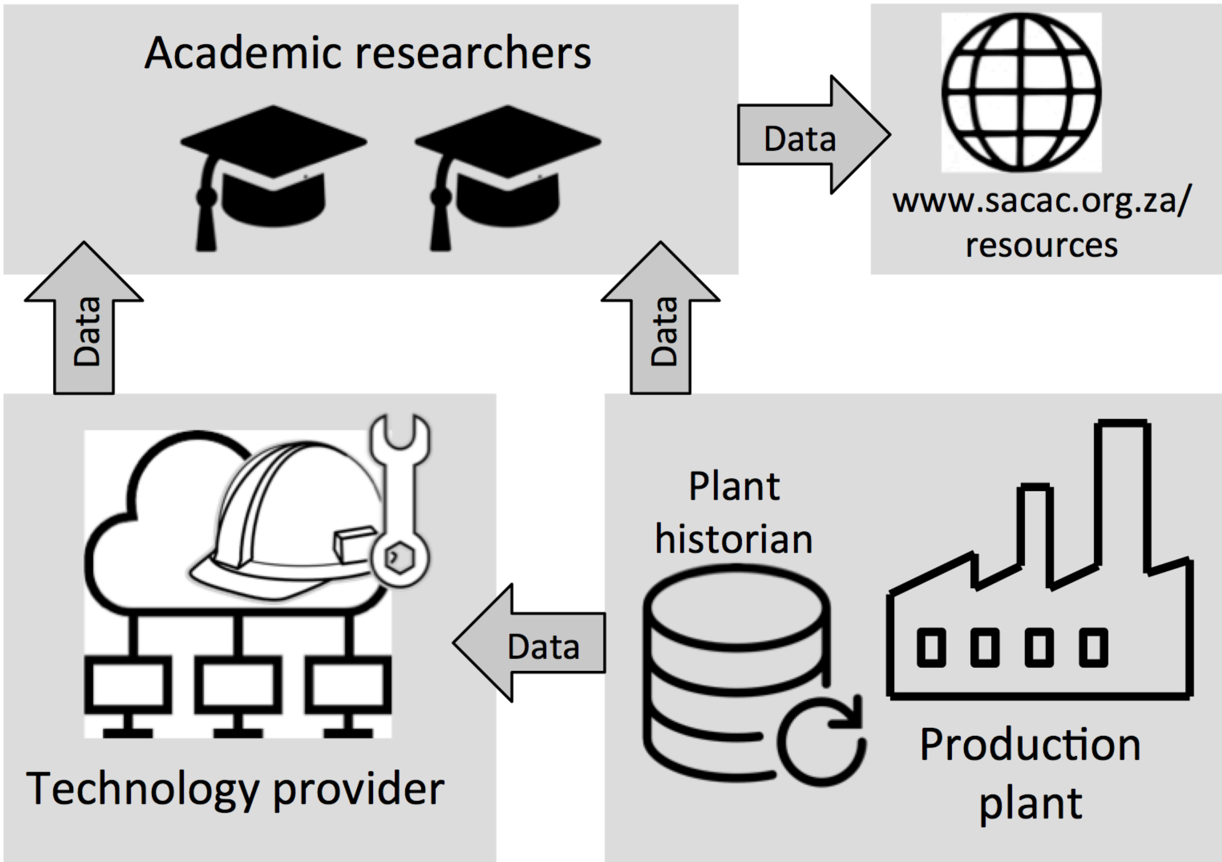
Academic researchers
Data
www.sacac.org.za/
resources
Data
Data
Plant
historian
Data
Technology provider
Production
plant
TOC Graphic

Industrial PID control loop data repository and comparison of fault detection methods

*Margret Bauer**

Department of Electrical, Electronic and Computer Engineering, University of Pretoria, Lynwood Road, 0002 Pretoria, South Africa (Email: margret.bauer@up.ac.za)

Lidia Auret

Department of Process Engineering, Stellenbosch University, Private Bag X1, Matieland, 7602, South Africa.

Riccardo Bacci di Capaci

Department of Civil and Industrial Engineering, Università di Pisa, Pisa, Italy

Alexander Horch

HIMA Paul Hildebrandt GmbH, Albert-Bassermann-Str. 28, 68782 Brühl, Germany

Nina F. Thornhill

Department of Chemical Engineering, Imperial College London, South Kensington Campus, London SW7 2AZ, UK

ABSTRACT: This paper presents the control loop data of industrial controllers that are recently made available online. All data is verified and some of it has been published previously to develop fault detection and diagnosis methods. Methods to detect faults that occur during the operation of an industrial process are important to improve profitability and have attracted attention previously. However, these methods are not always widely used in industry. One of the reasons is that any method needs to be robust and fully automated. The purpose of the data repository is to present data to test methods so that false positives and negatives are reduced to an insignificant number. Three previously published methods – oscillation detection based on the autocorrelation function, the idle index and a method for quantization detection – together with a simple, novel saturation detection method and one new detection methods are applied to all industrial data. The results are discussed and ways to improve the robustness and automation potential of these methods.

Keywords: Process control, PID, industrial processes, data, oscillation detection, idle index, actuator saturation, quantization.

1. INTRODUCTION

Fault detection and diagnosis is an important aspect of continuous improvement during the operation of industrial processes. Methods to detect various faults have been developed over the last three decades. The majority of these methods focuses on common problems that occur in industry in single input single output PID (proportional-integral-derivative) control loops and are often grouped under control loop performance monitoring (CPM). Many excellent overviews have been written over the last two decades[1, 2].

Despite the advances in the literature the problem of inadequate control performance still persists in many industries and loops continue to perform poorly[2]. Good progress has been made in specific areas such as valve stiction where numerous methods have been published[3,4].

However, the common approach is to focus on a particular type of fault and then develop a method that can detect the fault using simulated or industrial data.

However, there are still shortcomings when applying these established methods systematically and to all loops in a plant[5]. The challenges include avoiding false positives and false negatives, that is, detecting and diagnosing all faults reliably and not detecting faults where there are none. Another challenge is that often process knowledge is required for the different methods such as expert knowledge of the process dynamics and controller tuning parameters. Furthermore, there may be tunable parameters in the methods themselves. The aim is to automatically analyze all data and correctly identify the root cause of a poorly performing control loop.

Industrial data is critical when developing algorithms that can reliably identify faults. Furthering the area will help industrial production companies, control technology vendors and academic researchers alike. Once the problem of poorly performing loops has been solved, vendors can develop and sell reliable tools, production companies can improve their productivity and academic researchers can finally move on to the next topic.

Process data is usually confidential and industrial processing companies do not share the data. The academic contributors of this data repository have at some point overcome large organizational hurdles to gain access to the data with the permission to publish the data. And even then, only some data published in printed form can be shared in the form of the actual data. Sharing this kind of data usually requires a long standing trusting relationship between the academics and the industrial partners. Sharing data is to some extent culturally influenced. Countries with highly bureaucratic processes and hierarchies are usually less likely to share data with academic researchers or consultants. The confidentiality of data is also dependent on the industry. For example, minerals processing companies operate processes that vary little around the world. In addition, the core business of minerals and mining companies is centered around access to minerals resources and to a lesser extent on processing the materials once it is above ground. It is also important to note that mere access to data does not necessarily provide insight into the data. Two of the authors of this paper were given permission to publish normalized data of an entire process over a two months period. However, there was no access to process expertise because of organizational changes so that suspected root causes could not be confirmed. Thus, the plant data could be used as reference only and was effectively of no value.

In an industrial setting, every fault detection method is applied to any kind of data. When developing fault detection methods less attention is paid to false positives. **Figure 1** shows how false positives can occur when a fault detection method is applied to a different fault or normal operating conditions. False negatives that do not detect a fault present in the data are usually studied in detail.

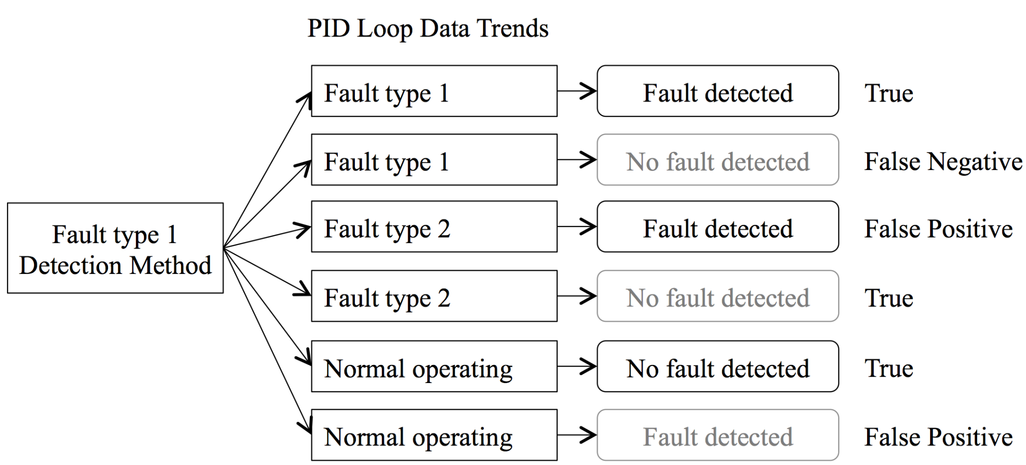

Figure 1: Fault detection method applied to all data of an industrial process.

In this paper, industrial data that has been made available online is presented, see www.sacac.org.za/resources. The concept of a repository was first described in[6] and industrial data has been added since. The authors encourage and would appreciate contribution of process data to grow the repository.

The paper is structured as follows. Section 2 describes 25 data sets of PID loops that have been gathered at the time of writing. The data is sorted into fault categories, which are described in detail. Section 3 describes four methods that address common problems in fault detection of control loop data. These methods are oscillation detection, sluggish tuning detection, actuator saturation detection and quantization detection. In Section 4, each of these methods is applied to all 25 data sets. Results are discussed in terms of robustness and the potential of complete automation of CPM.

The purpose of this paper is not to conduct an exhaustive study of these methods for all data sets, but instead to encourage researchers to conduct similar studies and to contribute their data to the repository so that methods can be successfully compared and introduced for successful industrial applications.

2. DESCRIPTION OF DATA SETS

The repository comprises industrial data that has been recorded during the operation of industrial processes in a number of industries such as chemicals and refining. The key focus of the repository is on recorded SISO PID loop data, which consists of time trends for the measurement - process variable (PV) - the setpoint (SP) and the controller output (OP). **Figure 2** shows a closed loop PID system indicating the measurements. Note that the OP data is recorded at the output of the controller and the PV after the sensor.

Great care has been taken to select each data set to ensure the relevancy. For confirmed root causes the data has been published previously the industrial partner has confirmed them independently.

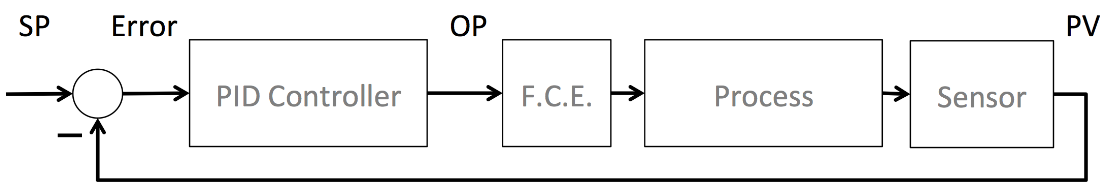

Figure 2: Control loop indicating the captured data trends – setpoint (SP), controller error (Error), controller output (OP) and process variable (PV). Note that the controller output is recorded before the final control element (F.C.E.) and is therefore equal to the input to the process. The process variable is measured by the sensor, which may distort the actual reading.

2.1. Data file description

The file format in which the data is contained is comma-separated values (CSV). This format was chosen because it is provided independently and can be easily imported into any common data analysis tool. The columns contain the time trend data, which are listed in **Table 1**. The first column contains a time vector, which is either a counter or the time/date information contained in the format YYYY-MM-DD hh:mm:ss. The following columns contain setpoint, process variable and controller output.

Table 1: Columns of .csv file for each SISO PID control

1 Date / Time	Time	The time is either available in a date/time format or in a counter from 1 to N.
2 Setpoint	SP	The setpoint, it may not be available (NaN) if the data is from an indicator, for example, for the problem of quantization.
3 Process variable	PV	The process variable in original or normalized form.

4 Controller Output	OP	The controller output should be scale to lie between 0 and 100. This may not be the case if the OP data is also normalized.
5 Error (optional)	Error	The controller error (SP-PV) cannot always be calculated from SP and PV for normalized data because these may be scaled differently. The controller error is available for some of this data.

2.2. Confidentiality and normalizing

The reason why industrial data is not commonly shared is that process companies are proprietorial about their processes and fear that confidential information can be extracted from it. In fact, data confidentiality is a key issue when it comes to new technologies such as cloud solutions and companies are very aware of cyber security. Here, this issue does not arise because the data is taken out of context and cannot be linked to a particular process. To further obscure any information, some companies prefer normalizing the data, that is, subtracting the mean and scaling the data to unit variance. This way, the data trend is still useful to identify the root cause. Caution has to be taken when normalizing the SP and PV trends individually. The control error – indicated as 'Error' in **Figure 2** – cannot be calculated by subtracting the PV from the SP. This is why some of the data contains another column labelled 'error' which contains the normalized control error.

2.3. Fault categories

The data sets of the PID loops are sorted firstly according to their fault categories, that is, there are folders for each fault category. These categories capture the most commonly found faults according to a survey conducted in 2016[7] and include:

- Incorrect tuning settings;
- Valve stiction;
- Other actuator faults;

- Sensor faults;
- Saturation;
- Other;
- Unknown.

A descriptive .txt file in the same folder of the database accompanies each CSV file containing a PID data set. The description contains information about the measurement and background of the fault. **Table 2** lists this information. It has been adapted from the initial publication and includes some minor adjustment from[6].

The data sets that were collected for this publication in alphabetical order – with a first emphasis on the fault categories – are listed in **Table 3** with a running index from 1 to 25. In the future, more data will be added to the repository. The time trends of these data sets are plotted in Fig. 2 with the first column showing the process variable, the second column the setpoint and the third column the controller output. Because of the different time dynamics and data lengths some of the trends appear differently in Fig. 2. For example, Data set 13 is in fact a fast oscillation but appears as random noise in the condensed version of Fig. 2. This is because the data length is larger than other sets (8641 samples) and it is sampled at 5 seconds whereas others are sampled at 1 second. Obviously, the process dynamics also affect the period of oscillation.

In the following the data trends that are comprised in the repository in each fault category are briefly discussed.

Table 2: Information contained in the .txt files associated with each data set

Information	**Values and examples**
Associated filename	e.g. unknown-F-chemicals-Thornhill-2003.csv. The name contains the fault category, measurement type, industry, contributor and year of origin.

Information	**Values and examples**
Type of measurement (PV)	Flow, Temperature, Pressure, Level or Other, Quality
Industry	Chemical, oil & gas including petrochemicals, minerals processing, paper, power generation, metals processing, manufacturing, food and beverages, other
Data length	Number of samples available
Sampling rate	The sampling rate may not be available from the time trend which can be in samples, not date/time format
Company	The company which has provided the industrial data. If no information is put forward the default value is anonymous.
Normalization	Yes/No, depending on whether the data has been scaled to zero mean and standard deviation.
Contributor	Name, surname and affiliation if available.
Year of origin	The year when the data was captured. This may not always be available and may be replaced by the year of publication.
Appears in publication	Authors, year, title, journal, volume, pages, if the data has been used in a publication previously

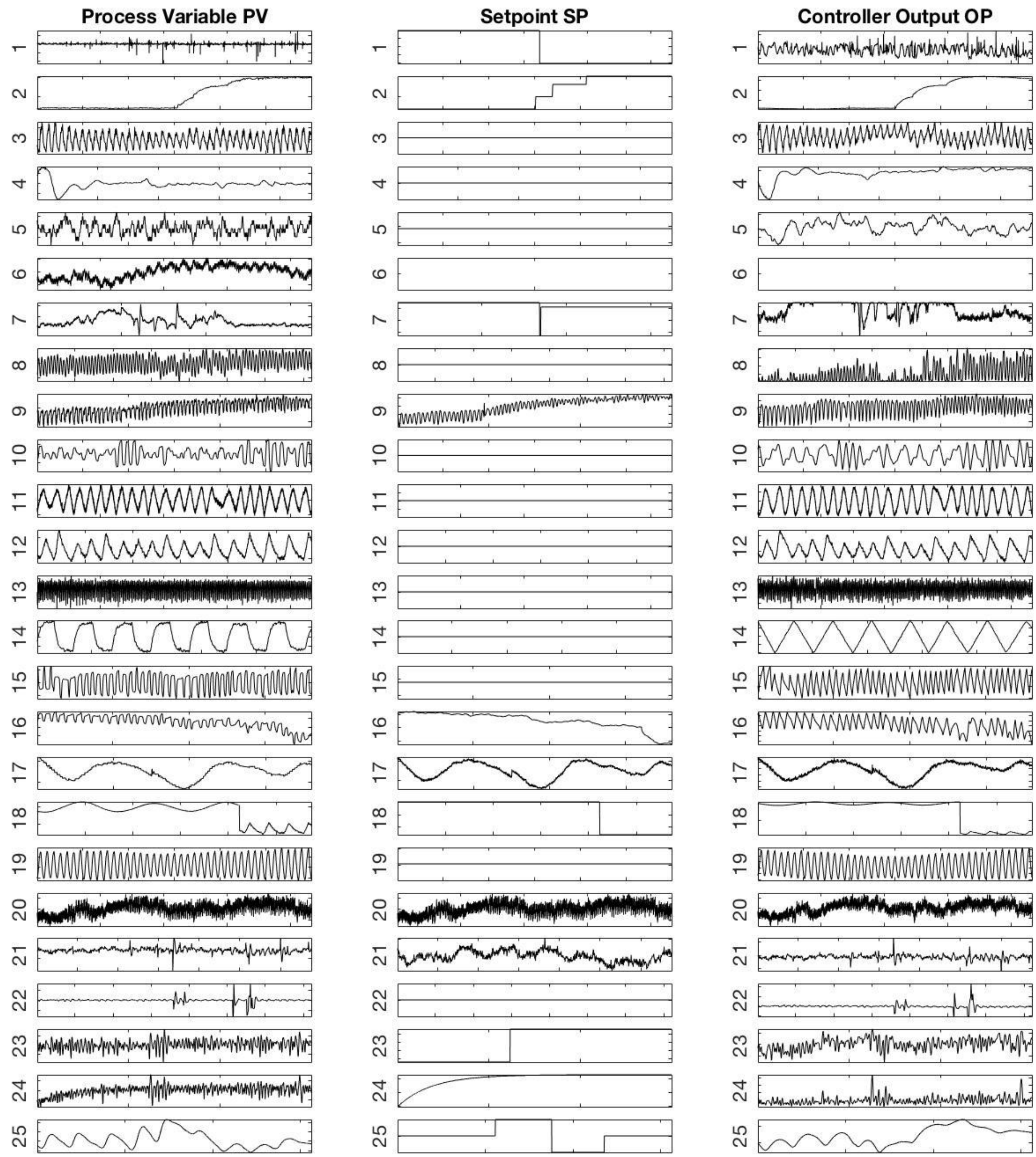

Figure 3: Time trends contained in repository data listed in **Table 3**.

Table 3: Data trends in the data repository at time of publications.

#	Fault-MT-industry-author-year	Length	Sampl. Rate	Norm.	Ref.
1	other-F-chemicals-thornhill-2003	8640	20	Yes	[9]
2	other-F-paper-horch-2003-2	1198	1	No	[8]
3	other-F-paper-horch-2003	1196	1	No	[8]
4	other-L-paper-horch-2003	904	2	No	[8]
5	quantisation-Q-paper-horch-2003	1196	1	Yes	[8]
6	quantisation-T-chemicals-Thornhill-2003	8640	20	Yes	[9]
7	saturation-L-minerals-bauer-2017	3241	10	Yes	-
8	saturation-T-oilgas-Thornhill-2002	1441	20	Yes	[10]
9	sensor-F-oilgas-thornhill-2007	1500	20	Yes	[11]
10	stiction-F-paper-horch-2003	1196	1	No	[8]
11	stiction-L-chemicals-thornhill-2003	8640	20	Yes	[9]
12	stiction-L-paper-horch-2003	1147	1	No	[8]
13	stiction-L-power-baccidicapaci-2018	8641	5	No	-
14	stiction-P-chemical-baccidicapaci-2018	1000	30	Yes	-
15	stiction-P-oilgas-baccidicapaci-2018	721	12	No	-
16	stiction-P-oilgas-DB-1-baccidicapaci-2018	721	12	No	-
17	tuning-F-chemical-DB-1-baccidicapaci-2018	7201	1	No	-
18	tuning-L-paper-horch-2003	1147	1	No	[8]
19	tuning-Q-paper-horch-2003.csv	1196	1	No	[8]
20	unknown-F-chemicals-thornhill-2003.csv	8640	20	Yes	[9]
21	unknown-F-paper-horch-2003.csv	1352	1	No	[8]
22	unknown-L-oilgas-thornhill-2002.csv	1441	20	Yes	[10]
23	unknown-P-oilgas-Thornhill-2007-1.csv	1500	20	Yes	[11]

#	Fault-MT-industry-author-year	Length	Sampl. Rate	Norm.	Ref.
24	unknown-P-oilgas-Thornhill-2007-2.csv	1500	20	Yes	[11]
25	unknown-T-oilgas-brooks-2018.csv	1561	15	No	-

2.3.1. Tuning

The recent industry survey[7] found that tuning problems are the most common cause for poorly performing loops. Wrong tuning settings can either cause an oscillation if the tuning is too tight, or of sluggish/slow behavior.

Three PID loop data trends are available (trends 17-19). Loop 17 by Bacci di Capaci is a cascade flow loop in a chemical plant, which has sluggish tuning settings. There are in fact twelve trends available for this particular loop in the repository. Each time trend was captured over a different period of time. The behavior of this loop was consistently sluggish.

Loops 18 and 19 were both contributed by Horch and are taken from a paper plant. In this case, the tuning settings were known to be too tight causing oscillations. A setpoint change occurs for level Loop 18, which changes the shape of the oscillation. The number of samples and sampling time of the two loops are the same but the oscillation in Loop 19 is much faster because of the different process dynamics.

2.3.2. Stiction

Valve stiction is extremely common in the process industries and is the cause of much consternation as it causes oscillations, which often propagate through the process. Many plant-wide disturbances have a sticky valve as a root cause. The detection of valve stiction has been studied in detail. The repository contains seven PID loops that were affected by valves stiction, namely Loops 10-16. The time trends in Fig. 2 show that the shape of the oscillation can vary significantly. The only common feature is that the oscillation is not sinusoidal but has higher frequency components altering the shape. Loop 16 is a cascade-controlled loop with a varying setpoint, highlighting the need to consider the controller error and not the process variable for stiction detection.

2.3.3. Sensor faults

Sensor faults are difficult to identify from visual inspection only. This is because the most common sensor faults are either drift – manifesting itself over a long period of time, bias, which cannot be seen in the normalized data, or complete failure.

There is one example of a sensor fault in the data repository that is of different nature. Loop 9 shows an oscillatory behavior. As the contributor of the data, N.F. Thornhill describes: 'It is known that there was a faulty steam sensor. It was an orifice plate flow meter but there was no weep-hole in the plate which had the effect that condensate collected on the upstream side until it reached a critical level, and the accumulated liquid would then periodically clear itself by siphoning through the orifice. The challenge for the analysis of this unit is to verify that the faulty steam flow loop is the root cause of the disturbance.' Because of the rareness of confirmed, interesting and observable sensor faults the authors of this paper are particularly grateful for this contribution.

2.3.4. Actuator saturation

The physical limitations of the actuator can cause the system to behave in a nonlinear way, causing oscillations. All physical actuators, pumps and valves alike, can potentially become saturated if the system design is inappropriate, a large external disturbance enters the process or the process is operated away from the designed specifications. Anti-windup schemes can deal with saturation but are not always employed in PID loops. Thus, saturation is a very common occurrence in the process industries and cited as one of the top three problems.

There are two time trends showing saturation of the actuator, Loops 7 and 8. The saturation is clearly visible in the controller output of the loop as the actuator is saturated and ‘sticks’ at its upper or lower limit for extended periods of time.

2.3.5. Quantization

The resolution of some sensors may not be sufficient or incorrectly adjusted resulting in measurements that only fall onto a very coarse grid. All readings are separated by a multiple of the quantization level. Quantization can be regarded as a type of a sensor fault where the sensor is scaled incorrectly. Temperature measurements are most commonly affected by quantization because an increased resolution is significantly more costly and not always required. There are two trends showing quantized data, Loops 5 and 6. Loop 6 is in fact not a PID loop but an indicator, which in this case is sufficient.

Quantization is easily artificially introduced by slotting the measurements on a specified grid. However, real industrial data is not as clear-cut and artefacts often occur when storing and compressing the data.

2.3.6. Other

There are many things that can go wrong when operating a process and faults can occur in the process as well as anywhere else in the control loop. The actuator is the moving part and the most common problems are valve stiction and saturation but there are other actuator issues that can arise. There are four loops that show other faults, Loop 1 to 4. Each will be described briefly here.

Loop 1 shows a non-period transient excursion that in fact caused a plant-wide upset. It originated from the injection steam into a distillation column. Problems with utilities such as steam supply are moderately common and particularly cumbersome because the origin can be difficult to trace as many measurements will be affected by the shared supply. Loop 2 has a similar problem though here the disturbance is of oscillatory nature. This shows that it can be difficult to discern between different root causes that all manifest in oscillatory form.

Loop 3 and 4 are in fact normal operating data. The authors felt that it is important to include this data in the repository as a baseline for detection and tuning of algorithms. Loop 3 includes setpoint changes that can have an impact on the analysis – some algorithms may pick up a disturbance or fault.

2.3.7. Unknown

All data that shows variations in the time trend has an unknown fault until the root cause is confirmed. Often, there are one or several hypotheses for the cause of a disturbance. The unknown category of faults is very important because here data can be stored that shows interesting features which were never identified. This is critical when testing detection algorithms because in the authors' experience industrial process data can take surprising shapes and sizes, causing false positive. Some disturbances may be industry specific and expected while other are complete surprises even to the process experts. There are six unknown disturbances stored in Loops 20 to 25, which have different behaviors from short, disruptive bursts to an oscillation in Loop 25.

2.3.8. Plant-wide data

The focus of this contribution is on PID single-input single-output data because many algorithms address particular problems. The data repository is therefore structured according to fault categories. Because industrial process are connected through the piping, control algorithms and shared utilities, disturbances travel from one measurement to others and the same features in the time trend can be seen in all data trends. Because a number and potentially all measurements in a process can be affected, these disturbances are referred to as plant-wide. Significant research effort has gone into finding the root cause of plant-wide disturbances and the contributors to this repository have worked on these disturbances, publishing the related data. The root cause loops have been extracted and added to the PID data repository.

The authors of this paper and custodians of the repository felt it would be valuable to include the entire data set so that researchers have access to this data. These complete data sets are added in the folder 'plant-wide' and an informal description with reference to relevant publications by the contributors are included.

3. DETECTION TECHNIQUES APPLIED TO ALL DATA SETS

Some faults have more distinct features and are therefore easier to detect than others. One example is valve stiction, which manifests itself as a nonlinear oscillation with distinctive shapes, particularly if the manipulated variable is available in addition to the controller output. There are numerous approaches to detect stiction, each with their own advantage and disadvantage. A literature search on valve stiction at the date of writing gave 1,300 results. A comparative study has been done in [12], on 93 industrial data sets comparing eleven different methods. The aim of the data repository is to encourage similar comparative studies but on all different types of data sets.

This contribution – supported by the data repository – takes a different approach than focusing on one single fault type. The control engineer and possibly an external service provider are confronted with hundreds if not thousands of single PID loops; many of them will be performing poorly. A 'scan' as described in[5] is applied to all loops in the plant and those with detected problems will be highlighted and are candidates for further investigation. It is therefore necessary to apply all methods to all data with any type of fault signatures.

The detection methods selected here address some of the most common problems in the process industry according to the study conducted in[7]. The purpose of this study is not to conduct an exhaustive comparison of methods, but to give an idea what kind of problems can be encountered. The methods selected are:

- Oscillation detection using the autocorrelation function[13];
- Idle index to detect sluggish tuned loops[14];
- Quantization detection[15];
- A newly proposed, simple saturation index.

The first two methods are established methods that address very common problems and have been widely adopted in industry and by fellow researchers. The second two methods address fault categories that are very common but sometimes thought to be too simple to detect.

The aim here is to indicate and discuss what can go wrong when e.g. a method to detect poor tuning is applied to time trend that shows the signature of valve stiction. The methods selected can potentially be fully automated and parameter settings are not optimized for specific data sets. Instead, standard parameters – if existing – suggested in the referenced publication are used.

The authors would like to encourage readers to apply their methods to the data given in the data repository, contribute their own test data, optimize parameter settings and conduct comparisons. This way, truly fully fault detection will be achievable in the near future.

3.1. Oscillation detection

The approach described in[13] for oscillation detection with zero crossings of auto covariance functions (ACF) is used in this work for detecting oscillatory controllers. An iterative approach for detecting multiple oscillations of different frequencies in a signal is described. The requirements for the oscillation index include the following:

- Data is available for continuous time periods;
- Raw PV values are used;
- Frequency bands of interest should be distinguished through pre-filtering if multiple oscillations affect the process variable.

The workflow for the calculation of the oscillation index takes the following form:

- Step 1: Using the spectral density of the time trend, the presence of multiple oscillations is determined for each data set from raw PV values.

- Step 2: For the cases where multiple, distinct peaks are observed in the spectrum (suggesting the presence of multiple oscillations), band pass filters are applied around each specific peak to generate filtered PV values; one filtered set for each clear peak frequency in the spectrum.

- Step 3: The autocorrelation function is calculated for the raw PV values and filtered PV values (where applicable). The length of the zero lag vector and lagged vector is set as a quarter of the total number of samples.

- Step 4: The intervals between zero crossings of the autocorrelation functions are determined, ignoring the first interval from zero-lag to the first crossing. If fewer than four intervals are present, then an oscillation index cannot be calculated, since derived statistics become unreliable.

- Step 5: The oscillation index r is calculated from the distribution of the intervals between zero crossings:

$$r = \frac{1}{3} \times \frac{\bar{T}_p}{\sigma_{T_p}}$$

Where $\bar{T}_p$ is the mean period (twice the mean value of the interval lengths) and σ_{T_p} is the standard deviation of the period (twice the standard deviation of the interval lengths).

For $r > 1$, the interpretation that regular oscillations are present in the PV.

3.2. Sluggish tuning detection (idle index)

The idle index proposed by[14] and further developed by[16] is used in this work for detecting sluggish controller tuning. The requirements for the idle index calculation include the following aspects:

- Data should be selected for continuous time periods where the set point was constant and load disturbances excited the controller;
- Noise-filtered PV and OP values are available;
- The sign of the controller gain parameter is known.

For cases the sample data of the repository where the sign of the controller gain parameter was not known it was assumed to be positive. The requirement of a known controller gain is a challenge to the automated application of the idle index to large industrial monitoring applications but feasible, if the analysis is done by a process and automation expert.

The workflow for the idle index calculation in this work follows the suggested workflow of[16]:

- Step 1: Filtering the PV and OP values with wavelet de-noising, with the same parameters and wavelet type as used by;

- Step 2: Detecting periods of steady-state data by means of[17] algorithm with a threshold based on a Type I error rate of 5%, and replacing PV and OP values for steady-state periods with their respective means during said periods;
- Step 3: Quantization of the filtered PV and OP values (for further noise removal, as suggested in[16] where the quantization increment is 5% of the range of the PV or OP values.

The idle index is calculated as follows:

$$I_i = \frac{t_{pos} - t_{neg}}{t_{pos} + t_{neg}}$$

where t_{pos} and t_{neg} represent periods when correlations between PV and OP value increments are positive and negative, calculated as follows updated every sampling instant, with h the sampling period:

$$t_{pos} = \begin{cases} t_{pos} + h & if\ \Delta OP \Delta PV > 0 \\ t_{pos} & if\ \Delta OP \Delta PV \leq 0 \end{cases}$$

$$t_{neg} = \begin{cases} t_{neg} + h & if\ \Delta OP \Delta PV < 0 \\ t_{neg} & if\ \Delta OP \Delta PV \geq 0 \end{cases}$$

For positive controller gains, the idle index can be interpreted as follows[14]: an idle index close to 1 suggests sluggish control; an idle index close to -1 suggests either well-tuned or oscillatory control; an idle index close to 0 suggests reasonably good tuning. Typically, an oscillation index would be used in conjunction with the idle index[14].

3.3. Quantization detection

Quantization in process measurements usually arises because of poor resolution of the sensor. It most commonly affects temperature measurements and sensors. In old process data historians, quantization was introduced through inadequate compression algorithms. The poor resolution of quantization can be seen even in the condensed form of Loop 5 in **Figure 3**, PV.

The detection algorithm described in[15] uses a histogram to identify the quantization levels. In the histograms, all measurements are 'binned' onto the amplitude axis. If quantization is present then the histogram will show discrete peaks at regular amplitude levels. The oscillation index considers the regularity of these bins in a quantization index

$$q = \frac{\mu_{\Delta x}}{\sigma_{\Delta x}}$$

where $\mu_{\Delta x}$ and $\sigma_{\Delta x}$ are the mean and standard deviation of the detected intervals between two peaks. A 3σ threshold is proposed in[15], that is, values of q larger than 3 are considered to be quantized signals. Quantization is comparatively easy to detect as it continuously present and does not depend on process dynamics etc. The main parameters that affect the detection are the number of bins of the histogram, the threshold setting for the peak detection to compute Δx and the detection threshold for q. Here, the number of bins was set to $N_{bins} = 100$, the threshold above which a peak Δx was detected was set to length of the time series as given in **Table 3** divided by the number of bins of the histograms (L/N_{bins}). If one or no peaks exceeded the threshold then the quantization index was set to zero. The detection threshold for q was set to 3 as recommended.

3.4. Saturation detection

Saturation or wind-up happens when the process dynamics require more control action than the actuator can deliver, as described in Section 2.3.4. Saturation can be clearly observed from the time trends of the controller output in **Figure 3**. The controller output of Loop 7 is often at – but obviously never above – the upper limit while the controller output of Loop 8 is often at the lower limit. Saturation can be so easily detected in a time trend that it is not considered to be worth developing a detection method worth reporting on. Industrial tools, however, do have features that detect saturation because actuator saturation is so common. The saturation index here is the percentage of measurement points that lie close to the upper and lower limit. Because the limits are unknown, the maximum and minimum value of the data series are considered

$$X_{min} = \min\{[x_1 \quad \cdots \quad x_L]\}$$

$$X_{max} = \max\{[x_1 \quad \cdots \quad x_L]\}$$

All measurements that are close to the evaluated period minimum and maximum are investigated where 'close' is defined to be within a band of width Δ_{sat} of the extrema. The band is set to a factor ϵ of the operating range of the selected time period of length L:

$$\Delta_{sat} = \epsilon(X_{max} - X_{min})$$

If the measurement stays for extended periods of time in this band then saturation is detected in sat_n. All extended periods are added together and divided by the total length L:

$$I_{sat} = \frac{\sum sat_n}{L}$$

The index is therefore scaled between 0 and 1 where 0 is no saturation and 1 indicates a constant value. Extended periods of time here are defined as $l = 10$ samples and the factor ϵ is set half a percent or 0.005.

4. RESULTS AND DISCUSSION

The four detection methods – oscillation, idle index, quantization and saturation – were applied to all 25 data sets currently collected in the data repository. The results are presented in **Table 4** and are discussed in this section for each index.

It is important to note that only simple methods were applied using standard settings. This was done deliberately to test the algorithms for their potential to automate them and applied them to plant data without using any expert knowledge.

One exception is made for the saturation detection index, which is very simple and was proposed in Section 3.4. Parameter l is optimized using the data in the data repository to demonstrate the usefulness of applying a detection method to all different types of process measurements. This procedure highlights the potential use of the data repository.

4.1. Oscillation detection

The oscillation detection r indicates that an oscillation is present if r is larger than one. Loops 3, 6, 8 to 19 and 23 to 25 all showed oscillations, caused by different faults. Oscillation detection is often the first step of fault diagnosis to establish the type of disturbance and if found to be oscillatory to determine the oscillation period. The oscillation index detected nearly all oscillations reliably. There were three exceptions where the oscillation was not detected. Firstly, the oscillation of Loop 24 had a changing setpoint, ramping up, during the data capture. The oscillation of Loop 24 can be easily detected when using the controller error instead, that is, setpoint minus process variable. The oscillation was correctly identified when using the controller error. The second instance was Loop 17 where the oscillation was less than two cycles in length. In order to detect an oscillation, more than three cycles are required. In the third instance, the oscillation of Loop 6 was masked by an external disturbance. Loop 6 is in fact only a measurement and not a PID loop, that is, only the PV exists. The method of oscillation detection could be improved with filtering prior to the analysis as suggested in[14] to also detect this oscillation correctly.

4.2. Sluggish tuning detection (Idle index)

The idle index as described in Section 3.1 addresses the difficult problem of detecting sluggish tuning of control loops. The results of applying the index to the data repository are presented in **Table 4**. The index is interpreted as follows: a value of one points to sluggish loops, -1 are well tuned or oscillatory loops and a value of 0 means reasonably good tuning.

The difficulty with the index is that the sign of the process gain needs to be obtained because it affects the sign of the idle index. The gain is generally unknown though companies may keep a record of it in the automation system from where it can be retrieved. For this reason, all gains were assumed to be positive as this is the more common setting. The difficulty is that getting the sign wrong changes interpretation completely: from a well-tuned controller to a sluggish controller. As a result, the results in **Table 4** are completely wrong. There is one (known) sluggishly tuned loop, Loop 17. It should be excluded from the analysis because it has a setpoint change. Loops 18 and 19 have poor tuning settings but are tightly tuned and therefore show oscillations. The design of the index range (-1 being well-tuned or oscillatory; 0 being tuned well; 1 being idle control) makes it difficult to interpret as well. Applied to the data repository, the idle index is prone to false positives and/or false negatives purely on its designed index range (-1 representing both good tuning and oscillatory behavior).

To summarize, the idle index gives confusing results. This does not mean that the idle index is not useful; rather that it requires expert knowledge for interpretation and cannot be automated as it is.

4.3. Quantization detection

The repository includes two Loops that are labeled as quantized, namely Loops 5 and 6. **Table 4** shows that indeed high values for the quantization, namely 52,083 and 1,581, which are both above the detection threshold of $q_{thresh} = 3$. There are another nine data sets that were detected to be quantized, Loops 8, 9, 12, 13, 15-17, 19 and 24. The ones with the highest number are 19 and 24 for which the histograms are plotted in **Figure 4**. The plots reveal that Loop 24 is indeed a quantized signal, which was not identified as such, while Loop 19 is not. The histogram of Loop 19 looks somewhat regular but there the bins between the regular spikes are filled. This can happen if the trend is deterministic to some degree because information is lost when computing the histograms. This is true for Loop 19, which is nearly sinusoidal. The recommendation for the quantization algorithm is to significantly increase the threshold from q_{thresh} = 3 to 300.

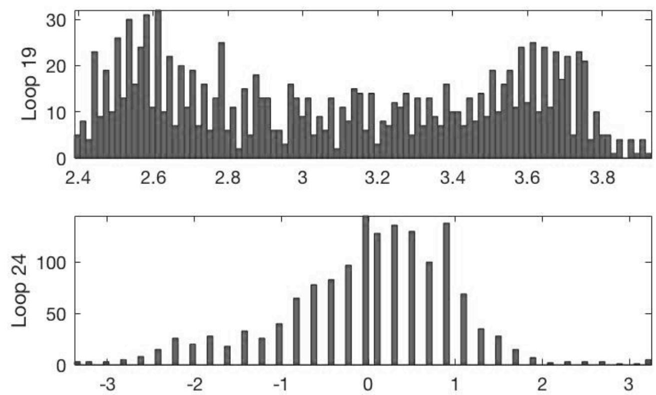

Figure 4. Histogram of data trends with high quantization detection values (Loop 19: q = 105; Loop 24: q = 845) that were not classified as quantized.

4.4.Saturation detection

The simple saturation detection algorithm has been proposed in this paper and no guidance for the threshold was given. The data trend in the repository will be used to fine tune and improve the method and highlight problems in the detection.

There were two Loops that were classified to be saturated, Loops 7 and 8. The results in **Table 4** show that a value for saturation was detected for five loops: 2, 7, 8, 18 and 25. In order to understand why the method detected some of the loops incorrectly, the controller output is plotted close up for all five loops in **Figure 5**. In Loops 2, 18 and 25, a setpoint change occurs during the data capture and the slow oscillation appears as saturation as the scaling factor changes using the maximum and minimum of the controller output. This is particularly prominent in these loops as they have a slow oscillation.

As a result, time periods when a setpoint change occurs should be excluded from the analysis because it will change the operating regime of the controller.

Loop 7 has a significantly higher saturation index than Loop 8. The reason is that Loop 8 has an oscillation so that the controller output moves to and away from the saturation. The parameter for saturation length *l* should be optimized. This is done using all controller output time trends that did not show saturation and the saturation time trends of Loops 7 and 8. The results are shown in **Figure 6**, exemplary for Loop 1 and Loops 7 and 8. Spurious saturation detection values were encountered for short lengths (in Loop 1 for values of *l* smaller than 4) because each trend is at its minima and maxima at some point during the data capture. The optimal length that avoids false negatives and false positives alike was found to be *l*=6.

The results or the improved method that excludes setpoint changes and sets *l* to 6 is given in the right hand column of **Table 4**.

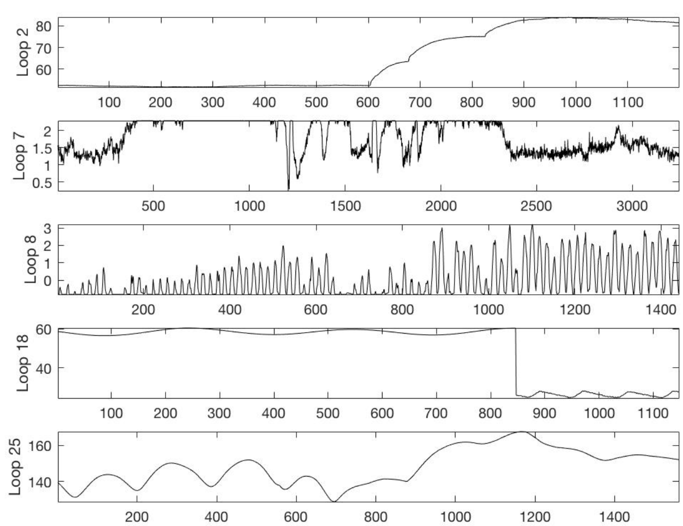

Figure 5. Controller output signals of loops that picked up saturation in the saturation detection method.

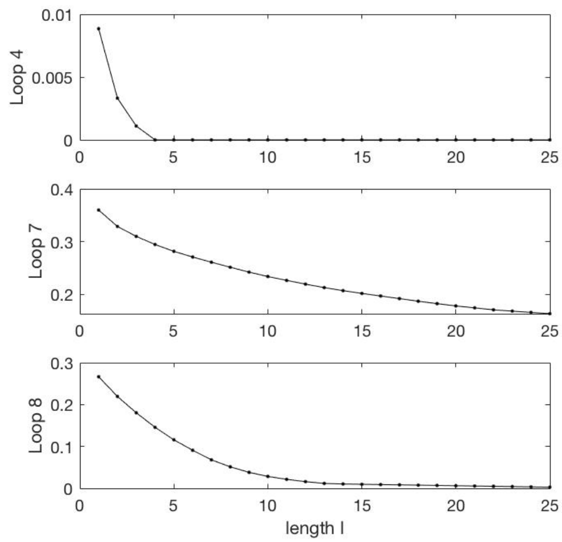

Figure 6. Saturation detection index as a function of the saturation detection length *l*.

Table 4: Fault detection indices for different data sets and different techniques. Bold numbers indicate that there was a fault to be detected. Red indicates incorrect results (bold red: false negative, normal red: false positive). *SP varying too much for assessment. **Optimized with *l*=6 and excluding setpoint changes.

#	**Fault**	**Oscillation index** r	**Idle index** I_i	**Quantization** q	**Saturation** I_{sat}	**Saturation opt.**** I_{sat}
1	Other	0.18	-0.20	0	0	0
2	Other	0.15	1.00	0	0.09	0
3	Other	**10.04**	-0.99	0	0	0
4	Other	0.73	-1.00	1	0	0
5	Quantization	0.55	0.76	**52,083**	0	0
6	Quantization	**NaN**	n.a.	**1,581**	0	0
7	Saturation	0.29	0.88	1	**0.23**	**0.27**
8	Saturation	**5.22**	0.97	5	**0.03**	**0.09**

#	**Fault**	**Oscillation index** *r*	**Idle index** I_i	**Quantization** *q*	**Saturation** I_{sat}	**Saturation opt.**** I_{sat}
9	Sensor	**1.32**	0.43*	4	0	0
10	Stiction	**1.43**	-0.26	2	0	0
11	Stiction	**2.37**	0.90	1	0	0
12	Stiction	**5.03**	1.00	7	0	0
13	Stiction	**3.66**	0.83	17	0	0
14	Stiction	**24.29**	-0.85	1	0	0
15	Stiction	**2.89**	-0.69	36	0	0
16	Stiction	**1.38**	-0.75*	27	0	0
17	Tuning	**NaN**	**-0.89***	64	0	0
18	Tuning	**2.52**	0.88	0	0.04	0
19	Tuning	**9.94**	1.00	105	0	0
20	Unknown	0.10	0.98*	0	0	0
21	Unknown	0.21	-0.70*	0	0	0
22	Unknown	0.84	0.92	0	0	0
23	Unknown	**1.14**	0.38	1	0	0
24	Unknown	**0.71**	0.19*	845	0	0
25	Unknown	**1.32**	0.70	0	0.01	0

5. CONCLUSIONS

In this contribution, industrial data of a new data repository was presented. The purpose of the repository is to facilitate critical assessment of existing fault detection methods. Many methods are perceived as mature in the academic environment but have not found their way into industry.

The reason is often robustness and resulting false positives, that is, detecting a fault where there is none (or a different kind of fault). Four detection methods were applied to the industrial data.

The methods were intentionally kept as simple as possible with no filtering or parameter optimization to improve the results. A new saturation detection index was proposed for which one parameter was optimized using all 25 data sets and to reliably detecting saturation in industrial data. The oscillation detection index proved to be robust and can be easily automated while the idle index requires expert knowledge that was not available and without gives very poor results.

The work presented here highlights the need for robust methods that are tested on a large variety of industrial data. The authors would like to encourage the use of the data repository to develop fault detection methods that are robust to false positives. Additional industrial PID data can be contributed by email to: piddata@sacac.org.za or directly to the corresponding author of this paper.

The authors thank the editors for the opportunity to prepare a paper for Industrial & Engineering Chemistry Research for the Sirish Shah Festschrift, and we offer Sirish the most sincere best wishes for his retirement

ACKNOWLEDGMENT

The authors would like to thank the SACAC, the South African Council of Automation and Control, which represent IFAC for the covering the cost of implementing and hosting the data repository under www.sacac.org.za.

REFERENCES

(1) Huang, B.; Shah, S. L. *Performance assessment of control loops: theory and applications*. Springer Science & Business Media: London, 1999.

(2) Jelali, M. *Control performance management in industrial automation: assessment, diagnosis and improvement of control loop performance*. Springer Science & Business Media: London, 2012.

(3) Choudhury, M. S.; Shah, S. L., Thornhill; N. F., & Shook D. S. Automatic detection and quantification of stiction in control valves. *Contr. Eng. Prac.* **2006**, 14, 1395-1412.

(4) Choudhury, M. A. A. S.; Shah, S. L.; Thornhill, N. F. *Diagnosis of process nonlinearities and valve stiction: data driven approaches*. Springer Science & Business Media: London, 2008.

(5) Starr, K. D.; Petersen, H.; Bauer, M. Control loop performance monitoring–ABB's experience over two decades. *IFAC-PapersOnLine* **2016**, 49, 526-532.

(6) Bauer, M.; Auret, L.; Le Roux, D.; Aharonson, V. An Industrial PID Data Repository for Control Loop Performance Monitoring (CPM). *IFAC-PapersOnLine* **2018**, 51, 823-828.

(7) Bauer, M.; Horch, A.; Xie, L.; Jelali, M.; Thornhill, N. The current state of control loop performance monitoring–A survey of application in industry. *J. Proc. Contr.* **2016**, 38, 1-10.

(8) Horch, A. Benchmarking control loops with oscillations and stiction. In *Process control performance assessment*; Ordys, A.W.; Uduehi D.; Johnson M.A. Eds.; Springer: London, 2007; pp 227-257.

(9) Thornhill, N.F.; Cox, J.W.; Paulonis, M. Diagnosis of plant-wide oscillation through data-driven analysis and process understanding. *Contr. Eng. Pract.* **2003**, 11, 1481-1490.

(10) Thornhill, N.F.; Shah, S.L.; Huang, B.; Vishnubhotla, A. Spectral principal component analysis of dynamic process data. *Contr. Eng. Pract.* **2002**, 10, 833-846.

(11) Thornhill, N.F., 2007, Locating the source of a disturbance, In *Process control performance assessment*; Ordys, A.W.; Uduehi D.; Johnson M.A. Eds.; Springer: London, 2007; pp 199-225.

(12) Jelali, M.; Scali, C. Comparative study of valve-stiction-detection methods. In *Detection and Diagnosis of Stiction in Control Loops*; Jelali, M.; Huang, B., Eds.; Springer: London, 2010; pp 295-358.

(13) Thornhill, N.F.; Huang, B.; Zhang, L. Detection of multiple oscillations in control loops. *J. Proc. Contr.* **2003**, 13, 91-100.

(14) Hägglund, T. Automatic detection of sluggish control loops. *Contr. Eng. Pract.* **1999**, 7, 1505-1511.

(15) Bauer, M.; Madolo, S. Detection and effect of quantisation in data-driven process analysis. Proceedings of the 9th International Symposium on Dynamics and Control of Process Systems, Cancun, Mexico, June 6-10, 2007.

(16) Kuehl, P.; Horch, A. Detection of sluggish control loops – experiences and improvements. *Contr. Eng. Pract.* **2005**, 13, 1019-1025.

(17) Cao, S.; Rhinehart, R. R. An efficient method for on-line identification of steady state. *J. Proc. Contr.* **1995**, 5, 363-374.