common types of compressor in industry are rotary, reciprocating and centrifugal machines. Reciprocating and rotary ma-

Table 1 Existing Compressed Air System Performance Management Methods

| Performance Management Method	Advantages	Disadvantages
Planned Maintenance Contracts	Security of asset reliability	Potential for unnecessary work
Periodic Audits	Likely to pick up on common opportunities for improvement	Dependent on skill level of auditor
Sequence Controllers	Can draw on manufacturer knowledge of system operation	Initial configuration may not be maintained due to system changes
BMS Monitoring	Desk-based site wide monitoring capability	Dependent on skill level of BMS reviewer. Unable to pick up on sensor errors
Computerised Maintenance Management Systems	Reduction in manpower required for decision-making	Requires substantial initial expert input
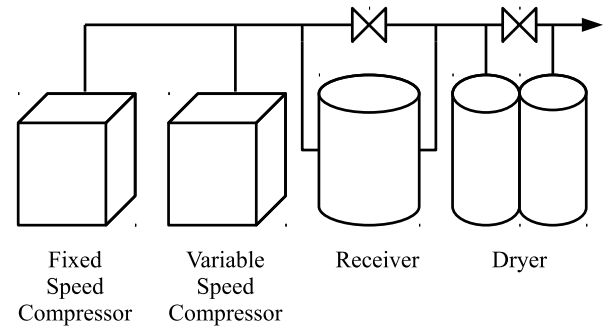
Fig. 1 Test Site Compressed Air System Layout

chines are both positive displacement type machines. They work through isolation of a quantity of air in a space which is then reduced in volume. Centrifugal machines are aerodynamic machines, which operate by imparting kinetic energy to air, which is then converted to pressure energy by stopping the moving air.

Research is being carried out to define the future of compressed air system performance management, with a representative sample comprising Dapeng Niu et al (2011), Ren et al (2012) and Kopanos et al (2015). In this paper the research considered is that of ongoing analysis of compressed air system data. This ongoing analysis could be designated as having any of the higher level goals outlined in Table 2.

This paper categorises industrial utility performance management methods into three high-level classifications, which are themselves subdivided into individual methods. These three categories are:
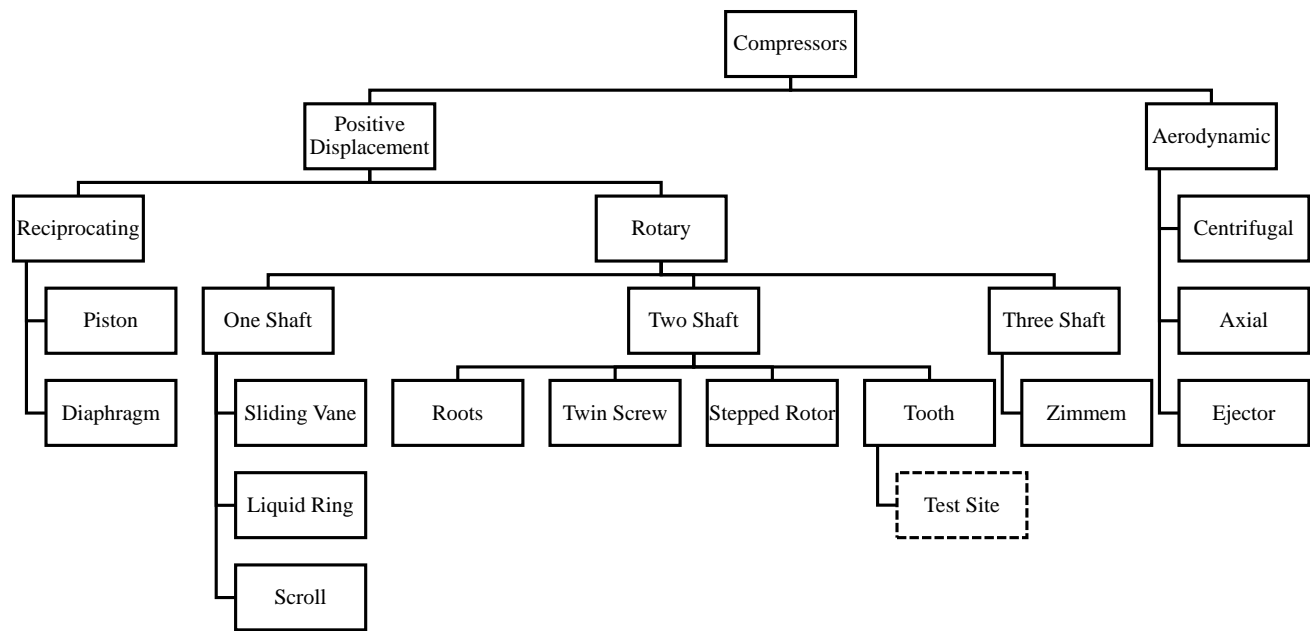
1. Quantitative model based methods
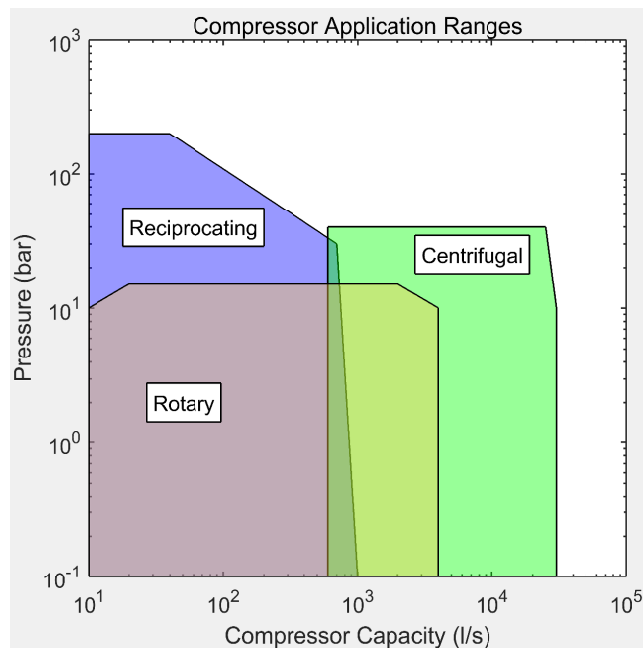
Fig. 3 Compressor Types

Fig. 2 Compressor Application Suitability

2. Qualitative model based methods
3. Process history based methods

These three categories are shown visually in Figure 4, which is adapted from previous works on system performance management and diagnostic approaches (Katipamula and Brambley (2005), Venkatasubramanian (2003), Venkatasubrama-nian et al (2003a), Venkatasubramanian et al (2003b), Gao et al (2015b), Gao et al (2015a), Bruton et al (2013)).

2.1 Quantitative model based methods

In the field of compressor performance management, one approach which may be used is that of the development of a quantitative model describing the compressor's operation, and analysing actual operation with respect to this modelled operation in order to achieve one of the goals outlined in Table 2. This comparison may lead to the generation of differences between measured and modelled variables, which are termed residuals. This concept of inconsistency between variables is known as redundancy.

Analytical or artificial redundancy is achieved through formalisation of the fundamental relationships between the states, inputs and outputs of a system. This inherent redundancy may take either a direct or a temporal approach.

A direct approach to analytical redundancy is to derive algebraic equations between different sensor measurements. If a sensor is available for a calculated value, the concept of redundancy may be used to generate a residual. If the residual exceeds a given threshold then a sensor fault may be present. In contrast, temporal redundancy is obtained by analysing the difference relationships between sensor outputs and actuator inputs. If an actuator input is intended to produce a difference between sensors, and this difference is
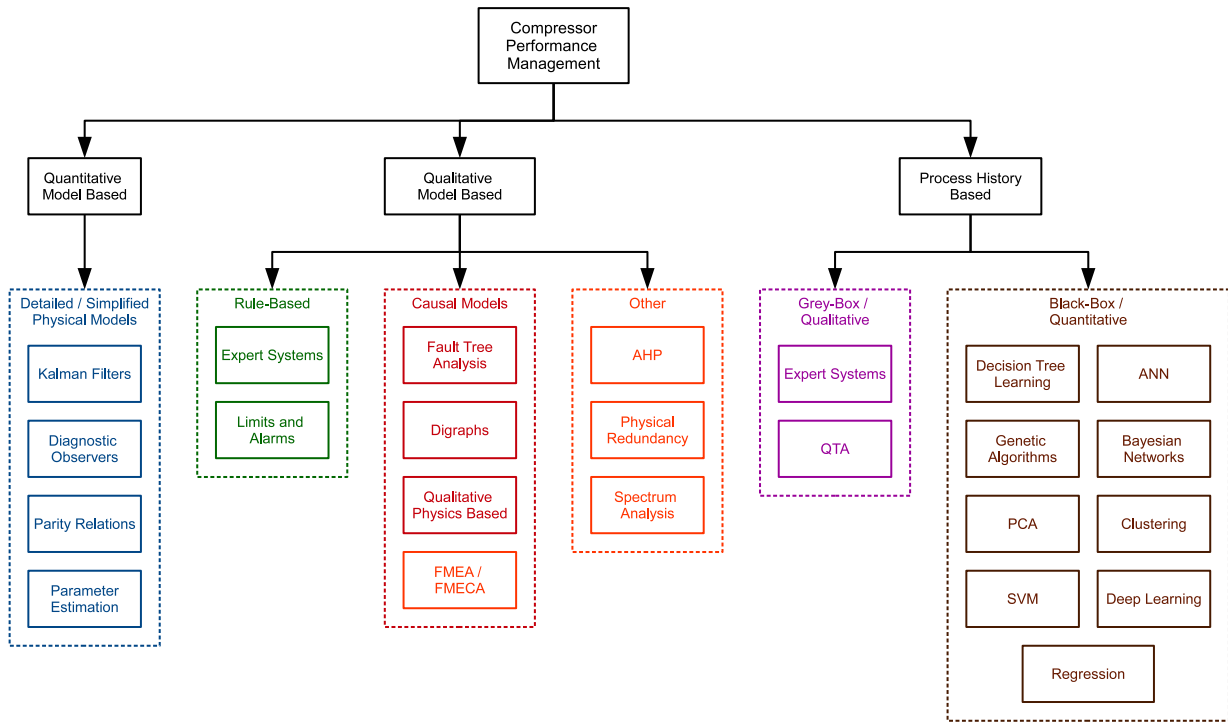
Fig. 4 Current research into performance management methods

not present, then either a sensor or actuator fault may be present. Various approaches to the generation and analysis of residuals are given in Figure 4. Quantitative model based methods have been used with success in the field of air compressors for surge control (Backi et al (2013)), state estimation (Nair et al (2011))and leakage detection (Krichel and Sawodny (2011)).

Numerical models of a system have also been used with success to analyse the potential benefits to be achieved by retrofitting measures (Murray et al (2012), Murray et al (2014)).

2.2 Qualitative model based methods

Qualitative model based methods may be distinguished from quantitative model based methods by their abstraction of the physical principles governing the operation of a system. Where quantitative methods seek precise numerical values for the parameters of a system, qualitative methods are generally satisfied with simplified models of a system.

To demonstrate this difference the example of an air compressor in operation is considered. If a quantitative model is used for analysis of this system, it can require inputs of all possible system and environmental variables (voltage, current, ambient air conditions) in order to make a calculation on what the compressed air flowrate should be. A

qualitative approach to this situation would be to hypothesise that with an increase in current drawn by the compressor, an increase in compressed air flowrate should also be observed. While the quantitative approach may highlight a slight decrease in performance of the machine if the expected flowrate is not met, the qualitative approach will immediately highlight a serious issue with the compressor if an increase in power does not correspond to an increase in flowrate. The time required to develop quantitative solutions is typically greater than that for qualitative solutions. This suggests that for rapid deployment of ad-hoc compressor performance management solutions, a qualitative approach may be more suitable. The ability to be deployed with minimal additional sensing also gives a high probability of required data availability, which is an additional benefit of this type of system. Of particular note in this set of performance management methods is rule-based expert systems, which have been selected as a method to trial in this paper. Rule-based expert systems have been applied with success in the area of HVAC (Bruton et al (2014a)). For compressed air systems, qualitative model based methods have been applied for general fault diagnosis (Liu (2001)), valve failure diagnosis (Ruilin Lin et al (2010)), reliability assessments (Ren et al (2012)), and safety evaluation (Zhu et al (2013)).

Table 2 Goals of Performance Management

Goal	Description	Example Work
Fault Detection and Diagnosis	Monitor system parameters to determine when system is in fault condition and the potential reasons for the identified fault	Using vibration, pressure and current signals to diagnose valve faults for a reciprocating compressor (Tran et al (2014))
Prognostics	Monitoring system parameters to determine when a component of a system will no longer perform its intended function (Vachtsevanos et al (2006))	Determining the remaining useful life of a gaseous circuit breaker based on gas pressure and ambient temperature (Catterson and Costello (2013))
Analytics	Monitoring system parameters to discover meaningful patterns which may advise on potential improvements to system operation	Determining abnormal appliance power consumption based on analysis of individual appliances acoustic noise (Pathak et al (2015))
Automated Commissioning	Achieving, verifying and documenting that the performance of a system satisfies the current user requirement	Automatically carrying out the normal testing procedure for an air compressor by replicating the tasks normally carried out during commissioning (Mazid and Martin (2008))
Optimisation	Improving system operation or design as measured against some defined criteria	Development of a tool which delivers an optimal design for a compressed air system based on energy and life cycle costing (Friden et al (2012))
Control	Managing the operation of a system in order that operating conditions remain in line with design states and undesirable states are avoided	Development of a control algorithm for fixed speed compressors that provides the pressure control capabilities of a variable speed system while limiting energy consumption (Facchinetti et al (2014))

2.3 Process history based methods

Process history based methods may be subdivided into qualitative and quantitative categories, or grey and black-box methods respectively. Grey box methods employ a degree of prior knowledge concerning the engineering operation of a system, whereas black box methods assume no such knowledge and focus purely on patterns in data.

Process history based methods have been applied to the area of compressed air system performance management for load forecasting (Liu et al (2013)), valve failure diagnosis (Wang et al (2010)), general fault detection (Petković et al (2012)), and overall system optimisation (Kopanos et al (2015)).

2.3.1 Qualitative process history based methods - expert systems

Expert systems as applied to the process history methodology as opposed to the qualitative model methodology are concerned with extracting useful features from historical data, and using qualitative knowledge of the relevant system to explain and make use of these features. One example of this methodology was employed during the creation of rules in section 3. While some rules were developed using hypotheses about the compressed air system under analysis, analysis of data gathered showed that under normal conditions, the oil pressure of the air compressor rose with increasing outlet pressure, and fell with decreasing outlet pressure. While this could have been hypothesised beforehand, it was not until the data was analysed that this feature was noticed. This lead to the development of a rule to highlight when oil pressure did not track compressor outlet pressure.

The limitations associated with process history based expert systems are similar to those associated with qualitative model based expert systems. In both cases a reliance is placed on the ability of the human expert to accurately determine rules and possible fault diagnoses. In the case of the oil pressure rule described above, it is hypothesised that a potential cause of fault should the oil pressure not rise when expected be that the oil pump has failed. However it is acknowledged that this fault may be equally symptomatic of a blocked valve or a sensor failure. As the fault detection system is developed further hypotheses may be drawn which will overcome multiple symptomatic faults such as this.

2.3.2 Quantitative process history based methods

Quantitative methods for system fault detection seek to extract useful features from historical data in a black-box fashion. The terminology of black-box is used to denote methods which are not influenced by fundamental engineering relationships between variables, but rely on statistical and machine learning methods to extract useful features. This is

distinguished from qualitative or gray-box methods which employ a modicum of understanding of the physical processes governing a systems operation.

Quantitative methods attempt to classify or group data into useful classes through pattern recognition. These methods are generally stochastic in nature, i.e. they do not assume that the future state of the system is necessarily influenced by past and present states. This gives such methods a probabilistic aspect, or a confidence rating in how accurately they are able to predict and classify system variables.

3 Rule base development and testing

3.1 Nomenclature for rule parameters

Before discussing the development of the expert rule set concerning air compressor performance assessment, it is useful to define the naming convention used. Table 3 gives the nomenclature used in rule development. The nomenclature was developed as part of this research, and could be deployed on similar two-stage compressor systems. The nomenclature was developed as the *de facto* standard for naming conventions in cases such as this (Project Haystack (2016)) was found to be limited in its application to air compressors. This nomenclature could be incorporated into Project Haystack should widespread deployment of the system be achieved.

3.2 Rule formulation methodology

One of the approaches for performance management of air compressors outlined in subsection 2.2 is that of Expert Systems. This is a qualitative method which can use rule sets to encode expert knowledge about a system for highlighting when a fault in operation is present. This method has been used with success in the case of HVAC (Bruton et al (2014b), House et al (2001)) and in the area of air compressors (Jiang et al (2013)). Rules for system analysis are typically derived by taking a fundamental engineering overview of the system, and hypothesising potential rules to determine particular faults.

Figure 5 shows a high level overview of the test site air compressor. The figure highlights the expected temperature changes as the compressed air is operated on by each fundamental component. These expected operations form the basis for an initial set of rules to highlight when the machine is not operating as expected. To illustrate this point the first element of compression is now analysed in detail.

Element 1 of the air compressor compresses air from atmospheric pressure (101 325 Pa) to 250 000 Pa. When the compressor is at minimum loading, it is designed to produce $41 \, \mathrm{L \, s^{-1}}$ of free air (i.e. it will draw in this quantity of air

Table 3 Nomenclature

| Reference	Item	Value
	Sensors	
T1	Plant Room Temperature	-
T2	Element 1 Outlet Temperature	-
T3	Element 2 Inlet Temperature	-
T4	Element 2 Outlet Temperature	-
T5	Final Delivery Temperature	-
P1	Compressed Air Pressure in Intercooler	-
P2	Compressed Air Final Delivery Pressure	-
P3	Compressed Air Receiver Pressure	-
P4	Oil Pressure	-
P5	Ambient Pressure	-
N1	Motor starts per 5 minutes	-
K1	Compressor power	
	Components	
C1	Element 1	-
C2	Intercooler	-
C3	Element 2	-
C4	After Cooler	-
C5	Motor	-
C6	Oil Pump	-
C7	Load/Unload Valve	-
	Expected Levels	
T6	Plant Room Temperature	25 °C
T7	Element 1 Outlet Temperature	140 °C
T8	Element 2 Inlet Temperature	22 °C
T9	Element 2 Outlet Temperature	100 °C
T10	Final Delivery Temperature	21 °C
P6	Compressed Air Pressure in Intercooler	2.5 bar
P7	Compressed Air Final Delivery Pressure	7.25 bar
P8	Compressed Air Receiver Pressure	7.25 bar
N2	Motor starts per 5 minutes	1
K2	Compressor power idle	1.5 kW
K3	Compressor maximum power unloaded	8.4 kW
	Warning Levels (High)	
T11	Plant Room Temperature	35 °C
T12	Element 1 Outlet Temperature	155 °C
T13	Element 2 Inlet Temperature	24 °C
T14	Element 2 Outlet Temperature	120 °C
T15	Final Delivery Temperature	21 °C
P7	Compressed Air Pressure in Intercooler	2.6 bar
P8	Compressed Air Final Delivery Pressure	7.6 bar
P9	Compressed Air Receiver Pressure	7.6 bar
N3	Motor starts per 5 minutes	2
	Error Thresholds	
E1	Plant Room Temperature	5 °C
E2	Element 1 Outlet Temperature	5 °C
E3	Element 2 Inlet Temperature	5 °C
E4	Element 2 Outlet Temperature	5 °C
E5	Final Delivery Temperature	5 °C
E6	Compressed Air Pressure in Intercooler	0.1 bar
E7	Compressed Air Final Delivery Pressure	0.1 bar
E8	Compressed Air Receiver Pressure	0.1 bar
E9	Oil Pressure	0.1 bar
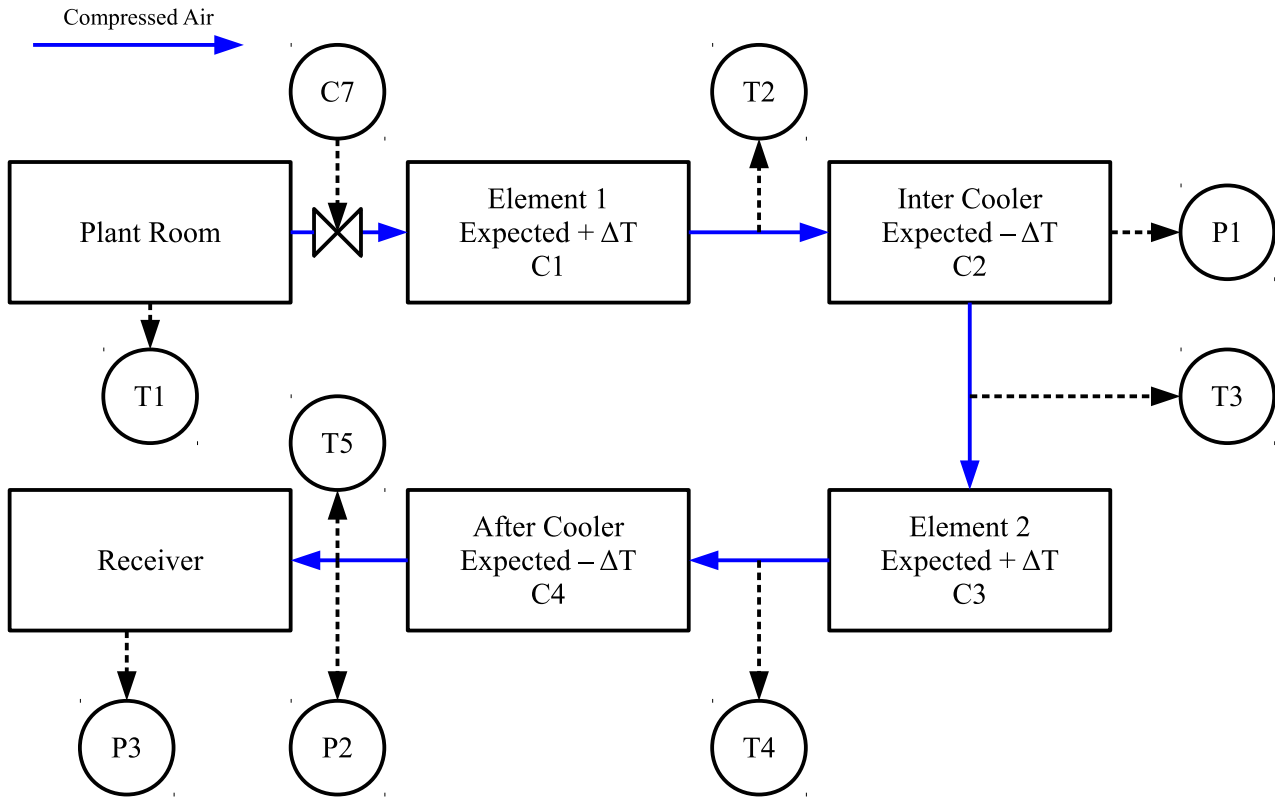
Fig. 5 Two stage compressor air flow

at atmospheric pressure from the plant room). It is useful to analyse the thermodynamics of compressing air at this point. Figure 6 shows the relevant thermodynamic diagrams for the compression of this volume of air isothermally, adiabatically and polytropically. The air is compressed in a space which is surrounded by ambient air, which is not an ideal heat reservoir. Therefore the isothermal compression case is not valid. The air is not however thermally isolated from its surroundings, therefore the adiabatic case is also not valid. Actual air compression typically follows the polytropic process model, with the polytropic exponent of Equation 1 ranging between 1 and 1.4.

$$P_1 V_1^n = P_2 V_2^n \tag{1}$$

where P = Absolute pressure (Pa)

V = Absolute volume (m^3)

n = Polytropic exponent

Given that the compression process taking place is not isothermal, it is logical to expect that the temperature of the compressed air will increase through the first element of compression, in accordance with Figure 6. If a decrease in temperature is observed, then either a failure in the first

stage of compression or in a temperature sensor can be determined. This forms the basis for the first rule of the rule set created for compressor performance analysis, which is summarised in Table 4.

Rule 1. $T1 - T2 > 0$ (2)

In a similar manner to Rule 1, Rules 2, 3 and 4 fire when the expected temperature of the compressed air does not change as expected across the intercooler, second element of compression and aftercooler respectively.

Rule 2. $T3 - T2 > 0$ (3)

Rule 3. $T3 - T4 > 0$ (4)

Rule 4. $T5 - T4 > 0$ (5)

Rules 5, 6, 7, 8 and 9 are threshold rules, in that they fire when defined thresholds for T1, T2, T3, T4 and T5 respectively are exceeded. These threshold values have been

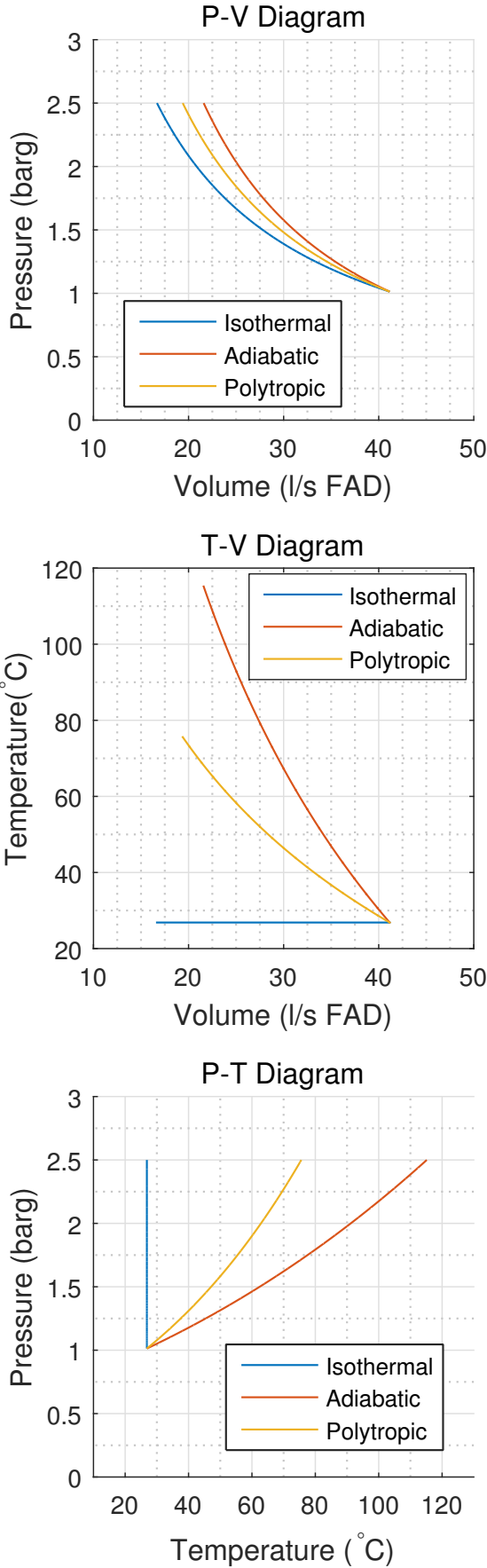## P-V Diagram

T-V Diagram

P-T Diagram

Fig. 6 Thermodynamics of Air Compression

heuristically developed through a combination of historical data analysis and experience of the operational parameters of the system.

Rule 5. $(T1 - E1) > T11$ (6)

Rule 6. $(T2 - E2) > T12$ (7)

Rule 7. $(T3 - E3) > T13$ (8)

Rule 8. $(T4 - E4) > T14$ (9)

Rule 9. $(T5 - E5) > T15$ (10)

The first element of compression in a two-stage compressor is designed to compress air to an intermediate pressure, before intercooling and compression to final delivery pressure by the second element of compression. If this intermediate pressure is not acheived, it is indicative of either a pressure sensor fault, or a fault in element 1 operation. If the first element of compression is faulty in operation, this will lead to excessive strain on the second element of compression to achieve final delivery pressure. Rule 10 highlights when this intermediate pressure is not met.

Rule 10. $(P1 + E6) < P6$ (11)

For a single compressor in isolation, if the final delivery pressure continues to rise when the compressor is running in unloaded mode, it is indicative of either a fault in the relevant pressure sensor, or that the load/unload valve of the compressor has failed. Rule 11 fires when this pressure rise in unloaded mode takes place.

Rule 11. $(P2(t) - (P2(t-1)) > 0$ (12)

If the motor driving the compressor is switching on and off excessively, it will lead to premature wear and tear on the compressor mechanical drive. Rule 12 fires when a threshold for the number of motor starts has been exceeded.

Rule 12. $\sum_{t=0}^{300s} N1 > N2$ (13)

Figure 7 shows the mechanical layout of a two stage air compressor. In an air compressor lubricated by oil, the oil should circulate when the compressor is loaded. If the oil
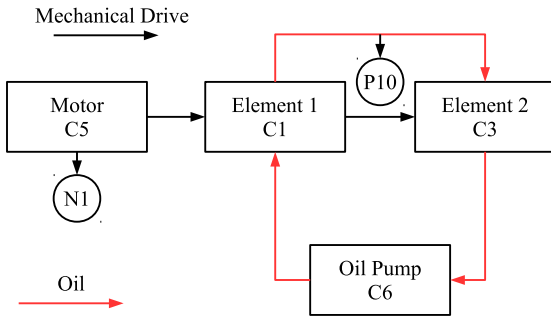
Fig. 7 Mechanical drive two stage compressor

pressure does not rise with the compressor loaded, it is indicative of a blockage or leakage in the oil circulation system, a failure of the oil pump, or a pressure sensor fault.

Rule 13. $(P4(t) + E9) < P4(t-1)$ (14)

The compression of air in both the first and second element of a two-stage air compressor follows defined thermodynamic principles. The temperature rise of compressed air for each of the three types of compression is given in Figure 6. The theoretical maximum temperature rise of air under compression is given by the adiabatic case. Therefore, while it is known that the air will compress polytropically, it can be assumed that if the outlet temperature of either the first or second element of compression is greater than that dictated by the adiabatic case (with the polytropic exponent of Equation 1 set to 1.4), then a fault is present in the system. This fault may either be in one of the relevant temperature sensors, or additional heat may be being supplied to the air under compression. This principle dictates when Rules 14 and 15 are fired.

Rule 14. $T2 > T1 * \left(\dfrac{P1}{P5} \right)^{\frac{\gamma-1}{\gamma}}$ (15)

Rule 15. $T4 > T3 * \left(\dfrac{P2}{P1} \right)^{\frac{\gamma-1}{\gamma}}$ (16)

where γ = heat capacity ratio of air (1.4)

Rules 1-15 are summarised in Table 4, along with the potential faults which would explain each rule being fired. The potential impact on system performance of each rule being fired is given in Table 5.

Table 5 Impact of rules on compressor performance

	Impact		
Rule	CA User Requirement	Energy	Maintenance / Equipment Life
1	X		X
2		X	X
3	X		X
4		X	
5		X	
6		X	X
7		X	X
8		X	X
9		X	X
10	X		
11		X	X
12		X	X
13			X
14		X	X
15		X	X

4 Operational mode identification

It was noted during rule set development that it is imperative to know the mode of operation of an air compressor when applying rules. This follows on from lessons learned in similar work relating to HVAC (Bruton et al (2014b)). With an air handling unit, the different modes of operation that are useful to determine which rules to apply are Heating, Cooling with Outdoor Air, Mechanical Cooling with 100% Air and Mechanical Cooling with Minimum Outdoor Air, with these modes further categorised into Occupied and Unoccupied Modes (House et al (2001)). Knowing the mode of operation of equipment ensures that rules are only applied where pertinent, reducing false positive occurences.

The modes of operation designated for a variable speed air compressor are given in Table 6. To determine which mode of operation the compressor is actually in, the power drawn by the machine was analysed. Prior information on the expected power drawn for each mode was not available from manufacturer documentation, and so a clustering methodology combined with some prior knowledge was employed.

The Compressed Air and Gas Institute (CAGI) specifies a standard data sheet for air compressors which major manufacturers adhere to (CAGI (2015)). The input power to the compressor as given on this sheet where available is also given in Table 6. Modes 1 and 3 were recognised as being useful modes to have knowledge of, despite not having prior knowledge of the associated input power from the CAGI data sheet. From visual observation of a power meter installed at the test site, it was noted that Mode 1 had an approximate power requirement of 1 kW, and Mode 3 an approximate power requirement of 20 kW

Table 4 Compressor performance assessment rule set

	Rules											Potential Faults						
ID	Description	T1	T2	T3	T4	T5	P1	P2	N1	P10	C1	C2	C3	C4	C5	C6	C7
1	Element 1 Temperature Decrease	X	X								X						
2	Intercooler Temperature Increase		X	X								X					
3	Element 2 Temperature Decrease			X	X								X				
4	Aftercooler Temperature Increase				X	X								X			
5	Plant Room Temperature High	X															
6	Element 1 Outlet Temperature High		X								X						
7	Element 2 Inlet Temperature High			X								X					
8	Element 2 Outlet Temperature High				X		X						X				
9	Final Outlet Temperature High					X								X			
10	Expected Intercooler Pressure not Achieved						X				X						
11	Increase in Outlet Pressure when Unloaded							X									X
12	High Stop-Start Frequency of Motor Observed								X						X		
13	Failure of Oil Pressure to Rise under Loading									X		X				X	
14	Theoretical Element 1 Temperature Rise Exceeded		X								X						
15	Theoretical Element 2 Temperature Rise Exceeded					X							X				

Table 6 VSD compressor operation modes

Mode	Description	CAGI Input Power (kW)
1	Idle	-
2	Unloaded	8.9
3	Minimally Loaded	-
4	VSD 0-20%	27.2
5	VSD 20-40%	31.2
6	VSD 40-60%	38.3
7	VSD 60-80%	42.6
8	VSD 80-100%	52

To monitor the power drawn by the compressor for mode identification, a clustering methodology was developed. This unsupervised learning method allows the compressor power meter data to be grouped into distinct clusters as per Table 6. An additional benefit to be gained from this approach is that a drift upward in the geometric mean or centroid of any cluster could diagnose a decrease in efficiency of the machine at that operating point, as more power is required to achieve the same output.

K-means clustering was used to group the electrical meter data into clusters. This technique groups unlabelled data into clusters where there is greater intra-group similarity than inter-group similarity. The general methodology of K-means clustering is shown in Figure 8. K-means clustering consists of two fundamental iterative steps preceded by one initial step. The number of clusters is set initially and denoted k. Centroids for the k clusters are then randomly initialised. All data points are then indexed according to the cluster centroid closest to them, as shown in Equation 17,

$$c^{(i)} := \min_{k} \|x^{(i)} - \mu_k\|^2 \quad \text{for} \quad i = (1, 2, ..., m) \tag{17}$$

where C is a vector of elements $c^{(i)}$ the same length as the data set X with m points $x^{(i)}$. The vector C gives the index of the cluster centroid closest to each point in X.
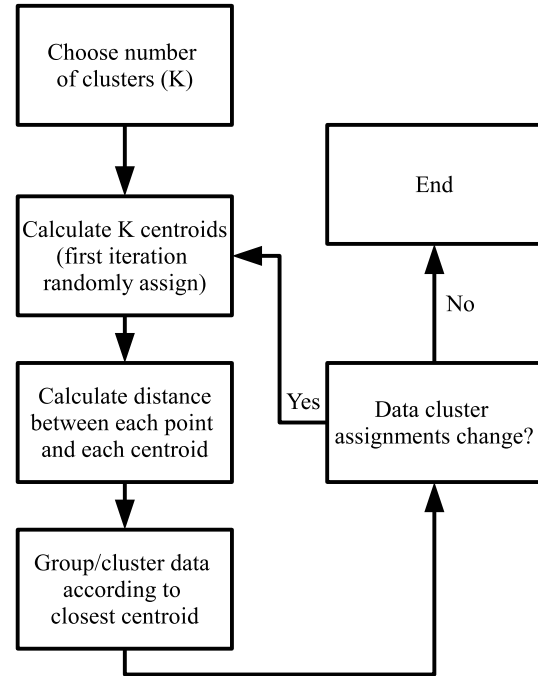
Fig. 8 K-means clustering methodology

Following this step there are k clusters in the dataset X, each of which is a subset of the overall dataset and is denoted by S_j. The centroids of each cluster are recomputed as the mean of all points assigned to each cluster, as shown in Equation 18,

$$\mu_j := \frac{1}{|S_j|} \sum_{x_j \in S_j} x_j \quad \text{for} \quad j = (1, 2, ..., k) \tag{18}$$

where x_j is a point in the dataset X which is assigned to the jth cluster. Equation 17 and Equation 18 are then iterated until the assignments of data points to different clusters no longer change.
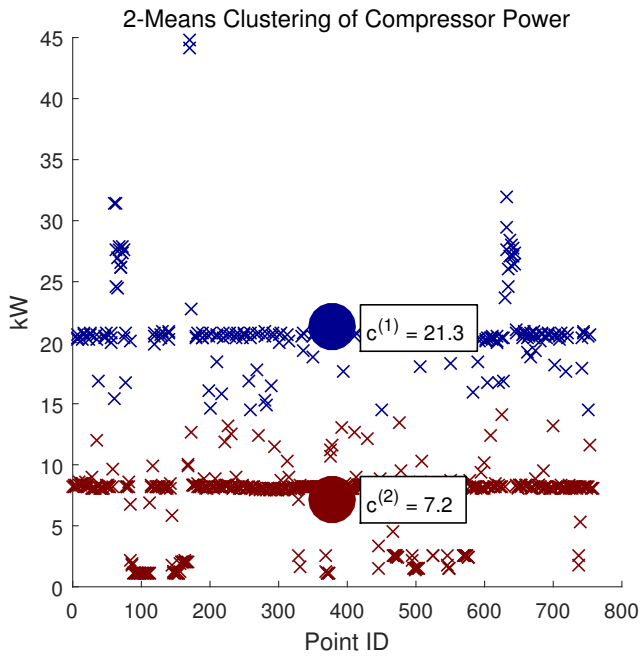
Fig. 9 2-Means Clustering Results

Fig. 10 3-Means Clustering Results

An initial attempt was made to cluster the power consumption data into two groups, i.e. Idle/Unloaded and Loaded or Modes 1-2 and 3-8 from Table 6 respectively. k was therefore initially set to 2. The results of this exercise are shown in Figure 9. When comparing the cluster centroids to the values in Table 6 and visual observation of the compressor power under operation, it can be determined that one of the two clusters formed is of use in further analysis, i.e. $c^{(1)} = 21.3$ kW. This is a reasonable representation of the power drawn by the compressor at minimal loading. However $c^{(2)} = 7.2$ kW is not as useful a value, as it is clear from Figure 9 that the cluster centroid is significantly below the clear cluster of compressor power operating in unloaded mode.

In order to try to gain more useful information about compressor power at different operating modes, k was increased to 3. The intention was that the clear distortion of $c^{(2)}$ by the grouping of compressor power in idle mode at approximately 1 kW be removed by forming a new cluster for idle mode or Mode 1 as per Table 6. The results of this exercise are shown in Figure 10. Mode identification is accurate with this approach for Modes 1, 2 and 3 as per Table 6, as the cluster centroids are consistent with both known and observed values for the different modes.

The ideal level of accuracy for mode identification would be to determine cluster centroids for each of Modes 1-8. However the test site air compressors are significantly oversized compared to the compressed air demand. It was known that at no point in the training set did the compressor operate in Mode 8, and very rarely operated in Modes 4-7. Therefore a 7-means clustering approach was implemented. The
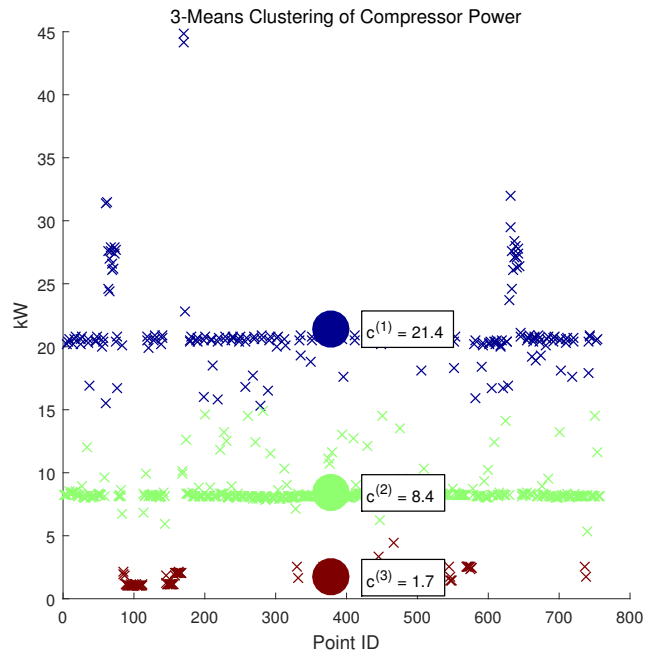
results of this analysis are shown in Figure 11. When comparing the results with Table 6, it is clear that there are two superfluous clusters in the region between Modes 2 and 3. That the analysis was unable to identify cluster centroids for Modes 5 and 6 is attributable to two factors. Firstly, there is a very low incidence of data points in these operating modes in the data set used. Secondly, the compressor takes a number of seconds to ramp up or down between Modes 2 and 3. The relatively high granularity of the data set (recording every 10 s) meant that a reasonably significant number of data points are recorded between these modes, and so the superfluous clusters have been identified.

While the 7-means clustering exercise identified superfluous clusters, it demonstrated that for the test case data set clustering was able to effectively identify five modes of operation. Therefore 5-means clustering was implemented on the data set to group the data into operational modes 1-4 and 7. The results of this implementation are shown in Figure 12. As expected this approach was able to correctly group the data into the required operational modes as per Table 6.

The clustering method for mode identification was used to apply rules only when pertinent. The applicability of rules to different modes of operation is given in Table 7.

4.1 Mode application to efficiency monitoring

A useful application of the 5-means clustering method utilised would be to detect when the efficiency of the air compressor deteriorates. This could be achieved by continually updating the cluster centroids with new data and monitoring
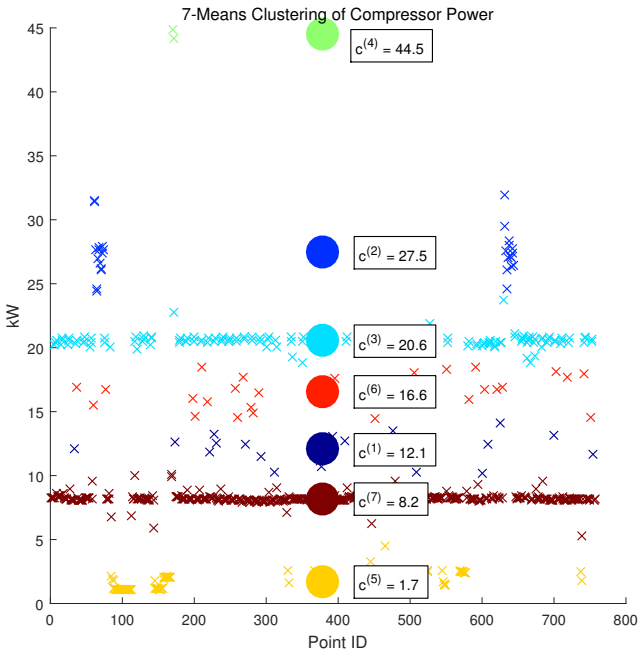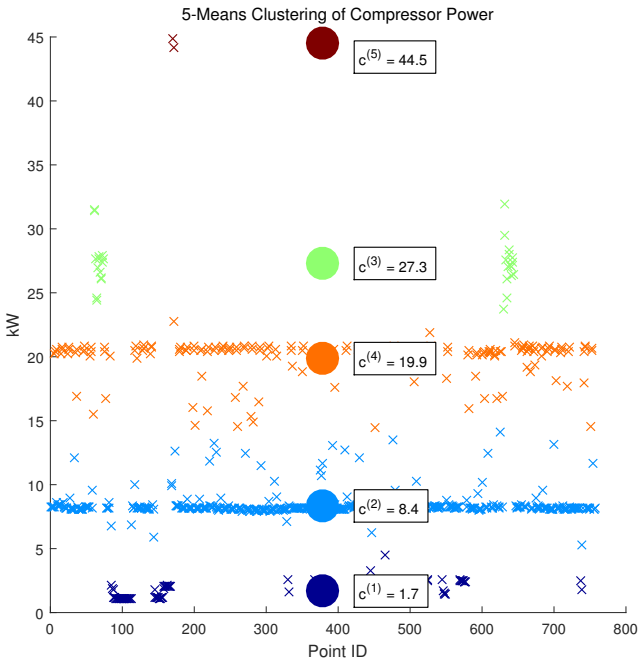
Fig. 11 7-Means Clustering Results

Fig. 12 5-Means Clustering Results

if there is an increase in power consumption for any given mode. The results of this approach are shown in Figure 13. It was observed visually during the recording period that the compressor did not cover the operational range from Modes 1-7 until approximately the 175th recording (approximately 0.5 h), which accounts for the clear initial training period with instability in cluster centroids.

There is a slight increase in power consumption for Mode 1 which may be attributable to an increase over time in the

Table 7 Rule applicability of modes

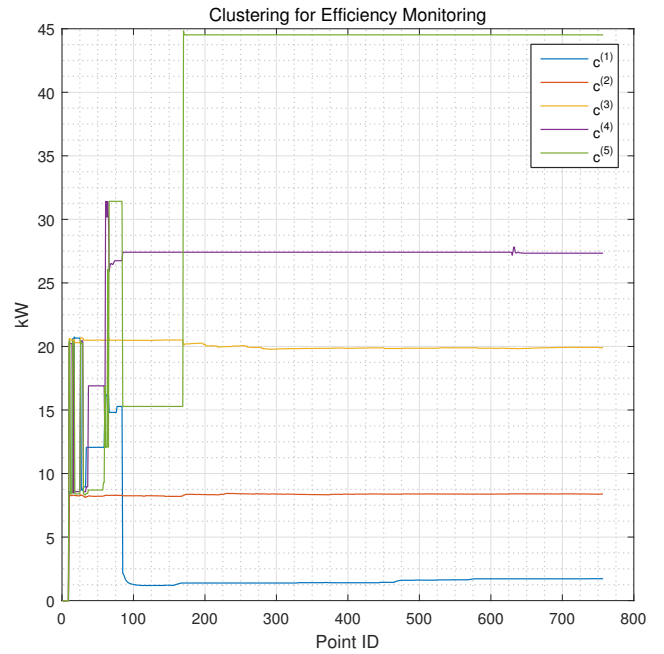| Rule	Mode		
	1	2	3:8
1			X
2			X
3			X
4			X
5	X	X	X
6			X
7			X
8			X
9			X
10			X
11		X	
12			X
13			X
14			X
15			X

Fig. 13 Clustering for efficiency monitoring

cooling of the power electronics associated with the variable speed drive controller. Future work will monitor these cluster centroids over a longer time span, allowing for concrete determination of when the power usage of the compressor is excessive compared to historical data. The effectiveness of tracking performance metrics such as this has been demonstrated as being powerful for improving a building's operation (OSullivan et al (2004)).

5 Fault detection implementation

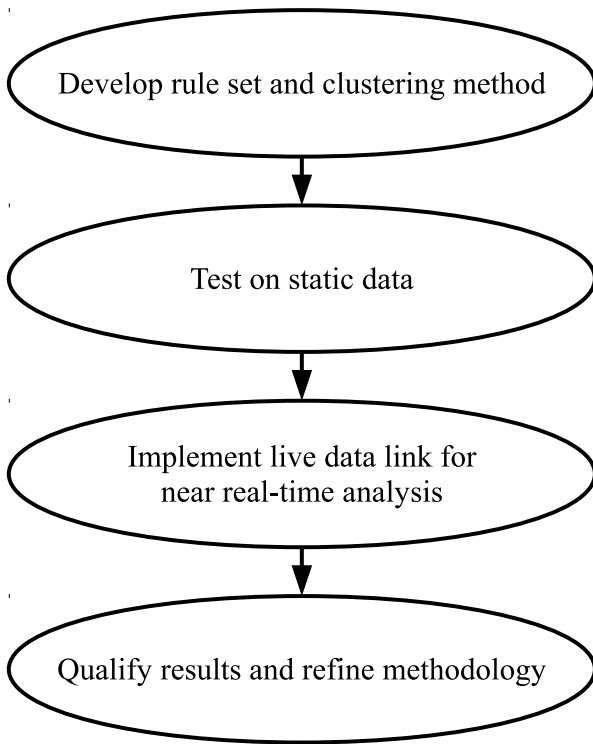In practice several key steps are taken to implement an operational performance management methodology. These steps

Fig. 14 Operational performance management implementation steps

Fig. 15 Data Flow Architecture

are outlined in Figure 14. Currently the initial two steps of the methodology have been implemented. Actual operational data is drawn manually and used as a static test for the system. Future work is planned to automate the data extraction process, and to refine the methods used for improved fault detection accuracy.

5.1 Data extraction

In order to apply the compressor performance rule set it was required to extract the operational data from the test site air compressors. A Modbus network was created to achieve this goal. The Modbus network daisy-chains the two air compressors on site with the compressed air dryer. A proprietary gateway exposes the Modbus registers of the 50 key performance parameters required for analysis. This gateway is supervised by a Tridium JACE box which acts as a Modbus master. The JACE box is connected to the internet using a 4G mobile modem, which allows for remote access and download of operational data. The overall flow of data extraction from the compressed air network to analysis by MATLAB is shown in Figure 15. A screenshot of the web interface of the JACE box is shown in Figure 16.
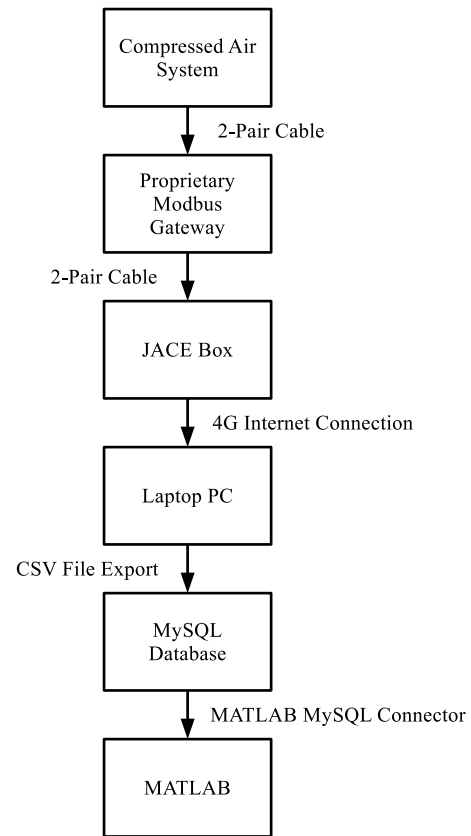
5.2 Rule set and clustering implementation

The data which was extracted from the compressed air monitoring system was stored in a MySQL database, to allow for analysis using MATLAB. Clustering was used to determine the operational mode of the air compressor at any given point. This knowledge was then used to apply pertinent rules to the data as per Table 7. To allow for visualisation of the knowledge gained from the analysis a GUI was created. A screenshot of this GUI is shown in Figure 17, with testing of rule 14 shown for demonstration purposes. The top section of the GUI displays operational data relevant to the rule being checked. The lower section shows the status of the rule being checked. A value of 1 denotes that the rule has fired, and 0 that there is no issue. Each rule was tested against the test case data set using this GUI. Of the 15 rules, nine (Rules 7-15) were found to recurrently highlight a fault in operation for the test data set. Each of these 9 fault symptoms indicates one of a possible two root causes. Manual verification of which is the actual cause of fault is therefore required at this stage. Future refinement of the rule set is planned to definitively diagnose the root cause for these multi-symptomatic faults. Rule 14 is now analysed in detail.
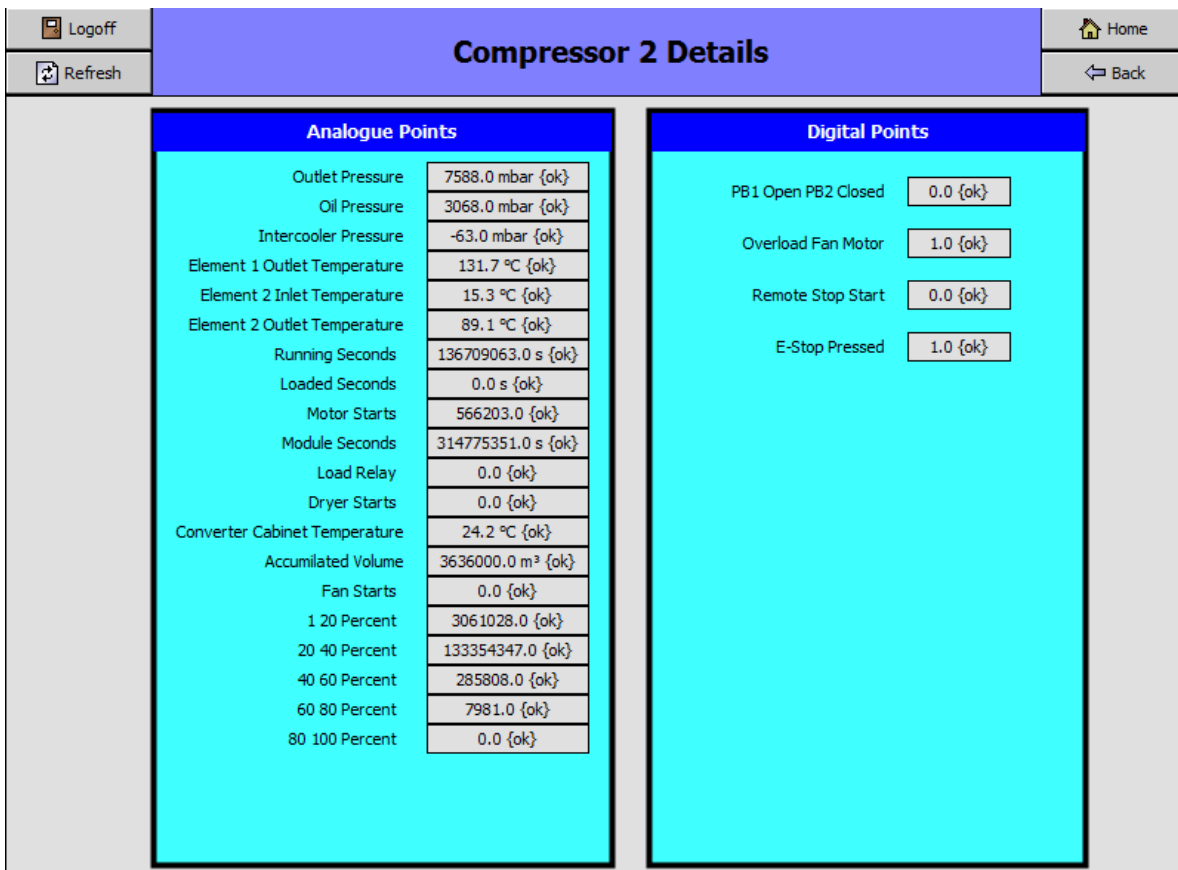
Fig. 16 Screenshot of compressed air monitoring system

It is clear that in the case of rule 14, the rule fires frequently. This suggests either that the compressed air is picking up heat in excess to the heat of compression, or that a sensor is faulty. Given that the rule does not fire continuously, a sensor fault may be ruled out. Therefore the compressed air is picking up excess heat during compression, likely from friction of moving parts close to the compression chamber. This could be verified by inspection of the machine by a service engineer. The GUI developed and shown for Rule 14 may be used for continual analysis of the compressor performance to allow for intelligent recommendations as to when the compressor is operating sub-optimally.

6 Conclusions and future work

A rule based expert system has been created for fault detection of an industrial compressed air system. The expert system employs k-means clustering for mode identification and improved energy efficiency monitoring. Implementation has been achieved using MATLAB, with results demonstrated using a custom-built GUI.

The methodology was able to successfully detect faults on a static data set. The threshold rules were developed heuristically from experience. Further refinement as more data be-

comes available would be advantageous for having greater confidence in these values.

It was found that the necessary data for implementation of the methodology was readily available using the Modbus capabilities of the air compressors. The test case air compressors are of a widely used make and are reasonably old. It can be expected therefore that this data will be readily available if the methodology were to be rolled out on a wider basis across industry.

The fact that the test site air compressors are oversized meant that their full operational range was not exercised. This limited the effectiveness of the K-means clustering methodology for mode identification. In practice it would be desirable to force the compressed air system to operate over a wide capacity range to improve the methodology's mode identification capabilities.

Currently each of the 15 rules has two or more potential diagnoses. Further refinement of the existing rule set of 15 rules would allow for definitive diagnosis of the root cause of faults. This refinement could take the form of subdivision of the existing rule set, to have a unique rule for each diagnosis.

Rules 1-6 did not detect any faults in the test data set. While this may be the case in reality, as more data becomes
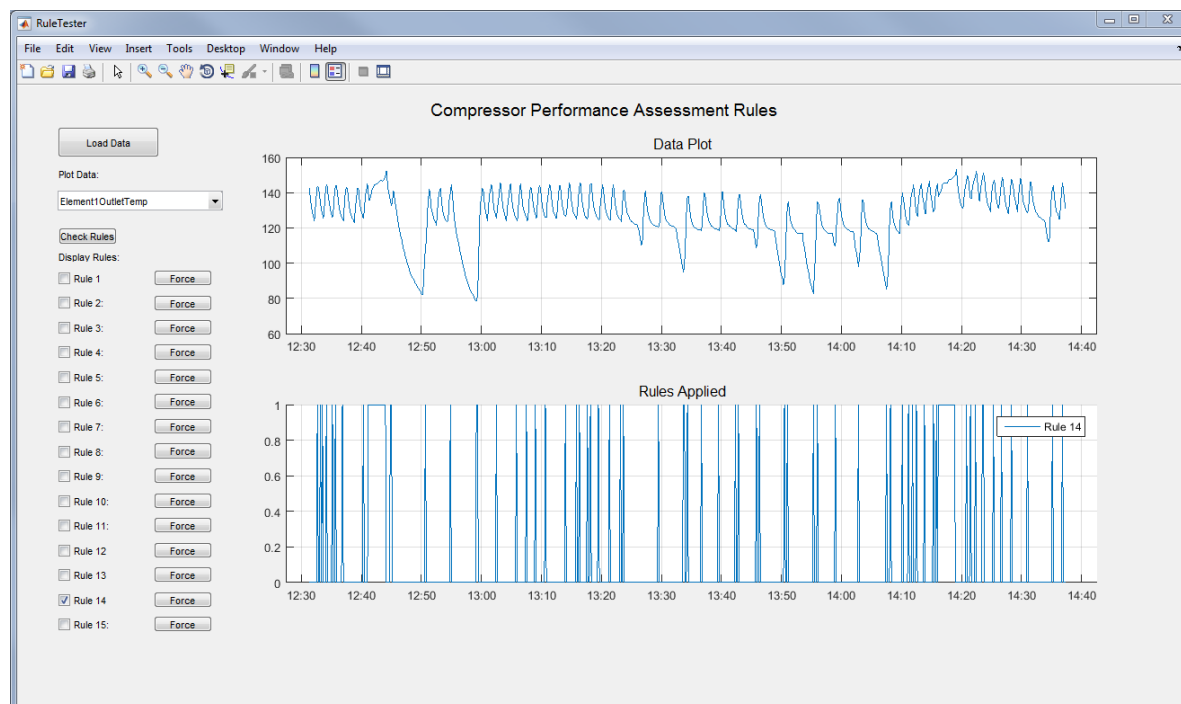
Fig. 17 Screenshot of compressor performance assessment GUI

available and manual inspections of the compressed air system are carried out it may be desirable to modify the error thresholds for these rules if actual faults are present in the system. The main goal of any refinement of the existing rule set is to reduce the presence of false positives or false negatives.

It is planned to automate the process of data extraction from the system Modbus network, as currently Excel sheets are required to be downloaded manually and added to the tool's MySQL database. As more data is added to the system, an analysis of the energy efficiency performance will be carried out using real-time updates to the operational mode cluster centroids. To ensure that this work is scalable across multiple facilities a best-practice approach will be taken, applying the methodology of an industrial big data pipeline (ODonovan et al (2015a)). Quantification of the financial savings achieved by implementing recommendations made by the expert system will also be carried out, taking a cost-benefit approach toward analysis (Walsh et al (2013)).

References

Backi CJ, Gravdahl JT, Grøtli EI (2013) Nonlinear observer design for a Greitzer compressor model pp 1457–1463

Blum H (2014) The economic efficiency of energy-consuming equipment: a DEA approach. Energy Efficiency 8(2):281–298

Bruton K, Raftery P, Kennedy B, Keane MM, OSullivan DTJ (2013) Review of automated fault detection and diagnostic tools in air handling units. Energy Efficiency

Bruton K, Coakley D, Raftery P, Cusack DO, Keane MM, OSullivan DTJ (2014a) Comparative analysis of the AHU InFO fault detection and diagnostic expert tool for AHUs with APAR. Energy Efficiency

Bruton K, Raftery P, O'Donovan P, Aughney N, Keane MM, O'Sullivan D (2014b) Development and alpha testing of a cloud based automated fault detection and diagnosis tool for Air Handling Units. Automation in Construction 39:70–83

CAGI (2015) Performance Verification Data Sheets

Catterson V, Costello J (2013) Increasing the Adoption of Prognostic Systems for Health Management in the Power Industry. Chemical Engineering Transactions 33:271–276

Dapeng Niu, Aiping Shi, Yuqing Chang, Fuli Wang (2011) Modelling of multistage centrifugal compressor. In: Proceedings of 2011 International Conference on Computer Science and Network Technology, IEEE, vol 973, pp 1144–1147, DOI 10.1109/ICCSNT. 2011.6182163

Duffuaa SO, Raouf A (2015) Planning and Control of Maintenance Systems. Springer International Publishing, Cham

Eurostat (2014) Energy balance sheets 2011-2012. Publications Office of the European Union, Luxembourg

Facchinetti T, Benetti G, Della Vedova ML (2014) Modeling and real-time control of an industrial air multi-compressor system. In: Proceedings of the 9th IEEE International Symposium on Industrial Embedded Systems (SIES 2014), IEEE, pp 67–76, DOI 10.1109/SIES.2014.6871189

Fairest S (2014) Delivering efficiency with compressed air the fourth utility

Fleiter T, Gruber E, Eichhammer W, Worrell E (2012) The German energy audit program for firmsa cost-effective way to improve energy efficiency? Energy Efficiency 5(4):447–469

Friden H, Bergfors L, Bjork A, Mazharsolook E (2012) Energy and LCC Optimised Design of Compressed Air Systems: A Mixed Integer Optimisation Approach with General Applicability. 2012 UKSim 14th International Conference on Computer Modelling and Simulation (Lcc):491–496

Gao Z, Cecati C, Ding SX (2015a) A Survey of Fault Diagnosis and Fault-Tolerant Techniques Part I: Fault Diagnosis. Industrial Electronics, IEEE Transactions on PP(99):1

Gao Z, Cecati C, Ding SX (2015b) A Survey of Fault Diagnosis and Fault-Tolerant Techniques Part II: Fault Diagnosis with Knowledge-Based and Hybrid/Active Approaches. Industrial Electronics, IEEE Transactions on PP(99):1

House JM, Vaezi-Nejad H, Whitcomb JM (2001) An expert rule set for fault detection in air-handling units / Discussion. In: ASHRAE Transactions, American Society of Heating, Refrigeration and Air Conditioning Engineers, Inc., Atlanta, vol 107, p 858

IEA (2012) International Energy Agency

Jiang Z, Zhang J, Jin M, Ma BO (2013) An expert system based on multi-source signal integration for reciprocating compressor 12(5):266–279

Katipamula S, Brambley M (2005) Review Article: Methods for Fault Detection, Diagnostics, and Prognostics for Building SystemsA Review, Part I. HVAC&R Research 11(2):169–187

Kopanos GM, Xenos DP, Cicciotti M, Pistikopoulos EN, Thornhill NF (2015) Optimization of a network of compressors in parallel: Operational and maintenance planning The air separation plant case. Applied Energy 146:453–470

Krichel SV, Sawodny O (2011) Analysis and optimization of compressed air networks with model-based approaches. Pneumatica pp 334–341

Laurijssen J, Faaij A, Worrell E (2013) Benchmarking energy use in the paper industry: A benchmarking study on process unit level. Energy Efficiency 6(1):49–63

Liu C, Kong D, Fan Z, Yu Q, Cai M (2013) Large flow compressed air load forecasting based on Least Squares Support Vector Machine within the Bayesian evidence framework. IECON Proceedings (Industrial Electronics Conference) (2011):7886–7891

Liu SYSC (2001) An Efficient Expert System for Air Compressor Troubleshooting. Expert Systems 18(4):203–214

Mazid AM, Martin R (2008) Automation of compressor test procedure using advantech data acquisition module. In: 2008 10th International Conference on Control, Automation, Robotics and Vision, IEEE, December, pp 2266–2271, DOI 10.1109/ICARCV.2008.4795885

Mills E, Shamshoian G, Blazek M, Naughton P, Seese RS, Tschudi W, Sartor D (2007) The business case for energy management in high-tech industries. Energy Efficiency 1(1):5–20

Murray SN, Rocher B, O'Sullivan D (2012) Static Simulation: A sufficient modelling technique for retrofit analysis. Energy and Buildings 47:113–121

Murray SN, Walsh BP, Kelliher D, O'Sullivan D (2014) Multi-Variable Optimization of Thermal Energy Efficiency Retrofitting of Buildings using Static Modelling and Genetic Algorithms - A Case Study. Building and Environment

Nair AT, Radhakrishnan TK, Srinivasan K, Rominus Valsalam S (2011) Kalman Filter Based State Estimation of a Thermal Power Plant. 2011 International Conference on Process Automation, Control and Computing pp 1–5

ODonovan P, Leahy K, Bruton K, OSullivan DTJ (2015a) An industrial big data pipeline for data-driven analytics maintenance applications in large-scale smart manufacturing facilities. Journal of Big Data 2(1):25

ODonovan P, Leahy K, Bruton K, OSullivan DTJ (2015b) Big data in manufacturing: a systematic mapping study. Journal of Big Data 2(1):20

OSullivan D, Keane M, Kelliher D, Hitchcock R (2004) Improving building operation by tracking performance metrics throughout the building lifecycle (BLC). Energy and Buildings 36(11):1075–1090

Pathak N, Khan M, Roy N (2015) Acoustic based appliance state identifications for fine-grained energy analytics. IEEE International Conference on Pervasive Computing and Communications (PerCom) pp 63–70

Petković M, Rapaić MR, Jeličić ZD, Pisano A (2012) On-line adaptive clustering for process monitoring and fault detection. Expert Systems with Applications 39(11):10,226–10,235

Project Haystack (2016) Project Haystack

Radgen P (2006) Efficiency through compressed air energy audits. In: Energy Audit Conference, www. audit06. fi

Ren Y, Zhang L, Ye Y, Liang W, Yang H (2012) Reliability Assessment of Anti-surge Control System in Centrifugal Compressor. 2012 Fourth International Conference on Computational and Information Sciences pp 1240–1243

Ruilin Lin, Boyun Liu, Qi Liu (2010) Study of the non-liner dynamic system theory for reciprocating compressor fault diagnosis. In: 2010 International Conference on Computer Application and System Modeling (ICCASM 2010), IEEE, Iccasm, pp V9–245–V9–248, DOI 10.1109/ICCASM.2010.5623041

Saidur R, Rahim N, Hasanuzzaman M (2010) A review on compressed-air energy use and energy savings. Renewable and Sustainable Energy Reviews 14(4):1135–1153

SEAI (2007a) Compressed Air Technical Guide. Tech. Rep. 1, SEAI, Dublin

SEAI (2007b) Special Working Group HVAC (SPIN I) 2007. Tech. Rep. Spin I

SEAI (2016) Compressed Air

Tran VT, Althobiani F, Ball A (2014) An approach to fault diagnosis of reciprocating compressor valves using Teager-Kaiser energy operator and deep belief networks. Expert Systems with Applications 41(9):4113–4122

Vachtsevanos G, Lewis F, Roemer M, Hess A, Wu B (2006) Intelligent Fault Diagnosis and Prognosis for Engineering Systems. John Wiley & Sons, Inc., Hoboken, NJ, USA

Venkatasubramanian V (2003) A review of process fault detection and diagnosis: Part III: Process history based methods. Computers & chemical . . . 27:293–311

Venkatasubramanian V, Rengaswamy R, Kavuri SN (2003a) A review of process fault detection and diagnosis: Part II: Qualitative models and search strategies. Computers & Chemical Engineering 27(3):313–326

Venkatasubramanian V, Rengaswamy R, Yin K, Kavuri SN (2003b) A review of process fault detection and diagnosis. Computers & Chemical Engineering 27(3):293–311

Walsh BP, Murray SN, O'Sullivan D (2013) Free-cooling thermal energy storage using phase change materials in an evaporative cooling system. Applied Thermal Engineering 59(1-2):618–626

Wang F, Song L, Zhang L, Li H (2010) Fault Diagnosis for Reciprocating Air Compressor Valve Using P-V Indicator Diagram and SVM. 2010 Third International Symposium on Information Science and Engineering (109047):255–258

Wang L (2008) Energy efficiency and management in food processing facilities. CRC Press

Wang L (2014) Energy efficiency technologies for sustainable food processing. Energy Efficiency 7(5):791–810

Wang Y, Brown Ma (2013) Policy drivers for improving electricity end-use efficiency in the USA: an economicengineering analysis. Energy Efficiency 7(3):517–546

Zhu Xp, Zhang LB, Liang W, Shi Gn (2013) A quantitative comprehensive safety evaluation method for centrifugal compressors using FMEA-fuzzy operations. In: 2013 2nd International Symposium on Instrumentation and Measurement, Sensor Network and

Automation (IMSNA), IEEE, pp 202–206, DOI 10.1109/IMSNA.
2013.6743251